



The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

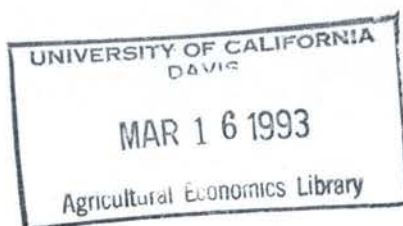
University of California, Davis
Department of Agricultural Economics

Working papers are circulated by the author without formal review. They should not be quoted without author's permission. All inquiries should be addressed to the author, Department of Agricultural Economics, University of California, Davis, California 95616-8512 USA.

COINTEGRATION AND SETTLEMENT OF COMMODITY
FUTURES CONTRACTS

by
Kenneth A. Foster and Arthur M. Havenner

Working Paper No. 92-09



1992
2384
Livestock trade -- Futures

Cointegration and Settlement of Commodity Futures Contracts

Kenneth A. Foster and Arthur M. Havenner¹

July 22, 1992

ABSTRACT

Intuitively, weekly live cattle prices in various markets would be expected to share common trends, *i.e.*, they cannot be driven by separate nonstationarities since at some point the prices will diverge sufficiently for it to be economic to cross-ship the cattle. This paper extends previous bivariate work to a multivariate analysis which is capable of modeling the transitive linkages in prices for many geographical regions. The empirical estimation represents the application of an innovation on Aoki's Linear Systems State Space model.

A nonparametric test is used to evaluate the value of arbitrage forecasts implied by the structure of the model. The arbitrage relationship is also employed to generate efficient discounts/premiums for either physical delivery or cash settlement of futures contracts. Because it is derived from a model with fewer unobservable states than prices, the proposed settlement mechanism accounts for spatial arbitrage opportunities and therefore will better represent the true geographical discounts faced by traders in individual markets.

¹ Ken Foster is an Assistant Professor of Agricultural Economics at Purdue University and Arthur Havenner is a Professor of Agricultural Economics at the University of California, Davis. The authors wish to thank Chris Hurt and J. William Uhrig for sharing their understanding of agricultural commodity futures markets.

1 Introduction

Commodity futures markets play a large role in U.S. agriculture and provide potential benefits to a wide range of society. These roles range from risk shifting for hedgers to price stabilization for consumers. The possibility of making delivery of the actual commodity as a means of offsetting a short futures position links these futures contracts to actual market conditions. In the case of all agricultural commodities except feeder calves, the mechanism for delivery on agricultural commodity futures contracts is physical exchange of the commodity. Delivery can be made at any of several spot markets designated as delivery sites with some of these sites being at par, and the others being at a discount or a premium. However, the bulky and/or perishable nature of agricultural commodities makes such transactions logistically difficult with potentially high transactions costs.

The mechanics of physical delivery can be confusing to the casual observer. Delivery becomes an option for the seller of a futures contract in the expiration month of the contract. However, it will seldom be chosen because offsetting the futures position is usually more profitable. Although delivery is infrequent, the opportunity to make or take delivery always exists and, if markets are not operating efficiently, delivery will occur to force the appropriate degree of convergence between the cash and futures prices. Identical cash and futures prices are unlikely due to transportation cost differences and market characteristics, but there is an equilibrium market relation between these prices, institutionalized in the discount and premium structure and arbitrated by actual delivery. Since delivery incurs substantial transactions costs, disproportionate amounts of delivery in a market are evidence of an improperly functioning market, often the result of an inflexible discount and premium structure unable to keep up with new developments. The net result is excessive costs to the industry and those who purchase its product.

Delivery on futures contracts is made by the seller providing a negotiable instrument to the buyer. This usually takes the form of a warehouse or stockyard receipt or a shipping certificate. Note that delivery does not require the seller to actually transport the commodity to the point of sale. The seller can purchase from a warehouse or stockyard the needed quantity of the commodity.² As a result, a group of animals or a bin of grain at a certified delivery point may be delivered many times without physically moving. Each delivery would generally be made at a different price until an equilibrium is reached.

