Journal of Agricultural and Resource Economics 21(2):187–198

Efficiency Tests of July Kansas City Wheat Futures

Terry L. Kastens and Ted C. Schroeder

Three procedures are used to test Fama *semistrong* form efficiency of harvesttime price of Kansas City July wheat futures from 1947 through 1995. The three methods are (*a*) testing for jointly significant parameter estimates on nonfutures explanatory variables in econometric forecasting models, (*b*) testing the relative accuracy between model-based forecasts and using deferred futures prices as forecasts, and (*c*) testing for abnormal profits associated with simulated futures trading signaled by the forecasts. Kansas City July wheat futures are generally efficient. Furthermore, relative to the efficiency associated with forecasts constructed one to two months before harvest, the efficiency associated with the five- to six-month period before harvest has increased, especially since the early 1980s.

Key words: futures efficiency, futures trading, wheat futures

Introduction

Hard red winter wheat is a major component of agricultural production in the Great Plains. The Kansas City Board of Trade (KCBOT) wheat futures market is closely associated with price discovery for hard red winter wheat. Market efficiency has important implications for producers, industry participants, and those involved in price outlook. If the market is generally efficient, the most accurate wheat price forecasts can be discerned from deferred futures. In addition, at most, small gains would be expected from increased emphasis on price outlook. Alternatively, if the market is generally inefficient, an economic return could be generated by using resources in forecasting wheat price. In spite of its importance, wheat futures efficiency has received little emphasis in the literature. This study addresses that by examining KCBOT July wheat futures efficiency over the last 50 years in a statistical, price forecasting, and trading simulation framework.

The requirement that abnormal profits must be demonstrated to reject market efficiency has been recognized at least since the time of Fama's work in 1970. Abnormal profits are profits sufficient to cover forecasting costs (Garcia et al.), transaction costs (Fama; Grossman and Stiglitz), as well as any risk premia required of futures traders (DeCoster, Labys, and Mitchell). Consequently, one line of futures efficiency research has focused directly on uncovering abnormal profits by constructing rule-based mechanical futures trading systems, typically based on price forecasts. While the type of information used in forecast construction distinguishes classes of futures efficiency, the emphasis has been on *additional* information, with the implicit assumption that traders who do not take account of this information are acting irrationally, leading to market failure. Examples

The authors are assistant professor and associate professor, respectively, Department of Agricultural Economics, Kansas State University. Helpful suggestions of three journal reviewers are appreciated.
This is contribution No. 96-460-J from the Kansas Agricultural Experiment Station.

of such studies are Helmuth; Palme and Graham; Peterson and Leuthold; Pluhar, Shafer, and Sporleder; Garcia et al.; and Kastens and Schroeder.

Other futures efficiency studies have focused on comparing deferred futures with alternative forecasts on the basis of forecast accuracy. An alternative forecast that is more accurate than deferred futures leads to efficiency rejection. The implicit assumption is that a more accurate forecast than deferred futures can be used to extract profits from that futures market. Examples of these studies include Just and Rausser; Martin and Garcia; Koppenhaver; Bessler and Brandt; and Kenyon, Jones, and McGuirk.

A third class of futures efficiency studies has focused on uncovering statistical anomalies in futures price data. The implicit assumption is that statistical anomalies, if persistent, can be used to construct statistically based forecasts of futures prices which can be used to earn abnormal trading profits. Included in this group are Kolb and Gay, Kolb, Vukina, and Dorfman. Closely related to the search for statistical anomalies are tests of rational price formation, as represented by Shonkwiler and Koontz, Hudson, and Hughes.

Livestock futures efficiency has been studied more often than grain futures efficiency. Twelve of the 18 studies reported above and all of those that involved trading simulations analyzed only cattle and/or hog futures. Futures efficiency studies have rarely focused on wheat futures. Only 4 of the 41 price series examined in 15 forecasting studies reported in the 1988 metanalysis of Garcia, Hudson, and Waller included wheat—only one involved KCBOT wheat futures. In that study (Just and Rausser) the time period analyzed was limited, involving only 1976 through 1978. They found wheat futures forecasted more accurately (mean squared error) than 5 large econometric models across 4 forecast horizons in 16 of 20 pairwise comparisons.

In the 18 efficiency studies reviewed here, besides Just and Rausser, only one involved wheat futures. In that study, Kolb examined 29 different commodities from the 1950s through 1988 primarily for normal backwardation (tendency to include a risk premium).[1] Especially notable among the results was the efficiency of wheat futures relative to other commodities. Among 29 commodities examined and four statistical tests conducted, wheat was the sixteenth most efficient commodity in one test, the seventh most efficient in another test, tied for first place in the third test, and the most efficient commodity in the fourth test. In addition, wheat futures were more efficient than either soybeans or corn (except for one test with soybeans more efficient than wheat).

