OPTIMAL LENGTH OF MOVING AVERAGES

TO USE WHEN FORECASTING BASIS

By

ROBERT BRADEN HATCHETT

Bachelor of Science and Arts in Agribusiness

The University of Arkansas

Fayetteville, AR

2007

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 2009

OPTIMAL LENGTH OF MOVING AVERAGES

TO USE WHEN FORECASTING BASIS

Thesis Approved:

Dr. B. Wade Brorsen
Thesis Adviser

Dr. Kim Anderson

Dr. Brian Adam

Dr. A. Gordon Emslie
Dean of the Graduate College

ACKNOWLEDGEMENT

I would like to take the opportunity to thank all the people who have supported me in reaching my goals. I would not have made it this far in life without the help of several people who I would like to recognize.

First of all I would I like to thank to thank my family for their continued love and support. To my parents, John and Karen Hatchett, who always allowed me to find my own way and have supported my every decision. Without such strong role models, I would be lost. To my loving grandparents, who have always been so interested in my accomplishments and stressed the importance of continued learning. I truly believe that without knowing where I have come from, I would not be confident in where I am going. To my brothers, Mason and Philip, who taught me maturity and that anything worth having is worth fighting for. Without the support of my family none of this would have been possible.

I would also like to recognize the entire Agricultural Economics Department at Oklahoma State University for providing me with the strong analytical skills I was searching for. I would like to recognize my committee members Drs. Wade Brorsen, Kim Anderson and Brian Adam. To my adviser, Dr. Brorsen: thanks for your guidance and for challenging me throughout this process. It was great to learn so much about statistics and commodities in your class and throughout my thesis program. Dr. Anderson and Debbie Wells, thank you for working so hard to compile wheat cash price data for this project.

Thank you too, Dr. Anderson for your insight into wheat markets. Dr. Adam, thank you for helping bring me into this department, and for all of your support while I was here. Several graduate students have also provided support throughout my program. I would like to especially thank Samarth Shah and Kate Brooks for their friendship and help throughout my time here. All of these people were integral in my success.

Last, but certainly not least, I would like to thank Gracie Teague for all her help with formatting my thesis. Also, all of the supports staff in the department for their hard work in helping make the success of students possible at Oklahoma State University.

.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION....1

Background 1
Objective 2
Theoretical Model 2

II. THE THEORY OF THE COST OF STORAGE....7

Explanatory Basis Models 11
Basis Forecasting Studies 14

III. MODEL 19

Data 19
Procedures 23

IV. MODEL RESULTS 26

Pooled Model Results 26
#2 Hard Wheat Model Results 29
Hard Wheat Changes over Time 32
#1 Hard Wheat Model Results 34
Soft Wheat Model Results 37
Corn Model Results 39
Soybean Model Results 41
Model Effects from Recent Years 43

V. CONCLUSIONS 49

VI. RECOMMENDATIONS FOR FUTURE STUDY 52

REFERENCES....54

APPENDIX....58

LIST OF TABLES

Table Page

Table II-1. Results from Previous Basis Forecasting Studies 18

Table IV-1. Number of Locations with a Given Length of Moving Average Having the Lowest Root Mean Squared Forecast Error, 1975-2008........ 26

Table IV-2. Absolute Error (cents/bu.) of Basis Forecasts as a Function of Number of Years in the Moving Average, Pooled Data, 1975-2008...................... 28

Table IV-3. Absolute Error (cents/bu.) of Hard Wheat Basis Forecasts as a Function of Number of Years in the Moving Average, 1978-2008.......... 29

Table IV-4. Absolute Error (cents/bu.) of Kansas City Ordinary Protein, #1 Hard Wheat Basis Forecasts as a Function of Number of Years in the Moving Average, 1976-2008 35

Table IV-5. Absolute Error (cents/bu.) of Kansas City 13% Protein, #1 Hard Wheat Basis Forecasts as a Function of Number of Years in the Moving Average, 1976-2008 36

Table IV-6. Absolute Error (cents/bu.) of Soft Wheat Basis Forecasts as a Function of Number of Years in the Moving Average, 1975-2008.......... 38

Table IV-7. Absolute Error (cents/bu.) of Corn Basis Forecasts as a Function of Number of Years in the Moving Average, 1980-2008[a] 40

Table IV-8. Absolute Error (cents/bu.) of Soybean Basis Forecasts as a Function of Number of Years in the Moving Average, 1980-2008[a] 42

Table IV-9. Absolute Error (cents/bu.) of Basis Forecasts as a Function of Number of Years in the Moving Average, Pooled Data, 1975-2005...................... 45

Table IV-10. Absolute Error (cents/bu.) of Basis Forecasts as a Function of Number of Years in the Moving Average, Pooled Data, 2006-2008...................... 46

Table IV-11. Descriptive Statistics of the Absolute Forecast Error for the Pooled Data over Different Time Periods 47

Table Page

Appendix Table 1. Locations and Time Periods Studied by Commodity 58

Appendix Table 2. 1978-2008 Hard Wheat Preharvest Basis Forecasting Model Comparisons 60

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons 67

Appendix Table 4. 1976-2008 Soft Wheat Preharvest Basis Forecasting Model Comparisons 76

Appendix Table 5. 1975-2008 Soft Wheat Storage Basis Forecasting Model Comparisons 77

Appendix Table 6. 1980-2008 Corn Preharvest Basis Forecasting Model Comparisons 78

Appendix Table 7. 1980-2008 Corn Storage Basis Forecasting Model Comparisons 80

Appendix Table 8. 1980-2008 Soybean Preharvest Basis Forecasting Model Comparisons 82

Appendix Table 9. 1980-2008 Soybean Storage Basis Forecasting Model Comparisons 84

LIST OF FIGURES

Figure Page

Figure II-1. Kansas and Oklahoma Hard red winter wheat locations studied.............. 20

Figure III-1. Number of minimum MAE forecasts produced by the previous year's basis vs. the 5-year moving average, 1979-2008...................................... 27

Figure III-2. 1975-1980 average Oklahoma wheat harvest basis.................................. 32

Figure III-3. 2008 Oklahoma wheat harvest basis... 33

Figure III-4. 1982-1986 Average Kansas wheat harvest basis...................................... 33

Figure III-5. 2007 Kansas wheat harvest basis.. 34

ABSTRACT

Futures prices when combined with a basis forecast provide a reliable way to forecast cash prices. The most popular method of forecasting basis is historical moving averages. Given the recent failure of longer moving averages proposed by previous studies, this research reassesses past recommendations about the best length of moving average to use in forecasting basis. This research compares practical preharvest and storage period basis forecasts for hard wheat, soft wheat, corn and soybeans to identify the optimal amount of historical information to include in moving average forecasts. Only with preharvest hard wheat forecasts are the best moving averages longer than 3 years. The structural changes over the period studied lead to the recommendation of shorter moving averages than have been found previously. The differences in forecast accuracy among the different moving averages are small and in most cases the differences are not statistically significant. The recommendation is to use longer moving averages during time periods (or at locations) when there have been no structural changes and use last year's basis when it appears that a structural change has occurred.

Keywords: Basis forecast, grain, Law of One Price, moving averages, structural change

CHAPTER I

INTRODUCTION

Background

Creating reliable preharvest price expectations and making postharvest storage decisions depend heavily on accurate basis forecasts. Without accurate forecasts of basis levels "it is impossible to make fully informed decisions about…whether to accept or reject a given price; (and) whether and when to store your crop" (CBT, 1990, p.23).

The most popular method of forecasting the basis is historical moving averages. The attractiveness of these models is their ease of application. Access to local prices is cheap and readily available, allowing basis values to be localized for specific markets. Studies have applied forecasts of various lengths in order to determine the optimal length of years to include. These models generally conclude that longer averages ranging from 3 to 7 years are optimal (Dhuyvetter and Kastens, 1998; Sanders and Manfredo, 2006). The idea is that these longer moving averages can smooth out temporary deviations in markets.

In stable market conditions, the longer historical average forecasts proposed by previous studies should form the most accurate basis expectations. These methods have failed recently as basis values have deviated greatly from previous levels, resulting in poor forecasts based on historical basis. Given this recent failure, there is a need to reassess past recommendations about the best length of moving average to use in forecasting the basis.

Objective

The specific objective of this study is to determine which length of moving average has been most accurate in forecasting basis in terms of mean absolute error.

Theoretical Model

One of the primary reasons futures markets were created was to provide market participants the opportunity to exchange cash price risk for more manageable basis risk. Basis risk is preferred to price risk because price levels are more variable than basis levels. This price variability can be shown mathematically as

(1) $$\sigma^2_{price} > \sigma^2_{basis},$$

where σ^2_{price} is the variance of the cash market price and σ^2_{basis} is the variance of the basis. Basis forecasting seeks to reduce σ^2_{basis} by reducing forecast error (ε_t):

(2) $$\varepsilon_t = Basis_t - Ba\hat{s}is_t$$

where $Basis_t$ is the actual basis at time t, and $Ba\hat{s}is_t$ is basis forecast, and $\varepsilon_t \sim N(0, \sigma^2_{basis})$ assuming unbiased forecasts.

The most popular practical approach to forecasting basis is historical moving averages (FarmDoc, AgManager). Moving average models use the simple average of the previous N years:

(3) $$Ba\hat{s}is_t(N) = \frac{1}{N}\sum_{i=1}^{N} Basis_{t-i}.$$

By substituting (3) into (2) we can define how the optimal moving average length is selected to minimize basis forecast error

(4) $$\min_N E(\hat{e}_t^2) \;=\; \min_N E(Basis_t - \;\frac{1}{N}\sum_{i=1}^{N} \; Basis_{t-i})\,.$$

.
Rather than take the partial derivative of (4) with respect to N, this equation must be solved through enumeration due to the choice variable N being discrete. Once these individual forecasts are aggregated, the optimal forecast minimizes the error for the entire sample, T by

(5) $$\min_N \sum_{t=1}^{T}(\,Basis_t - \;\frac{1}{N}\sum_{i=1}^{N} \; Basis_{t-i})$$

The variance minimizing moving average length depends on the underlying stochastic process. Under normality and homoskedasticity the stochastic process for basis is

(6) $$Basis_t \sim N(\mu_t, \sigma^2)$$

where μ_t is the time varying mean and σ^2 is variance. The optimal moving average forecast length depends on μ_t.

Without structural change in the mean basis is $\mu_t = \mu$, and the longest moving average (largest N) would result in the minimum variance forecast. Basis forecast error variance in this case is

(7) $$\sigma_{forecast}^2 = \frac{\sigma^2}{N} + \sigma^2\,.$$

These two sources of error originate in equations (5) and (6), in the variance of the moving average forecast, and in the current basis variance. So long as $\mu_t = \mu$, then as

$N \to \infty, \frac{\sigma^2}{N} \to 0$, and the primary source of basis forecast error is σ^2. Therefore markets that are not prone to structural changes would find longer moving average forecasts optimal.

Structural changes within grain markets can change the dynamics of price relationships, and the resulting basis values. An extreme example of a stochastic process that could explain changes in markets is a random walk:

$$\mu_t = Basis_{t-1} \ . \tag{8}$$

An example of a random walk process would be a permanent increase in transportation costs, which would widen the basis. With a random walk, as (8) shows, the optimal forecast is with *N*=1.

A more general stochastic process that includes both the constant mean and random walk models as special cases is a variation in a normal jump process. Diffusion-jump processes that combine a normal and a Poisson jump process are popular processes for modeling stochastic volatility in equity, stock and options markets (Anderson et al. 2002; Chernov et al. 2003; Bates 1996). With this model, the mean is constant and then occasionally changes as

$$\mu_t = \mu_{t-1} + J_t P_t \tag{9}$$

where $J_t \sim N(\theta, \delta^2)$ and P_t is the jump process that is often assumed to follow a Poisson distribution. The difficulty in measuring this process is that the jump parameters and probability of the jumps occurring varies over time. Equation (9) could result in a random walk if $P_t = 1$ and $\delta^2 = 0$ in (6), and it gives a constant mean if $P_t = 0$. Ethanol plants are a major source of new demand in corn markets and cause the basis levels near the plant to strengthen. The structural change reflected by the jump affects prices initially, making the previous year's basis the optimal predictor for the year following the jump. The size of the shock in basis drastically changes the current basis levels so that all data before the change no longer reflect the current market. As the supply feeds the plant and markets adjust, bids will gradually decrease and the effects from the initial jump will result in a new mean and longer moving averages will then become optimal.

Mean-reverting models can also be used to model changes from historical basis levels (Jiang and Hayenga 1998; Sanders and Manfredo 2006). The basic mean-reverting model is the autoregressive moving average, or ARMA(p, q),

$$Basis_t = \alpha + \varepsilon_t + \sum_{i=1}^{p} \emptyset_i Basis_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{10}$$

where α is an intercept, $\emptyset_1 \ldots \emptyset_p$ are the autoregressive parameters, p is the number of autoregressive terms, q is the number of moving average terms, $\theta_1 \cdots \theta_q$ are the moving average parameters, and $\varepsilon_t \sim N(0, \sigma^2)$. If ($\emptyset_i = 1, p = 1$) and ($q = 0$) then it is a random walk, and if ($\emptyset_i = 1/p$) and ($q = 0$) then it is a simple moving average.

If the ARMA model in (10) is stationary, then the basis will converge toward its long-run mean of $\alpha / \sum_{i=1}^{p} \phi_i$. If the ARMA model is nonstationary (has a unit root) then the long-run mean will change over time. While Tomek and Wang (2007) argue that cash prices do not have unit roots, it is hard to argue that the mean of the basis is constant over time.

If plenty of observations are available, estimating an ARMA model should outperform the simple moving average of basis. But time series are often too short or structural changes are too frequent to estimate an ARMA model. Even if ARMA models could provide slightly more accurate forecasts, ARMA models may still not be preferred because of the difficulty in estimating and explaining them to producers.

ARMA (p, q) models, and another generalization, a seasonal autoregressive integrated moving average or SARIMA(p, d, q), have been used to forecast the basis (Sanders and Manfredo, 2006; Jiang and Hayenga, 1998). These studies found little improvement in forecast accuracy when compared to the moving average models. In order to identify the correct level of $\emptyset_i$, the appropriate covariance function of the process must be identified by the partial autocorrelation and autocorrelation plots. This econometric technique is too complicated for producers to understand, and is not modeled in this study for that reason. Instead, this research focuses on simple moving average forecasts, which are ARMA $(p, 0)$ processes where $\emptyset_i = 1/p$ and $\theta_i = 0$.

The optimal length of moving average to forecast the basis is expected to depend on the size and frequency of structural changes. When conditions are static, longer moving averages are optimal. However, after a structural change occurs, the optimal length of a moving average is one.

CHAPTER II

THE THEORY OF THE COST OF STORAGE

The first attempts to explain the difference between cash and futures market prices focused on the components of the futures market price not contained in the cash markets. In his explanation of inverse carrying charges Vance (1946) states that cash and futures prices, though related, are not equivalents. Although market prices are primarily formed in futures markets, cash prices differ, even at delivery, from these levels. This is the earliest explanation of a lack of convergence in delivery markets, due to the form differences in what the two prices reflect.

A narrow interpretation of this divergence in market prices by Working (1948) disagrees with Vance's position that price differences arise from differences in quality or location of the commodity or due to uncertainty as to time of delivery. Working believes that the true carrying charge reflects the difference between identical commodities, at the same location, separated only by differences in time of delivery. However, since the quality quoted in most cash wheat contracts exceeds those actually delivered on the futures contract, the basis usually reflects both time and quality differences. Since Working believes the true carrying charge does not include time and quality differences, these two components are in addition to the basis. Working believes that efficient arbitrage between cash and futures markets merge the two markets into one. He admits that while the two markets may differ due the differing expectations of traders in each market, the practice of basing cash quotations through the basis relationship of the futures markets makes the two inseparable. To treat the two markets as separate, according to Working, is to incorrectly imply a level of independence that does not actually exist.

Another difference in form between cash and futures proposed by Keynes in his "Treatise on Money" explains why negative carrying charges in cash and futures markets might occur through downward biased futures prices. If futures are indeed downward biased, then cash prices exceed the futures price by the risk premium paid by hedgers to speculators. This premium is paid by risk-averse hedgers, who participate in futures markets to transfer risk. If this is the case, then hedgers sell contracts below the expected futures price, and create a downward bias in the price levels. Since uncertainty is a decreasing function of the time to contract maturity, the risk premium is a form difference that diminishes with time.

