An Examination of Spatial and Intertemporal Aspects of Basis Determination

by

Maximilian Strobl, T. Randall Fortenbery, and Paul Fackler

Suggested citation format:

Strobl, M., T. R. Fortenbery, and P. Fackler. 1992. “An Examination of Spatial and Intertemporal Aspects of Basis Determination.” Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL. [http://www.farmdoc.uiuc.edu/nccc134].

An Examination of the Spatial and Intertemporal Aspects of Basis Determination

Maximilian Strobl, T. Randall Fortenbery, and Paul Fackler*

Basis describes the relationship between a commodity's cash price at a particular location and its simultaneously quoted futures price. Both cash and futures prices are influenced by local and national changes in supply and demand. The cash price reflects local conditions and transportation differentials, and the futures price reflects the consensus of a multitude of different market agents about the current value of a commodity. Basis is generally defined as the cash price minus the nearby futures price. Thus, a negative basis implies a cash price below the futures price, and a positive basis a cash price above futures.[1]

Understanding and anticipating basis changes is an important aspect in the pricing of agricultural products. Basis information can aid in both the timing and the selection of specific pricing strategies. An understanding of basis allows economic agents to evaluate the attractiveness of various forward pricing opportunities. For example, if an elevator offers a weaker basis in a forward cash contract for corn than a producer expects will exist at harvest, then the forward cash contract may not be as attractive a pricing opportunity as a direct futures market hedge.[2] Conversely, if the elevator offers a basis which is stronger than that expected to exist later, than the forward cash contract is a more attractive pricing opportunity than the futures hedge. The challenge to the producer is to formulate an expectation, or forecast, of the harvest basis. Without a reasonable forecast, an objective comparison of the forward contract with the hedge is not possible.

This paper focuses on the forecasting problem. Specifically, we wish to identify a basis forecasting model which can be applied with relative ease to a local market. We proceed with a short discussion of previous basis work, outline our objectives, data, and empirical procedure, and then present the results and conclusions.

---

* The authors are a Graduate Student, and Assistant Professors, respectively, in the Department of Agricultural and Resource Economics at North Carolina State University, Raleigh.

[1] The academic literature defines the basis either as futures minus cash or as cash minus futures price, the former following the convention established by H. Working. The commercial practice is to define the basis as cash minus futures price (Tomek & Robinson, 1990).

[2] A weak basis is one which exhibits a low cash price relative to futures, while a strong basis has a high cash price relative to futures. The terms weak and strong imply nothing about the actual sign of the basis: A basis, for example, of -0.10 (meaning the cash price is 10 cents below the futures) is said to be stronger than a basis of -0.20, but weaker than a basis of -0.05.

Literature Review

Theoretical studies of basis for seasonally produced, storable commodities have identified several components critical to basis determination. Two major determinants of basis relate to the intertemporal and spatial relationships between the cash and futures markets being considered. As such, basis is, in large part, a reflection of the costs associated with moving a physical commodity through time (i.e. storage) and the costs of moving a commodity from one market to another (transportation costs).

Working's paper, followed by related work from both Brennan and Telser, provided a theoretical foundation for the intertemporal aspect of basis. Working argued that intertemporal basis reflects the price of storage, and is determined by the supply and demand for storage space. The price of storage was hypothesized to include three components, 1) physical storage costs (warehouse charges, insurance, and opportunity costs), 2) a risk premium, and 3) a convenience yield. The risk premium represented the reward paid to futures speculators for offsetting hedger's risk, and the convenience yield explained why some minimum level of stocks might be held even in the absence of a positive return to the storage activity.[3]

The spatial aspect of basis has been explicitly recognized in several applied studies (Kahl and Curtis, and Martin, Groenewegen, and Pidgeon, for example), but has not induced as much theoretical discussion as the intertemporal component. This may be due in part to the less controversial nature of spatial basis. Most analysts can agree that spatial basis is influenced by transportation costs, and the subject is not encumbered by less clear issues such as risk premiums and convenience yields.

Most applied work directed at quantifying the parameters of basis determination handle the intertemporal and spatial components of basis simultaneously, while also introducing other basis arguments. These have included a market liquidity measure (first introduced by Gray), and a myriad of factors intended to account for the unique characteristics of individual markets.

Examples of such work include Ward and Dasse. They introduced a freeze bias variable in their structural modeling of the basis for frozen concentrated orange juice. They found the freeze bias variable to be significant, and concluded that basis models need to account for the unique characteristics of individual markets being considered.