Of interest is whether Kolb's substantial evidence in support of statistical efficiency in wheat futures has manifested itself in findings of economic efficiency. That is, when attempts to extract profits from the futures markets have been made, have they been successful? One possible answer to that question has come from the popular press. *Top Producer* magazine has routinely tracked and occasionally reported on the marketing advice of 7 to 10 private marketing advisory firms for 1988–96 crop years (information provided by Merrill Lynch). With each observation representing profits from trading 5,000 bushels with one advisory firm and for one crop year, examination of the reported results for suggested wheat futures strategies yielded 56 observations, −3.48¢/bu. mean profit, 16.4¢/bu. standard deviation. The same statistics for corn are 65, 1.79, 17.1 and

[1] Although not explicitly stated, the context of Kolb's study suggested that Chicago wheat futures were examined, not Kansas City.

for soybeans 61, 5.89, 39.7.[2] With smaller wheat (relative to corn or soybeans) futures profits, this popular press evidence is consistent with Kolb's findings that wheat futures are relatively more efficient than corn or soybean futures.

General Analytical Procedures

Futures efficiency tests constructed in a forecasting/trading simulation framework must be judged on the basis of plausibility in several dimensions (Kastens and Schroeder). The study period must be of sufficient length to ensure results are not merely spurious and can be generalized. Forecasting methods, models, and trading systems must be simple (to ensure low construction costs) and historically feasible. Little confidence is instilled in results from procedures whose discovery or implementation would have been unlikely during the time period of study. Finally, if positive simulated trading profits are used to reject futures efficiency, they should be of sufficient magnitude to account for risk.

Time Period Studied

The July KCBOT wheat futures price for the twenty-fifth week of each calendar year, beginning in 1947 and ending in 1995, is forecasted each week during the preceding six months.[3] A simulated futures trade is constructed corresponding to each forecast.

Forecasting Technique

Forecasts are constructed with simple econometric modeling techniques. Econometric forecasts have fallen into some disfavor over the last two decades due to the advance of slightly more accurate extrapolative methods such as various autoregressive models (see Allen).[4] Recently, in spite of lower accuracy, causal econometric forecasts have been shown to be more economically valuable than extrapolative methods when used to determine futures trades (see Gerlow, Irwin, and Liu; and Leitch and Tanner). Thus, when the goal is trading profits generation, causal econometric models are appropriate.

Explanatory Variables

Explanatory variables used in econometric forecasting models should be selected based on underlying economics affecting harvesttime price of July wheat futures. Such vari-

[2] The 1995/96 and 1996/97 marketing year results were only through 1 February 1996. Crop year marketing gains associated with a specific advisory firm and for a given 5,000-bushel quantity of grain were the cumulative profits from several futures transactions. That is, futures positions were placed and removed several times in a year and not always in the same contract month. No adjustments were made for trading commissions. Allowance for commissions and advice subscription costs would have substantially diminished trading profits.

[3] A week is represented by Wednesday's closing price. Week 1 is defined as the week containing the first Monday of the year. Successive weeks are numbered consecutively. The Wednesday of week 25 is associated with an average date around 23 June, which corresponds to harvesttime for much of Kansas—an important point for which price forecasts are often needed [Kansas Agricultural Statistics shows that a disproportionate share—41% for 1991–95—of the state's wheat crop is marketed during June and July, the two months encompassing its harvest (Kansas Department of Agriculture)]. The first forecast is made in week 51 of the previous calendar year.

[4] *Econometric* is used here in its broad sense, meaning that explanatory variables are included in regression based forecasts because of their expected economic relationships with wheat prices. This is distinguished from the more narrow definition surrounding the multiequation structural econometric models popular in the 1960s and 1970s. More important, *econometric* is used to distinguish the models from extrapolative models, which tend to rely on the statistical properties of a single price series over time.

ables represent additional market information to be used in price forecasting, thus futures efficiency testing—information that may not be adequately captured in deferred futures prices only. Accordingly, focus is on expected demand and supply of U.S. wheat as represented by publicly available direct and indirect projections by the United States Department of Agriculture (USDA) for domestic usage, exports, and production. Data are taken directly or calculated from USDA's *Wheat Situation and Outlook, Crop Production,* or *World Agricultural Supply and Demand Estimates* (WASDE) reports.[5]