Later tests of futures price bias and the existence of risk premiums met with mixed results. Kolb (1992) identified risk premia in livestock and lumber markets, but not in many other markets, while Telser (1958, 1960) found no risk premium in wheat and cotton markets. Cootner (1960a, 1960b) used the same data as Telser and found that risk premia did exist once the data were divided into pre- and postharvest months. The mixed results of these studies show how differences in data and model specification can lead to conflicting conclusions.

Keynes' explanation for the existence of a risk premium relies on the hedger's motivation to participate in futures markets to transfer their price risk. Unlike Keynes, Working believes that hedgers enter futures markets not solely to transfer risk, but to profit from changes in the relative cash and futures prices. Working argues that a hedge is arbitrage, through a double transaction in the futures market based on the relation of the cash and futures prices (1948). The effectiveness of the transaction is determined primarily by the first contract, along with the price difference between the first and second contracts (1953). This price difference, along with the initial futures price and the final cash price gives the final return from the hedged transaction. Working (1953) identifies four reasons for hedging in futures markets: “(1) (to) facilitate buying and selling decisions, (2) (to) give greater freedom for business action, (3) (to) give a reliable basis for conducting storage of commodity surpluses, and (4) (to) reduce business risk.” Although reducing price risk may be an effect of these actions, it is not a primary incentive to hedging.

Working (1949, 1953) does not include a risk premium in his carrying cost, but instead argues that basis reflects net carrying cost (including storage costs, insurance, opportunity costs, and a convenience yield). The physical cost of storage, insuring the grain and the opportunity cost are accepted components of the basis, measured by the quotations for commercial storage, insurance and short term interest rates, respectively. The final component of Working’s price of storage is Kaldor’s (1939) “convenience yield.” To Working, the value of the convenience yield helps explain why stockholders hold surplus inventories during times of backwardation, thus relieving the constraints of Keynes’ theory described by the risk premium. Rather than having to pay speculators to take on the price risk, hedgers may hold their surplus stocks beyond harvest in order to gain a return from holding stocks. This stock holding process is important in allocating inventories over time, and helps ensure that processors will have raw inputs available throughout the year.

Brennan (1958) develops the convenience yield as a necessary business cost incurred by producers, merchandisers and processors who store inventory in support of their primary business. These participants can remove stocks from storage in order to meet sudden and unexpected increases in demand resulting from day-to-day fluctuations in the market. Thus, the convenience yield lowers the cost of keeping regular customers satisfied and provides the advantage of capturing rising demand and prices without drastically changing production schedules.

The presence of risk premia and convenience yields in futures prices are two conflicting components used to explain price spreads below the full cost of carry. Empirical work has shown that both of these proposed components of the futures price are used to explain basis levels below the full cost of storage (Working, 1953; Cootner, 1960a). Each theory supports a difference in form that exists in the basis to explain the difference between empirical findings and the full carry of the market.

More recent explanations of the "storage at a loss phenomenon" cite mismeasurement as the source of storage at a loss, not a difference between empirical results and studies. Wright and Williams (1989) account for spatial and grade differences of stocks in the measurement of their supply-of-storage curve. Their results support the ability of greater precision in defining relevant prices and stocks to reduce the occurrence of holding stocks under backwardation. By studying locations within a spatially dispersed market, Benirschka and Binkley (1995) show that optimal storage for a firm depend upon the site's distance from the terminal market. Since transportation costs lower the realized price as distance increases from the terminal market, firms farther from the market experience lower opportunity costs and less pressure to liquidate stocks. By discouraging storage closer to the terminal markets, through higher opportunity costs, markets efficiently supply stocks from storage. Instead of incurring storage at a loss, the basis efficiently allocates storage and marketing over space and time. Wright and Williams (1989) and Benirschka and Binkley (1995) show that, when modeled with disaggregate data, the storage at a loss paradox disappears. Klumpp, Brorsen, and Anderson (2007) using local prices, however, find that storage at a loss does occur and so the mismeasurement hypothesis is not sufficient to explain the occurrence of holding stocks at returns below full carry.

Explanatory Basis Models

Several variables have been used to explain the basis. Most of these variables correspond to differences in time, form, and space, but the theoretical basis for some of these variables is not as clear. Differences in form are explained through components of the futures price not reflected in the cash market price. Cost of storage and transportation measures are accepted components of the basis from literature that explain the transformation of prices over time and space, but the theoretical support for supply and demand variables used to explain the basis over space is not as clear.

Naik and Leuthold (1991) empirically examined differences in form in the corn basis using components of cost of storage theory apart from storage costs. Evidence of a risk premium, a speculative component, and an expected basis level at maturity is tested on the underlying assumption of constant storage costs. According to the authors, if the absolute value of the correlation coefficient between cash and futures prices is one during maturity then no risk premium exists. The presence of a speculative component in the maturity basis is supported when, by regressing the cash price on the futures price during the maturity month, the resulting coefficient is 1. The third component, the expected basis level at maturity, was regressed using lagged basis, cash prices, lag export, and contract dummy variables. These three components of the basis are used to explain the basis apart from the physical storage costs, opportunity costs, and convenience yield.

Seasonality in the basis has been identified throughout the explanatory literature (Martin, Groenewegen, and Pigeon, 1980; Jiang and Hayenga, 1997). Since this seasonality has identified certain supply and demand variables as only being significant during certain periods, dummy variables have been used to indicate different periods within the marketing year (Martin et al., 1980; Jiang and Hayenga, 1997; Dykema, Klein, and Taylor, 2002). Monthly dummy variables (Martin, Groenewegen, and Pigeon, 1980), futures contract maturity variables (Jiang and Hayenga, 1997), and quarterly dummy variables (Dykema, Klein, and Taylor, 20002) are included in explanatory models.

Spatial differences are explained in different ways by explanatory models. Martin, Groenewegen, and Pigeon (1980) subtract transportation costs, tariffs, and loading fees from their explanatory model before estimation. Spatial differences between the various markets can also be measured in the model using barge rates (Jiang and Hayenga, 1997) and a seasonally adjusted producer price index for intermediate energy has also been used as a proxy for transportation costs (Dykema, Klein, and Taylor, 20002).

Differences in space are also measured using supply and demand variables at local markets. Supply variables for markets include crop production levels, a dummy variable for the presence of loan deficiency payments (LDP), the ratio of Eastern Canadian corn production to consumption, and Western feed grain availability (Dykema, Klein, and Taylor, 2002; Martin, Groenewegen, and Pigeon, 1980; Jiang and Hayenga, 1997). Soybean crushing levels, animal units consuming grain (corn), corn usage estimates, and export volumes were all used as demand variables to identify the differences in markets (Jiang and Hayenga, 1997, Dykema, Klein, and Taylor, 2002). These supply and demand variables represent proxy variables used to identify the factors that constitute the basis at a particular location.

Various attempts to explain the basis have identified several variables used to explain the basis. A risk premium and a speculative component existed in just over 50 percent of corn contracts, and lagged variables explained 49-63 percent of the maturity basis (Naik and Leuthold, 1991). By removing the spatial aspects of the basis, and studying only nearby futures prices for each month of the year, Martin, Groenewegen, and Pigeon (1980) were able to explain 66-82 percent of the basis residual through variables that reflect differences in form between Chicago futures markets and the cash prices at Chatham, Ontario. All three aspects of the Law of One Price are used to explain 50 to 80 percent of the corn and soybean basis (Jiang and Hayenga, 1997). The futures price, local supply and demand variables, the dummy variable to account for LDP, and seasonal dummies explained 75.7 percent of the South Dakota corn basis (Dykema, Klein, and Taylor, 2002).

From these explanatory models, we can see how a wide variety of variables are used to explain the basis. Some of these variables lacked any clear theoretical basis, but correspond to differences in the cash and futures price over time, form, and space. All of these variables correspond to aspects of the Law of One Price, and explaining the basis through time, form, and space supports these variables in accepted theory. Structural changes in grain markets affect the impact of these variables on the basis. These changes can be explained when they occur by the Law of One Price through changes in terms of time, form, or space.

Basis Forecasting Studies

Historical moving average models are the most popular method of forecasting the basis. The attractiveness of these models lies in their simplicity. No advanced modeling or econometric techniques are necessary, only historical basis values. Jiang and Hayenga (1997) compared more advanced time series techniques against the simple 3-year moving average. Although the advanced techniques were more accurate in most cases, the simple moving average was optimal for 51 percent of corn contracts, and 46 percent of the soybean models studied. These findings support the use of historical average forecasts in producing basis expectations.

Several studies have applied moving averages of various lengths to identify the most accurate method of forming basis expectations. Hauser et al. (1990) compared several naïve models in forming their soybean basis expectations for ten Illinois elevators. Models included: expected basis is current basis, expected basis is previous year's expiration basis, and the expected basis is the average of the previous 3 years' expiration basis. Optimal forecast methods differed over periods, but these simple models provided reliable forecasts.

Dhuyvetter and Kastens (1998) forecast nearby basis for wheat, corn, soybeans and milo for multiple Kansas locations. Some models included 1-7 years in the historical average, and some incorporated futures price spreads and a 3-year average with the current basis deviation from the 3-year average. Sanders and Manfredo (2006) tested models of varying complexity in forecasting basis within the soybean complex in Central Illinois. A 5-year moving average, previous year's basis, and the expected nearby basis is the ending basis models are compared against more advanced times series methods. Taylor, Dhuyvetter, and Kastens (2004) revisited Dhuyvetter and Kastens (1998), and included models to determine the optimal amount (weight) of current market information, i.e. the current basis deviation from the moving average, needed to improve forecast accuracy.

Forecast horizon is another important determinant of forecast accuracy. Dhuyvetter and Kastens (1998) forecast basis over 4 week increments from 4 weeks to 32 weeks before contract expiration, and the results indicated that the horizon to expiration dictated the optimal forecast method. Over the shorter horizons, models that included current market information outperformed historical average methods, but longer term forecasts did not benefit from the additional information. Taylor, Dhuyvetter, and Kastens (2004) forecast the harvest basis in 4 week increments from 4 to 32 weeks prior to harvest, and the nearby basis 24 weeks after harvest in 4 week increments up to 20 weeks prior to expiration. The benefit of the additional information varied within crops and over periods. When the additional current market information increased accuracy, the optimal amount to include increased as the forecast length shortened. The results of these studies show the influence of uncertainty over time on forecast accuracy. However, a lack of any clear pattern over time indicates that time is not the only determinant of optimal forecasts methods.

Ward and Dasse (1977) have shown that different factors determine the basis at delivery and nondelivery locations. If the basis is effective in pricing at nondelivery points, it should reflect the value of the commodity at the local market (Martin, Groenewegen, and Pigeon, 1980). Forecasting models study the impact of spatial differences on the basis over different locations. To better represent the U.S. corn and soybean markets, Jiang and Hayenga (1997) forecast the basis at both delivery and nondelivery locations. Spatially dispersed locations allow Dhuyvetter and Kastens (1998) and Taylor, Dhuyvetter, and Kastens (2004) to study delivery and nondelivery locations within Kansas. By including multiple locations in each study, these models can determine patterns in the accuracy of basis forecasting over separated markets.

Practical forecasting approaches to forecasting the basis use current market information to identify any additional accuracy through differences in basis form. Current basis deviations from historical levels are used to determine whether the current basis reflects any differences from historical levels (Dhuyvetter and Kastens, 1998; Taylor, Dhuyvetter and Kastens, 2004). If the difference in the current basis from the historical level for a particular location can increase forecast accuracy, then what the basis reflects has changed and the models can benefit from the additional market information.

Table II-1 lists the results from these forecasting studies. These results do not provide a clear pattern in what forecast performs the best. From the table we can see that practical forecasts perform comparably to more complex forecasts. The optimal amount of historical data included in the forecasts does not follow any rule of thumb. And the inclusion of current information is shown to increase forecast accuracy over short horizons, but its effectiveness diminishes greatly with time. No clear patterns in the amount nor kind of current information to consistently improve basis forecasts exists. These inconsistent findings reveal that no clear patterns exist in forming optimal forecasts.

Table II-1. Results from Previous Basis Forecasting Studies

Study	Optimal Forecasts	Conclusions
“Forecasting Crop Basis: Practical Alternatives” -Dhuyvetter and Kastens (1997)	• 4-year moving average for wheat. • 7-year moving average for corn. • 7-year moving average for soybeans. • 5-year moving average for milo.	Futures price spreads and current nearby basis increased accuracy, but futures price spreads were best. The benefit from incorporating current market information diminished beyond 4-12 weeks.
“Incorporating Current Information into Historical-Average-Based Forecasts to Improve Crop Price Basis Forecasts” – Taylor, Dhuyvetter, and Kastens (2004)	• 3-year moving average for wheat. • 2-year moving average for corn. • 3-year moving average for soybeans. • 2-year moving average for milo.	Futures price spreads and current basis deviations from historical levels helpful in post-harvest and harvest (only 4 weeks prior to harvest). As the post harvest horizon approached, the optimal amount of current market information increased.
“An Analysis of Anticipatory Short Hedging Using Predicted Harvest Basis” - Kenyon and Kingsley (1980)	• Regression equation using initial local cash and Chicago futures market prices, the Chicago cash price at planting, and the residual of open interest.	The regression estimates predicted 73-81% of the change in corn basis, and 95%-97% of the change in soybean basis as harvest approached using initial basis and the difference between actual and predicted open interest.
“Basis Expectations and Soybean Hedging Effective” – Hauser, Garcia, and Tumblin (1990)	• 1 or 3-year historical basis during preharvest. • Futures price spreads after the harvest.	Forecasts that include the implied return to storage outperform historical averages in 2 of the 3 contract periods. Historical average models perform comparably to models incorporating current market information.
“Corn and Soybean Basis Behavior and Forecasting: Fundamental and Alternative Approaches” - Jiang and Hayenga (1998)	• 3-year moving average plus current market information best for corn. • Seasonal ARIMA best for soybeans.	Although the 3-year moving average performs relatively well, it is out performed by models that include current market information and seasonal ARIMA models.
“Forecasting Basis Levels in the Soybean Complex: A Comparison of Time Series Methods” - Sanders and Manfredo (2006)	• ARMA model best for soybeans. • VAR model best for soybean meal. • Previous year’s basis best for soybean oil.	Over time, the accuracy of the 1 and 5-year moving averages do not diminish. Even within closely related markets there is no rule-of-thumb for producing the most accurate forecasts.

CHAPTER III

MODEL

Data

The commodities considered are corn, soybeans, soft wheat, and hard wheat. To create the basis data, futures prices must be subtracted from their corresponding cash price.

Two basis values are used for each year. One is selected to represent the basis for a preharvest hedge and the other for a storage hedge. For corn, the December contract in October represents the harvest basis, while the May contract in April represents storage hedges. For soybeans, the November contract in October represents the harvest basis, while the May contract in April represents storage hedges. The basis values used for soft and hard wheat are the July contract in June and the December contract in November.

Cash and futures prices consist of second Wednesday or Thursday prices for corn, soybeans and wheat, and when unavailable, monthly-average prices are used. Daily #2 corn and #1 soybean cash prices are from the Illinois Agricultural Marketing Service, and reflect the midrange of elevator bids for each region on the second Thursday of each month from 1975-2008 (FarmDoc, 2009). When the second Thursday fell on a holiday, the third Thursday was used. Second Wednesday daily Oklahoma reported prices paid to producers for #2 hard red winter wheat were taken from the Oklahoma Department of

Agriculture, Food and Forestry's weekly "Oklahoma Market Report" from 1974 through 2008. This report also provides the Galveston Gulf Port prices. When a holiday prevented the release of the report, the third Wednesday was used. Second Wednesday prices from an additional Oklahoma location, the Port of Catoosa, are for 1988-2008 (Peavey Grain, 1988-2008). Second Wednesday Kansas cash prices cover 1982-2007 (Dhuyvetter, 1982-2007). Simple average monthly wheat prices were taken from the USDA AMS "Grain and Feed Market News" for #2 soft red winter wheat at Chicago, IL, Toledo, OH, and St. Louis, MO, along with #1 hard red winter wheat at Kansas City, MO over 1970-2008. Figure 1 shows the Kansas and Oklahoma hard red winter wheat locations studied.

