NCCC-134
APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

Futures-Based Price Forecasts for Agricultural Producers and Businesses

by

Terry Kastens, Rodney Jones, and Ted Schroeder

Futures-Based Price Forecasts for Agricultural Producers and Businesses

Terry L. Kastens, Rodney Jones, and Ted C. Schroeder*

This study examines the accuracy of five competing naive and futures-based localized cash price forecasts. The third week's price for each month from 1987-1996 is forecasted from vantage points one to 11 months preceding the observed price. Commodities examined included those relevant to Kansas producers: the major grains, slaughter steers and slaughter hogs, several classes of feeder cattle, cull cows, and sows. Information about relative forecasting accuracy across forecast methods was collapsed into regression models of forecast error. The traditional forecast method of deferred futures plus historical basis had the greatest accuracy—even for commodities that are substantially different from those specified in related futures contracts. Adding complexity to forecasts, such as including regression models to capture nonlinear bases or biases in futures markets, did not improve accuracy.

Introduction

Futures prices are regularly used to construct agricultural commodity price forecasts. Both grain elevators and livestock packer buyers forward price "off the board," generally using a formula. Even commodities that are not deliverable on the underlying futures contract—such as milo (grain sorghum) are often priced this way. However, if futures/cash differentials (bases) are not stable over time, gains in predictive accuracy may result from using bases which have proportional as well as differential components. Further, if deferred futures prices are biased estimates of future prices, modeling cash/deferred futures relationships may provide greater forecast accuracy than just adjusting futures prices for expected basis.

This research examines the accuracy associated with using deferred futures prices, along with historical average bases, to predict future cash prices of various crop and livestock commodities important to the Midwest. Several forecast horizon lengths, up to a year, are considered. Futures-plus-basis price forecasts are compared to naive cash price forecasts and to other futures-based forecasts. Simple, regression-based forecasts are also included. Regression analysis is used to determine which factors affect forecast errors of competing models, and to test which forecast methods are most accurate.

Background

Agricultural production is becoming increasingly differentiated in physical characteristics, time, and/or space. For example, corn is becoming segregated into several

---

* Assistant professors and professor, Department of Agricultural Economics, Kansas State University.

classes such as high oil or high lysine, and wheat is increasingly segregated according qualities, especially protein. Livestock are becoming increasingly differentiated, with premiums and/or discounts associated with various characteristics. In addition, profit maximizing cropping decisions now rely more on price projections because virtually r constraints are imposed by the most recent farm legislation (Federal Agricultural Impr and Reform Act of 1996)—which means the accuracy of crop price projections is bec important. Together, these observations imply producers and agricultural businesses r price forecasts that are more product-, location-, and time-specific.

Extension outlook price forecasts have not traditionally been product-, location specific. Rather, they have focused on broad-based price forecasts, such as quarterly o national commodity prices. In part, this may be because Extension models regularly ir fundamental supply/demand data that would be prohibitively expensive to obtain at fir and space distinctions. Also, it may be due to Extension forecasters' attempts to maxi audience around each forecast provided. In addition, recent research has shown that E: price forecasts have typically been less accurate than those of the USDA (Kastens, Sch and Plain 1996). Considering that Extension regularly forecasts many of the same pric USDA, and that producers will demand more specific forecasts in the future, this is an opportunity for Extension economists, who by definition are more localized than USD. refocus their outlook efforts.

Grain and livestock businesses regularly forward price based on deferred future futures prices *are* price expectations (Eales et al. 1990). Futures prices are inexpensive and are at least as accurate as commercial and public providers of price forecasts (Just : Rausser 1981; Marines-Filho and Irwin 1995; Kastens, Schroeder, and Plain 1996). Be they are virtually continuously available, futures prices could provide an essential com Extension's development of more specific price forecasts, and in the development of sy forecasting procedures that could be adopted by users. However, to assure timeliness, availability, and the potential for user-development, futures-based cash price forecasts simple to construct and easy to understand.

Brorsen and Anderson (1994) have challenged Extension forecasters by arguing the efficient market hypothesis and the law of one price should be the cornerstone of ex marketing programs" (p.90). This research builds on their challenge by embodying tho economic concepts in procedures that can be used in real-time forecasting. Using futur to construct cash price forecasts depends on futures market efficiency. If a futures marl efficient then a deferred futures price will, on average, be an unbiased estimate of deliv price of the underlying commodity. That means a cash price forecast can be made by a futures price for expected basis. On the other hand, tieing delivery-time basis to a biase deferred futures price will result in a biased cash price forecast.

The futures efficiency literature is large, with diverse procedural approaches tak diverse conclusions. Overall, the evidence favors futures efficiency. However, there is tendency to find livestock futures inefficient than grain futures (Garcia, Hudson, and Waller 1988; Kolb 1992 and 1996). In some cases, most notably live cattle futures, reported inefficiencies were in *biases*, meaning that economically significant trends persisted in futures prices (Kastens and Schroeder 1995). Thus, it may be important that simple futures-based price forecasting procedures allow for possible underlying biases.

For some agricultural commodities, especially grains, locational price differences are more important than differences between cash commodity characteristics and related futures contract specifications. Hence, in developing futures-based cash grain price forecast procedures, it is important to test historical data from many locations. For other commodities, especially livestock, where products vary by type of animal, weight, or sex, departures from futures contract specifications are especially important. Thus, in developing futures-based cash livestock price forecasts it is important to incorporate historical data from several animal classes, weights, and from both sexes. Finally, to be of general value, forecasts need to provide information for numerous points in the future.