Deferred futures prices contain considerable information regarding market expectations of future prices (Eales et al.; Gardner). Therefore, deferred futures prices are perhaps the most important information to begin with in developing a wheat price forecasting model. In addition to deferred futures prices, other market information may also affect future prices. The government loan program influences wheat acreage, in addition to expected cash price (Lidman and Bawden; Morzuch, Weaver, and Helmberger). Therefore, wheat loan price is included to capture the influence of changing government farm programs. Major demand and supply wheat balance sheet information could also influence future price (especially if the wheat futures market is not efficient as discussed later). Thus, wheat exports and usage relative to stocks are included as well as production to account for the primary fundamental information affecting wheat price over time. Temporal price spreads between adjacent futures contracts may also affect future price. A large carrying charge between September and July futures may induce more wheat storage at harvest supporting the harvesttime price. Alternatively, inverse carrying charges would induce less storage (Benirschka and Binkley). The September-to-July wheat futures price spread was included to capture this effect.

Model specification, including explanatory variable selection, is always questionable, especially in the realm of futures efficiency testing. Conclusions regarding efficiency are inherently entangled with, and tainted by, model specification. To minimize this problem, a literature search of wheat price forecasting models was conducted to assist in explanatory variable selection. The variables used here and the form in which they entered models (e.g., ratios or levels) represent a generalization from the literature search. Although it was not uncommon to find variables entering models at higher orders [one USDA model (Barr) included the ratio of normal food use to ending stocks, raised to the fifth power], only first-order terms have been included here to ensure plausibility of the models.

Econometric Procedures

In developing and updating price forecasts, econometric techniques should be driven by appropriate statistical methods. Attempts should be made to deal with well-understood econometric difficulties, including estimation sample size, model specification, heteroskedasticity, and autocorrelation. When data availability is not the primary determinant of estimation sample length, a forecaster might consider a time period when data relationships are stable. Parameter stability persisting at least into the forecast time period is an assumption underlying all econometric forecasting. Under the mild assumption that

[5] When production projections are absent, ancillary forecasts of wheat yield are constructed by regressing the last 10 years' yields on time. Planted acres, domestic usage, and exports, if not yet projected, are set equal to the most recent USDA estimate of last year's planted acres, domestic usage, or exports, respectively. That is consistent with naive forecasting in the absence of better information. In all cases, forecasts use only information available at the time of forecasting.

the most recent data are always relevant, a forecasting model's estimation sample size was chosen by stepping backwards in time, beginning with the minimum sample size that ensures sufficient degrees of freedom to conduct the required tests, incrementing by one, and stopping when parameter stability breaks down. Harvey and Collier's recursive residuals test, as described by Greene, determined parameter stability. The estimation sample size was one less than the size associated with the stopping point.

Model functional form was chosen to be either linear or double-log according to MacKinnon, White, and Davidson's P_E model specification test. Heteroskedasticity, when found by the Breusch-Pagan test, was dealt with by using a double-log specification (a crude but effective solution because the log transformation tends to scale down variables and hence the residuals). When first-order autocorrelation was uncovered by a chi-squared test for independent residuals, stationarity was imposed with a maximum likelihood joint estimation of the parameters and a serial correlation coefficient constrained between -1 and 1 (first observations were transformed using the Prais-Winsten procedure).

Forecast Models

The linear specification for estimation models underlying forecast models is depicted by the expression:

$$(1)\quad \begin{aligned} JUL_{harvest,t} = {} & \alpha_i + \beta_{i,1}JUL_{harvest-i,t} + \beta_{i,2}LOAN_{harvest-i,t} + \beta_{i,3}EXPORT_{harvest-i,t} \\ & + \beta_{i,4}USAGE_{harvest-i,t} + \beta_{i,5}SEPJUL_{harvest-i,t} \\ & + \beta_{i,6}PRODUCTION_{harvest-i,t} + \epsilon_{i,t}; \end{aligned}$$

$$\text{for } T = 1946 \text{ to } 1994, \text{ for } i = 1 \text{ to } 26, \text{ and for } t = T - SAMPSIZ + 1 \text{ to } T;$$

where T represents the year associated with the last observation in an estimation sample of length *SAMPSIZ*; the forecast horizon, i, ranges from 1 to 26 weeks; and *harvest* is a scalar equal to 25, representing the twenty-fifth week of the calendar year. *JUL* represents the price of KCBOT July wheat futures ($/bu.), and *LOAN* is the ratio of the government loan price of wheat to current July futures price. *EXPORT* is the ratio between projected exports for the crop year ending 31 May and projected beginning stocks for the year beginning 1 June. *USAGE* is the ratio between projected domestic usage and projected beginning stocks for the same time periods used in *EXPORT*. *SEPJUL* is the spread between September and July futures (September less July + $0.20—the additive constant precludes logging negative numbers when the double-log specification is used), and *PRODUCTION* is projected current year wheat production (bil. bu.). The earliest t available is 1937. The αs, βs, and ϵs are model-specific parameters to be estimated and error terms, respectively. Assuming parameters have been estimated, and using the expected-value notation to make it clear what a model is forecasting, the forecast model associated with each linear estimation model in (1) is