The exchange designates which cash market locations can be used as delivery sites. Geographical discounts and premiums are often set by the exchange. That is, the settlement price for physically delivered contracts is the short futures established by the seller less a discount or plus a premium for the place where delivery is made. Some or all of the sites may trade on par with the futures, which is itself a discount/premium structure. The premiums and discounts are set based on historical evidence and do not account for factors such as seasonality and the distance of the delivery point from the center of production; in general, they are flat rates that remain constant for relatively long periods of time. Thus, for example, as production shifts geographically this institutionalized discount/premium structure tends to encourage arbitrage even if it was initially set optimally. Also, the cash and futures represent distinctly different assets. While they are highly correlated, they differ stochastically rather than by a constant as the exchange determined premiums and discounts would suggest. The *ad hoc* nature of the current discounts, premiums, and pars may result in a market failure: undesirable delivery will sometimes occur because the premium or discount is smaller than the market would dictate.

² This method has an added advantage. Purchasing the commodity reduces the risk of not meeting the quality standards set down in the futures contract. This eliminates quality discounts which would undoubtedly arise if commodity were shipped from a farm or other point of origin.

Delivery represents an attempt to arbitrage in two dimensions, space and time. The spatial considerations are easier to conceptualize because they are related to the basis between the delivery market and the futures market. That is, delivery is triggered when the basis (futures minus cash) exceeds the *ad hoc* discount plus the associated transactions costs. The time aspect is related to the fact that the short futures position of the seller was established some time in the past. Therefore, the decision to deliver is also based on the current futures and cash prices relative to the price established previously in the futures transaction.

Commodity prices in some spatially different cattle markets have been shown by Goodwin and Schroeder to be cointegrated, i.e., they follow a common trend. Intuitively, this must be so because these markets generate prices for nearly homogeneous commodities and will generally only differ by grading differences, local demand and supply, transportation, and other transactions costs of moving the commodity from one point to another. Arbitrage opportunities would ensure that these prices ultimately move together. Goodwin and Schroeder found that, in general, the distance between two markets was inversely related to the level of relatedness in the prices. However, their bivariate models were incapable of modeling any transitive linkages. Therefore, in order to model the full spatial linkages between spot markets it is necessary to formulate a multivariate model which includes many price series. In fact, the commodity futures price should also be included in the list of cointegrated prices since it represents a highly correlated asset and, because delivery is possible, its price will not infinitely diverge from those of the spot markets. Heuristically, we would expect the futures price to be highly correlated with the dominant common trend in spot prices because the cash markets will tend to converge to the futures as contract expiration nears.

An arbitrary set of discounts, premiums, and pars cannot represent the cointegrated nature and the spatial arbitrage opportunities until after the market has begun to exploit them: excess deliveries trigger the adjustment of the discounts and premiums by the exchanges. In order for the futures market, and by association the spot markets, to operate efficiently, any deliveries should be made on a price which reflects the current spatial relationships and not those which existed in the past.

This paper suggests a possible method of utilizing the cointegration between a subset of spot prices and the nearby futures price to develop price indices which reflect the arbitrage opportunities of the various markets. These indices are used to generate weekly estimates of the appropriate discounts or premiums associated with individual markets. The next section develops the methodology used to obtain parameter estimates and to derive the arbitrage price indices. In the third section an application to live cattle is presented. The accuracy of this model is assessed from the standpoint of predictive accuracy using uncompromised post-sample observations and a set of non-parametric tests of forecast value. The final section of the paper is reserved for a conclusion and summary.