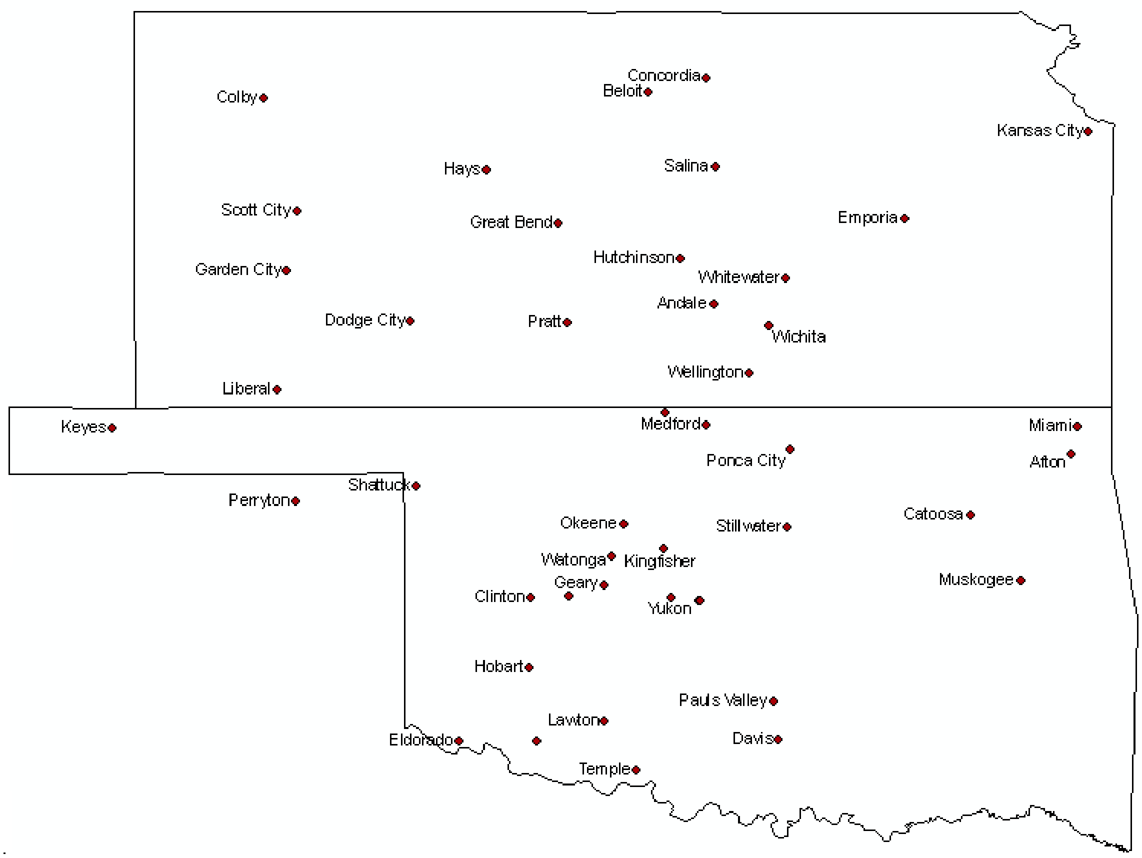

Figure III-1. Kansas and Oklahoma Hard red winter wheat locations studied

Futures prices reflect daily closing prices at the CBT and KCBT for each commodity (R & C Data), and match the same days as the cash prices. When only monthly cash prices were available, average monthly futures prices were used. Corn, soybeans, and soft wheat futures prices are reflected by CBT contracts, while KCBT wheat contracts reflect hard wheat. These futures prices, along with their corresponding cash prices, provide the nearby basis values used in this study.

In order to create accurate basis values based upon the available data, several stages of data cleansing were necessary. The Oklahoma Market Report is a weekly newsletter containing agricultural prices for Oklahoma producers, but did not previously exist in electronic format. Prices were compiled by location in a single spreadsheet, and were checked against the original reports to ensure accuracy. Numerous mistakes (20-50) in the original report were found and corrected, but none of these corrections were in the data used in this study. Missing bids accounted for approximately 0.3% (3 of 945 observations) of the Oklahoma time series, and were substituted with the third Wednesday prices.

The data series was checked to ensure that none of the days studied happened to fall on days when the futures price hit the daily limit. The earliest reported historical daily price limits for the CBT were found to be 30 cents per bushel for soybeans, 10 cents per bushel for corn, and 20 cents per bushel for both soft and hard wheat as of 1982 (CBT, 1982). The earliest change to KCBT daily price limits occurred when the limit increased from 10 cents per bushel in 1973, and it is assumed that these levels rose to the CBT limit of 20 cents per bushel. These values were assumed to have remained constant in the preceding years. Price limits remained stable until March 12, 1992 when CBT corn price limits increased from 10 to 12 cents per bushel, while soybean and wheat limits remained at 30 and 20 cents per bushel, respectively (Park, 2000). On August 14, 2000 daily price limits increased at the CBT from 12 to 20 cents per bushel for corn, from 30 to 50 cents per bushel for soybeans, and from 20 to 30 cents per bushel for wheat (CFTC). The KCBT limit changed when the wheat price limit was raised from 25 to 30 cents on October 9, 2000 (KCBT). On March 28, 2008 the KCBT and CBT both doubled the 30 cent price limit for wheat futures to 60 cents, while the CBT also expanded trading limits from 50 to 70 cents for soybeans and 20 to 30 cents for corn (CMEGroup). None of the limit days occurred on one of the days of interest to this study.

Procedures

Basis values were created by taking the cash market price less the futures market price. Basis forecasts were created using equation (3), where N=1,… ,5. The resulting forecast errors from each model were then evaluated.

Dhuyvetter and Kastens (1998) compare forecast accuracy with mean absolute error:

(11) $$MAE = \frac{1}{T}\sum_{t=1}^{T}|Basis_t - Ba\hat{s}is_t|$$

where the absolute value of each forecast error is averaged over the forecast period. This measure of forecast accuracy will be used in this study to identify the optimal historical period to include in basis forecasts.

Another popular determinant of forecast accuracy is the root mean squared error (RMSE) and is calculated as:

(12) $$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Basis_t - Ba\hat{s}is_t)^2}\,.$$

Jiang and Hayenga (1997) identified the RMSE as a popular measure of forecast accuracy as it penalizes the cost of larger errors with the square of the forecast error. The RMSE is included in the appendix because optimal forecasts may not only minimize forecast error, but also minimize the size of any potential errors. The RMSE identifies these models through its increased penalty of large errors.

The complex nature of modeling time-series, cross sectional data makes misspecification a concern when modeling basis forecast errors and interpreting their results. Econometric problems prevalent with this type of data include spatial autocorrelation, cross correlations, and heteroskedasticity. Failing to correct for these correlations and unequal error variance can lead to misleading standard errors and hypothesis testing. Dhuyvetter and Kastens (1998) tested for heteroskedasticity, and identified groupwise heteroskedasticity amongst forecast methods and time horizon variables for corn, soybeans, and wheat forecasting models. To correct for this heteroskedasticity, interaction terms of methods and forecast time horizon squared were included in each of their separate models. Although the dependence of the errors amongst competing forecast models could not be corrected, Dhuyvetter and Kastens (1998) conclude that a 4-year moving average was more accurate than the 3-year benchmark at 0.01 significance, while acknowledging that their significance levels are overstated. When independence across observations is incorrectly assumed, the standard errors and their ensuing t-tests can lead to overstated significance (Irwin, Good, and Martines-Filho, 2006).

A variation of the Dhuyvetter and Kastens (1998) approach to correct for heteroskedasticity was attempted with both the aggregate dataset and the individual commodities in this study. The pooled data set contains 15,180 observations. To correct for unequal variance using random effects, combinations of variables such as *period*location* and *location*year*, where *period* represented the preharvest or storage contract, *location* identified the market, and *year* identified the year of the forecast, were considered. However, these interaction terms resulted in too many parameters, which prevented the model from converging. As an alternative, we follow Irwin, Good, and Martines-Filho (2006) and pool the data, leaving only the absolute error, the dependent variable, the forecast length *N*, the only independent variable, and the year for the random effect and regressed using PROC MIXED. This model was also run for the individual commodities by period to identify any patterns that would be lost in the pooled model. The final mixed model is:

$$AE_{it} = \beta_0 + \sum_{j=1}^{4} \beta_j D_{ij} + v_t + \varepsilon_{it} \tag{13}$$

where AE_{it} is the absolute error of the ith forecast, at time t, β_0 is an intercept term created for the 5-year moving average to serve as a benchmark for model comparison, and β_j , $j = 1, \dots, 4$, are the coefficients for moving averages of *j* length, where D_{ij}=1 when $i = j$, v_t is the random-effects vector for years at time *t* and ε_{it} is the stochastic error term for the observation i at time t. The random-effects vector and stochastic error term are uncorrelated, and are distributed $v_t{\sim}N(0, \sigma_v^2)$ and $\varepsilon_{it}{\sim}N(0, \sigma_\varepsilon^2)$.

CHAPTER IV

MODEL RESULTS

Pooled Model Results

Table IV-1 shows the optimal forecast length by year for the pooled data. From this table we can see that the previous year's basis provides the optimal forecast for 37.51% (1144/3,050) of the values. The 5-year moving average produces the second most optimal forecasts at 25.77%, while the 2, 3, and 4-year moving averages account for 14.59, 11.64, and 10.49% of the sample, respectively.

Table IV-1. Number of Locations with a Given Length of Moving Average Having the Lowest Root Mean Squared Forecast Error, 1975-2008

Commodity	Period	*N*=1	*N*=2	*N*=3	*N*=4	*N*=5
Hard wheat	Preharvest	25	2	5	7	6
	Storage	34	2	4	1	4
Soft wheat	Preharvest	3	0	0	0	0
	Storage	0	2	0	0	1
Corn	Preharvest	0	0	0	0	7
	Storage	7	0	0	0	0
Soybeans	Preharvest	2	5	0	0	0
	Storage	7	0	0	0	0

Figure IV-2 graphs the number of optimal forecasts produced by the previous year's basis and 5-year moving average for the pooled data. The one-period forecast is usually close to the 5-year forecast, but following periods of structural change like the early 1980's (inflation, collapse of land prices, oil price shocks, etc.), 1988 (US-Canada free trade) and 2006 (lack of convergence at contract expiration) there are many more optimal forecasts using the one-period forecast. These large gaps in the amount of optimal forecast methods identified show the inferiority of basing expectations on longer period models after times of structural change.

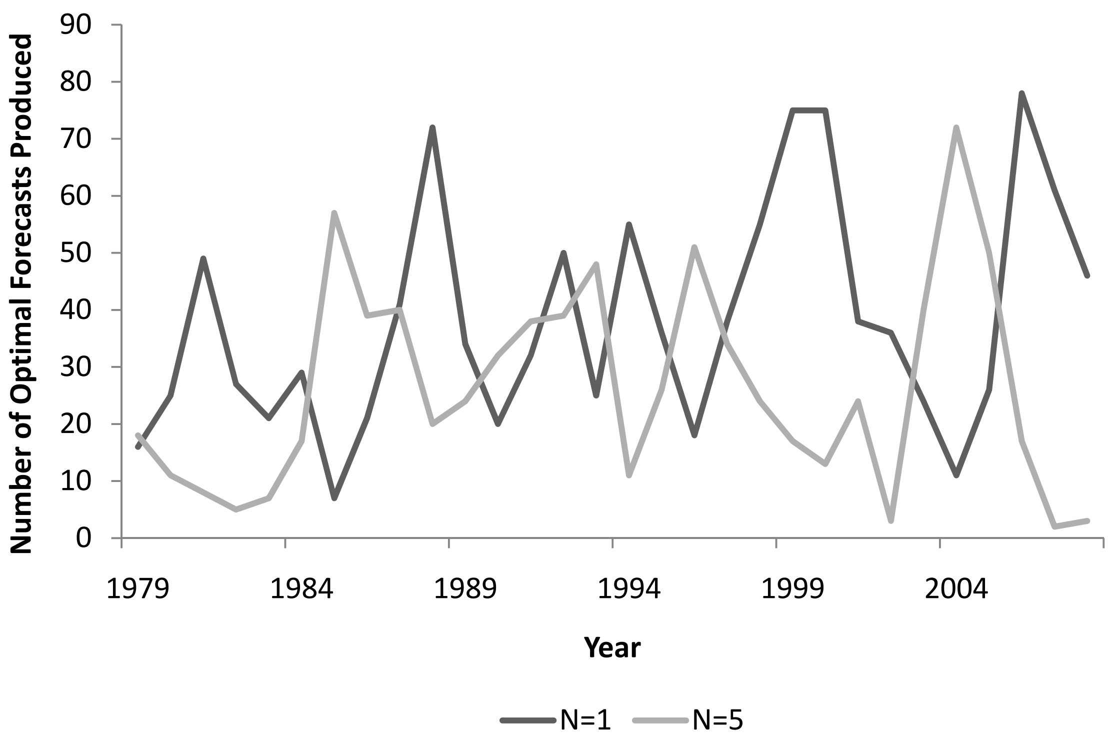

Figure IV-1. Number of minimum MAE forecasts produced by the previous year's basis vs. the 5-year moving average, 1979-2008

Table IV-2 shows the results from the pooled model of absolute forecast errors for the entire study. The F-test value of 0.92 fails to reject any difference amongst the competing forecast methods. The intercept term is the benchmark in the error model, the 5-year moving average, and is 12.89 cents/bu. Forecast accuracy increases as the amount of historical information used decreases, with the previous year's basis providing the lowest pooled MAE at 12.34 cents/bu. These results are generally within the range of the MAE's found by previous studies. Dhuyvetter and Kastens (1998) find the pooled MAE's of moving average forecasts to be between 10-13 cents/bu. for wheat, corn, and soybeans. The individual t-tests show that all shorter moving averages outperform the 5-year moving average, although the parameter estimates only differ by 0.50 cents/bu. between the previous year and 4- year averages.

Table IV-2. Absolute Error (cents/bu.) of Basis Forecasts as a Function of Number of Years in the Moving Average, Pooled Data, 1975-2008

Effect	Estimate	t-value	p-value
Intercept	12.34	12.06	0.000
N=1	-0.57	-2.06	0.040
N=2	-0.22	-0.79	0.427
N=3	-0.16	-0.58	0.562
N=4	-0.05	-0.18	0.858
N=5	-	-	-
F-statistic[a]	1.31	-	0.263

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

#2 Hard Wheat Model Results

Preharvest and storage hard wheat basis forecasting model results are listed in Table IV-3. For the preharvest forecasts, only the 2-year moving average is statistically different from the 5-year benchmark. The only preharvest model to produce a lower MAE than the benchmark is the 4-year moving average, which improves by only 0.04 cents/bu. These results indicate that, over the sample, any of the 5 preharvest models considered would result in a forecast error of approximately 15 cents/bu.

Table IV-3. Absolute Error (cents/bu.) of Hard Wheat Basis Forecasts as a Function of Number of Years in the Moving Average, 1978-2008

Period	Effect	Estimate	t-value	p-value
Preharvest	Intercept	12.77	8.71	0.000
	N=1	0.35	1.06	0.291
	N=2	0.68	2.06	0.040
	N=3	0.41	1.24	0.216
	N=4	-0.06	-0.19	0.853
	N=5	-	-	-
	F-statistic[a]	1.73	-	0.141
Storage	Intercept	13.03	9.06	0.000
	N=1	-1.94	-5.90	0.000
	N=2	-1.09	-3.32	0.001
	N=3	-0.77	-2.33	0.020
	N=4	-0.23	-0.70	0.481
	N=5	-	-	-
	F-statistic[a]	10.85	-	0.000

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

The storage model results for hard wheat support a significant difference in results from competing basis forecasting models with an F-statistic of 10.85. Individual t-tests of no difference from the 5-year benchmark are rejected for all but the 4-year moving average. The previous year's basis lowers the benchmark MAE from 13.03 cents/bu. to 11.09 cents/bu. The improvement in accuracy as the historical period shortens supports using shorter moving averages used to forecast the hard wheat storage basis.