$$(2)\quad \begin{aligned} \mathrm{E}[JUL_{harvest,T+1}] = {} & \alpha_i + \beta_{i,1}JUL_{harvest-i,T+1} + \beta_{i,2}LOAN_{harvest-i,T+1} \\ & + \beta_{i,3}EXPORT_{harvest-i,T+1} + \beta_{i,4}USAGE_{harvest-i,T+1} \\ & + \beta_{i,5}SEPJUL_{harvest-i,T+1} + \beta_{i,6}PRODUCTION_{harvest-i,T+1}. \end{aligned}$$

Because T denotes years, the $T+1$ subscript components in (2) make it clear that forecasts are one step ahead, where the step is one year. The system described by (1) and (2) represents 1,274 models and 1,274 point forecasts of harvesttime price (49 years forecasted—1947 to 1995—coupled with 26 forecast horizons).[6]

Trading Simulations

The selected trading scheme was decidedly simple. When a price forecast was higher (lower) than the current futures price, a long (short) position was entered. All futures positions were closed in week 25 each year. Using the notation in (2), a trade was initiated in week *harvest-i* and closed at *harvest.*

Futures Efficiency

A joint F-test on the parameter estimates for the nonfutures explanatory variables in (1) (*LOAN, EXPORT, USAGE,* and *PRODUCTION*) constitutes a statistical test of futures efficiency for each forecasting model. If the test fails to reject the hypothesis that the estimates are jointly zero, then futures efficiency is not rejected, because the nonfutures variables contribute no useful additional information for explaining harvest price over that already provided in the deferred futures and futures spread series.[7] The futures efficiency test can be represented in discrete form (rejection or nonrejection) or in continuous form by considering the p-value associated with the test (*FPVALUE*) as an indication of the relative efficiency of futures across the 1,274 statistical test values provided.

With the second way of examining futures efficiency, the relative forecast accuracy between the forecast provided by deferred futures price and that provided by models designed to predict harvesttime futures price is determined. If models forecast no more accurately than deferred futures, then harvesttime futures market efficiency cannot be rejected. Mathews and Diamantopoulos placed forecast accuracy measures into four distinct classes. They suggested at least one measure from each class be included in a comprehensive analysis of forecast accuracy. Representing their bias-type class, mean error (ME) is used, and their ratio-type class, mean absolute percentage error (MAPE). Root mean squared forecast error (RMSE) is used to represent their volume-type class and the squared linear correlation coefficient between the actual and forecast series (RSQ) is used to represent their pattern-matching class. High RSQ but low ME, MAPE, or RMSE is desirable.

Forecast accuracy measures provide discrete tests of futures efficiency during selected time periods. Absolute percentage error (APE) can be used to examine relative futures efficiency over time. The difference constructed by subtracting, for each point forecast, the APE for deferred futures from the APE for the model forecast (*APEDIF*) provides a measure of relative efficiency at that point (large $\rightarrow$ efficient).

The third approach to futures efficiency examination involves testing for sufficiently positive profits resulting from trading simulations based upon the forecasts. This is the most important test because it tests the necessary condition underlying futures efficiency.

[6] Because a double-log model involves a nonlinear transformation of the dependent variable, an unbiased forecast requires that the inverse functional transformation of the model-predicted left-hand side be multiplied by the factor, $\exp(\frac{1}{2} s^2)$, where s^2 is the MSE of the underlying regression (see Neyman and Scott). These adjustments were made for all predictions from double-log models.

[7] Following Fama's delineation, this is a test for *semistrong* form efficiency, which is a test of the value of using publicly available information in a futures price forecast model.

As in the first two methods of testing efficiency, profits provide a discrete test of efficiency for a selected time period (either the profits are sufficiently positive or they are not). So, profit magnitude can be used to measure relative futures efficiency over time (small → efficient). Because the time period studied is long and associated with widely varying nominal prices, the measurement of profits used here is returns (*RETURN*), defined as trading profit divided by trade entry price, both in cents per bushel. The interpretation is rate of return on investment, where the investment is the full face value of a futures contract.