Similarly, Martin, Groenewegen,and Pidgeon hypothesized that Ontario corn basis was influenced by different economic factors at different times of the year. To account for this they estimated separate models of basis for each month. In addition, they included both the local and U.S. crop size as explanatory variables, and found these to be the dominating factors determining basis in harvest months.

Garcia and Good explored the determinants of corn basis in different areas in Illinois.

---

[3] Empirical and theoretical studies on basis relationships have questioned the existence of both a risk premium (e.g. Telser, Gray), and a convenience yield (Wright & Williams). These topics are still controversial.

They also used the local crop size as an explanatory variable. In addition, they introduced the price level to help explain basis. They argued that farmers market more of their crop when the price level is high, and this would tend to lead to a weaker basis.

Kahl and Curtis (1986, 1988) investigated corn basis in South Carolina and Illinois. They also used the price level as an explanatory variable, but in contrast to Garcia and Good, they argued that a high local cash price is associated with a small local crop and large storage capacity relative to stocks. This leads elevators to narrow their margins and bid up the cash price relative to the futures price to attract additional stocks, i.e. the basis becomes stronger.

Explicit models for forecasting basis have not generally been based on the econometric specifications discussed above, but rather on time series analysis. One reason is the ease with which simple time series models can be employed by market agents. The Chicago Board of Trade (CBOT) has suggested that historical basis is an important tool in anticipating basis in future years, and recommends using moving averages of historical basis as future forecasts. This approach is common in marketing education programs conducted through Cooperative Extension Services. Examples include Curtis, Harris, and Miller, and Fortenbery. However, Kahl and Curtis noted that despite the annually repeating pattern in basis movements, there can be considerable variance from year to year and from location to location. Kenyon and Blakely made a similar observation. This suggests that more rigorous time series models may improve forecasting performance.

Lacking from the basis forecasting literature has been an empirical evaluation of the time series models generally employed, and the specifications of alternate time series models has not been explored.

Objectives and Data

The objective of this study is to improve on the forecasting accuracy of basis models by explicitly exploring the time series properties of basis movement. We use standard diagnostics to identify appropriate time series structures, and compare forecast performance with the traditional simple moving average models. A secondary objective is to keep the time series specification as simple as possible in order to insure it has value to market participants as a basis predictor. We do not consider an econometric forecasting model here for two reasons. First, econometric models can be very difficult to employ by market agents due to the data requirements and statistical skills necessary for implementation. Second, previous econometric specifications have involved independent variables which are determined simultaneously with basis. Attempts to forecast the independent variables would introduce an additional source of error in the econometric specification, and thus potentially reduce the accuracy of the forecast.

We use two different approaches to specifying the time series model. The first measures basis as a relationship between the cash market and nearby Chicago futures. This approach treats the spatial and intertemporal components simultaneously, and is consistent with previous work. Our second approach separates the spatial and intertemporal components, resulting in two separate models. The forecasts from these models are then combined to generate a forecast of the local cash versus futures market basis. This allows an evaluation of the forecast performance of the joint basis models relative to the combined forecasts of disaggregated models. The approach is utilized to specifically address the question of whether storage and transportation influence basis differently, and aggregating them in model specification results in a loss of potentially valuable information. The criteria used to evaluate the different models are out-of sample mean absolute deviation and root mean square forecast error statistics.

The models are used to generate forecasts of the Greenville, North Carolina soybean basis. We utilize cash soybean prices in Greenville, and futures prices for the nearby soybean contract traded at the CBOT. The disaggregated models use Greenville cash prices relative to Toledo, Ohio cash prices for the spatial component. Toledo was chosen because it is a soybean cash delivery point for the CBOT soybean futures market and because its more active cash market should provide more accurate cash price quotes.[4] It represents the St. Lawrence export market. The intertemporal component is modeled using the Toledo cash price and the nearby Chicago soybean futures price.

The price data for initial estimation consists of monthly observations from January 1978 to December 1987 (120 observations). The monthly futures prices are the arithmetic mean of the Thursday closing price for each week of a month.[5] The Greenville cash prices were obtained from the Market Grain Report which reports cash prices at different locations in North Carolina and is published weekly by the North Carolina Department of Agriculture[6]. The Toledo cash price series was collected from a grain elevator. The Toledo monthly average price is calculated as the average of the Thursday closing cash prices.[7]

Empirical Analysis

The econometric analyses discussed earlier attempted to identify and quantify the economic factors which determine basis (local and delivery point) to assist economic agents in their marketing decisions. But they are difficult to apply for basis forecasting. In previous work there is a gap between very simple moving average forecasting models and theoretically sound but difficult to apply econometric models. This study attempts to fill this gap by developing ARIMA models for local basis prediction.