Table IV-3 shows a pattern consistent throughout the results of these forecasts. By studying the preharvest and storage basis separately, we can see that MAEs are greater for preharvest than storage models. One possible explanation of this difference comes from Dhuyvetter and Kastens (1998), who found that forecast errors peak during critical production periods. Local conditions are much more variable around harvest, and spatial differences between cash and futures markets do not reflect the same supply and demand.

Modeling forecast accuracy for individual locations may prevent any significant findings from being lost in the aggregation of the larger models. Dhuyvetter and Kastens (1998) identified differences in forecast accuracy over several of the locations studied, admitting that their significance levels were overstated, but were unable to determine the effect of location on forecast errors. Absolute error was modeled for each location to identify any differences in accuracy across space.

Appendix Table 2 shows the model results by location for the preharvest forecasts. Consistent with Dhuyvetter and Kastens (1998), there is a tendency for Kansas location forecasts to generate optimal forecasts with longer moving averages. Of all the models, only the preharvest basis models for Hays, KS showed a difference in forecast accuracy with an F-statistic of 2.89. The previous year's basis is significantly worse than the 5-year benchmark at 0.05% for Hays and Great Bend, KS. MAE for Hays would increase by 5.51 cents/bu., while the MAE for Great Bend would increase from 10.73 to 14.06 cents/bu.

Preharvest forecasts for Oklahoma locations, on the other hand, tend to benefit from models based on shorter historical periods. The greatest reduction in the MAE by any model occurs when the previous year's basis at Muskogee, OK reduces the benchmark from 21.52 to 15.08 cents/bu. using the previous year's forecast. Oklahoma markets are farther from delivery points, and grain does not flow to a delivery point. This allows more structural changes across space in Oklahoma markets than Kansas.

Storage forecasts in Appendix Table IV-3 indicate that the previous year's basis is the most accurate method for Oklahoma. Significant improvement is identified from using the previous year's basis instead of the 5-year benchmark for 16 of the 26 Oklahoma locations. Davis, OK experiences the greatest reduction in the storage MAE, 4.87 cent/bu., when the previous year's basis is used instead of the benchmark.

The same pattern exists in Kansas markets. Using the previous year's storage basis forecast lowered the benchmark MAE in 14 of the 19 Kansas locations.

Hard Wheat Changes over Time

Figure IV-2 is a map of the 1975-1980 average harvest basis values from the beginning of the Oklahoma Market Report. Basis values tend to be weakest in the northern part of the state, and grow stronger when moving south.

Figure IV-3 shows the 2008 harvest basis values. The trend from the first map is now reversed, with basis strengthening from the southern to the northern part of the state. A major shift in the primary market for Oklahoma wheat occurred over the period studied. Oklahoma wheat was shipped via rail to the Gulf Port at Houston, but now travels by barge to New Orleans. This change in the transportation of Oklahoma wheat over the time period studied explains why Oklahoma wheat basis changed over the 30 plus years studied.

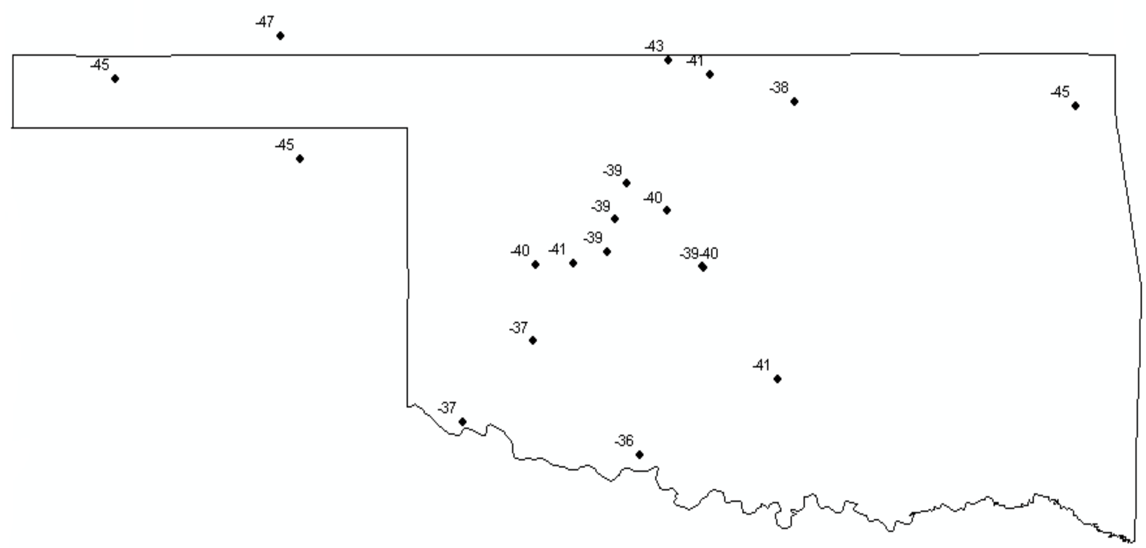

Figure IV-2. 1975-1980 average Oklahoma wheat harvest basis

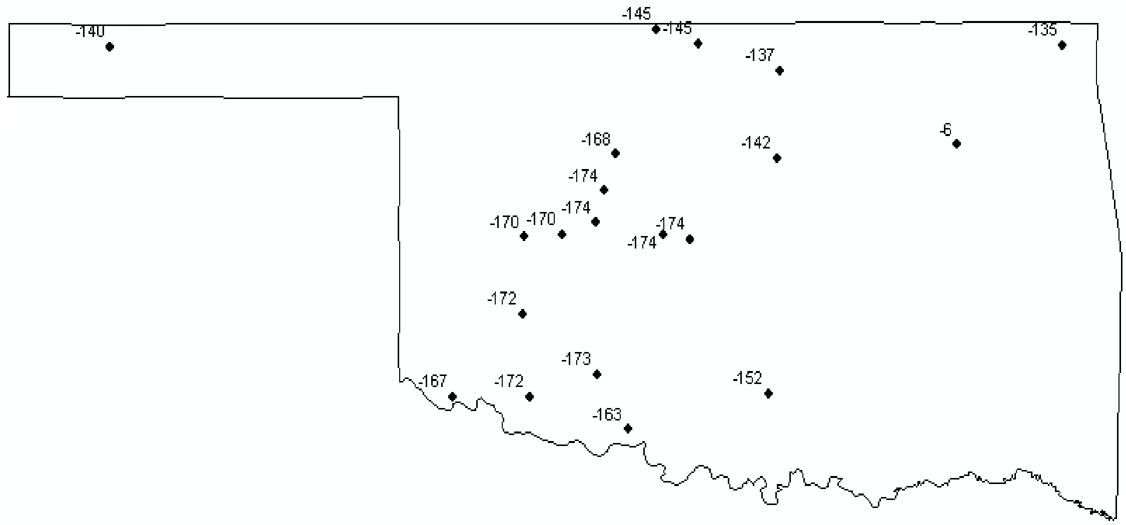

Figure IV-3. 2008 Oklahoma wheat harvest basis

Figure III-4 shows the 5-year average basis for Kansas locations over 1982-1986. The trend in this map is that the basis weakens the further south and west you move away from Kansas City.

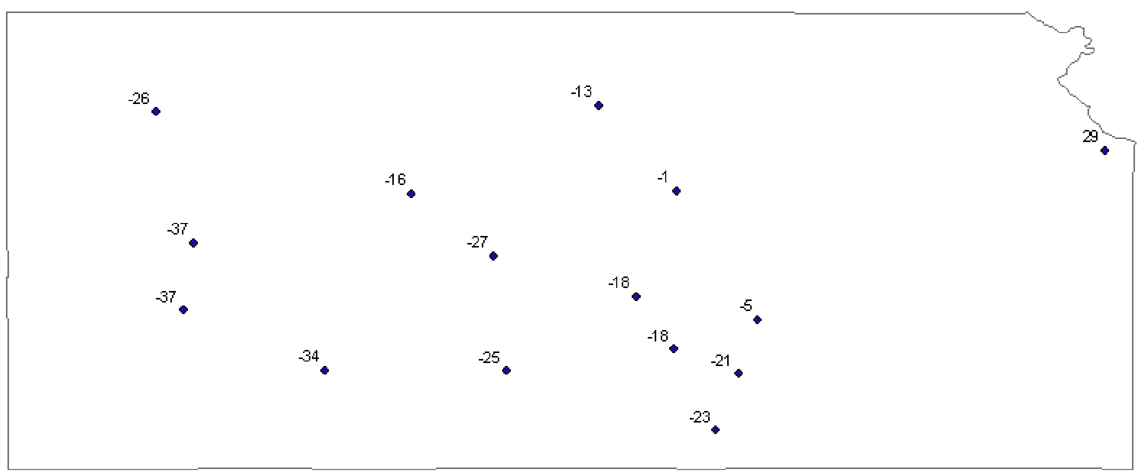

Figure IV-4. 1982-1986 Average Kansas wheat harvest basis

The 2007 Kansas harvest basis is shown in Figure III-5. Similar to the relationships in Figure 3, the harvest basis tends to weaken as you move from Kansas City southwest.

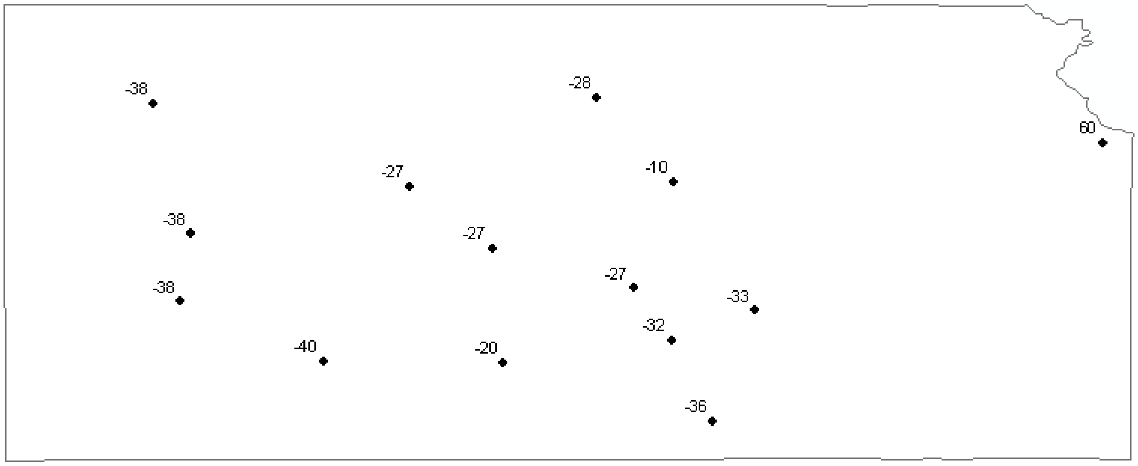

Figure IV-5. 2007 Kansas wheat harvest basis

The greatest difference between the two time periods is that most locations seem to be aligned with the markets surrounding them. In Figure III-3 there were isolated markets that experienced much stronger basis than their closest neighbors. Figure III-4 shows that nearly all of the locations are within a few cents of their surrounding locations.

#1 Hard Wheat Model Results

The Kansas City price data allows this study to compare the differences in forecasting both the regular protein #1 hard red wheat, and 13% protein #1 hard red wheat. Table IV-4 shows the model results for the Kansas City ordinary protein #1 hard wheat basis models.

Table IV-4. Absolute Error (cents/bu.) of Kansas City Ordinary Protein, #1 Hard Wheat Basis Forecasts as a Function of Number of Years in the Moving Average, 1976-2008

Period	Effect	Estimate	t-value	p-value
Preharvest	Intercept	15.72	8.33	0.000
	N=1	0.47	0.29	0.770
	N=2	0.21	0.13	0.897
	N=3	-0.12	-0.07	0.942
	N=4	-0.96	-0.60	0.552
	N=5	-	-	-
	F-statistic[a]	0.23	-	0.923
Storage	Intercept	12.99	5.24	0.000
	N=1	2.17	1.35	0.180
	N=2	1.89	1.17	0.244
	N=3	1.83	1.14	0.259
	N=4	0.76	0.47	0.638
	N=5	-	-	-
	F-statistic[a]	0.65	-	0.629

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

Table IV-5 reports the results of the model for the 13 percent #1 hard wheat. When compared to the results of Table IV-3, we can see how forecasting two subclasses of the same commodity affect forecast accuracy with little variation due to space. The benchmark intercept for the 13 percent protein model is 3.22 cents/bu. higher than the ordinary protein forecast model. This best preharvest forecast is still 1.22 cents/bu. more than the worst ordinary protein forecast model.

Table IV-5. Absolute Error (cents/bu.) of Kansas City 13% Protein, #1 Hard Wheat Basis Forecasts as a Function of Number of Years in the Moving Average, 1976-2008

Period	Effect	Estimate	t-value	p-value
Preharvest	Intercept	18.94	4.11	0.000
	N=1	0.26	0.18	0.856
	N=2	-1.37	-0.95	0.344
	N=3	-1.53	-1.07	0.289
	N=4	-0.32	-0.22	0.824
	N=5	-	-	-
	F-statistic[a]	0.64	-	0.636
Storage	Intercept	19.34	3.75	0.001
	N=1	5.52	1.88	0.064
	N=2	3.55	1.20	0.231
	N=3	1.06	0.36	0.721
	N=4	0.77	0.26	0.793
	N=5	-	-	-
	F-statistic[a]	1.21	-	0.310

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

Comparing the forecast results of ordinary and 13% protein #1 hard wheat shows the effect of differences in grain form on forecast accuracy. Forecast errors are lower in both periods for ordinary protein. Higher forecast errors for 13% protein are likely the result of changes in the variable premiums for protein content at KCBT. Rather than using a fixed premium similar to what exists between #1 and #2 grade wheat, the market posts a protein premium scale that allows for market adjustments to premiums (KCBT).

These differences in supply and demand for wheat qualities differentiate the form of ordinary and 13% protein hard wheat markets.

Soft Wheat Model Results

Table IV-6 displays the model results for the soft wheat basis forecasting models. Using the previous year's basis to predict soft wheat preharvest basis would lead to an average forecast error of 25.95 cents/bu., while the most accurate method, the 2-year moving average, only lowers the MAE to 23.42. Only the 2-year moving average proves to be a better forecast of the storage basis than the benchmark for soft wheat. Although it decreases the MAE to nearly 13 cents/bu., the 2-year moving average is not significantly different from the benchmark. At nearly 10 cents/bu. below the preharvest forecast intercept, the storage model intercept helps support the ability to forecast the storage basis more accurately than the preharvest basis.

Appendix Table 4 lists the basis forecast error model for the soft wheat preharvest period. When looking at the models by location, some interesting results become apparent. The Chicago, IL MAEs range from 11.81 to 13.73 cents/bu., while the St. Louis, MO and Toledo, OH MAEs more than double to 35.13 to 41.33 and 28.60 to 32.78 cents/bu., respectively. The only forecast significantly different from the 5-year moving average over these locations is the previous year's basis for St. Louis, which is 6.20 cents/bu. worse, and statistically different than the benchmark at 0.10 significance.

Table IV-6. Absolute Error (cents/bu.) of Soft Wheat Basis Forecasts as a Function of Number of Years in the Moving Average, 1975-2008

Period	Effect	Estimate	t-value	p-value
Preharvest	Intercept	23.45	4.64	0.000
	N=1	2.50	0.78	0.434
	N=2	-0.03	-0.01	0.991
	N=3	0.20	0.06	0.951
	N=4	0.48	0.15	0.882
	N=5	-	-	-
	F-statistic[a]	0.22	-	0.926
Storage	Intercept	13.91	8.52	0.000
	N=1	0.59	0.45	0.654
	N=2	-0.93	-0.71	0.479
	N=3	0.16	0.13	0.900
	N=4	0.51	0.39	0.696
	N=5	-	-	-
	F-statistic[a]	0.43	-	0.788

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

Appendix Table 5 shows the individual model comparisons by location for the soft wheat storage absolute error models. The only location to reject to the null hypothesis of difference in forecast accuracy is Toledo, with an F-statistic of 4.34. If the previous year's basis is used to forecast the storage basis instead of the 5-year benchmark, the MAE will increase by over 50 percent from 9.27 to 14.38 cents/bu. Only the Toledo model does not increase accuracy by applying a 3-year or shorter moving average to the basis forecast.