2 State Space Model with Cointegration

There are several time series methodologies by which a model can be determined and estimated. For this study, the linear system state space representation due to Aoki was chosen³. The state space representation differs from the original development by Granger (1983) and Engle and Granger (1987) demonstrating the potential loss of information from differencing individual series in a multivariate time series setting and leading to the notion of cointegration, which was developed in an ARMA framework. Even so, the state space technique is well suited to model-

³ See also Aoki and Havenner (1989, 1991).

ing cointegrated series compared to other time series techniques, because the state space representation provides a methodology which identifies the number and form of the common trends of the cointegrated series, while the more commonly used time series methods lack this capacity.

Arrange the centered and scaled series of spot and futures prices into an $(m \times 1)$ vector y_t . Let the $(n \times 1)$ vector z_t denote the unobservable states while the m -element vector e_t indicates serially uncorrelated error terms. Now Aoki's state space model can be written in two matrix equations called the state and observation equations respectively:

$$z_{t+1|t} = A z_{t|t-1} + B e_t \quad (1)$$

$$y_t = C z_{t|t-1} + e_t. \quad (2)$$

In the above equation pair $z_{t+1|t}$ is a vector of conditional means of unobservable states and A , B , and C are matrices of coefficients to be estimated. This state space representation and the more commonly specified ARMA representation are well known to be equivalent, see, e.g., Havenner and Aoki (1991). In fact, any state space representation has an equivalent ARMA representation and vice versa. Stationarity is not imposed on the state equation. Consequently, unit (or even larger magnitude) roots may arise in A so that the states are nonstationary. Cointegration suggests that the m series be characterized with fewer than m nonstationary states; in this event anything in the null space of C is a cointegrating matrix since it annihilates the nonstationary components of $z_{t|t-1}$. The data-driven determination of the nonstationarity, and the model specification used below, guarantees that cointegration will be found naturally, if it is there. The state space technique also estimates roots in A off the point null, dropping the sharp distinction between unit roots and large roots. While the point null hypothesis of real unit roots is of interest in various settings, the primary focus of this endeavor is to enhance modeling capability and esti-

mate the cointegrating vectors. Thus, the real interest lies in characterizing the spectrum of slow and fast dynamics in the system. Aoki and Havenner (1991) detail this aspect of the modeling procedure in section four of their paper.

In order to separate the slow and fast dynamic components, decompose y_t into long run (trend) and short run (cycle) dynamics. The first can be represented by a set of states $\tau_{t|t-1}$ and is associated with a set of large eigenvalues denoted $\lambda_1(A)$. The second, denoted y_t^* , is a stationary transform of y_t which retains the fast dynamics (small eigenvalues) and takes the place of the error term in the state and observation equation pair. The state space model for the described decomposition is

$$\tau_{t+1|t} = A_\tau \tau_{t|t-1} + B_\tau y_t^* \quad (3)$$

$$y_t = C_\tau \tau_{t|t-1} + y_t^* \quad (4)$$

The above model will be referred to as the trend model because it specifies the slow dynamics, which would include any common trends. The residual from the trend model is used as the input variable into another state space representation designed to model the high frequency cycles. Thus, it will be referred to throughout this paper as the cycle model. The cycle model is

$$\eta_{t+1|t} = A_\eta \eta_{t|t-1} + B_\eta e_t \quad (5)$$

$$y_t^* = C_\eta \eta_{t|t-1} + e_t \quad (6)$$

The state space procedure arranges the autocovariances (or autocorrelations) in the form of a Hankel matrix (block band counterdiagonal) and uses the singular value decomposition to determine the rank of this matrix. The rank of the Hankel matrix thus determined can be shown to be equal to the number of states (minimum sufficient statistics) required to characterize the series. The number of autocovariances included in the Hankel matrix is a function of the lag

parameter N , which balances sampling error at longer lags against model generality. Thus, in state space time series analysis, model specification consists of selecting the series to be modeled, setting the lag parameter N , and determining the number of states required to characterize the dynamics.