Two different methods are applied to develop a forecasting model. The first method is a "pure", univariate ARIMA time series approach (Box-Jenkins model). This type of model uses

---

[4] Investigations of basis convergence during the delivery months suggest that Toledo, not Chicago, is the more relevant cash market from the perspective of the futures contract (Peck and Williams).

[5] Futures price data was obtained from the "Technical Tools" data base.

[6] The data is obtained in a weekly survey of elevators and processors. The cash prices are obtained after the closing of trade at CBOT, and are quoted for immediate delivery.

[7] The data was obtained from the Grain Marketing Basis Analysis Program (GMBA) from the Ohio Cooperative Extension Service at Ohio State University.

only weighted lagged values of the dependent variable and a weighted sum of current and lagged random disturbances to explain current basis. In the second method seasonality in basis is modeled by using a set of monthly dummy variables. Subsequently the residuals of the dummy variable model are analyzed in the same way as the "pure" ARIMA models for the first method. Finally the seasonal dummy variables and the ARIMA model for the residuals are combined in one model. This approach is a good alternative to "pure" Box-Jenkins models if the seasonality is highly regular (Brocklebank & Dickey).

Both models are applied to estimate the so-called joint basis (local nearby basis), and the decomposed basis parts, the spatial basis and the intertemporal basis.

Time Series Analysis

The classical Box-Jenkins approach is applied to develop the ARIMA models. It consists of the identification, estimation, and diagnostic checks of the models. If the diagnostics do not support the selected model the whole process is repeated. In the identification process plots of the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) are used to identify the underlying correlation structure in the data and to select an appropriate model. Subsequently the model is estimated using the maximum-likelihood method. To check if the model is adequate, t values for the coefficients of the estimated model and a Q test of autocorrelation among the residuals of the fitted model for lags 6, 12, 18, and 24 are used as diagnostic tools.[8] If several models adequately represent the identified correlation structure, Schwartz's Bayesian criterion (SBC) is applied to select the more parsimonious model.[9] As a final check overfitting through adding additional moving average (MA) or autoregressive (AR) terms is used to test if the fitted model is appropriate.

For the seasonal dummy variable model (SEASON), OLS is used to estimate a preliminary model. The preliminary model is estimated with an intercept. October is chosen as the base month for the dummy variable structure and is therefore omitted. The residuals of the preliminary model are then analyzed in the same way as the ARIMA models. After identification of the correlation structure the model is re-estimated with the monthly dummy variables and the identified ARIMA structure as explanatory variables. For the final model an F test is conducted to test the significance of the monthly dummy variable set. For the decomposed basis models a linear time trend proved to be significant in the preliminary models and is thus included. The estimated models and relevant statistics are presented in table 1.

---

[8] The Q statistic has a qui-squared distribution and is used to test for serial correlation in the model residuals. If a model is not appropriate the overall value of Q will be inflated. The high value indicates that an inappropriate model has been identified. This test statistic has been used by Box and Jenkins to test how well estimated models fit data. Ljung and Box modified the statistic for the case of small samples. The Q definition of Ljung and Box used in this study differs in the weighing which is applied to the autocorrelations, giving more emphasis to later correlations (Nicholls, 1979).

[9] SBC penalizes for overfitting. Over-parameterization increases the numerical value of SBC and urges one to choose a relatively more parsimonious model.

Table 1. Estimation Results

* significance at the 95% level.
** significance at the 99% level.