Of further interest is how the three alternative measures of futures efficiency (*FPVALUE, APEDIF,* and *RETURN*) compare, especially regarding the directional changes in efficiency across time associated with each of the three measures. To uncover these changes three regression models are estimated, one for each measure of futures efficiency. Efficiency is likely to have changed over the years studied due in part to large improvements in informational technology. Efficiency may be different across different forecast horizons. Although deferred futures may be less accurate at distant relative to near horizons, they may still provide the best available forecast at distant horizons, making the expected efficiency implications unclear. Furthermore, the relative difficulty in predicting delivery time futures price between long and short horizons may have changed over the years (Kenyon, Jones, and McGuirk). To capture these temporal effects *HORIZ* (forecast horizon, in weeks 1–26), *YEAR* (year whose harvest price is forecast, 1947–95), and *HORIZYEAR* (an interaction term between *HORIZ* and *YEAR*) are considered as explanatory variables. Because the three efficiency measures (especially *FPVALUE*) are closely tied to econometric features of forecast models, the following variables are also included: *SAMPSIZ,* estimation sample size; *PETEST,* 1 if P_E test selected the double-log specification, else 0; *HET,* 1 if the heteroskedasticity test selected the double-log model, else 0; and *AUTO,* 1 if the estimation model was corrected for autocorrelation, else 0.

Results and Discussion

Although not essential, a brief discussion of the forecasting models' parameter estimates may enhance interpretation for the reader. Because some models were double-log, partial elasticities should be more relevant than parameter estimates. Furthermore, because models often involved many years, years in which underlying explanatory variables have substantially trended, elasticity estimates may be more understandable when they are evaluated at the most recent independent variable values (those used in making the actual forecasts). Across the 1,274 estimated models, for each of the independent variables abbreviated in equation (1), the mean of the partial elasticity estimates, along with the median of the p-values associated with testing the underlying parameter estimates against 0 (reported in parentheses), are as follows: deferred July futures price, 0.96 (0.00); loan price ratio, 0.13 (0.16); exports ratio, 0.004 (0.55); domestic usage ratio, −0.01 (0.49); September/July futures price spread, −0.12 (0.35); and U.S. wheat production, 0.13 (0.20). The average R^2 for the 1,274 models was 0.96. The associated elasticity estimates mean depicts the strong relationships between deferred and harvesttime July futures prices (near 1 and typically highly significant). Although the typical loan price ratio statistical confidence level was 84%, less confidence was associated with the remaining estimates. The potential difficulty associated with using such models to predict harvesttime prices of July futures more accurately than do deferred July futures prices is immediately apparent from the large median p-values.

Overall, the F-tests for efficiency resulted in 505 of 1,274 (40%) futures efficiency rejections.[8] Although futures efficiency was rejected using this test in a larger percentage of models than might be expected by chance, given that a null of efficiency were true, this limited result alone is not particularly convincing or illuminating. Over the 1,274 point forecasts, the forecast accuracy measures associated with deferred futures prices are as follows: ME 0.0061, MAPE 0.0630, RMSE 0.2912, RSQ 0.9045. For the model forecasts the measures are ME 0.0101, MAPE 0.0854, RMSE 0.3660, RSQ 0.8517. By each measure, model forecasts are nominally less accurate than deferred futures. In addition, they are statistically less accurate according to the MAPE and RMSE statistics.[9] Economically more important, the no-commissions returns mean for the 1,274 trades (at -0.0073, or -0.73%) is statistically less than zero. The forecast error comparisons and mean returns are all indicative of market efficiency in the KCBOT July wheat futures market.

Not every forecast should necessarily result in a trade, because some forecasts of harvesttime price may not be sufficiently different from deferred futures price to cover commissions. Therefore, a filter of 0.003 on forecasted returns was used to screen out trades associated with insufficient returns (0.003 represents commissions of $45 for a 5000-bushel contract trading at $3). This filter, excluding 87 trades, resulted in mean returns of -0.008 (no commissions, 1,187 trades). Increasing the filter at increments of 0.003 resulted in nominally positive mean returns at a filter value of 0.036 (611 trades, mean returns 0.0003). However, mean returns did not become statistically positive until using a filter value of 0.168 (only 65 trades, mean returns 0.026).

Historical trading scheme optimization routines are of little value in efficiency studies unless they are robust into the future. Accordingly, the filter optimizing routine was made operational by using the filter which had generated the largest mean returns during 1947 in 1948. In 1949 the one which had generated the largest mean returns during 1947 through 1948 was used, and so on. Overall, trading the 1948 through 1995 period (1,248 possible trades) resulted in mean returns of 0.057. However, only nine trades were made in the 48 years![10] Furthermore, no trades were made until 1964. In spite of the potentially abnormal mean returns, it is doubtful that any trader would have the patience to wait 17 years before taking a system-based trade. Thus, these results cannot be used to reject efficiency in the KCBOT July wheat futures market.