General Analytical Procedures

Five approaches are used to forecast future cash prices. The procedures are presented in order of increasing complexity. The first approach, referred to as *NAIVE1*, uses last year's price to forecast price in the same week this year. Formally, in a model framework, this approach states that the cash price for commodity i, in location j, for week w of year T, $CP_{i,j,w,T}$, is equal to the cash price observed for the same commodity, location, and week in year $T\text{-}1$, plus some error $\varepsilon_{i,j,w,T}$:

$$\text{NAIVE1:} \quad CP_{i,j,w,T} = CP_{i,j,w,T-1} + \varepsilon_{i,j,w,T} \,. \tag{1a}$$

The specification in (*1a*) yields a one-step-ahead forecast of price in week w of year $T+1$, with that expectation taken h (for *h*orizon) weeks prior to when the actual price is observed:

$$\text{NAIVE1:} \quad \mathrm{E}_{w-h}\left[CP_{i,j,w,T+1}\right] = CP_{i,j,w,T} \,. \tag{1b}$$

That $w\text{-}h$ is not included on the right-hand side of (*1b*) makes it clear the price forecast for a particular week of year $T+1$ is the same for all forecast horizons.

The second approach, referred to as *NAIVE5*, assumes cash price regresses to its multiple-year average. However, because policy and other changes can fundamentally alter long-term prices, the number of years considered is only five. Formally, this approach is:

$$\text{NAIVE5:} \quad CP_{i,j,w,T} = \frac{1}{5} \sum_{t=T-5}^{t=T-1} CP_{i,j,w,t} + \varepsilon_{i,j,w,T} , \tag{2a}$$

with the corresponding forecast specification:

$$\text{NAIVE5:} \quad \mathrm{E}_{w-h}\left[CP_{i,j,w,T+1}\right] = \frac{1}{5} \sum_{t=T-4}^{t=T} CP_{i,j,w,t} . \tag{2b}$$

As in (*1b*), the forecasts from (*2b*) are the same for all forecast horizons.

The third forecast approach (*FUTLBAS*) incorporates futures and basis, with basis a fixed *level* (or, differential), as in cents per bushel. Basis is defined here as cash price less nearby futures price, implying cash price equals nearby futures price plus basis. If basis does not trend over time, cash price can be defined as nearby futures price plus historical average basis plus some error. As in *NAIVE5*, five years are used to generalize historical basis information. The formal specification is:

$$\text{FUTLBAS:} \quad CP_{i,j,w,T} = FP^{w,T}_{i,w,T} + \frac{1}{5} \sum_{t=T-5}^{t=T-1} \left(CP_{i,j,w,t} - FP^{w,t}_{i,w,t}\right) + \varepsilon_{i,j,w,T} , \tag{3a}$$

where the subscript, i, on the futures price variable, FP, refers to the contract nearest in specification to, or most likely to be used in hedging, cash commodity i; the j is omitted because it is assumed that the pertinent futures contract does not change across cash price locations. As for cash price, the remaining two subscripts of FP denote the week (w) and year (T). The superscripts on FP further specify the futures contract represented. Namely, w, T specifies that the futures contract is the nearby contract in week w of year T.

If futures are unbiased, deferred futures price provides a reasonable forecast of delivery-time futures price. Consequently, the forecast specification associated with (*3a*) is:

$$\text{FUTLBAS:} \quad \mathrm{E}_{w-h}\left[CP_{i,j,w,T+1}\right] = FP^{w,T+1}_{i,w-h,T+1} + \frac{1}{5} \sum_{t=T-4}^{t=T} \left(CP_{i,j,w,t} - FP^{w,t}_{i,w,t}\right) . \tag{3b}$$

Equation *3b* reads as follows. The expectation (or forecast) taken in week w-h, for the cash price of commodity i in location j that will be observed in week w of year $T+1$, is equal to the price, observed in week w-h of year $T+1$, for the futures contract corresponding to commodity i that will be the nearby in week w of year $T+1$, plus the respective 5-year moving average basis. Unlike *NAIVE1* and *NAIVE5*, *FUTLBAS* forecasts are unique for each forecast horizon—because these forecasts incorporate current deferred futures prices.

The fourth forecast approach retains the 'futures plus basis' idea embodied in *FUTLBAS*. However, it allows more flexibility by specifying basis in *level* and *proportional* components. This forecast method is called *FUTLPBAS*. The increased basis flexibility comes about by assuming that cash price equals some proportion of nearby futures price, plus an additive constant, plus an error. As with *FUTLBAS*, relationships in *FUTLPBAS* are assumed to hold over only the most recent 5 years:

$$\text{FUTLPBAS:} \quad CP_{i,j,w,t} = \alpha_{i,j,w,T} + \beta_{i,j,w,T} FP^{w,t}_{i,w,t} + \varepsilon_{i,j,w,t} \; ; \quad \text{for } t = T-4,\ldots,T \; . \tag{4a}$$

Equation *4a* may be thought of as a first-order approximation of some higher-ordered underlying cash/futures price relationship (without the error term).

Values for $\alpha_{i,j,w,T}$ and $\beta_{i,j,w,T}$ in (*4a*) are estimated using ordinary least squares regression. Each commodity, location, week, and year has its own unique regression and corresponding α and β estimates. Regressions are estimated over $t = T\text{-}4$ to $t = T$. Like *FUTLBAS*, *FUTLPBAS* assumes unbiased futures, using deferred futures prices, along with regression estimates that are unique across week forecasted but not across forecasting horizon, to develop horizon-specific forecasts:

$$\text{FUTLPBAS:} \quad \mathrm{E}_{w-h}\left[CP_{i,j,w,T+1}\right] = \hat{\alpha}_{i,j,w,T} + \hat{\beta}_{i,j,w,T} FP^{w,T+1}_{i,w-h,T+1} \; . \tag{4b}$$

FUTLPBAS is inherently more complex than *FUTLBAS*, or the two naive methods, in that regression models must be estimated. However, because the models are not horizon-specific, the total number of models required are not excessive, and the potential forecasting accuracy gains could be large.