The procedure for fitting⁴ the above trend/cycle model is a two step process. First, fit the trend model to the slow system dynamics, obtaining the residuals \hat{y}_t^{*5} . The second step is to again use the methods of model selection discussed above to fit the cycle model to the high frequency dynamics remaining in \hat{y}_t^{*} .

To facilitate further analysis and forecasting, the trend and cycle models can be stacked to form a combined state space model of the full spectrum of dynamics. To form the combined model substitute equation (5) into (3) and (4) rearrange and stack them with (5) and (6). The result is

$$\begin{pmatrix} \tau_{t+1|t} \\ \eta_{t+1|t} \end{pmatrix} = \begin{pmatrix} A_\tau & B_\tau C_\eta \\ 0 & A_\eta \end{pmatrix} \begin{pmatrix} \tau_{t|t-1} \\ \eta_{t|t-1} \end{pmatrix} + \begin{pmatrix} B_\tau \\ B_\eta \end{pmatrix} e_t \quad (7)$$

$$y_t = (C_\tau \ C_\eta) \begin{pmatrix} \tau_{t|t-1} \\ \eta_{t|t-1} \end{pmatrix} + e_t \quad (8)$$

which can be written [in the form originally introduced in equations (1) and (2)] as

$$z_{t+1|t} = A^* z_{t|t-1} + B^* e_t \quad (9)$$

$$y_t = C^* z_{t|t-1} + e_t \quad (10)$$

⁴ Havenner and Aoki show that the estimators, based on the singular value decomposition, can be interpreted as instrumental variables estimators.

⁵ It is easier to see the detrending inherent in the trend model by rearranging to solve for the error. This yields: $y_t^* = y_t - C_\tau \tau_{t|t-1}$ which has the long run trends removed.

3 State-Space Arbitrage Relationships

Cerchi and Havenner demonstrated that with cointegrated series it is possible to exploit the structure of C to construct an arbitrage portfolio of stocks. A similar approach will be used in this analysis with the addition that the arbitrage relationships can be used to construct price indices on which to base efficient settlement of delivered futures contracts or alternatively for cash settlement.

This application requires the construction of a new asset which is a linear combination of the existing assets. The weights of the linear combination are determined by the coefficient matrix C^* which has dimension $(m \times n)$. This matrix is partitioned horizontally such that the upper submatrix, call it C_1 , is square with dimensions $(n \times n)$, and the lower submatrix, subsequently denoted as C_2 , is $(m - n \times n)$, where m is the number of series and n is the number of states chosen in the specification of equations (1) and (2). Partition y_t conformably, into y_{1t} and y_{2t} as follows:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} z_{t|t-1} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$$

which implies

$$z_{t|t-1} = C_1^{-1} y_{1t} - C_1^{-1} e_{1t} \quad (11)$$

so that

$$\begin{aligned} y_{2t} &= C_2 C_1^{-1} y_{1t} + \epsilon_t \\ &\equiv D y_{1t} + \epsilon_t. \end{aligned} \quad (12)$$

Whenever the number of states (n) is less than the number of series (m) an arbitrage relationship can be formed between n of the markets and the remaining $n-m$ markets⁶. This relationship is analogous to those utilized in the Arbitrage Pricing Theory (APT) in that it represents a multi-factor return-generating process where the factors (or indices) are the states. Analogous to the factors in traditional APT, the states in the above model also affect returns on more than one asset and are the sources of covariance between the asset prices in the model. The traditional APT factors are static, however, while these are dynamic. It is not uncommon in state space models for the number of series to exceed the number of states so that these dynamic arbitrage relations exist and have meaning in a variety of contexts; see Cerchi and Havenner for details and another application, and see Cerchi for relation to alternative financial models.

The coefficients in D represent unique information about pairs of markets. This information is attributable to relative buyer concentrations, volumes of trade, transportation costs, *etc.* The elements of D represent the weights used to construct the new portfolio of spot and/or futures markets holdings. If we partition y_t such that the futures price is in y_{1t} , then each row of D contains the vector of weights associated with an arbitrage portfolio of $m-n$ spot markets and the futures against each of the n remaining spot markets. This provides the starting point for constructing the settlement price indices which incorporate the implied arbitrage opportunities and result in the optimal premiums and discounts.