All three of these locations represent current CBT delivery locations, but have not been delivery points over the entire period studied. In 1982, the only locations accepting CBT wheat were Chicago and Toledo, at par and a 2 cent/bu. discount, respectively. As of 1997, Toledo, OH received delivery at a discount of 2 cent/bu., and St. Louis, MO at an 8 cent/bu. premium to Chicago (CFTC).

A market undergoes structural change when it is made a delivery point on a futures contract. Delivery point cash prices are set by a set premium or discount aligned with the Chicago cash market. Grain qualities above the contract minimum no longer flow to delivery points, and the market price adjusts to reflect the cheapest to delivery grains specified in the futures contract. These changes over time help explain the higher MAEs for Toledo and St. Louis found by this study.

Corn Model Results

Table IV-7 shows the results for the corn models across all regions of Illinois. Results from the preharvest model indicate that using the previous year's basis outperforms the 5-year benchmark over all Illinois locations. The F-statistic and individual t-tests both fail to indicate any significant differences in forecast choice. The F-statistic of 4.10 for the storage models rejects the null hypothesis, and concludes that model forecast accuracy does differ over the sample for corn storage basis. Significant differences from the 5-year benchmark exist in every model except the 4-year moving average at a 0.05 level. This result indicates that shorter moving averages can outperform the 5-year moving average at forecasting the corn storage basis. The best model, using the previous year's basis, lowers the MAE from the 5-year moving average of 7.59 cents/bu. to 6.32 cents/bu.

Table IV-7. Absolute Error (cents/bu.) of Corn Basis Forecasts as a Function of Number of Years in the Moving Average, 1980-2008[a]

Period	Effect	Estimate	t-value	p-value
Preharvest	Intercept	11.74	9.60	0.000
	N=1	-0.12	-0.21	0.836
	N=2	0.63	1.07	0.286
	N=3	0.55	0.94	0.349
	N=4	0.52	0.87	0.385
	N=5	-	-	-
	F-statistic[b]	0.70	-	0.594
Storage	Intercept	7.49	7.59	0.000
	N=1	-1.17	-3.63	0.000
	N=2	-0.68	-2.12	0.034
	N=3	-0.66	-2.05	0.041
	N=4	-0.18	-0.56	0.574
	N=5	-	-	-
	F-statistic[b]	4.10	-	0.003

[a] Storage model forecasts begin in 1981, due to the time-series available.
[b] The null hypothesis is that all values of *N* have the same forecast accuracy.

Corn preharvest forecasting results by region are listed in Appendix Table 6. These results show that the 5-year moving average is the optimal forecast method. The range of MAEs between the competing models is consistently below 2 cents/bu. across each region. Individual t-tests do not find any significant differences in model accuracy for any region.

Appendix Table 7 lists the corn storage basis forecast comparisons by region. The results of the individual regions indicate that the previous year's basis is significantly more accurate than the 5-year benchmark for the South-Central region of Illinois at 0.05 significance. Using the previous year to forecast the storage basis results in a 1.69 cent/bu. reduction in the MAE of the benchmark. Three other regions, the Northern, Western, and West-Southwest regions all showed significant reductions from the 5-year benchmark using the previous year's basis at 0.10 significance.

Taylor, Dhuyvetter, and Kastens (2004) also found that shorter moving average models resulted in lower MAEs based on their sample of the nearby Kansas corn basis forecasts 24 weeks after preharvest. The previous year's basis resulted in a MAE of 10.57 cents/bu. for Kansas compared to the 6.32 cents/bu. for Illinois found in this study.

Soybean Model Results

Table 7 shows the results from the absolute error models for the Illinois soybean basis. The preharvest 5-year benchmark MAE is 11.23 cents/bu., and can be improved by all of the shorter moving-average models. The most improvement comes from the 2-year moving average, which lowers the MAE to 10.62 cents/bu. Although the benchmark can be improved upon, the improvement is not enough to be statistically different based on the t-test comparisons. The narrow range (< 0.61 cents/bu.) of MAEs shows that little difference exists across preharvest basis models over the period studied.

It is clear from the results of the storage basis error model that the choice of forecasting models results affects the accuracy of the Illinois soybean storage basis. The F-statistic of 8.58 indicates that the choice of models can result in different forecasting accuracies. While all of the shorter moving average models outperform the benchmark, both the previous year's basis and the 2-year moving average result in 1.98 and 1.16 cents/bu. lower forecasts, respectively. Compared to the soybean preharvest model, the storage basis forecasts result in decreased MAEs of over 1.6 cents/bu.

Table IV-8. Absolute Error (cents/bu.) of Soybean Basis Forecasts as a Function of Number of Years in the Moving Average, 1980-2008[a]

Period	Effect	Estimate	t-value	p-value
Preharvest	Intercept	11.23	8.58	0.000
	N=1	-0.47	-0.78	0.438
	N=2	-0.61	-1.00	0.318
	N=3	-0.50	-0.82	0.410
	N=4	-0.20	-0.32	0.748
	N=5	-	-	-
	F-statistic[b]	0.43	-	0.852
Storage	Intercept	9.61	8.25	0.000
	N=1	-1.98	-4.99	0.000
	N=2	-1.16	-2.92	0.004
	N=3	-0.66	-1.66	0.100
	N=4	-0.08	-0.19	0.846
	N=5	-	-	-
	F-statistic[b]	8.58	-	0.000

[a] Storage model forecasts begin in 1981, due to the time-series available.
[b] The null hypothesis is that all values of *N* have the same forecast accuracy.

According to Appendix Table 8, the shorter moving average models perform the best at forecasting soybean preharvest basis for each region in Illinois. The 2-year moving average produces optimal forecasts in 5 of the 7 regions, while the previous year's basis performs the best in the remaining 2. The lowest MAE from the models is 8.78 cents/bu. using the previous year's basis in North Central Illinois. Not every region benefits from using shorter moving averages in the MAE models. While coefficients for some models indicate that one of the shorter forecast lengths would result in greater MAEs, the South-Central region's benchmark forecast accuracy improves with the 4-year moving average. However, none of the competing models for the Illinois regions show a significant difference from the 5-year benchmark.

Appendix Table 9 shows the results when absolute error is modeled by location for soybean storage basis forecasting. Only one region, Western Illinois, rejects the F-test of no difference in forecast accuracy at 0.10 significance. In this region, forecasting the soybean storage basis with the previous year's basis results in a 2.24 cent/bu. reduction in the MAE from the 5-year benchmark. Similar reductions of over 1 cent/bu. occur across all locations when choosing the previous year's basis over the 5-year benchmark. The overall trend across every region is an improvement in forecast accuracy as the length of the moving average model decreases.

Model Effects from Recent Years

Recent events have lead to erratic basis levels across commodity markets over the last three years. A combination of inconsistent basis levels at expiration, weather complications, and an abnormal shock in oil prices has created abnormal basis conditions.

The erratic basis at expiration was first identified in the CBT July 2006 wheat contract, when inconsistent convergence between cash and futures markets began.

Poor soybean contract convergence first occurred in the March 2007 contract. Corn contract convergence had also been erratic, but was generally better than wheat and soybeans over the same period (Irwin et al. 2009).

Irwin et al. (2009) propose that this lack of convergence is due to: 1) commercial storage rates below CBT maximum storage rates, 2) the presence of “long-only” index funds in commodity markets, and 3) an a larger risk premium in futures prices due to increased uncertainty. Another explanation of this can be tied to differences in form. In some cases, the wheat specified on the futures warehouse receipt was not equal to the quality of the market available in the market. These four factors can explain the convergence problems that have occurred in futures markets over the last 2-3 years.

World wheat and feed stocks experienced tight stocks in the 2007/2008 marketing year, and reacted sharply to small changes in supply and demand (Anderson, 2007). U.S. crops could have been affected by the dramatic increase in fuel costs during the first half of 2008 resulting in wide basis values. Commodity markets might have also been affected by the credit crunch and recession that occurred in the second half of 2008.

These irregular market conditions could affect the results of this study. Table 9 shows the results of the pooled model without the 2006-2008 data. Compared to the model of the entire dataset, the model that does not include the recent years shows a 1.78 cent/bu. improvement in the benchmark model. This shows the impact of the higher forecasts errors common in recent years. Parameter estimates for the 2-4 year moving averages no longer improve the accuracy of the benchmark, and the improvement of the previous year's basis on the benchmark is reduced. Excluding 2006-2008 does not change the most important finding of the full model, that no significant difference exists in forecast accuracy over the period studied.

Table IV-9. Absolute Error (cents/bu.) of Basis Forecasts as a Function of Number of Years in the Moving Average, Pooled Data, 1975-2005

Effect	Estimate	t-value	p-value
Intercept	11.01	17.14	0.000
N=1	-0.12	-0.50	0.615
N=2	0.17	0.71	0.476
N=3	0.08	0.33	0.740
N=4	0.06	0.24	0.813
N=5	-	-	-
F-statistic[a]	0.40	-	0.807

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

From Table III-10, we can see the error model for the 2006-2008 period. Parameter estimates indicate that the 5-year benchmark can be outperformed by shorter forecasts, with the improvement proportional to the reduction in years included in the forecast. Only using the previous year's forecast is significantly different from the benchmark, and reduces the MAE by 4.89 cents/bu. to 27.6 cents/bu. Even with the reduction in forecast error by models including 1-4 years of historical data, the reduction does not lead to a significant difference in forecast model accuracy.

Table IV-10. Absolute Error (cents/bu.) of Basis Forecasts as a Function of Number of Years in the Moving Average, Pooled Data, 2006-2008

Effect	Estimate	t-value	p-value
Intercept	26.02	3.56	0.070
N=1	-4.87	-2.64	0.009
N=2	-3.92	-2.12	0.034
N=3	-2.44	-1.32	0.187
N=4	-1.05	-0.57	0.569
N=5	-	-	-
F-statistic[a]	2.33	-	0.054

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

Table III-11 shows the descriptive statistics of the absolute forecast error over the time period studied. These results allow us to see how including the 2006-2008 crop years affects the level of absolute forecast error. From this table we can see that the absolute errors for the 2006-2008 crop years experience considerably larger forecast errors. The recent forecasts constitute only 9.5 percent of the total forecast absolute errors, but are large enough to raise the mean absolute error by 1.52 cents/bu. and increase the standard deviation by 5.30 cents/bu.

Table IV-11. Descriptive Statistics of the Absolute Forecast Error for the Pooled Data over Different Time Periods

Period	Optimal Forecasts	N	Mean	S.D.	Minimum	Maximum
1975-2005		13740	11.13	9.37	0.00	146.76
	N=1		10.97	9.62	0.00	146.76
	N=2		11.26	9.34	0.00	135.80
	N=3		11.17	9.36	0.00	129.97
	N=4		11.15	9.26	0.00	126.95
	N=5		11.09	9.26	0.00	127.52
2006-2008		1440	27.13	34.64	0.00	234.72
	N=1		24.71	34.42	0.02	223.88
	N=2		25.67	33.31	0.00	174.81
	N=3		27.14	35.17	0.08	211.29
	N=4		28.53	34.91	0.50	218.71
	N=5		29.60	35.36	0.10	234.72
1975-2008		15180	12.65	14.67	0.00	234.72
	N=1		12.27	14.56	0.00	223.88
	N=2		12.63	14.20	0.00	174.81
	N=3		12.68	14.77	0.00	211.29
	N=4		12.80	14.79	0.00	218.71
	N=5		12.85	15.00	0.00	234.72

However, these results do not affect the main conclusion that can be drawn from the 1975-2008 pooled model results, which indicates that shorter moving averages decrease the forecast MAE.

CHAPTER V

CONCLUSIONS

The most popular method of forecasting the basis is historical moving averages. Given the recent failure of longer moving averages proposed by previous studies, this research reassesses past recommendations about the best length of moving average to use in forecasting basis. Our study uses a longer time series with more locations and crops than these previous studies to determine the optimal length of historical data to forecast basis. The hypothesis testing procedure using the pooled data is valid in the presence of cross correlations.

Basis values for hard wheat, soft wheat, corn and soybeans were used to create basis forecasts. The methods considered included the previous year's basis and moving averages of the previous 2-5 years. Mean absolute error was modeled for the pooled data following Irwin, Good, and Martines-Filho (2006). The mean absolute error was the dependent variable, the forecast length was the independent variable, and year was the random effect. This model was also run for the individual commodities by period to identify any patterns that would be lost in the pooled model.

This research found the size of most MAEs to be consistent with previous studies (Dhuyvetter and Kastens, 1998; Taylor, Dhuyvetter and Kastens, 2002). These values were generally between 10 and 17 cents/bu.

The optimal forecast length found for each commodity is generally shorter than previous recommendations. Using a 4-year moving average produced the minimum MAE preharvest wheat forecast, consistent with Dhuyvetter and Kastens (1998), but the optimal storage forecast model has lower forecast error using shorter historical information. This study finds that the optimal amount of historical data included in corn and soybean forecasts have shortened to one or two years for both preharvest and storage periods. Significant differences in forecast accuracy among the different models are rare, and in most cases the differences are not statistically significant.

Another important component of this study is a synthesis of basis literature, which explains the basis through the Law of One Price. Given the recent structural changes in basis, there is a need to better understand what causes structural changes even when using moving averages to forecast the basis. Explaining the basis in terms of time, form, and space can help identify structural changes, and helps select the correct amount of historical information to include in historical moving averages.

Structural changes over the time period studied have led to recommending shorter historical moving average to forecast the basis. Markets within this study undergo varying amounts of structural change for different reasons. Kansas wheat markets, for example, maintained consistent basis relationships over space, which may be due to their proximity to the KCBT hard wheat market delivery points. Toledo, OH and St. Louis, MO experienced more structural change when they became futures contract delivery points. Prices at delivery points are more sensitive to changes in transportation costs, and change from being determined by local supply and demand to reflecting the cheapest to deliver commodity on futures contracts. The structural changes apparent in the basis data in this study identifies that shorter moving averages produce the most accurate basis forecast in terms of mean absolute error.

Although our individual models produced varied results, the general rule of thumb supported by this research is: When a location or time period does not undergo structural change longer moving averages produce optimal forecasts, but when it appears that a structural change has occurred, the previous year’s basis should be used.

CHAPTER VI

RECOMMENDATIONS FOR FUTURE STUDY

The narrow scope of the time series methods used in this study allows future projects to expand these results. Advanced time series modeling techniques have their advantages, and are not considered in this research. Future studies may find a benefit in forecasting accuracy from techniques such as ARIMA and other time series models.

Another potential avenue of further study would be to develop a hypothesis testing approach that could use disaggregate data. One of the restrictions in modeling forecast error in this study was the inability to get forecast error covariance matrices to converge. Correcting for random effects in disaggregate data might lead to more powerful statistical tests.

These opportunities provide some, but not all, of the possible extensions that can be made to this comparison of practical basis forecasting methods.

REFERENCES

Anderson, T.G., L. Benzoni, and J. Lund. 2002. "An Empirical Investigation of Continuous-Time Equity Return Models." *Journal of Finance* 57:1239-1284.

Bates, D.S. 1996. "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options." *Review of Financial Studies* 9:69-107.

Benirschka, M., and J. Binkley. 1995. "Optimal Storage and Marketing over Space and Time." *American Journal of Agricultural Economics* 77:521-24.

Breeden, D.T. 1980. "Consumption Risks in Futures Markets." *Journal of Finance* 35:503-20.

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

CBT. 1982. "Statistical Annual: Futures Data, Part 1." Chicago Board of Trade.

Commodity Futures Trading Commission. 2000. "CFTC Approves CBT Proposal to Increase Maximum Daily Price Fluctuation Limits for its Agricultural and Precious Metal Futures and Option Contracts." Release: #4435-00.

CME Group. February 8, 2008. "CME Group, KCBT and MGEX Announce Price Limit Expansion for Wheat Futures."

Chernov, M., A.R. Gallant, E. Ghysels, and G. Tauchen. 2003. "Alternative Models for Stock Price Dynamics." *Journal of Econometrics* 116:225-257.

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

Cootner, P.H. 1960b. "Rejoinder." *Journal of Political Economy* 68:415-418.

— 1960a. "Returns to Speculators: Telser Versus Keynes." *Journal of Political Economy* 68:396-404.

Dhuyvetter, K. July 10, 2008. "WheatCash—Wade." Kansas State University Extension.