Where basis may be unstable or difficult to predict, and where futures prices have a tendency to be biased, it may be helpful to circumvent the concept of basis altogether, and model cash price directly as a function of deferred futures price (not nearby). Thus, the fifth forecasting approach, *MODFUT*, is specified as:

$$\text{MODFUT:} \quad CP_{i,j,w,t} = \alpha_{i,j,w-h,T} + \beta_{i,j,w-h,T} FP^{w,t}_{i,w-h,t} + \varepsilon_{i,j,w-h,t} \; ; \quad \text{for } t = T-4,\ldots,T \; . \tag{5a}$$

MODFUT in (*5a*) looks a lot like *FUTLPBAS* in (*4a*), with one important difference: the subscripts on the right-hand side include the letter *h*. That means a separate model is estimated for each price forecasted and each forecasting vantage point, *w-h*. The forecast specification associated with (*5a*) is:

$$\text{MODFUT:} \quad \mathrm{E}_{w-h}\left[CP_{i,j,w,T+1}\right] = \hat{\alpha}_{i,j,w-h,T} + \hat{\beta}_{i,j,w-h,T} FP^{w,T+1}_{i,w-h,T+1} \; . \tag{5b}$$

MODFUT involves many more regressions than does *FUTLPBAS*. The additional computation time and data basing required, though seemingly small in a research setting, could be enough to preclude real-time forecasters from using this approach. However, if forecasting accuracy gains are large, then the additional burden may be worthwhile.

Data Used and Forecasts Developed

Weekly prices for various cash commodities and locations were collected from the first week of 1982 through the last week of 1996. Locations selected were those relevant for Midwestern (with focus on Kansas) producers and businesses. Commodities examined were wheat, corn, milo, soybeans, slaughter steers, cutter cows, 7-8 cwt. steers, 4-5 cwt. steers, 7-8 cwt. heifers, 4-5 cwt. heifers, slaughter hogs, and sows. Price data were structured on the basis of 4 weeks per month (if a month had 5 weeks the 4th and 5th week's prices were averaged). Nearby futures price data corresponding to the cash price series were also collected, with nearby defined as nearest to delivery but not in the delivery month. For some commodities, deferred futures prices were consistently available up to 11 months prior to the nearby period. For others, they were only available for shorter time periods.

NAIVE1, *NAIVE5*, *FUTLBAS*, *FUTLPBAS*, and *MODFUT* forecasts were developed for each commodity and location.[1] Because all but one method (*NAIVE1*) required 5 years of historical data, all forecasts were for weeks in the years 1987 through 1996. Because of the large volume of data, prices from only selected weeks were forecasted and only at selected horizons. Prices were forecasted for the third week of each month in each year. The vantage points from which these prices were forecasted (the forecast horizons) were 4 weeks prior, 8 weeks prior, and so on, stepping back in time as long as deferred futures prices were available. Because of the weeks selected, both forecasted periods and forecast horizons are one month apart. Missing data were extrapolated to ease the computational burden (an appendix describes missing data procedures and other data details). Table 1 provides a description of the cash price series forecasted, the associated underlying futures markets, the number of forecast horizons considered, and the total number of forecasts constructed.

Forecast Evaluation Procedures

A series of forecasts is associated with a series of forecast errors. For evaluation, the information embodied in a forecast error series is routinely condensed into a single test statistic, such as sum of squared errors or mean absolute error, so that competing forecasts can be compared using their associated test statistic values. Unfortunately, this approach is limited to

---

[1] OLS was used in estimating underlying regressions for regression-based forecasts. We recognize that potential cointegration between cash and futures prices may cause underlying parameter estimate standard errors to be unreliable. However, cointegration is not useful in these models that are estimated only over 5 observations ($t=T-4$ to $t=T$) each year.

pairwise comparisons. To extract information of interest requires numerous pairwise comparisons, making it difficult to generalize results.

An alternative forecast comparison approach, that generalizes large amounts of information, collapses the information in a forecast error series into a regression model where forecast error is the dependent variable. In that framework, forecast errors from competing forecasts across time and space can be stacked, so that partial effects of interest can be isolated using appropriate independent variables. For an example of this method of forecast comparison see Kastens, Schroeder, and Plain (1996).

Because the number of forecasts examined here was large, varying across years, weeks within the year, horizon length, location, and commodity, the forecast error regression model approach to forecast comparison was selected. This approach considers that cash price forecast errors for a commodity are affected by forecast method, forecast horizon, time period forecasted, and the cash price location:

$$Forecast\ Error = f\,(method, horizon, time\ period, location)\ . \tag{6}$$

A goal of this research was to reach general conclusions about relative accuracy for alternative cash price forecasting methods. The effect of forecast horizon on the accuracy of competing forecast methods is expected to vary widely. For example, forecasts using the two naive methods are constant across horizon. Thus, it is important to specify (6) so that the effects of horizon by method, on relative accuracy, can be measured—suggesting an interaction term. Prices for some time periods within the year, and for some locations, are likely to be inherently more difficult to forecast than other times or locations. It is important to isolate these inherent forecast accuracy differences, so that they do not mask information sought: comparing relative accuracy across competing forecast methods. However, in order to generalize the results into usable forecast procedure recommendations, no interactions with method were considered for the time and location effects.