(1) Joint Basis

(a) AR(1) x $SAR_{12}(1)$:

$$\underset{(0.0723)}{(1 - 0.6197\,B)}\ \underset{(0.0759)}{(1 - 0.5573\,B^{12})}\ BAS_t = -0.0126^{*} + e_t$$

$Q_6 = 6.25$
$Q_{12} = 16.58$
$Q_{18} = 23.66$
$Q_{24} = 32.48$

(b) Seasonal dummy variables + AR(1) :

$$\underset{(0.0349)}{(1 - 0.6579\,B)}\ BAS_t = -0.0698^{*}$$

$$+ \underset{(0.0287)}{0.0821\,NOV} + \underset{(0.0368)}{0.0692\,DEC} + \underset{(0.0414)}{0.1003\,JAN}$$

$$+ \underset{(0.0440)}{0.1664\,FEB} + \underset{(0.0452)}{0.1363\,MAR} + \underset{(0.0455)}{0.1538\,APR}$$

$$+ \underset{(0.0451)}{0.1410\,MAY} + \underset{(0.0438)}{0.1831\,JUN} + \underset{(0.0413)}{0.2250\,JUL}$$

$$+ \underset{(0.0369)}{0.2343\,AUG} + \underset{(0.0288)}{0.1720\,SEP} + e_t$$

$Q_6 = 7.23$
$Q_{12} = 14.54$
$Q_{18} = 19.63$
$Q_{24} = 32.48$

$F = 7.805^{**}$

(2) Spatial Basis

(a) ARMA(1,1) :

$$\underset{(0.0763)}{(1 - 0.8911\,B)}\ SPBAS_t = 0.0045 + \underset{(0.1268)}{(1 - 0.5990\,B)}\ e_t$$

$Q_6 = 5.60$
$Q_{12} = 6.41$
$Q_{18} = 12.71$
$Q_{24} = 17.39$

Table 1. Estimation Results (continued)

(b) seasonal dummy variables + AR(1) x $SAR_{12}(1)$:

$$\underset{(0.0913)}{(1 - 0.3641\,B)}\ \underset{(0.1075)}{(1 + 0.2687\,B^{12})}\ SPBAS_t = 0.0719^{*}$$

$$+ \underset{(0.0282)}{0.0070}\,NOV + \underset{(0.0329)}{0.0119}\,DEC + \underset{(0.0346)}{0.0635}\,JAN$$

$$+ \underset{(0.0352)}{0.0732}\,FEB + \underset{(0.0353)}{0.0885}\,MAR + \underset{(0.0354)}{0.0618}\,APR$$

$$+ \underset{(0.0353)}{0.0672}\,MAY + \underset{(0.0351)}{0.0463}\,JUN + \underset{(0.0347)}{0.0273}\,JUL$$

$$+ \underset{(0.0330)}{0.0141}\,AUG + \underset{(0.0282)}{0.1402}\,SEP + \underset{(0.0003)}{0.0016}\,T + e_t$$

$Q_6 = 5.10$
$Q_{12} = 6.32$ $\qquad F = 3.157^{**}$
$Q_{18} = 14.81$
$Q_{24} = 18.98$

3. Intertemporal Basis

(a) AR(2) x $SAR_{12}(1)$:

$$(1 - \underset{(0.0911)}{0.5057}\,B \ \underset{(0.0905)}{-0.1845}\,B^2)\ (1 - \underset{(0.0930)}{0.4144}\,B^{12})\ INTBAS_{t,d}$$

$$= -0.0210^{*} + e_t$$

$Q_6 = 6.52$
$Q_{12} = 12.98$
$Q_{18} = 14.90$
$Q_{24} = 21.58$

(b) Seasonal dummy variables + AR(2) :

$$(1 - \underset{(0.0955)}{0.4707}\,B \ \underset{(0.0963)}{- 0.2582}\,B^2)\ INTBAS_{t,d} = -0.0841^{*}$$

$$- \underset{(0.0342)}{0.0877}\,NOV + \underset{(0.0373)}{0.0568}\,DEC + \underset{(0.0425)}{0.0449}\,JAN$$

$$+ \underset{(0.0445)}{0.0968}\,FEB + \underset{(0.0459)}{0.0511}\,MAR + \underset{(0.0462)}{0.0965}\,APR$$

$$+ \underset{(0.0457)}{0.0734}\,MAY + \underset{(0.0442)}{0.1371}\,JUN + \underset{(0.0420)}{0.2049}\,JUL$$

$$+ \underset{(0.0373)}{0.2212}\,AUG + \underset{(0.0342)}{0.0297}\,SEP - \underset{(0.0008)}{0.0018}\,T + e_t$$

$Q_6 = 6.34$
$Q_{12} = 11.61$ $\qquad F = 6.790^{**}$
$Q_{18} = 18.13$
$Q_{24} = 21.58$

Forecasting and Evaluation

The main objective of the forecasting models is to get accurate short term forecasts of nearby basis. These will provide a valuable tool for marketing decisions. To accomplish this, forecast horizons from 1 to 6 months are chosen, i.e. in each stage of the forecasting process forecasts from one to six steps ahead are made. For each forecasting horizon and each model a set of 30 forecasts is estimated. These 30 forecasts provide the basis for evaluation of the different forecasting models.