To enhance understanding of the three alternative measures of futures efficiency described in the Futures Efficiency section (*FPVALUE, APEDIF,* and *RETURN*), the three associated explanatory models described there were estimated. Model results are reported in table 1. The *FPVALUE* and *RETURN* Pearson linear correlation coefficient and Spearman rank correlation coefficient are -0.04 and -0.26, respectively. For *APEDIF* and

[8] Unless indicated, all statistical tests in this research are at the 0.05 significance level.

[9] Forecast error MAPE and mean returns were tested with a signed-rank Wilcoxon test and RMSE was tested with the Ashley, Granger, and Schmalensee test. The forecast ME for deferred futures, although nominally lower than for model forecasts, was not statistically lower. The RSQ difference was not tested.

[10] The 0.057 mean is statistically larger than zero at a p-value of 0.047. Considering a trading investment to be the face value of a contract, subtracting 0.003 from a return to allow for commissions, and subsequently multiplying by 52 and dividing by the associated forecast horizon (number of weeks that the position was held), yields an *annualized* rate of return on investment. The mean annualized rate of return for the nine trades is 0.132, or 13.2%.

Table 1. Parameter Estimates of Futures Efficiency Test Value Determinants: KCBOT July Wheat Futures, 1947–95

Parameter Estimates	Efficiency Model		
	FPVALUE	*APEDIF*	*RETURN*
Intercept	−12.8658**	−3.4322**	1.5067
	(2.2988)	(0.6062)	(0.9456)
HORIZ	0.1743	0.0176	0.1115*
	(0.1443)	(0.0334)	(0.0519)
YEAR	0.0068**	0.0018**	−0.0008
	(0.0012)	(0.0003)	(0.0005)
HORIZYEAR	−0.0001	−0.000009	−0.00006*
	(0.0001)	(0.000017)	(0.00003)
SAMPSIZ	0.0004	0.0018**	0.0004
	(0.0005)	(0.0002)	(0.0003)
PETEST	−0.0233	0.0162*	0.0098
	(0.0176)	(0.0071)	(0.0111)
HET	−0.0052	−0.0017	0.0443**
	(0.0133)	(0.0056)	(0.0089)
AUTO	−0.0888	0.0322**	0.0051
	(0.0175)	(0.0082)	(0.0129)
Model R^2	0.77	0.08	0.05
F-Value	34.5118**	16.2023**	10.2909**
No. obs.	1,274	1,274	1,274

Notes: *FPVALUE* and *APEDIF* models were estimated with autocorrelation-correction models. Significance at 0.05 and 0.01 is denoted with one and two asterisks, respectively. Standard errors are in parentheses.

RETURN, corresponding coefficients are −0.43 and −0.67. These values are as expected, reflecting larger trading profits associated with periods of less efficiency.

Table 1 shows that parameter estimates representing econometric modeling features (*SAMPSIZ, PETEST, HET,* and *AUTO*) were not significant in the *FPVALUE* model. That is surprising considering that both larger estimation sample sizes and using econometrically correct models should make it easier to generate more reliable parameter estimates (required for the underlying futures efficiency F-tests). On the other hand, *SAMPSIZ* was negative in the *APEDIF* model. That is likely because large sample sizes reflect long periods of parameter stability, which translate into more accurate model forecasts (smaller APEs), which imply smaller *APEDIF*s. Choosing the double-log specification because of the P_E test (*PETEST*) and correcting for autocorrelation (*AUTO*), both appear to be associated with a decrease in forecasting accuracy of the econometric model relative to deferred futures, as judged by the positive relationships with *APEDIF.* Unfortunately, results are not segregated between *finding* and *correcting* the econometric problems. So, it should not be inferred that implementing econometric corrections leads to less accurate forecasting. Profits rise when heteroskedasticity is found and corrected (positive *HET* in *RETURN* model). This suggests that correcting for heteroskedasticity by selecting the double-log specification may be beneficial in constructing forecasts for generating trading profits. That is consistent with the negative sign on *HET* in the *APEDIF* model, which suggests improvement in forecast accuracy, and the negative sign on

HET in the *FPVALUE* model, which suggests better use is being made of the nonfutures information in the forecasting models.