Focusing on error magnitude, forecast errors were measured as absolute values. Since the scale of cash price varies substantially across time and location, errors were computed as percent errors (actual less predicted, divided by actual, multiplied by 100). The end result was dependent variables that are absolute percentage forecast error (APE) series. The final model estimated separately for each commodity is:

$$\begin{aligned} APE_{j,w,T,w-h} &= \alpha + \beta_1 NAIVE1_{j,w,T,w-h} + \beta_2 NAIVE5_{j,w,T,w-h} + \beta_3 FUTLPBAS_{j,w,T,w-h} \\ &+ \beta_4 MODFUT_{j,w,T,w-h} + \beta_5 HORIZON_h + \beta_6 NAIVE1H_{j,w,T,w-h} \\ &+ \beta_7 NAIVE5H_{j,w,T,w-h} + \beta_8 FUTLPBASH_{j,w,T,w-h} + \beta_9 MODFUTH_{j,w,T,w-h} \\ &+ \beta_{10} JAN_w + \ldots + \beta_{20} NOV_w + \beta_{21} LOC_1 + \ldots + \beta_{J+19} LOC_{J-1} + \varepsilon_{j,w,T,w-h}\ . \end{aligned} \tag{7}$$

In equation 7, j represents location (1. . .J); w is the week (3, 7, ... , 47) of year T (1987–1996) corresponding to the period forecasted (thus, the *APE*); h represents forecast horizon length in weeks, so that w-h denotes a forecast made in week w-h. *NAIVE1*, *NAIVE5*, *FUTLPBAS*, and *MODFUT* are forecast dummies that equal 1 when the forecast was generated by that respective method, else 0 (default method is *FUTLBAS*). $HORIZON_h$ is a variable equal to h. $NAIVE1H_{j,w,T,w\text{-}h}$ through $MODFUTH_{j,w,T,w\text{-}h}$ are slope dummies equal to the product of $HORIZON_h$ and the corresponding forecast dummy (default is *FUTLBASH*). JAN_w through NOV_w are 1 if week w corresponds to the month specified, else 0; and LOC_j is 1 if the underlying forecast corresponds to the cash price in location j, else 0.

Results

To provide a general background, table 2 shows mean absolute percentage error (MAPE) and maximum absolute percentage error (maxAPE) by forecast method and commodity. The minimum APE was near zero in all cases, so not reported. As judged by the average MAPEs at the bottom, *FUTLBAS* (futures plus level basis) and *FUTLPBAS* (futures plus level and proportional basis) generally provide the greatest accuracy across the forecast methods. On average, their average MAPEs and average maxAPEs are similar. *NAIVE1* (1-year naive) had the second-highest average MAPE, yet an average maxAPE similar to the two futures-plus-basis methods. Forecasts based on last year's price, while not particularly accurate, did not diverge too far from actual price either. The relatively more complex forecast method, *MODFUT* (where cash price is modeled as a function of deferred futures price) was in the middle of the group in terms of average MAPE, but the worst method by maxAPE. That suggests *MODFUT* performs especially poorly in some forecasts, causing occasional large errors. Observing the rows in table 2 shows that *MODFUT* had the highest maxAPE in 8 of 12 commodities. It performed especially poorly in the grains, where worst-case errors were consistently above 100%.

Overall, in terms of MAPE, table 2 shows that *NAIVE5* (5-year naive) was generally the least accurate forecast method. For the 6 cattle price series, *NAIVE5* was the single worst method for MAPE, and had the highest maxAPE for 4 out of 6 cattle series. Underlying cattle price cycles may be to blame for diminishing the accuracy of *NAIVE5*, causing the 5-year average price to repeatedly be a poor predictor of future price. The rightmost column of table 2 shows the grains to have the least accuracy across the 12 commodities. The average MAPE for wheat, corn, milo, and soybeans is 14.45%—contrasted with an average MAPE for the 6 cattle series of 10.88%. Slaughter steer price forecasting, with the lowest average MAPE and lowest average maxAPE, stands out among the commodities as having the greatest accuracy. On the other hand, with an average MAPE over 17%, sow prices are the most difficult to forecast.

Results of models explaining forecast errors are in table 3 (grains, slaughter steers, and cutter cows) and table 4 (feeder cattle, slaughter hogs, and sows). The regression models are used to condense information in a systematic manner and to statistically compare the accuracy of competing methods by forecast horizon. Thus, to conserve space, coefficient estimates for binary month and location variables are not reported.[2] The models do not have particularly high explanatory power, as R^2s range from a high of 0.24 for 7-8 cwt. feeder steers to a low of 0.05 for cutter cows.

The *HORIZON* estimate depicts the change in accuracy for a one month increase in forecast horizon for the default forecast method, *FUTLBAS* (futures plus level basis). All *HORIZON* estimates are significant and positive, confirming that forecasting further into the future is less accurate than forecasting prices close to the present. *FUTLBAS* APEs increase more with lengthening horizons for grain forecasts than for livestock forecasts. *FUTLBAS* wheat price forecast accuracy diminished 1.34% for each 1-month increase in horizon, the most among all commodities. With naive forecasts, last year's price (for *NAIVE1*), or the last 5-years' average price (for *NAIVE5)*, is the same price forecast for all horizons. Thus, horizon does not impact accuracy for the naive models. This can be seen by noting that *NAIVE1H* and *NAIVE5H* parameters are each the negative of the *HORIZON* parameter.

Generally, both naive forecast methods yield substantially reduced accuracy relative to the default method *FUTLBAS* (all *NAIVE1* and *NAIVE5* estimates are statistically positive). However, this is only consistently true for sufficiently short forecast horizons. As noted, *FUTLBAS* forecast accuracy deteriorates with increased horizon length while naive accuracy does not. Dividing values in either the *NAIVE1* or *NAIVE5* rows by same-column values in the corresponding *NAIVE1H* or *NAIVE5H* row yields the forecast horizon where naive accuracy equals *FUTLBAS* accuracy. In all but three cases, the horizon where this occurs is either associated with a non-significant slope dummy estimate or is otherwise greater than the maximum horizon length tested. The three cases are *NAIVE5* for soybeans (at 8.4 months), *NAIVE1* for slaughter steers (7.6 months), and *NAIVE5* for slaughter hogs (10.3 months).