Because the arbitrage relations determine only relative portfolio weights, D must be normalized so that the elements in each row of D sum to unity. This is accomplished by dividing each weight by the sum of the weights in its row, i.e., creating

⁶ This relationship does not depend on the cointegration of y_t (i.e., the trend-cycle model). However, cointegration preserves the unforecastable trend(s).

$$d_{ik}^* = \frac{d_{ik}}{\sum_{j=1}^m d_{ij}} \quad (13)$$

where d_{ik}^* is the ik^{th} element of a new matrix D^* . The scaled weights are the inputs for the proposed price indices.

In practice, the relationship between y_{1t} and y_{2t} implied by (12) and D^* is

$$\hat{y}_{2t} = \hat{D}^* y_{1t}. \quad (14)$$

Since the regional markets are sometimes thin and can be dominated by particular individual transactions, it is advantageous to average over multiple geographic prices when setting a settlement price for a specific market. This is possible whenever the number of states is less than $m-1$, since in that case there will be multiple partitions of y_t , each assigning some markets to y_{1t} to be used in determining settlement indices for the markets in y_{2t} . Specifically, there will be $(m-2)!/[(n-1)!n!]$ different partitions which each yield an index for a given market price holding the futures in y_{1t} at all times,⁷ allowing use of the average of the calculated indices to construct the discounts and premiums. By doing so, we are averaging over the different events in various markets, including such things as large transactions and different regional qualities (in contrast to the single hypothetical quality of the futures).⁸ Averaging over several regional markets, as well as the futures, also virtually eliminates the potential for local events, which may be

⁷ Averaging is possible unless the number of states is exactly one less than the number of series. Then y_{2t} becomes a one element vector and there is only a single way the partition y_t .

⁸ Ross follows a similar tack in developing the APT model. That is, in his step 2, he invokes the law of large numbers to suggest that averaging over a large number of assets all with white noise random errors leads to the elimination of the stochastic component. Here we are averaging over the prices of the different constructed portfolios for each geographical market.

potentially manipulable, to create unfair settlements.

The main advantage of the proposed settlement price index is that it reflects the multimar-
ket arbitrage opportunities which exist between a spot market, other spot markets, and the
futures, and therefore saves the economic system the costs of undesirable delivery due to
legitimate basis differences not being taken into account in setting premiums, discounts, and
pars. The discounts and premiums implied by the arbitrage indices are the difference between
the index for the spot market and the futures. This time-varying structure more directly reflects
the character of the markets and the particular industry than the previous fixed relations to the
futures. While the state space model is statistically somewhat sophisticated, given the parameter
estimates it is a simple matter for all interested parties to validate the settlement index.

4 State Space Model of Live Cattle Prices

In order to demonstrate the implementation of the above model, series for weekly Choice
yield grade 1 and 2 (900 - 1100 #) steer prices for seven markets and the nearby Chicago Mer-
cantile Exchange (CME) futures price for live cattle were used. The cash prices are Friday clos-
ings and the futures is the Friday settlement price for the next contract due to expire. Thus, as a
contract expires the next chronological contract price enters the data. The sample is from
January 1973 through May 1987 for a total of 752 weeks and was compiled by the Chicago
Board of Trade. The model was estimated using the first ten years of data (520 weeks) with the
remaining 232 weeks kept uncompromised for model validation. The seven spot markets chosen
were the seven designated delivery points as of 1987 when the sample terminates.⁹ In practice it
would not be necessary for these markets to be used; they simply represent a convenient set of

⁹ The seven delivery points were: Amarillo, Dodge City, Greeley, Joliet, Omaha, Peoria, and
Sioux City. In 1987, these all traded at par, but this was not the case for earlier years when some
discounts were in effect.