The *FPVALUE* and *APEDIF* models show that futures efficiency has increased over the years. These results are generally corroborated by the *RETURN* model, where a change in *RETURN* for a change in *YEAR* is negative across all forecast horizons. That is, trading returns deteriorated over the years, as efficiency improved. In addition, the *RETURN* model shows that trading returns deteriorated even faster for longer horizons relative to shorter ones, as judged by the negatively significant *HORIZYEAR* estimate. In that regard, if the increase in efficiency was due to information technology improvements (e.g., speed of information transmission, or faster computers for building forecast models and implementing trading systems), then the improvements must have impacted the longer term more than the shorter term forecasts. The change in *RETURN* for a change in *HORIZ* is positive in 1947, falls to zero around 1983, and becomes negative after that. This suggests that the purported increased impact of technology changes on longer term relative to shorter term forecasts may have primarily come only in the last 12 years.

Conclusions

Econometric models were constructed using the most current USDA projections for wheat exports, domestic usage, and production, along with the government loan price and current price of July KCBOT wheat futures, to forecast harvesttime futures price. Models were constructed for each week in the six months preceding harvest, which are periods of interest to Kansas farmers who may consider forward pricing their crops prior to harvest. Joint significance of the nonfutures explanatory variable parameter estimates provides a test of Fama *semistrong* form efficiency in each model. The point accuracy of each model's postsample forecast, relative to the accuracy associated with using current deferred futures price as a forecast of future price, provided a second measure of efficiency. Examination of simulated futures trading returns based upon the forecasts provided the third and most important measure of efficiency.

Each of the three efficiency measures showed that July KCBOT wheat futures efficiency has generally increased over the last 50 years—for all forecast horizons in the six-month period preceding harvest. Overall, simulated trading returns were not abnormally positive. In fact, they were typically negative over the period studied, even without considering trading commissions. If July wheat futures are to be deemed efficient or inefficient over the last 50 years, this research is consistent with efficient. Although forecast model/trading simulation tests of futures efficiency, such as those conducted in this research, are always joint tests of efficiency and model specification. Unless outlook economists can devise better forecast models than those developed here, deferred futures price is the best estimate of harvesttime price from six months prior to harvest up to harvest.

[*Received November 1995; final version received May 1996.*]

References

Allen, P. G. "Economic Forecasting in Agriculture." *Internat. J. Forecasting* 10(1994):81–135.

Ashley, R., C. W. J. Granger, and R. Schmalensee. "Advertising and Aggregate Consumption: An Analysis of Causality." *Econometrica* 48(1980):1149–167.

Barr, T. N. "Demand and Price Relationship for the U.S. Wheat Economy." *Wheat Situation-226,* pp. 15–25. U.S. Department of Agriculture, Economic Research Service, November 1973.

Benirschka, M., and J. K. Binkley. "Optimal Storage and Marketing Over Space and Time." *Amer. J. Agr. Econ.* 77(1995):512–24.

Bessler, D. A., and J. A. Brandt. "An Analysis of Forecasts of Livestock Prices." *J. Econ. Behavior and Organization* 18(1992):249–63.

DeCoster, G. P., W. C. Labys, and D. W. Mitchell. "Evidence of Chaos in Commodity Futures Prices." *J. Futures Mkts.* 12(1992):291–305.

Dorfman, J. H. "Bayesian Efficiency Tests for Commodity Futures Markets." *Amer. J. Agr. Econ.* 75(1993): 1206–210.

Eales, J. S., B. K. Engle, R. J. Hauser, and S. R. Thompson. "Grain Price Expectations of Illinois Farmers and Grain Merchandisers." *Amer. J. Agr. Econ.* 72(1990):701–8.

Fama, E. F. "Efficient Capital Markets: A Review of Theory and Empirical Work." *J. Finance* 25(1970): 383–417.

Garcia, P., M. A. Hudson, and M. L. Waller. "The Pricing Efficiency of Agricultural Futures Markets: An Analysis of Previous Research Results." *S. J. Agr. Econ.* 20(1988):119–30.

Garcia, P., R. Leuthold, T. R. Fortenbery, and G. Sarassoro. "Pricing Efficiency in the Live Cattle Futures Market: Further Interpretation and Measurement." *Amer. J. Agr. Econ.* 70(1988):162–69.

Gardner, B. L. "Futures Prices in Supply Analysis." *Amer. J. Agr. Econ.* 58(1976):81–84.

Gerlow, M. E., S. H. Irwin, and T. Liu. "Economic Evaluation of Commodity Price Forecasting Models." *Internat. J. Forecasting* 9(1993):387–97.

Greene, W. H. *Econometric Analysis,* 2nd ed. New York: Macmillan, 1993.