It would not have been surprising to find naive distant-horizon forecasts to be more accurate than futures-based forecasts for commodities that are undeliverable or substantially removed from the specifications in the underlying futures contract (for example, cutter cows). However, it is unexpected that naive forecasts at longer horizons are as accurate as futures-based forecasts for soybeans, slaughter steers, and slaughter hogs. For hogs it could be explained by noting that in 10.3 months a lot can change in hog production to negate futures-anticipated profits associated with hog feeding. That is, futures information 10 months prior to slaughter may be no more reliable than some simple long-run measure, such as naive price. Even for slaughter steers, with a feeding period of 4 to 5 months, production decisions may be substantially altered in 7.6 months. These findings for livestock are consistent with conclusions of Koontz, Hudson, and Hughes (1992). No explanation immediately emerges for soybeans.

---

[2] A total of 64 cash price locations were considered in grain price forecasts (wheat, 23; corn, 11; milo, 17; soybeans, 13). Among the 60 related location dummies, 21 had parameter estimates significant at the 0.05 level. Among the 132 total monthly dummies (12 commodities times 11 months), 82 were significant at the 0.05 level. Non-reported parameter estimates are available from the authors.

However, Kenyon, Jones, and McGuirk (1993) have noted that futures forecast accuracy has been especially poor for soybeans, and especially since 1973—due partly to yield uncertainty.

Forecasts that characterized basis in a regression of cash price on nearby futures price, *FUTLPBAS*, were not more accurate than the default method, *FUTLBAS*, where expected basis was the simple 5-year historical average level basis. For two commodities, corn and milo, *FUTLPBAS* performed worse. Nor did *FUTLPBAS* gain in relative accuracy over *FUTLBAS* as horizons lengthened (none of the *FUTLPBASH* estimates are significant). Thus, nothing was gained by including a proportional component to basis. The simple deferred futures plus level basis forecast method was more accurate.

MODFUT forecasts were based directly on regressions of cash price on deferred futures, not relying on the concept of basis. That increased complexity (a separate regression model is required for each horizon, point forecast combination) did not improve accuracy over the default *FUTLBAS* model. For one half of the commodities (wheat, corn, milo, soybeans, slaughter hogs, and sows) *MODFUT* resulted in statistically less accurate forecasts than *FUTLBAS*. *MODFUT* did not gain in relative accuracy over *FUTLBAS* as horizons lengthened. For one half of the commodities (corn, milo, slaughter steers, cutter cows, 7-8 cwt. steers, and 7-8 cwt. heifers) accuracy relative to *FUTLBAS* actually deteriorated with longer horizons (see the positively significant *MODFUTH* estimates). Thus, for corn and milo, *MODFUT* starts less accurate than *FUTLBAS* at short horizons and becomes increasingly less accurate as horizons lengthen.

Soybeans are somewhat anomalous. *MODFUT* forecasts at short horizons are less accurate than *FUTLBAS* counterparts. Model-predicted APE is 1.77 greater (1.9536−0.1791). Yet, beyond 10.9 month horizons (1.9536/0.1791), *MODFUT* soybean forecasts are more accurate than *FUTLBAS* counterparts. Earlier it was noted that, beyond horizons of 8.4 months, *NAIVE5* soybean forecasts are more accurate than *FUTLBAS* counterparts. Why did the default approach, *FUTLBAS*, forecast so poorly at distant horizons? Neither *NAIVE5* nor *MODFUT* depend on basis, but *FUTLBAS* does. Therefore, one possibility is that basis is less predictable for soybeans than other commodities. However, a broad look at basis variability (not shown) does not confirm that. For example, taking the standard deviation of weekly basis over 1987-1996 for each location, dividing by the average nearby futures price for the same time period, and averaging the quotients across all cash price locations, results in soybean basis variability that is 3% of futures price. Yet, comparable computations for wheat locations results in 4% basis variability.

As noted, the soybean anomaly does not appear to lie in difficulties with the cash/futures relationship. Rather, it has to do with difficulties in predicting delivery-time futures using deferred futures, as suggested by Kenyon, Jones, and McGuirk (1993). In an exercise where nearby soybean futures were treated as the only cash price series, forecast accuracy results were similar to those displayed in table 3. In short, deferred futures are merely poor predictors of eventual nearby futures when time gaps are large (favoring *NAIVE5*). Furthermore, biases in distant soybean futures persist long enough that historical regressions of nearby on deferred futures can capitalize on them (favoring *MODFUT*).

The default forecast method, *FUTLBAS*, which uses deferred futures price plus historical basis, is typically superior to other methods reported in tables 3 and 4. Among the 210 horizon-by-commodity combinations for *NAIVE1* and *NAIVE5*, only 6 of those combinations involved a naive forecast that was superior to *FUTLBAS*. The case was even stronger for comparisons involving more sophisticated futures-based forecasts. Among the 210 horizon-by-commodity combinations for *FUTLPBAS* and *MODFUT*, only 1 of those combinations (*MODFUT* for soybeans at a 11 month horizon) involved a sophisticated forecast that was superior to *FUTLBAS*. Together, these results make a strong case for using deferred futures plus historical basis for forecasting future cash commodity prices—at least among the relatively simple forecast methods considered here. Even where commodities are substantially different from those specified in futures contracts, using a related contract to aid forecasting appears beneficial. Virtually nothing was gained by assuming futures market biases (inefficiencies) are systematic enough to be picked up in historical regressions of cash price on deferred futures price.

The forecast method *FUTLPBAS*, where basis involved both level and proportional components, was never statistically more accurate than the simpler *FUTLBAS* method (for corn and milo it was statistically worse). In general, tables 3 and 4 show that it was difficult to statistically distinguish the forecasting accuracy of *FUTLPBAS* from that of *FUTLBAS*. Thus, in real-time forecasting, there is little reason to expend the extra effort in constructing regressions of cash price on nearby futures.