One assumption of the forecasting process is that the basic structure of the fitted ARIMA and seasonal dummy variable models does not change during the forecasting period. Thus, it is assumed that the identified and modeled correlation structure represents the true model and does not change over time. However, we allow for the coefficients of the models to change over time. This is accomplished through a continuous updating of the sample period for the fitted model and a re-estimation of the model coefficients after each update. We also test the relevance of old information in the forecasting process as new information becomes available. The question is whether old data still makes an important contribution to the forecasting process or whether it should be discounted. A simple filter is introduced in the prediction to address this question. Each forecasting model is estimated both with the full data set including the update of the data after each forecast, and with a data set which discounts old information by replacing the oldest observation in a time series with the newest observation obtained after each forecast. This is similar to what Garcia et. al referred to as full information and limited information modeling. The out-of-sample forecasting period is June 1988 through November 1990. This period is used for all forecast horizons from 1 through 6 months. For example, the first one-step-ahead forecast uses information available in May 1988, whereas the first six-step-ahead forecast uses information available in December 1987. This forecasting scenario was chosen to make the forecasting performance of different models comparable not only within a forecasting horizon, but also among different forecasting horizons.

A recursive procedure is used to get forecasts for the different horizons. This means that forecasts of future basis are used in the forecast process if the lag of the variable in a model is shorter than the forecast horizon. In the case of an AR(1) model and a forecast horizon of two months the one-step-ahead forecast value of the dependent variable is used as an independent variable for the forecast of the two-steps-ahead value of the dependent variable.

Two evaluation criteria are used to evaluate and rank the different prediction models:[10] (1) mean absolute forecast error (MAFE), and (2) root mean squared forecast error (RMSFE). The forecast error $e_t^{(i)}$ for a prediction made at time t for a point in time i periods ahead is hereby defined as the actual basis minus the basis forecast. An overestimation of basis results in a negative forecast error and an underestimation in a positive forecast error. We focus on the performance of the different prediction models in comparison to a naive model. The naive

[10] The difference between these two evaluation criteria lies in the different weighing of the size of a forecast error. The mean absolute forecast error weighs each forecast error equal, whereas the root mean square forecast error uses larger weights for larger forecast errors.

model uses a 3-year moving average of basis to generate forecasts.[11] Different criteria were chosen because the evaluation of a forecasting model relative to an alternative model depends on the cost function $C(e_t^{(i)})$ which an individual forecaster associates with a wrong forecast. The functional form of this cost function depends on the objective of the forecaster. The functional form chosen is generally an approximation of the true cost function[12]. The choice criterion for the selection of the best among a set of alternate models is the one which results in the lowest expected (average) cost for a given cost function $C(e_t^{(i)})$ (Granger).

The procedure FORECAST of PROC ARIMA in SAS is used to derive individual basis forecasts. The evaluation criteria based on the sample of 30 forecasts are presented in table 2, (MAFE) and table 3, (RMSFE).

The evaluation criteria shows that the forecasts from the fitted models outperform the naive (i.e. 3 year seasonal moving average) model with a few exceptions for longer forecast horizons. The combination of seasonal dummy variables with an ARIMA model (SEASON in tables 2 and 3) yields better results than the more parsimonious pure ARIMA models. This result remains consistent through all forecasting horizons. It indicates that the seasonality in basis is highly regular.

The forecast error for the naive model is not sensitive to an increasing forecast horizon as long as the forecast horizon is shorter than one year. The ARIMA and SEASON model results, however, prove to be highly sensitive to an increasing forecast horizon. Note that the greatest deterioration in forecast performance occurs between the one-step-ahead and two-steps-ahead forecasts. A further decline in forecasting accuracy for the longer forecast horizons two through six months is visible but not conclusive over all model types.

An analysis of the joint basis predictions versus the combination of separately forecast basis components generally does not show an improvement in forecasting accuracy. This is especially obvious for the RMSFE statistic which penalizes larger forecast errors more than small ones. Thus, decomposition of basis does not yield the a-priori expected improvement in forecasting accuracy.

The results for all models which consider the whole available information set compared to the models which completely discount old information do not exhibit any conclusive evidence for a superior performance of either one of the different approaches.