Dhuyvetter, K. C., and T. L. Kastens. 1998. “Forecasting Crop Basis: Practical Alternatives.” Paper presented at the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL, April.

Dykema, A., N.L. Klein, and G. Taylor. 2002. “The Widening Corn Basis in South Dakota: Factors Affecting and the Impact of the Loan Deficiency Payment.” Paper presented at WAEA annual meeting, Long Beach, California, 28-31 July.

Farm.Doc. 2009. “Illinois Regional Cash Price Database.” University of Illinois.

Hauser, R.J., P. Garcia, and A.D. Tumblin. 1990. “Basis Expectations and Soybean Hedging Effectiveness.” *North Central Journal of Agricultural Economics* 12:125-136.

Irwin, S.H., D.L. Good, and J. Martines-Filho. 2006. “The Performance of Agricultural Market Advisory Services in Corn and Soybeans.” *American Journal of Agricultural Economics* 88:162-181.

Irwin, S.H., P. Garcia, D.L. Good, and E.L. Kunda. 2009. “Poor Convergence Performance of CBOT Corn, Soybean and Wheat Futures Contracts: Causes and Solutions." Marketing and Outlook Research Report 2009-02, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, March [http://www.farmdoc.uiuc.edu/marketing/reports]

Jiang, B., and M. Hayenga. 1997. “Corn and Soybean Basis Behavior and Forecasting: Fundamental and Alternative Approaches.” Paper presented at the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL, April.

Kaldor, N. 1939. “Speculation and Economic Stability,” *Review of Economic Studies* 7:1-27.

KCBT. 1974-1976. “Historical Data.” Accessed at [http://www.kcbt.com/historical_data.asp] in June, 2008.

KCBT. March 10, 2009. Via e-mail “60 Cent Limit.” Sheila Summers. VP of Marketing,

Kenyon, D.E. and S.E. Kingsley. 1973. “An Analysis of Anticipatory Short Hedging Using Predicted Harvest Basis.” *Southern Journal of Agricultural Economics* 5:199-203.

Kolb, R.W. 1992. “Is Normal Backwardation Normal?” *Journal of Futures Markets* 12:75-91.

Klumpp, J.M., B.W. Brorsen, and K.B. Anderson. 2007. *Journal of American Agricultural Economics* 39:571-579.

Martin, L. J.L. Groenewegen and E. Pidgeon. 1980. “Factors Affecting Corn Basis in Southwestern Ontario.” *American Journal of Agricultural Economics* 62:107-112.

Naik, G. and R.M. Leuthold. 1991. “A Note on the Factors Affecting Corn Basis Relationships.” *Southern Journal of Agricultural Economics* 23:147-153.

Oklahoma Department of Agriculture, Food and Forestry. 1974-2008. “Oklahoma Market Report.”

Park, C.L. 2000. “Examining Futures Price Changes and Volatility on the Trading Day After a Limit-lock Day.” *Journal of Futures Markets* 20:445-466.

Peavey Grains. 1988-2008. “Peavey Catoosa Basis Bids.”

R & C Data. 1970-2008. “Financial Price Data.”

Sanders, D.R. and M.R. Manfredo. 2006. “Forecasting Basis Levels in the Soybean Complex: A Comparison of Time Series Methods.” *Journal of Agricultural and Applied Economics* 38:513-523.

Taylor, M., K.C. Dhuyvetter, and T.L. Kastens. 2004. “Incorporating Current Information into Historical-Average-Based Forecasts to Improve Crop Price Basis Forecasts.” Paper presented at NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, St. Louis, MO, 19-20 April.

Telser, L.G. 1958. “Futures Trading and the Storage of Cotton and Wheat.” *Journal of Political Economy* 66:233-55.

—. 1960. “Reply.” *Journal of Political Economy* 68:404-415.

U. S. Department of Agriculture. 1970-2008. “Grain and Feed Market News.” Agricultural Market Service.

Vance, L.L. 1946. “Grain Market Forces in the Light of Inverse Carrying Charges.” *Journal of Farm Economics* 28:1036-1040.

Wang, D., and W.G. Tomek. 2007. “Commodity Prices and Unit Root Tests.” *American Journal of Agricultural Economics* 89:873-889.

Ward, R.W. and F.A. Dasse. 1977. “Empirical Contributions to Basis Theory: The Case of Citrus Futures.” *American Journal of Agricultural Economics* 59: 71-80.

Working, Holbrook. 1953. “Hedging Reconsidered.” *Journal of Farm Economics* 35:544-61.

— 1934. “Price Relationships Between May and New-Crop Futures at Chicago Since 1885.” *Wheat Studies* 10:183-228.

— 1949. “Theory of the Price of Storage.” *American Economic Review* 39:1254-1262.

Wright, B.D. and J.C. Williams. 1989. “A Theory of Negative Prices for Storage.” *Journal of Futures Markets* 20:1-13.

APPENDIX

Appendix Table 1. Locations and Time Periods Studied by Commodity

Commodity	Location	Time Period
#2 Soft Red Winter Wheat	Chicago, IL	1970-2008
	St. Louis, MO	1970-2008
	Toledo, OH	1970-2008
#1 Hard Red Winter Wheat	Kansas City, MO	1976-2008
#2 Hard Red Winter Wheat	Andale, KS	1982-2007
	Beloit, KS	1982-2007
	Colby, KS	1982-2007
	Dodge City, KS	1982-2007
	Emporia, KS	1983-2004
	Garden City, KS	1982-2004
	Great Bend, KS	1982-2007
	Hays, KS	1982-2007
	Hutchinson, KS	1982-2007
	Liberal, KS	1974-1999
	Pratt, KS	1982-2007
	Salina, KS	1982-2007
	Scott City, KS	1982-2007
	Wellington, KS	1982-2007
	Whitewater, KS	1982-2007
	Wichita, KS	1982-1999
	Afton, OK	1974-2005
	Banner, OK	1976-2008
	Catoosa, OK	1993-2008
	Clinton, OK	1974-2008
	Davis, OK	1984-2008
	El Reno, OK	1974-2008
	Eldorado, OK	1976-2008
	Frederick, OK	1980-2008
	Geary, OK	1974-2008
	Hobart, OK	1974-2008

Appendix Table 1. Locations and Time Periods Studied by Commodity

Commodity	Location	Time Period
	Keyes, OK	1974-2008
	Kingfisher, OK	1974-2005
	Lawton, OK	1977-2008
	Manchester, OK	1974-2008
	Medford, OK	1974-2008
	Miami, OK	1982-2008
	Muskogee, OK	1975-2008
	Okeene, OK	1974-2008
	Pauls Valley, OK	1975-2008
	Ponca City, OK	1975-2008
	Stillwater, OK	1988-2008
	Temple, OK	1980-2008
	Watonga, OK	1975-2008
	Weatherford, OK	1974-2008
	Yukon, OK	1974-2005
	Perryton, TX	1974-1999
Corn	Northern, IL	1970-2008
	Western, IL	1970-2008
	North Central, IL	1970-2008
	South Central, IL	1970-2008
	Wabash, IL	1970-2008
	West-Southwest, IL	1970-2008
	Little Egypt, IL	1970-2008
Soybeans	Northern, IL	1970-2008
	Western, IL	1970-2008
	North Central, IL	1970-2008
	South Central, IL	1970-2008
	Wabash, IL	1970-2008
	West-Southwest, IL	1970-2008
	Little Egypt, IL	1970-2008

Appendix Table 2. 1978-2008 Hard Wheat Preharvest Basis Forecasting Model Comparisons

Location	Optimal MA[a]	Effect	Estimate	t-Value	P-Value
Afton, OK	1	Intercept	10.89	5.15	0.000
		N=1	1.02	0.71	0.482
		N=2	1.66	1.15	0.254
		N=3	1.41	0.98	0.332
		N=4	0.35	0.24	0.810
		N=5	0.00	-	-
		F-statistic[b]	0.47	-	0.758
Andale, KS	1	Intercept	11.55	5.30	0.000
		N=1	0.68	0.39	0.698
		N=2	1.11	0.64	0.527
		N=3	0.90	0.52	0.605
		N=4	-0.12	-0.07	0.946
		N=5	0.00	-	-
		F-statistic[b]	0.20	-	0.939
Banner, OK	1	Intercept	15.09	3.24	0.003
		N=1	-0.21	-0.17	0.863
		N=2	0.85	0.70	0.488
		N=3	1.07	0.88	0.380
		N=4	0.53	0.44	0.664
		N=5	0.00	-	-
		F-statistic[b]	0.40	-	0.807
Beloit, KS	4	Intercept	12.61	6.06	0.000
		N=1	3.12	1.82	0.073
		N=2	1.16	0.68	0.500
		N=3	1.37	0.80	0.428
		N=4	0.02	0.01	0.990
		N=5	0.00	-	-
		F-statistic[b]	1.10	-	0.360
Catoosa, OK	5	Intercept	10.09	5.51	0.000
		N=1	2.16	1.53	0.132
		N=2	0.14	0.10	0.919
		N=3	0.01	0.01	0.995
		N=4	0.62	0.44	0.663
		N=5	0.00	-	-
		F-statistic[b]	0.84	-	0.507
Clinton, OK	1	Intercept	15.13	3.53	0.001
		N=1	-1.22	-1.04	0.300
		N=2	0.56	0.48	0.634
		N=3	0.22	0.19	0.852
		N=4	-0.43	-0.37	0.713
		N=5	0.00	-	-
		F-statistic[b]	0.69	-	0.603
Colby, KS	5	Intercept	12.63	7.58	0.000
		N=1	1.19	0.78	0.436
		N=2	0.40	0.26	0.792
		N=3	-0.09	-0.06	0.955
		N=4	0.39	0.26	0.798
		N=5	0.00	-	-
		F-statistic[b]	0.22	-	0.927

Appendix Table 2. 1978-2008 Hard Wheat Preharvest Basis Forecasting Model Comparisons

Location	Optimal MA[a]	Effect	Estimate	t-Value	P-Value
Davis, OK	4	Intercept	12.37	2.29	0.033
		N=1	2.06	1.62	0.109
		N=2	1.51	1.19	0.239
		N=3	0.87	0.68	0.498
		N=4	0.09	0.07	0.945
		N=5	0.00	-	-
		F-statistic[b]	0.99	-	0.420
Dodge City, KS	5	Intercept	9.66	6.02	0.000
		N=1	1.35	1.04	0.304
		N=2	1.02	0.78	0.437
		N=3	0.33	0.26	0.799
		N=4	0.39	0.30	0.766
		N=5	0.00	-	-
		F-statistic[b]	0.36	-	0.839
El Reno, OK	1	Intercept	15.43	2.84	0.009
		N=1	0.55	0.50	0.618
		N=2	1.62	1.46	0.148
		N=3	0.52	0.47	0.641
		N=4	0.13	0.12	0.903
		N=5	0.00	-	-
		F-statistic[b]	0.66	-	0.625
El Dorado, OK	1	Intercept	15.20	3.29	0.003
		N=1	-1.08	-0.93	0.357
		N=2	0.57	0.49	0.623
		N=3	0.52	0.45	0.657
		N=4	-0.10	-0.09	0.930
		N=5	0.00	-	-
		F-statistic[b]	0.65	-	0.627
Emporia, KS	3	Intercept	13.65	6.25	0.000
		N=1	2.81	1.34	0.185
		N=2	0.85	0.41	0.685
		N=3	1.05	0.50	0.617
		N=4	0.61	0.29	0.771
		N=5	0.00	-	-
		F-statistic[b]	0.50	-	0.733
Frederick, OK	1	Intercept	16.93	3.12	0.005
		N=1	-0.54	-0.39	0.699
		N=2	0.49	0.35	0.729
		N=3	0.18	0.13	0.898
		N=4	-0.35	-0.25	0.805
		N=5	0.00	-	-
		F-statistic[b]	0.17	-	0.952
Garden City, KS	5	Intercept	10.90	5.44	0.000
		N=1	1.64	1.14	0.258
		N=2	0.43	0.30	0.768
		N=3	0.15	0.11	0.916
		N=4	0.14	0.10	0.922
		N=5	0.00	-	-
		F-statistic[b]	0.44	-	0.783

Appendix Table 2. 1978-2008 Hard Wheat Preharvest Basis Forecasting Model Comparisons

Location	Optimal MA[a]	Effect	Estimate	t-Value	P-Value
Geary, OK	1	Intercept	14.23	3.22	0.003
		N=1	0.09	0.09	0.925
		N=2	1.30	1.32	0.188
		N=3	0.53	0.54	0.589
		N=4	-0.21	-0.22	0.827
		N=5	0.00	-	-
		F-statistic[b]	0.75	-	0.561
Great Bend, KS	4	Intercept	10.73	6.52	0.000
		N=1	3.33	2.19	0.031
		N=2	2.28	1.50	0.137
		N=3	1.50	0.99	0.326
		N=4	0.34	0.23	0.822
		N=5	0.00	-	-
		F-statistic[b]	1.64	-	0.173
Gulf of Mexico	1	Intercept	13.72	4.92	0.000
		N=1	-1.85	-1.50	0.136
		N=2	-0.14	-0.11	0.912
		N=3	-0.57	-0.46	0.648
		N=4	-0.40	-0.32	0.747
		N=5	0.00	-	-
		F-statistic[b]	0.72	-	0.583
Hays, KS	5	Intercept	14.21	6.11	0.000
		N=1	5.51	3.02	0.003
		N=2	1.33	0.73	0.467
		N=3	1.03	0.57	0.573
		N=4	0.61	0.33	0.740
		N=5	0.00	-	-
		F-statistic[b]	2.89	-	0.027
Hobart, OK	1	Intercept	16.35	3.72	0.001
		N=1	-1.69	-1.42	0.157
		N=2	0.13	0.11	0.911
		N=3	-0.18	-0.15	0.878
		N=4	-0.51	-0.43	0.666
		N=5	0.00	-	-
		F-statistic[b]	0.76	-	0.551
Hutchinson, KS	1	Intercept	9.47	4.98	0.000
		N=1	1.16	0.83	0.408
		N=2	0.94	0.67	0.505
		N=3	0.94	0.67	0.505
		N=4	-0.07	-0.05	0.961
		N=5	0.00	-	-
		F-statistic[b]	0.34	-	0.847
Keyes, OK	1	Intercept	19.20	5.32	0.000
		N=1	-1.91	-1.29	0.201
		N=2	-1.64	-1.10	0.272
		N=3	-0.73	-0.49	0.623
		N=4	-0.20	-0.14	0.892
		N=5	0.00	-	-
		F-statistic[b]	0.65	-	0.625

Appendix Table 2. 1978-2008 Hard Wheat Preharvest Basis Forecasting Model Comparisons

Location	Optimal MA[a]	Effect	Estimate	t-Value	P-Value
Kingfisher, OK	4	Intercept	9.98	6.75	0.000
		N=1	-0.01	-0.01	0.996
		N=2	1.33	1.25	0.214
		N=3	0.32	0.31	0.760
		N=4	-0.36	-0.34	0.736
		N=5	0.00	-	-
		F-statistic[b]	0.74	-	0.568
Lawton, OK	1	Intercept	15.96	3.47	0.002
		N=1	-1.71	-1.29	0.200
		N=2	-0.09	-0.07	0.943
		N=3	-0.05	-0.04	0.971
		N=4	-0.40	-0.30	0.764
		N=5	0.00	-	-
		F-statistic[b]	0.59	-	0.669
Liberal, KS	1	Intercept	13.89	6.38	0.000
		N=1	-0.24	-0.15	0.879
		N=2	-0.13	-0.08	0.934
		N=3	-0.55	-0.35	0.726
		N=4	-0.28	-0.18	0.858
		N=5	0.00	-	-
		F-statistic[b]	0.03	-	0.998
Manchester, OK	1	Intercept	12.78	3.57	0.001
		N=1	-0.77	-0.78	0.435
		N=2	0.65	0.66	0.512
		N=3	0.13	0.13	0.897
		N=4	-0.46	-0.46	0.645
		N=5	0.00	-	-
		F-statistic[b]	0.62	-	0.651
Medford, OK	1	Intercept	12.38	3.47	0.002
		N=1	0.83	0.84	0.404
		N=2	1.45	1.47	0.145
		N=3	0.62	0.62	0.534
		N=4	-0.38	-0.38	0.701
		N=5	0.00	-	-
		F-statistic[b]	1.05	-	0.386
Miami, OK	4	Intercept	11.90	2.36	0.028
		N=1	1.32	1.44	0.154
		N=2	0.84	0.91	0.364
		N=3	0.83	0.91	0.365
		N=4	0.06	0.07	0.945
		N=5	0.00	-	-
		F-statistic[b]	0.76	-	0.555
Muskogee, OK	1	Intercept	21.52	4.40	0.002
		N=1	-6.44	-1.76	0.088
		N=2	-2.51	-0.68	0.499
		N=3	-1.20	-0.33	0.745
		N=4	-0.24	-0.07	0.947
		N=5	0.00	-	-
		F-statistic[b]	1.03	-	0.407