An interesting question revolves around *MODFUT*. Why was that method typically less accurate than *FUTLBAS*? After all, *MODFUT* accounts for persistent biases that may be present in the underlying futures market, and should not be unduly hampered if biases are not present. Furthermore, it should simultaneously account for cash prices that are consistently below futures (that is, basis). However, the relatively large maximum APEs reported for *MODFUT* in table 2 suggested this method forecasts some prices especially poorly. Regressions could impose too much structure on the data. That is, the relationship between a futures contract's current price and its price several months prior may be highly unstable. This points to an age-old problem of empirical economists: How can historical data best be generalized for making future decisions? Or, how can the real-time forecaster be restrained from making too much of historical data? There is, of course, no simple answer. Here, at least, combining the concept of futures efficiency with the simplest of models, the mean (of 5-year historical basis), resulted in more accurate forecasts of cash commodity prices than did using more complex models involving regressions.

Conclusions

This study examined the accuracy of five competing naive and futures-based localized cash price forecasts. The third week's price for each month of 1987 through 1996 was forecasted out-of-sample from vantage points one to 11 months preceding the observed price. Commodities examined were wheat, corn, milo, soybeans, slaughter steers, cutter cows, 7-8 cwt. steers, 4-5 cwt. steers, 7-8 cwt. heifers, 4-6 cwt. heifers, slaughter hogs, and sows. Locations selected w important to Midwestern producers and businesses. Only simple-to-construct forecasting methods were considered—methods that could easily be adopted for real-time forecasting by practitioners, producers, and businesses. Naive methods involved one-year lagged price and most-recent 5-year average price. Futures-based methods involved the traditional deferred futures plus historical basis (the most-recent 5-year average), deferred futures plugged into th estimates from a regression of cash price on nearby futures (assumes basis has both proportio and level components), and deferred futures plugged into the estimates from a regression of c price on deferred futures (captures persistent futures trends and historical cash/futures relationships directly in a model).

Information about relative forecasting accuracy across forecast methods was collapsed into regression models of forecast error. Results supported the traditional deferred futures plu historical basis method. That method was either statistically more accurate or not statistically less accurate in 413 of 420 commodity-by-forecast horizon combinations. Even for commodit substantially different from those specified in related futures contracts, such as cutter cows or cwt. heifers, the added sophistication of regression models was not merited. Although considering other forecast methods, or other historical data lengths, may have altered conclusions, the best models were those that used the economic principle of futures market efficiency along with one of the simplest models—the mean of historical basis. The implicati is that forecasters would do well to provide historical localized basis values directly to produce and businesses, and instruct them to simply add current deferred futures.

References

Brorsen, B.W. and K. Anderson. "Cash Wheat Marketing: Strategies for Real People." *Journ of Agribusiness*. 12-2(Fall 1994):85-94.

Eales, J.S., B.K. Engle, R.J. Hauser, and S.R. Thompson. "Grain Price Expectations of Illinois Farmers and Grain Merchandisers." *American Journal of Agricultural Economics*. 72(1990):701-708.

Garcia, P., M. A. Hudson, and M. L. Waller. "The Pricing Efficiency of Agricultural Futures Markets: An Analysis of Previous Research Results." *Southern Journal of Agricultural Economics*. 20(1988):119-30.

Just, R.E. and G.C. Rausser. "Commodity Price Forecasting with Large-Scale Econometric Models and the Futures Market." *American Journal of Agricultural Economics*. 63(1981):197-208.

Kastens, T.L. and T.C. Schroeder. "A Trading Simulation Test for Weak-Form Efficiency in Live Cattle Futures." *Journal of Futures Markets*. 15(September 1995):649-675.

Kastens, T.L., T.C. Schroeder, and R. Plain. "Evaluation of Extension and USDA Price and Production Forecasts." *NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*, Chicago, Illinois, April 22,23 1996.

Kenyon, D., E. Jones, and A. McGuirk. "Forecasting Performance of Corn and Soybean Harvest Futures Contracts." *American Journal of Agricultural Economics*. 75(May 1993):399-407.

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

Kolb, R.W. "The Systematic Risk of Futures Contracts." *Journal of Futures Markets*. 16(1996):631-654.

Koontz, S.R., M.A. Hudson, and M.W. Hughes. "Livestock Futures Markets and Rational Price Formation: Evidence for Live Cattle and Live Hogs." *Southern Journal of Agricultural Economics*. 24(July 1992)233-249.

Marines-Filho, J. and S.H. Irwin. "Pre-Harvest Hedging Behavior and Market Timing Performance of Private Market Advisory Services." *NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*, Chicago, Illinois, April 24,25 1995.

Appendix: Additional Data Details

A number of missing data points were approximated to expedite computations. Futures problems were limited to feeder cattle, where a few missing points would have precluded considering horizons beyond 14 weeks. Thus, in weeks 23 and 24 of 1983, January 1984 feeder cattle futures were not yet trading and were replaced with corresponding averages over 1982, 1984, 1985, and 1986 (only used in forecast model initialization). In week 19 of 1992 the January 1993 feeder cattle futures, which was not yet trading, was assumed to be 0.987 times the week 22 price (when it was trading), which was the same proportion observed in the November 1992 contract over the same time span.

For cash series, missing data problems were more severe, although typically less than 2% over the entire 1982-1996 time period for a particular commodity in a location, and typically less than 1% for the period forecasted, 1987-1996. Missing data were filled in using proportional changes in corresponding nearby futures prices before and after the missing points. For example, if a cash price in week 2 were missing, but weeks 1 and 3 were present, then the cash price was the average: [(week 2 fut/week 1 fut * week 1 cash)+(week 2 fut/week 3 fut * week 3 cash)]/2. If contiguous cash prices were absent, the adjustment process was iterated until convergence within $0.00001. In one case, cutter cow prices, missing data were severe during the forecast initialization period (1982-1986), where 72% of the data were missing. However, during the period forecasted (1987-1996) only 0.6% were missing. Consequently, because we wished to be consistent in both series length and in procedures, we used the same missing data computations. We recognize that this may introduce error in the cutter cow price forecasts, at least early in the 1987-1996 time period.