Grossman, S. J., and J. E. Stiglitz. "On the Impossibility of Informationally Efficient Markets." *Amer. Econ. Rev.* 70(1980):393–408.

Harvey, A., and G. Collier. "Testing for Functional Misspecification in Regression Analysis." *J. Econometrics* 6(1977):103–19.

Helmuth, J. "A Report on the Systematic Downward Bias in Live Cattle Futures Prices." *J. Futures Mkts.* 1(1981):347–58.

Just, R. E., and G. C. Rausser. "Commodity Price Forecasting with Large-Scale Econometric Models and the Futures Market." *Amer. J. Agr. Econ.* 63(1981):197–208.

Kansas Department of Agriculture. Kansas Agricultural Statistics. *Crops* 95,12(12 December 1995):2.

Kastens, T. L., and T. C. Schroeder. "A Trading Simulation Test for Weak-Form Efficiency in Live Cattle Futures." *J. Futures Mkts.* 15(1995):649–75.

Kenyon, D., E. Jones, and A. McGuirk. "Forecasting Performance of Corn and Soybean Harvest Futures Contracts." *Amer. J. Agr. Econ.* 75(1993):399–407.

Kolb, R. W. "Is Normal Backwardation Normal?" *J. Futures Mkts.* 12(1992):75–91.

Kolb, R. W., and G. D. Gay. "The Performance of Live Cattle Futures as Predictors of Subsequent Spot Prices." *J. Futures Mkts.* 3(1983):55–63.

Koontz, S. R., M. A. Hudson, and M.W. Hughes. "Livestock Futures Markets and Rational Price Formation: Evidence for Live Cattle and Live Hogs." *S. J. Agr. Econ.* 24(1992):233–49.

Koppenhaver, G. D. "The Forward Pricing Efficiency of the Live Cattle Futures Market." *J. Futures Mkts.* 3(1983):307–19.

Leitch, G., and J. E. Tanner. "Professional Economic Forecasts: Are They Worth Their Costs?" *J. Forecasting* 14(1995):143–57.

Lidman, R., and D. L. Bawden. "The Impact of Government Programs on Wheat Acreage." *Land Econ.* 50(1974):327–35.

MacKinnon, J. G., H. White, and R. Davidson. "Tests for Model Specification in the Presence of Alternative Hypotheses." *J. Econometrics* 21(1983):53–70.

Martin, L., and P. Garcia. "The Price-Forecasting Performance of Futures Markets for Live Cattle and Hogs: A Disaggregated Analysis." *Amer. J. Agr. Econ.* 63(1981):209–15.

Mathews, B. P., and A. Diamantopoulos. "Towards a Taxonomy of Forecast Error Measures." *J. Forecasting* 13(1994):409–16.

Morzuch, B. J., R. D. Weaver, and P. G. Helmberger. "Wheat Acreage Supply Response under Changing Farm Programs." *Amer. J. Agr. Econ.* 62(1980):29–37.

Neyman, J., and E. L. Scott. "Correction for Bias Introduced by a Transformation of Variables." *Annals of Mathematical Statis.* 31(1960):643–61.

Palme, L. A., Jr., and J. Graham. "The Systematic Downward Bias in Live Cattle Futures: An Evaluation." *J. Futures Mkts.* 1(1981):359–66.

Pluhar, D. M., C. E. Shafer, and T. L. Sporleder. "The Systematic Downward Bias in Live Cattle Futures: A Further Evaluation." *J. Futures Mkts.* 5(1985):11–20.

Peterson, P. E., and R. M. Leuthold. "Using Mechanical Trading Systems to Evaluate the Weak Form Efficiency of Futures Markets." *S. J. Agr. Econ.* 14(1982):147–51.

Prais, S., and C. Winsten. "Trend Estimation and Serial Correlation." Discus. Pap. 383, Cowles Commission, Chicago, 1954.

Shonkwiler, J. S. "Are Livestock Futures Prices Rational Forecasts?" *West. J. Agr. Econ.* 11(1986):123–28.

Top Producer. Philadelphia PA: Farm Journal Inc. Various issues, 1988–96.

U.S. Department of Agriculture. *Crop Production.* Washington DC: National Agricultural Statistical Service. Several issues, 1942–95.

———. *The Wheat Situation.* Washington DC: Economic Research Service. Several issues, 1936–73.

———. *World Agricultural Supply and Demand Estimates* (WASDE). Washington DC: World Agricultural Outlook Board. Several issues, 1973–95.

Vukina, T. "Hedging with Forecasting: A State-Space Approach to Modeling Vector-Valued Time Series." *J. Futures Mkts.* 12(1992):307–27.