Appendix Table 2. 1978-2008 Hard Wheat Preharvest Basis Forecasting Model Comparisons

Location	Optimal MA[a]	Effect	Estimate	t-Value	P-Value
Okeene, OK	1	Intercept	13.56	3.22	0.003
		N=1	0.23	0.24	0.811
		N=2	1.57	1.64	0.103
		N=3	0.54	0.57	0.569
		N=4	-0.38	-0.40	0.691
		N=5	0.00	-	-
		F-statistic[b]	1.20	-	0.316
Pauls Valley, OK	4	Intercept	11.46	3.67	0.004
		N=1	2.27	1.13	0.263
		N=2	1.69	0.84	0.405
		N=3	0.87	0.43	0.666
		N=4	-0.15	-0.08	0.939
		N=5	0.00	-	-
		F-statistic[b]	0.54	-	0.705
Perryton, TX	3	Intercept	15.29	7.42	0.000
		N=1	-0.06	-0.04	0.969
		N=2	-0.61	-0.37	0.713
		N=3	-0.81	-0.49	0.626
		N=4	0.01	0.00	0.997
		N=5	0.00	-	-
		F-statistic[b]	0.11	-	0.979
Ponca City, OK	1	Intercept	13.36	3.86	0.001
		N=1	1.25	1.19	0.237
		N=2	1.76	1.68	0.096
		N=3	0.69	0.66	0.513
		N=4	-0.04	-0.03	0.972
		N=5	0.00	-	-
		F-statistic[b]	1.11	-	0.354
Pratt, KS	3	Intercept	9.94	6.27	0.000
		N=1	0.31	0.22	0.824
		N=2	-0.01	-0.01	0.995
		N=3	-0.14	-0.10	0.923
		N=4	0.23	0.17	0.868
		N=5	0.00	-	-
		F-statistic[b]	0.04	-	0.998
Salina, KS	2	Intercept	12.92	5.72	0.000
		N=1	2.33	1.21	0.230
		N=2	0.45	0.23	0.816
		N=3	1.75	0.91	0.367
		N=4	0.68	0.35	0.724
		N=5	0.00	-	-
		F-statistic[b]	0.50	-	0.735
Scott City, KS	3	Intercept	10.26	5.01	0.000
		N=1	2.11	1.49	0.141
		N=2	0.60	0.42	0.676
		N=3	0.26	0.18	0.854
		N=4	0.45	0.31	0.754
		N=5	0.00	-	-
		F-statistic[b]	0.68	-	0.608

Appendix Table 2. 1978-2008 Hard Wheat Preharvest Basis Forecasting Model Comparisons

Location	Optimal MA[a]	Effect	Estimate	t-Value	P-Value
Stillwater, OK	1	Intercept	15.71	2.40	0.030
		N=1	0.07	0.05	0.961
		N=2	1.21	0.84	0.407
		N=3	1.32	0.91	0.368
		N=4	0.20	0.14	0.888
		N=5	0.00	-	-
		F-statistic[b]	0.40	-	0.809
Temple, OK	1	Intercept	16.83	4.12	0.000
		N=1	-0.86	-0.62	0.536
		N=2	-0.05	-0.03	0.974
		N=3	0.21	0.15	0.877
		N=4	-0.35	-0.26	0.799
		N=5	0.00	-	-
		F-statistic[b]	0.18	-	0.948
Watonga, OK	1	Intercept	14.23	3.13	0.004
		N=1	0.12	0.12	0.905
		N=2	1.32	1.33	0.185
		N=3	0.42	0.42	0.672
		N=4	-0.36	-0.36	0.719
		N=5	0.00	-	-
		F-statistic[b]	0.82	-	0.515
Weatherford, OK	1	Intercept	14.97	3.49	0.002
		N=1	-0.85	-0.74	0.461
		N=2	0.77	0.67	0.502
		N=3	0.46	0.40	0.691
		N=4	-0.12	-0.11	0.915
		N=5	0.00	-	-
		F-statistic[b]	0.58	-	0.678
Wellington, KS	1	Intercept	10.29	5.23	0.000
		N=1	-0.55	-0.32	0.751
		N=2	0.14	0.08	0.937
		N=3	-0.01	-0.01	0.996
		N=4	-0.40	-0.23	0.819
		N=5	0.00	-	-
		F-statistic[b]	0.06	-	0.994
Whitewater, KS	2	Intercept	12.79	5.89	0.000
		N=1	0.79	0.46	0.647
		N=2	-0.12	-0.07	0.943
		N=3	0.23	0.13	0.896
		N=4	-0.05	-0.03	0.976
		N=5	0.00	-	-
		F-statistic[b]	0.09	-	0.984
Witchita, KS	1	Intercept	9.78	5.33	0.000
		N=1	-1.07	-0.75	0.459
		N=2	-0.39	-0.27	0.788
		N=3	-0.02	-0.02	0.986
		N=4	-0.28	-0.19	0.848
		N=5	0.00	-	-
		F-statistic[b]	0.18	-	0.946

Appendix Table 2. 1978-2008 Hard Wheat Preharvest Basis Forecasting Model Comparisons

Location	Optimal MA[a]	Effect	Estimate	t-Value	P-Value
Yukon, OK	5	Intercept	10.67	7.35	0.000
		N=1	0.34	0.28	0.779
		N=2	1.14	0.94	0.349
		N=3	1.01	0.84	0.402
		N=4	0.36	0.30	0.766
		N=5	0.00	-	-
		F-statistic[b]	0.32	-	0.862

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Afton, OK	Intercept	12.75	8.22	0.000
	N=1	-1.77	-1.34	0.182
	N=2	-0.84	-0.64	0.525
	N=3	-0.03	-0.02	0.981
	N=4	0.18	0.14	0.891
	N=5	0	-	-
	F-statistic[a]	0.77	-	0.549
Andale, KS	Intercept	13.83	6.97	0.000
	N=1	-0.81	-0.55	0.586
	N=2	-0.35	-0.24	0.815
	N=3	-0.91	-0.61	0.541
	N=4	-0.46	-0.31	0.760
	N=5	0	-	-
	F-statistic[a]	0.12	-	0.974
Banner, OK	Intercept	12.39	7.34	0.000
	N=1	-1.67	-1.46	0.147
	N=2	-1.53	-1.34	0.183
	N=3	-0.70	-0.61	0.543
	N=4	-0.07	-0.06	0.951
	N=5	0	-	-
	F-statistic[a]	0.95	-	0.439
Beloit, KS	Intercept	15.11	5.90	0.000
	N=1	-0.58	-0.29	0.771
	N=2	1.02	0.52	0.606
	N=3	0.65	0.33	0.742
	N=4	-0.12	-0.06	0.950
	N=5	0	-	-
	F-statistic[a]	0.21	-	0.933
Catoosa, OK	Intercept	18.43	4.88	0.000
	N=1	-4.25	-2.27	0.027
	N=2	-1.80	-0.96	0.341
	N=3	-0.29	-0.15	0.878
	N=4	0.22	0.11	0.909
	N=5	0	-	-
	F-statistic[a]	1.98	-	0.109

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Clinton, OK	Intercept	12.12	5.70	0.000
	N=1	-3.15	-2.81	0.006
	N=2	-1.92	-1.71	0.090
	N=3	-1.10	-0.98	0.327
	N=4	-0.14	-0.12	0.904
	N=5	0	-	-
	F-statistic[a]	2.73	-	0.032
Colby, KS	Intercept	15.49	4.83	0.000
	N=1	0.30	0.14	0.892
	N=2	1.83	0.81	0.418
	N=3	1.33	0.59	0.554
	N=4	0.36	0.16	0.874
	N=5	0	-	-
	F-statistic[a]	0.24	-	0.915
Davis, OK	Intercept	17.07	4.24	0.000
	N=1	-4.87	-3.39	0.001
	N=2	-2.79	-1.94	0.056
	N=3	-1.03	-0.72	0.475
	N=4	0.22	0.15	0.881
	N=5	0	-	-
	F-statistic[a]	4.43	-	0.003
Dodge City, KS	Intercept	17.22	6.12	0.000
	N=1	-3.11	-1.54	0.129
	N=2	-1.97	-0.98	0.331
	N=3	-2.63	-1.31	0.197
	N=4	-1.07	-0.53	0.598
	N=5	0	-	-
	F-statistic[a]	0.76	-	0.554
El Reno, OK	Intercept	11.68	6.90	0.000
	N=1	-1.17	-1.07	0.289
	N=2	-1.02	-0.93	0.353
	N=3	-0.46	-0.42	0.677
	N=4	0.06	0.06	0.956
	N=5	0	-	-
	F-statistic[a]	0.53	-	0.711

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Eldorado, OK	Intercept	12.00	6.31	0.000
	N=1	-2.69	-2.64	0.009
	N=2	-2.04	-2.01	0.047
	N=3	-1.21	-1.19	0.236
	N=4	-0.03	-0.02	0.980
	N=5	0	-	-
	F-statistic[a]	2.78	-	0.030
Emporia, KS	Intercept	16.94	6.26	0.000
	N=1	-2.38	-1.09	0.281
	N=2	-0.40	-0.18	0.856
	N=3	-1.42	-0.65	0.519
	N=4	-1.20	-0.55	0.586
	N=5	0	-	-
	F-statistic[a]	0.36	-	0.835
Frederick, OK	Intercept	13.06	4.96	0.000
	N=1	-3.29	-2.26	0.026
	N=2	-1.51	-1.04	0.300
	N=3	-0.52	-0.36	0.722
	N=4	0.07	0.05	0.964
	N=5	0	-	-
	F-statistic[a]	1.86	-	0.124
Garden City, KS	Intercept	16.98	6.16	0.000
	N=1	-1.93	-0.92	0.364
	N=2	-0.83	-0.39	0.695
	N=3	-1.92	-0.91	0.367
	N=4	-1.48	-0.70	0.487
	N=5	0	-	-
	F-statistic[a]	0.3	-	0.875
Geary, OK	Intercept	10.79	6.75	0.000
	N=1	-1.98	-2.06	0.042
	N=2	-1.48	-1.54	0.126
	N=3	-0.71	-0.74	0.461
	N=4	0.09	0.10	0.924
	N=5	0	-	-
	F-statistic[a]	1.78	-	0.137

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Great Bend, KS	Intercept	16.25	5.87	0.000
	N=1	-1.93	-0.97	0.334
	N=2	-1.05	-0.53	0.599
	N=3	-1.61	-0.81	0.420
	N=4	-1.39	-0.70	0.487
	N=5	0	-	-
	F-statistic[a]	0.28	-	0.890
Gulf of Mexico	Intercept	11.86	7.39	0.000
	N=1	-1.99	-1.66	0.100
	N=2	-1.34	-1.12	0.266
	N=3	-1.22	-1.02	0.312
	N=4	-0.60	-0.50	0.619
	N=5	0	-	-
	F-statistic[a]	0.8	-	0.529
Hays, KS	Intercept	15.84	5.92	0.000
	N=1	-2.59	-1.38	0.172
	N=2	-0.79	-0.42	0.676
	N=3	-1.71	-0.91	0.365
	N=4	-1.07	-0.57	0.568
	N=5	0	-	-
	F-statistic[a]	0.54	-	0.704
Hobart, OK	Intercept	11.33	5.32	0.000
	N=1	-1.96	-1.89	0.061
	N=2	-1.36	-1.31	0.192
	N=3	-0.63	-0.61	0.542
	N=4	-0.08	-0.07	0.942
	N=5	0	-	-
	F-statistic[a]	1.33	-	0.264
Hutchinson, KS	Intercept	13.96	5.62	0.000
	N=1	-1.34	-0.79	0.435
	N=2	-0.21	-0.12	0.904
	N=3	-0.67	-0.40	0.694
	N=4	-0.64	-0.37	0.710
	N=5	0	-	-
	F-statistic[a]	0.18	-	0.946

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Keyes, OK	Intercept	12.94	7.91	0.000
	N=1	-2.97	-2.57	0.011
	N=2	-1.69	-1.46	0.146
	N=3	-0.33	-0.29	0.774
	N=4	0.08	0.07	0.947
	N=5	0	-	-
	F-statistic[a]	2.61	-	0.039
Kingfisher, OK	Intercept	9.07	6.36	0.000
	N=1	-0.03	-0.03	0.976
	N=2	-0.43	-0.42	0.679
	N=3	-0.04	-0.04	0.967
	N=4	0.24	0.23	0.821
	N=5	0	-	-
	F-statistic[a]	0.11	-	0.980
Lawton, OK	Intercept	11.64	4.89	0.000
	N=1	-2.30	-1.81	0.073
	N=2	-1.19	-0.94	0.351
	N=3	-0.83	-0.66	0.514
	N=4	-0.10	-0.08	0.936
	N=5	0	-	-
	F-statistic[a]	1.08	-	0.369
Liberal, KS	Intercept	11.11	7.55	0.000
	N=1	-2.34	-1.99	0.050
	N=2	-1.43	-1.21	0.228
	N=3	-0.75	-0.64	0.523
	N=4	-0.33	-0.28	0.778
	N=5	0	-	-
	F-statistic[a]	1.26	-	0.294
Manchester, OK	Intercept	12.73	6.92	0.000
	N=1	-1.39	-1.07	0.288
	N=2	-0.95	-0.73	0.466
	N=3	-0.83	-0.64	0.523
	N=4	0.00	0.00	1.000
	N=5	0	-	-
	F-statistic[a]	0.45	-	0.774

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Medford, OK	Intercept	12.06	7.48	0.000
	N=1	-2.20	-1.93	0.055
	N=2	-1.46	-1.29	0.201
	N=3	-1.05	-0.92	0.360
	N=4	-0.01	-0.01	0.993
	N=5	0	-	-
	F-statistic[a]	1.40	-	0.238
Miami, OK	Intercept	15.76	7.83	0.000
	N=1	-3.51	-2.36	0.021
	N=2	-2.37	-1.59	0.115
	N=3	-1.30	-0.87	0.385
	N=4	-0.24	-0.16	0.873
	N=5	0	-	-
	F-statistic[a]	1.96	-	0.108
Muskogee, OK	Intercept	11.19	3.61	0.004
	N=1	4.52	1.92	0.060
	N=2	4.35	1.85	0.070
	N=3	1.91	0.81	0.420
	N=4	0.27	0.11	0.909
	N=5	0	-	-
	F-statistic[a]	1.69	-	0.168
Okeene, OK	Intercept	11.84	6.51	0.000
	N=1	-2.82	-2.54	0.013
	N=2	-2.01	-1.81	0.074
	N=3	-0.90	-0.81	0.419
	N=4	-0.09	-0.08	0.939
	N=5	0	-	-
	F-statistic[a]	2.45	-	0.050
Pauls Valley, OK	Intercept	6.28	5.82	0.000
	N=1	1.93	2.46	0.017
	N=2	0.23	0.29	0.772
	N=3	0.17	0.22	0.827
	N=4	0.02	0.02	0.981
	N=5	0	-	-
	F-statistic[a]	2.20	-	0.079