Both evaluation criteria, MAFE and RMSFE, yield similar results. Differences are only observable in the decline of the forecasting accuracy from one-step-ahead to two-step-ahead forecasts. The decline in the forecasting accuracy is greater for the RMSFE. This indicates not only a general decline in the forecasting performance but, particularly, a strong increase in the forecast error for forecasting horizons larger than one month.

---

[11] This is the basis model used by Fortenbery to forecast joint basis for Greenville, North Carolina.

[12] The commonly used cost function according to Granger (1989) is $C(e_t^{(i)}) = e_t^{(i)\,2}$, the squared forecast error.

Table 2. Mean Absolute Forecast Error*

BASIS	MODEL	CONST/INCR**	FORECASTING HORIZON (in months) 1	2	3	4	5	6
joint***	ARIMA	const	0.09711	0.11435	0.12041	0.12994	0.13276	0.13471
		incr	0.09670	0.11652	0.12272	0.13068	0.13441	0.13522
	SEASON	const	0.08917	0.10731	0.11145	0.12138	0.12353	0.12353
		incr	0.08811	0.10782	0.11212	0.12201	0.12543	0.12530
separated****	ARIMA	const	0.11962	0.14395	0.15345	0.15386	0.14920	0.14755
		incr	0.11564	0.13713	0.14626	0.14453	0.14059	0.14168
	SEASON	const	0.10061	0.10935	0.12011	0.12713	0.13083	0.13303
		incr	0.09468	0.10239	0.11355	0.12059	0.12530	0.12900
	NAIVE	-	0.13868	0.13868	0.13868	0.13868	0.13868	0.13868

* The forecast error was calculated in dollars per bushel.

** *const* means that the number of observations used to estimate the model was held constant. After each forecast the data used to estimate the model was updated by adding a new observation (the new observed value for the basis) and omitting the oldest observation.
incr means that the number of observations used to estimate the model was increased. After each forecast the new, available observation was added to re-estimate the model and get the new forecast.

*** *joint* indicates the models in which basis was estimated jointly.

**** *separated* indicates the model in which the spatial and intertemporal parts of the basis were estimated separately and summed up.

Table 3. Root Mean Square Forecast Error*

			FORECASTING HORIZON (in months)					
BASIS	MODEL	CONST/INCR**	1	2	3	4	5	6
joint***	ARIMA	const	0.12943	0.15662	0.16342	0.16703	0.16971	0.17059
		incr	0.12864	0.15657	0.16326	0.16679	0.16966	0.17043
	SEASON	const	0.12631	0.14835	0.14677	0.15021	0.15251	0.15178
		incr	0.12291	0.14789	0.14770	0.15114	0.15352	0.15325
separated****	ARIMA	const	0.15947	0.18553	0.18909	0.18472	0.18703	0.18948
		incr	0.15621	0.18228	0.18457	0.17921	0.18141	0.18563
	SEASON	const	0.13802	0.15576	0.15620	0.16150	0.16818	0.17425
		incr	0.13070	0.14943	0.15162	0.15465	0.16063	0.16822
	NAIVE	-	0.16636	0.16636	0.16636	0.16636	0.16636	0.16636

* The forecast error was calculated in dollars per bushel.

** *const* means that the number of observations used to estimate the model was held constant. After each forecast the data used to estimate the model was updated by adding a new observation (the new observed value for the basis) and omitting the oldest observation.
incr means that the number of observations used to estimate the model was increased. After each forecast the new, available observation was added to re-estimate the model and get the new forecast.

*** *joint* indicates the models in which basis was estimated jointly.

**** *separated* indicates the model in which the spatial and intertemporal parts of the basis were estimated separately and summed up.

Note that the joint SEASON model consistently outperforms all other models, including the naive specification. For the one month ahead forecast, the joint SEASON model improves the forecast of basis by an average of 5 cents per bushel, and for the six month ahead by nearly 2 cents per bushel. While we have not explicitly accounted for the relative costs of utilizing the naive model relative to the joint SEASON model, it appears that rejection of the naive model in favor of the joint SEASON model seems reasonable.

Conclusions

This paper investigates the extent to which traditional basis forecast models based on simple moving averages of historical basis can be improved. Results indicate that a relatively simple time series model combined with monthly dummy variables can substantially increase forecast accuracy for planning horizons from one to six months. While explicit cost/benefit analysis is not conducted, the potential savings resulting from improved basis forecasts appear to justify use of the models identified. Basis forecast accuracy is improved by 2 cents for six month forecasts up to 5 cents for one month forecasts.