Hog futures contracts changed exclusively to lean hogs with the February 1997 contract. This involved 4 weeks of nearby futures in December 1996, as well as the deferred futures prices associated with the various forecast horizons. To be consistent with the preceding data, prices for the lean hog contract were converted to old contract equivalents by multiplying by 0.74.

Table1. Cash Price Forecast Description. Forecasts are for third week in each month, 1987-1996.

Cash Commodity[a]	Number of or location	Futures Market[b]	Forecast Horizons (months)	Total Forecasts
Wheat	23[c,d]	KCBT Wheat	1, ..., 8	110,400[g]
Corn	11[c]	CBOT Corn	1, ..., 11	72,600
Milo	17[c,e]	CBOT Corn	1, ..., 11	112,200
Soybeans	13[c,f]	CBOT Soybeans	1, ..., 11	85,800
Slaughter Steers	Western Ks. Direct	CME Live Cattle	1, ..., 9	5,400
Cutter Cows	Sioux City, Iowa	CME Live Cattle	1, ..., 9	5,400
7-8 cwt. Steers	Dodge City, Ks.	CME Feeder Cattle	1, ..., 6	3,600
4-5 cwt. Steers	Dodge City, Ks.	CME Feeder Cattle	1, ..., 6	3,600
7-8 cwt. Heifers	Dodge City, Ks.	CME Feeder Cattle	1, ..., 6	3,600
4-5 cwt. Heifers	Dodge City, Ks.	CME Feeder Cattle	1, ..., 6	3,600
Slaughter Hogs	St. Joseph, Mo.	CME Live Hogs	1, ..., 11	6,600
Sows	St. Joseph, Mo.	CME Live Hogs	1, ..., 11	6,600

[a] All grain prices are for Wednesday (or Thursday if no market on Wednesday). Slaughter steers, hogs, and sows are weekly averages. Other livestock prices are market day prices.

[b] All futures prices are Wednesday's close (or Thursday if no market on Wednesday).

[c] All grains share these Kansas markets: Colby, Dodge City, Emporia, Garden City, Great Bend, Hutchinson, Kansas City, Pratt, Scott City, Topeka, and Whitewater.

[d] Other Kansas wheat locations: Andale, Beloit, Concordia, Hays, Hoxie, Liberal, Marysville, Russell, Salina, St. Francis, Wellington, and Wichita.

[e] Other Kansas milo locations: Andale, Beloit, Hays, Liberal, Salina, and Wichita.

[f] Other Kansas soybean locations: Andale and Beloit

[g] Total forecasts are obtained by taking the number of forecast methods (5: *NAIVE1*, *NAIVE5*, *FUTLBAS*, *FUTLPBAS*, and *MODFUT*) times the number of weeks forecasted each year (12, or one for each month) times the number of years forecasted (10) times the number of locations (for wheat, 23) times the number of horizons considered (for wheat, 8).

Table 2. Mean and Maximum APE's by Commodity and Forecast Method, 1987-1996.

Commodity		Forecast Method					Average by Commodity
		NAIVE1	*NAIVE5*	*FUTLBAS*	*FUTLPBAS*	*MODFUT*	
Wheat	MAPE:	20.25	18.99	10.73	10.89	12.95	14.76
	maxAPE:	77.42	54.42	57.83	57.53	132.00	45.84
Corn	MAPE:	19.33	18.48	11.58	12.32	15.23	15.39
	maxAPE:	79.56	104.39	58.89	66.94	107.50	83.46
Milo	MAPE:	20.43	20.03	12.47	13.13	16.39	16.49
	maxAPE:	73.17	95.56	72.45	65.25	133.20	87.93
Soybeans	MAPE:	15.42	11.51	9.41	9.15	10.29	11.16
	maxAPE:	54.48	32.36	84.93	79.16	142.32	78.65
Slaughter Steers	MAPE:	6.87	9.67	5.82	6.35	7.98	7.34
	maxAPE:	20.78	30.17	19.47	22.75	49.17	28.47
Cutter Cows	MAPE:	12.60	18.74	11.22	10.77	13.96	13.46
	maxAPE:	59.62	85.00	67.27	66.89	82.36	72.23
7-8 cwt. Steers	MAPE:	9.57	15.47	6.12	5.83	8.15	9.03
	maxAPE:	28.80	53.43	24.05	24.29	59.99	38.11
4-5 cwt. Steers	MAPE:	12.19	19.73	10.87	9.19	9.87	12.37
	maxAPE:	53.22	81.61	51.81	46.05	50.39	56.62
7-8 cwt. Heifers	MAPE:	9.80	16.51	6.75	6.76	9.02	9.77
	maxAPE:	29.76	64.23	27.34	22.11	57.98	40.28
4-5 cwt. Heifers	MAPE:	13.07	22.01	11.64	9.29	10.64	13.33
	maxAPE:	53.95	94.04	58.39	39.38	54.62	60.08
Slaughter Hogs	MAPE:	15.84	12.76	10.22	10.45	13.32	12.52
	maxAPE:	49.73	56.37	65.95	65.30	79.35	63.34
Sows	MAPE:	20.93	18.90	13.66	14.17	18.94	17.32
	maxAPE:	79.02	89.21	103.75	104.02	117.96	98.79
Avg. by Method	MAPE:	14.69	16.90	10.04	9.86	12.23	
	maxAPE:	54.96	70.07	57.68	54.97	88.90	

Table 3. Selected Parameter Estimates for Determinants of Absolute Percentage Errors Associated with Forecast Models for Grains, Slaughter Steers, and Cutter Cows, 1987-1996.[a]