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Perryton, TX	Intercept	12.65	6.29	0.000
	N=1	-1.90	-1.38	0.173
	N=2	-0.99	-0.72	0.476
	N=3	-0.77	-0.56	0.580
	N=4	-0.54	-0.39	0.695
	N=5	0	-	-
	F-statistic[a]	0.51	-	0.729
Ponca City, OK	Intercept	12.14	7.29	0.000
	N=1	-2.25	-1.91	0.059
	N=2	-1.23	-1.04	0.300
	N=3	-0.68	-0.58	0.566
	N=4	-0.18	-0.15	0.879
	N=5	0	-	-
	F-statistic[a]	1.18	-	0.321
Pratt, KS	Intercept	14.57	5.88	0.000
	N=1	-2.62	-1.41	0.163
	N=2	-1.29	-0.69	0.490
	N=3	-1.98	-1.07	0.291
	N=4	-1.31	-0.71	0.483
	N=5	0	-	-
	F-statistic[a]	0.55	-	0.698
Salina, KS	Intercept	13.83	6.28	0.000
	N=1	-2.25	-1.34	0.184
	N=2	-1.01	-0.61	0.547
	N=3	-1.98	-1.18	0.241
	N=4	-1.08	-0.64	0.522
	N=5	0	-	-
	F-statistic[a]	0.57	-	0.687
Scott City, KS	Intercept	15.36	5.93	0.000
	N=1	-2.24	-1.14	0.259
	N=2	-0.84	-0.43	0.670
	N=3	-1.96	-1.00	0.322
	N=4	-1.24	-0.63	0.529
	N=5	0	-	-
	F-statistic[a]	0.42	-	0.795

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Stillwater, OK	Intercept	16.98	7.22	0.000
	N=1	-3.91	-1.94	0.057
	N=2	-1.61	-0.80	0.426
	N=3	-0.27	-0.13	0.895
	N=4	0.26	0.13	0.897
	N=5	0	-	-
	F-statistic[a]	1.47	-	0.221
Temple, OK	Intercept	12.33	4.66	0.000
	N=1	-3.60	-2.75	0.007
	N=2	-2.34	-1.79	0.077
	N=3	-1.24	-0.94	0.347
	N=4	0.02	0.01	0.991
	N=5	0	-	-
	F-statistic[a]	2.83	-	0.029
Watonga, OK	Intercept	11.00	6.41	0.000
	N=1	-1.79	-1.72	0.088
	N=2	-1.26	-1.21	0.228
	N=3	-0.48	-0.46	0.649
	N=4	0.03	0.03	0.978
	N=5	0	-	-
	F-statistic[a]	1.19	-	0.319
Weatherford, OK	Intercept	12.43	6.28	0.000
	N=1	-3.51	-3.08	0.003
	N=2	-2.07	-1.82	0.072
	N=3	-1.08	-0.95	0.346
	N=4	-0.22	-0.19	0.849
	N=5	0	-	-
	F-statistic[a]	3.21	-	0.015
Wellington, KS	Intercept	16.12	4.65	0.002
	N=1	-2.56	-0.95	0.351
	N=2	-1.81	-0.67	0.508
	N=3	-1.31	-0.48	0.632
	N=4	-1.03	-0.38	0.706
	N=5	0	-	-
	F-statistic[a]	0.25	-	0.909

Appendix Table 3. 1978-2008 Hard Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Whitewater, KS	Intercept	14.50	5.69	0.000
	N=1	-0.40	-0.21	0.836
	N=2	-0.45	-0.23	0.817
	N=3	-2.23	-1.15	0.256
	N=4	-1.48	-0.76	0.451
	N=5	0	-	-
	F-statistic[a]	0.45	-	0.775
Witchita, KS	Intercept	12.96	5.48	0.000
	N=1	0.64	0.37	0.710
	N=2	0.78	0.46	0.648
	N=3	-0.42	-0.25	0.807
	N=4	-0.65	-0.38	0.705
	N=5	0	-	-
	F-statistic[a]	0.27	-	0.893
Yukon, OK	Intercept	9.48	7.23	0.000
	N=1	0.66	0.65	0.517
	N=2	-0.16	-0.16	0.873
	N=3	-0.03	-0.03	0.979
	N=4	0.24	0.23	0.816
	N=5	0	-	-
	F-statistic[a]	0.20	-	0.936

Appendix Table 4. 1976-2008 Soft Wheat Preharvest Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Chicago, IL	Intercept	13.73	4.81	0.000
	N=1	-1.92	-1.64	0.104
	N=2	-1.18	-1.00	0.317
	N=3	-0.42	-0.36	0.718
	N=4	-0.02	-0.02	0.988
	N=5	0	-	-
	F-statistic[a]	1.00		0.410
St. Louis, MO	Intercept	35.13	4.02	0.000
	N=1	6.20	1.73	0.087
	N=2	0.47	0.13	0.897
	N=3	0.27	0.07	0.941
	N=4	0.73	0.20	0.840
	N=5	0	-	-
	F-statistic[a]	1.07		0.374
Toledo, OH	Intercept	28.60	4.56	0.000
	N=1	4.18	1.11	0.270
	N=2	0.85	0.23	0.822
	N=3	0.89	0.23	0.815
	N=4	0.82	0.22	0.827
	N=5	0	-	-
	F-statistic[a]	0.37		0.828

[a] The null hypothesis is that all values of N have the same forecast accuracy.

Appendix Table 5. 1975-2008 Soft Wheat Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Chicago, IL	Intercept	13.62	8.41	0.000
	N=1	-0.73	-0.52	0.604
	N=2	-1.58	-1.12	0.264
	N=3	-0.43	-0.30	0.761
	N=4	0.67	0.48	0.634
	N=5	0	-	-
	F-statistic[a]	0.71	-	0.587
St. Louis, MO	Intercept	16.37	7.28	0.000
	N=1	-1.06	-0.60	0.547
	N=2	-1.60	-0.91	0.365
	N=3	-0.17	-0.10	0.924
	N=4	0.39	0.22	0.823
	N=5	0	-	-
	F-statistic[a]	0.43	-	0.784
Toledo, OH	Intercept	9.27	3.68	0.001
	N=1	5.11	3.71	0.000
	N=2	1.09	0.79	0.431
	N=3	1.57	1.14	0.256
	N=4	0.45	0.33	0.745
	N=5	0	-	-
	F-statistic[a]	4.34	-	0.003

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

Appendix Table 6. 1980-2008 Corn Preharvest Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Northern, IL	Intercept	13.21	8.97	0.000
	N=1	-1.90	-1.46	0.148
	N=2	-0.62	-0.47	0.636
	N=3	-0.25	-0.20	0.845
	N=4	0.06	0.05	0.962
	N=5	0	-	-
	F-statistic[a]	0.76	-	0.551
Western, IL	Intercept	11.54	7.69	0.000
	N=1	-0.47	-0.36	0.718
	N=2	0.55	0.42	0.673
	N=3	0.47	0.36	0.716
	N=4	0.43	0.33	0.742
	N=5	0	-	-
	F-statistic[a]	0.22	-	0.927
North Central, IL	Intercept	10.08	7.37	0.000
	N=1	-0.44	-0.40	0.688
	N=2	0.38	0.34	0.732
	N=3	0.23	0.21	0.836
	N=4	0.48	0.44	0.659
	N=5	0	-	-
	F-statistic[a]	0.22	-	0.924
South Central, IL	Intercept	10.37	8.11	0.000
	N=1	-0.11	-0.10	0.923
	N=2	0.80	0.69	0.489
	N=3	0.62	0.54	0.592
	N=4	0.53	0.46	0.646
	N=5	0	-	-
	F-statistic[a]	0.24	-	0.914
Wabash, IL	Intercept	12.90	7.09	0.000
	N=1	0.26	0.18	0.858
	N=2	1.01	0.70	0.482
	N=3	1.21	0.84	0.401
	N=4	0.71	0.49	0.622
	N=5	0	-	-
	F-statistic[a]	0.25	-	0.910

Appendix Table 6. 1980-2008 Corn Preharvest Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
West-Southwest, IL	Intercept	12.34	6.74	0.000
	N=1	0.59	0.44	0.659
	N=2	1.09	0.82	0.413
	N=3	0.47	0.35	0.726
	N=4	0.57	0.43	0.669
	N=5	0	-	-
	F-statistic[a]	0.17	-	0.953
Little Egypt, IL	Intercept	11.77	6.59	0.000
	N=1	1.21	0.92	0.357
	N=2	1.22	0.93	0.352
	N=3	1.14	0.87	0.385
	N=4	0.82	0.63	0.531
	N=5	0	-	-
	F-statistic[a]	0.31	-	0.869

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

Appendix Table 7. 1980-2008 Corn Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Northern, IL	Intercept	8.48	6.53	0.000
	N=1	-1.53	-1.85	0.067
	N=2	-1.18	-1.43	0.156
	N=3	-0.99	-1.20	0.232
	N=4	-0.21	-0.25	0.802
	N=5	0.00	-	-
	F-statistic[a]	1.25	-	0.296
Western, IL	Intercept	7.40	5.88	0.000
	N=1	-1.28	-1.87	0.064
	N=2	-1.03	-1.49	0.138
	N=3	-0.84	-1.23	0.222
	N=4	-0.24	-0.35	0.726
	N=5	0.00	-	-
	F-statistic[a]	1.24	-	0.300
North Central, IL	Intercept	7.62	6.81	0.000
	N=1	-0.68	-0.92	0.360
	N=2	-0.78	-1.05	0.295
	N=3	-0.94	-1.26	0.209
	N=4	-0.37	-0.50	0.620
	N=5	0.00	-	-
	F-statistic[a]	0.51	-	0.731
South Central, IL	Intercept	7.52	7.06	0.000
	N=1	-1.69	-2.44	0.016
	N=2	-0.76	-1.09	0.278
	N=3	-0.49	-0.71	0.480
	N=4	-0.15	-0.22	0.827
	N=5	0.00	-	-
	F-statistic[a]	1.86	-	0.122
Wabash, IL	Intercept	7.11	6.47	0.000
	N=1	-1.06	-1.36	0.175
	N=2	-0.11	-0.14	0.891
	N=3	-0.28	-0.36	0.717
	N=4	-0.03	-0.04	0.970
	N=5	0.00	-	-
	F-statistic[a]	0.64	-	0.632
West-Southwest, IL	Intercept	7.34	6.99	0.000
	N=1	-1.18	-1.67	0.097
	N=2	-0.80	-1.13	0.261
	N=3	-0.55	-0.78	0.436
	N=4	-0.24	-0.34	0.736
	N=5	0.00	-	-

Appendix Table 7. 1980-2008 Corn Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
	F-statistic[a]	0.86	-	0.490
Little Egypt, IL	Intercept	6.96	6.75	0.000
	N=1	-0.72	-1.03	0.305
	N=2	-0.12	-0.17	0.869
	N=3	-0.50	-0.72	0.475
	N=4	-0.02	-0.03	0.972
	N=5	0.00	-	-
	F-statistic[a]	0.42	-	0.792

[a] The null hypothesis is that all values of *N* have the same forecast accuracy.

Appendix Table 8. 1980-2008 Soybean Preharvest Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Northern, IL	Intercept	10.53	7.76	0.00
	N=1	-1.30	-1.17	0.24
	N=2	-0.09	-0.08	0.94
	N=3	0.01	0.01	0.99
	N=4	-0.07	-0.06	0.95
	N=5	0	-	-
	F-statistic[a]	-	0.52	0.72
Western, IL	Intercept	11.57	6.23	0.00
	N=1	0.06	0.04	0.97
	N=2	-0.74	-0.59	0.56
	N=3	-0.24	-0.19	0.85
	N=4	-0.24	-0.19	0.85
	N=5	0	-	-
	F-statistic[a]	-	0.12	0.97
North Central, IL	Intercept	9.48	7.34	0.00
	N=1	-0.70	-0.73	0.46
	N=2	-0.52	-0.55	0.59
	N=3	-0.40	-0.42	0.68
	N=4	-0.20	-0.21	0.83
	N=5	0	-	-
	F-statistic[a]	-	0.16	0.96
South Central, IL	Intercept	9.65	7.02	0.00
	N=1	0.04	0.04	0.97
	N=2	0.05	0.05	0.96
	N=3	0.07	0.06	0.95
	N=4	-0.34	-0.31	0.75
	N=5	0	-	-
	F-statistic[a]	-	0.05	1.00
Wabash, IL	Intercept	13.07	7.98	0.00
	N=1	-1.38	-1.16	0.25
	N=2	-1.46	-1.23	0.22
	N=3	-1.51	-1.27	0.21
	N=4	-0.34	-0.29	0.77
	N=5	0	-	-
	F-statistic[a]	-	0.72	0.58

Appendix Table 8. 1980-2008 Soybean Preharvest Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
West-Southwest, IL	Intercept	11.76	5.46	0.00
	N=1	0.37	0.26	0.80
	N=2	-0.09	-0.06	0.95
	N=3	-0.18	-0.13	0.90
	N=4	0.24	0.17	0.87
	N=5	0	-	-
	F-statistic[a]	-	0.05	0.99
Little Egypt, IL	Intercept	12.54	6.59	0.00
	N=1	-0.40	-0.35	0.73
	N=2	-1.43	-1.23	0.22
	N=3	-1.28	-1.10	0.27
	N=4	-0.42	-0.36	0.72
	N=5	0	-	-
	F-statistic[a]	0.56	-	0.69

[a] The null hypothesis is that all values of N have the same forecast accuracy.

Appendix Table 9. 1980-2008 Soybean Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
Northern IL	Intercept	10.11	7.91	0.00
	N=1	-1.94	-2.20	0.03
	N=2	-1.30	-1.47	0.14
	N=3	-0.80	-0.90	0.37
	N=4	-0.13	-0.15	0.88
	N=5	-	-	-
	F-statistic[a]	-	1.69	0.16
Western IL	Intercept	8.84	6.99	0.00
	N=1	-2.24	-2.51	0.01
	N=2	-0.96	-1.08	0.28
	N=3	-0.72	-0.81	0.42
	N=4	-0.02	-0.02	0.98
	N=5	-	-	-
	F-statistic[a]	-	2.11	0.08
North Central IL	Intercept	10.02	7.71	0.00
	N=1	-2.14	-2.34	0.02
	N=2	-1.19	-1.30	0.19
	N=3	-0.82	-0.90	0.37
	N=4	-0.15	-0.17	0.87
	N=5	-	-	-
	F-statistic[a]	-	1.79	0.14
South Central IL	Intercept	9.50	6.84	0.00
	N=1	-1.86	-1.97	0.05
	N=2	-0.74	-0.78	0.44
	N=3	-0.18	-0.19	0.85
	N=4	-0.04	-0.04	0.96
	N=5	-	-	-
	F-statistic[a]	-	1.37	0.25
Wabash IL	Intercept	10.27	6.95	0.00
	N=1	-2.11	-2.21	0.03
	N=2	-1.34	-1.41	0.16
	N=3	-0.78	-0.81	0.42
	N=4	-0.14	-0.15	0.88
	N=5	-	-	-
	F-statistic[a]	-	1.68	0.16
West-Southwest IL	Intercept	9.04	6.01	0.00
	N=1	-2.19	-2.32	0.02
	N=2	-1.08	-1.15	0.25
	N=3	-0.58	-0.62	0.54
	N=4	0.04	0.04	0.97
	N=5	-	-	-

Appendix Table 9. 1980-2008 Soybean Storage Basis Forecasting Model Comparisons

Location	Effect	Estimate	t-value	p-value
	F-statistic[a]	-	1.90	0.12
Little Egypt IL	Intercept	9.47	7.62	0.00
	N=1	-1.40	-1.60	0.11
	N=2	-1.49	-1.71	0.09
	N=3	-0.74	-0.84	0.40
	N=4	-0.09	-0.10	0.92
	N=5	-	-	-
	F-statistic[a]	-	1.30	0.28

[a] The null hypothesis is that all values of N have the same forecast accuracy.

VITA

Robert Braden Hatchett

Candidate for the Degree of

Masters of Science

Thesis: LENGTH OF OPTIMAL MOVING AVERAGES TO USE WHEN FORECASTING BASIS.

Major Field: Agricultural Economics

Biographical:

Personal Data: Born in Bartlesville, Oklahoma on September 13,1984, son of John and Karen Hatchett, and brother of Mason and Philip Hatchett.

Education: Graduated from Bartlesville High School, Bartlesville, Oklahoma in May 2003, received a Bachelor of Science and Arts degree in Agribusiness from the University of Arkansas in May, 2007, and completed the requirements for the Master of Science in Agricultural Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2009.

Experience: Data Analysis Intern, The Corporate Executive Board; June-August 2006, Graduate Research Assistant, Oklahoma State University, Department of Agricultural Economics, August 2007 to present.

Professional Memberships: None.