We also investigate the extent to which basis forecasts can be improved by estimating separate models for the spatial and intertemporal basis components. Results indicate disaggregated models yield no increase in forecast efficiency. In addition, forecasts do not seem to be affected by the decision to keep or delete old observations when models are updated.

Selected References:

Brennan, M.J. (1958). "The supply of storage". American Economic Review, 47, 50-72.

Brocklebank, J.C. & Dickey, D.A. (1986). SAS System for Forecasting Time Series. SAS Series in Statistical Applications, 1986 edition. SAS Institute Inc., Box 8000, SAS Circle, Cary, NC 27511-8000

Chicago Board of Trade. "Introduction to Hedging". Chicago, 1982.

Curtis, C.E. Jr., Harris, H.M Jr., & Miller, S.E. (1988). "Corn, soybean, and wheat basis in South Carolina". Extension Economics Report, EER 104, Clemson University, August 1988.

Fortenbery, T.R. (1991). "A Look at North Carolina Basis". Cooperative Extension Service, North Carolina State University, AG-451, 6/91.

Garcia, P. & Good, D.L. (1983). "An analysis of the factors influencing the Illinois corn basis, 1971-1981". Applied Commodity Price Analysis, Forecasting, and Market Risk Management, proceedings of the NCR-134 conference, Des Moines, Iowa, April 28-29, 1983, 306-326.

Garcia, P., Leuthold, R.M., Fortenbery, T.R. & Sarassoro, G.F. (1988). "Pricing efficiency in the live cattle futures market: further interpretation and measurement". American Journal of Agricultural Economics, 70, 162-169.

Granger, C.W.J. (1989). "Forecasting in Business and Economics". 2nd edition. San Diego, CA: Academic Press Inc.

Gray, R.W. (1960). "The characteristic bias in some thin futures markets". Food Research

Institute Studies, 1, 296-313.

Gray, R.W. (1961). "The search for a risk premium". Journal of Political Economy, 64, 250-260.

Gray, R.W. (1967). "Price effects of a lack of speculation". Food Research Institute Studies, Supplement 7, 177-194.

Kahl, K.H. & Curtis, C.E. Jr. (1986). "A comparative analysis of the corn basis in feed grain deficit and surplus areas". Review of Research in Futures Markets, 5, 220-232.

Kahl, K.H. & Curtis, C.E. Jr. (1988). "An analysis of the South Carolina Corn Basis". South Carolina Agricultural Experiment Station Bulletin, Bulletin 665, March 1988.

Kenyon, D.E. & Blakeley, P.K. (1974). "Alternative basis estimation procedures and hedging strategies for corn and soybean producers in selected Virginia markets". Department of Agricultural Economics Bulletin, AE 18, Virginia Polytechnic Institute and State University, September 1974.

Malick, W.M. & Ward, R.W. (1987). "Stock effects and seasonality in the FCOJ futures basis". The Journal of Futures Markets, 7, 157-167.

Martin, L., Groenewegen, J.L. & Pidgeon E. (1980). "Factors affecting corn basis in Southwestern Ontario". American Journal of Agricultural Economics, 62, 107-112.

Nicholls, D.F. (1979). "The analysis of time series - the time domain approach". Australian Journal of Statistics, 21, 93-120.

Peck, A.E. & Williams, J.C. (1991). "An Evaluation of the Performance of the Chicago Board of Trade Wheat, Corn, and Soybean Futures Contract During Delivery Periods from 1964-65 Through 1988-89". Report to the National Grain and Feed Association, April 1991, 1201 New York Ave., N.W., Suite 830, Washingtion, D.C. 20005.

Telser, L.G. (1958). "Futures trading and the storage of cotton and wheat". Journal of Political economy, 66, 233-255.

Tomek, W.G. & Robinson, K.L. (1990). "Agricultural Product Prices". 3rd edition. Ithaca, N.Y.: Cornell University Press.

Ward, R.W. & Dasse, F.W. (1977). "Empirical contribution to basis theory: the case of citrus futures". American Journal of Agricultural Economics, 59, 71-80.

Working, H. (1948). "The theory of inverse carrying charge in futures markets". Journal of Farm Economics, 30, 1-28.

Working, H. (1949). "The theory of the price of storage". American Economic Review, 31, 1254-62.11