	Forecast Error Models					
Estimate	Wheat	Corn	Milo	Soybeans	Slaughter Steers	Cutter Cows
Intercept	5.2614** (0.2555)	7.7412** (0.3550)	9.3694** (0.3032)	4.4593** (0.2127)	2.7531** (0.4165)	9.5092** (1.0818)
Forecast dummies; default is futures plus basis method, FUTLBAS						
NAIVE1	15.5642** (0.2555)	13.3189** (0.3629)	13.1984** (0.2909)	11.2415** (0.2127)	3.0725** (0.4870)	2.9346* (1.2648)
NAIVE5	14.2967** (0.2555)	12.4735** (0.3629)	12.7951** (0.2909)	7.3268** (0.2127)	5.8709** (0.4870)	9.0836** (1.2648)
FUTLPBAS	0.3060 (0.2555)	0.9174* (0.3629)	0.8457** (0.2909)	-0.0358 (0.2127)	0.7244 (0.4870)	-0.0368 (1.2648)
MODFUT	2.5961** (0.2555)	1.4360** (0.3629)	2.3809** (0.2909)	1.9536** (0.2127)	0.9408 (0.4870)	0.3143 (1.2648)
HORIZON	1.3413** (0.0358)	0.9288** (0.0378)	0.8732** (0.0303)	0.8716** (0.0222)	0.4045** (0.0612)	0.3121* (0.1589)
Forecast/horizon interactions						
NAIVE1H	-1.3413** (0.0506)	-0.9288** (0.0535)	-0.8732** (0.0429)	-0.8716** (0.0314)	-0.4045** (0.0865)	-0.3121 (0.2248)
NAIVE5H	-1.3413** (0.0506)	-0.9288** (0.0535)	-0.8732** (0.0429)	-0.8716** (0.0314)	-0.4045** (0.0865)	-0.3121 (0.2248)
FUTLPBASH	-0.0326 (0.0506)	-0.0288 (0.0535)	-0.0309 (0.0429)	-0.0376 (0.0314)	-0.0399 (.0865)	-0.0837 (0.2248)
MODFUTH	-0.0823 (0.0506)	0.3695** (0.0535)	0.2572** (0.0429)	-0.1791** (0.0314)	0.2440** (0.0865)	0.4840* (0.2248)
No. of Obs.	111400	72600	112200	85800	5400	5400
Mean of dep. var. (MAPE)	14.7612	15.3880	16.4892	11.1574	7.3404	13.4572
R^2	0.1362	0.0781	0.0814	0.1105	0.1182	0.0502

[a] Significance at the 0.10 and 0.05 levels denoted by * and **, respectively. Standard errors are in parentheses. All models included 11 monthly dummies. Grains models included these numbers of location dummies: wheat, 22; corn, 10; milo, 16; soybeans, 12. Livestock series each have only one location.

Table 4. Selected Parameter Estimates for Determinants of Absolute Percentage Errors Associated with Forecast Models for Feeder Cattle, Slaughter Hogs, and Sows, 1987-1996.[a]

	Forecast Error Models					
Estimate	7-8 cwt. Strs.	4-5 cwt. Strs.	7-8 cwt. Hfrs.	4-5 cwt. Hfrs.	Slaughter Hogs	Sows
Intercept	2.6497** (0.7154)	8.3366** (1.1322)	3.4652** (0.7562)	9.0451** (1.2110)	3.7200** (0.6827)	6.9678** (1.0368)
Forecast dummies; default is futures plus basis method, FUTLBAS						
NAIVE1	6.1302** (0.8481)	3.0554* (1.3422)	5.7775** (0.8964)	3.6503* (1.4356)	9.2055** (0.7941)	10.9587** (1.2059)
NAIVE5	12.0306** (0.8481)	10.5915** (1.3422)	12.4840** (0.8964)	12.5862** (1.4356)	6.1281** (0.7941)	8.9361** (1.2059)
FUTLPBAS	0.3740 (0.8481)	-1.6356 (1.3422)	0.3381 (0.8964)	-2.3562 (1.4356)	0.1795 (0.7941)	0.4848 (1.2059)
MODFUT	0.3789 (0.8481)	-2.3676 (1.3422)	0.6018 (0.8964)	-2.6824 (1.4356)	3.3053** (0.7941)	3.6804** (1.2059)
HORIZON	0.7666** (0.1540)	0.4955* (0.2437)	0.7799** (0.1628)	0.6352* (0.2607)	0.5966** (0.0828)	0.6147** (0.1257)
Forecast/horizon interactions						
NAIVE1H	-0.7666** (0.2178)	-0.4955 (0.3447)	-0.7799** (0.2302)	-0.6352 (0.3686)	-0.5966** (0.1171)	-0.6147** (0.1778)
NAIVE5H	-0.7666** (0.2178)	-0.4955 (0.3447)	-0.7799** (0.2302)	-0.6352 (0.3686)	-0.5966** (0.1171)	-0.6147** (0.1778)
FUTLPBASH	-0.0950 (0.2178)	-0.0141 (0.3447)	-0.0938 (0.2302)	0.0009 (0.3686)	0.0084 (0.1171)	0.0056 (0.1778)
MODFUTH	0.4700* (0.2178)	0.3897 (0.3447)	0.4751* (0.2302)	0.4810 (0.3686)	-0.0334 (0.1171)	0.2677 (0.1778)
No. of Obs.	3600	3600	3600	3600	6600	6600
Mean of dep. var. (MAPE)	9.0291	12.3682	9.7699	13.3308	12.5183	17.3198
R^2	0.2389	0.1190	0.2220	0.1413	0.1138	0.0918

[a] Significance at the 0.10 and 0.05 levels denoted by * and **, respectively. Standard errors are in parentheses. All models included 11 monthly dummies. Livestock series each have only one location.