markets. In theory, any set of markets should be cointegrated. The eight series were subjected to Dickey-Fuller tests with the result that the random walk hypothesis could not be rejected at any reasonable significance level. The calculated Dickey-Fuller statistics were: -1.50, -1.44, -1.48, -1.48, -1.44, -1.59, -1.54, and -1.70. Test statistics smaller than -3.12 would be necessary in order to reject the null hypotheses of a random walk at the .05 level of significance. Thus, individually the series would not be easily forecasted and conventional wisdom of several years ago would have dictated first differencing each series.

With eight series, even setting the lag parameter to one allows complex model dynamics since the implied univariate models for each series potentially include lags up to eight. The lag parameter was set to one for the trend model since this implies the lowest possible order determinantal polynomial for each series and to model the slow dynamics requires only immediate lags. To determine the rank of the Hankel matrix, and thus the number of states required in the model, an examination is made of the ratios of successive singular values to the first singular value (a scale-free measure analogous to a condition number). Including the first, the ratios for the trend model were: 1.0, 0.0014, 0.0013, 0.0002, 0.0001, 0.0001, 0.0000, and 0.0000. Obviously there is a large break between the first and second singular values. As a result a trend model with a single state was chosen, i.e., the data support one set of slow dynamics common to all eight series.

The lag parameter was set to two for the cycle model specification. This means that the implied univariate model for each detrended price series can include lags up to sixteen periods. This allows the identification of relatively faster dynamics remaining in the detrended series. The ratios of the singular values to the first singular value for the cycle model were: 1.0, .7357, .5417, .1449, 0.1053, 0.0959, 0.0502, and 0.0480. One large change occurs between the third

and fourth singular values. This would suggest a cycle model with three states. Models with more states were investigated using only the in-sample data but were found to be inferior. Two of the cycle states gave rise to complex roots in A .

The trend and cycle models were joined to form the model depicted in equations (7) and (8). This combined model has four states, one from the trend and three from the cycle model. The eigenvalues of A^* for the combined model show two complex roots inherited from the cycle model. Estimates of the coefficient matrices A^* , B^* , and C^* for the combined model are presented in Table 1. The matrix A^* determines the dynamic response of the model. An examination of the moduli of the eigenvalues of A^* (.9893, .7682, .7682, and .3834) reveals a spectrum of dynamics ranging from very slow to relatively fast. It is not necessary for us to categorize the dynamics into stationary and nonstationary classes to produce accurate forecasts with the state space method, so we will not attempt to test formally for unit roots. Informally, it appears that one nonstationarity exists associated with the eigenvalue 0.9893 which clearly tracks very slow dynamics. This is followed by some intermediate dynamics associated with the two complex eigenvalues and finally a high frequency component reflected in the smallest eigenvalue. This full range of responses is unusual in our experience; one practical consequence is that it would be difficult indeed to successfully model these series by individually differencing them first.

The discovery of only one potential nonstationarity for the eight series leads to the conclusion that the prices are in fact cointegrated. However, this comes as no surprise. To reiterate earlier assertions, the price series of interest represent closely related commodities, and the arbitrage opportunities between locations will tend to force the prices together. Therefore, the cattle price series cannot infinitely diverge and must be cointegrated. Dickey-Fuller tests were conducted on the trend model residuals to confirm that the nonstationary components of price

series were removed. These tests are analogous to cointegration test presented by Granger (1986, p. 219). The eight test statistics computed from this set of regressions were: -18.8, -16.4, -18.7, -17.6, -20.5, -17.8, -17.3, -13.9. In all cases, there is an overwhelming rejection of the random walk hypothesis suggesting that the price series are indeed cointegrated¹⁰.

The in-sample summary statistics for this model are presented in Table 2. Backcasting was performed to obtain reasonable estimates of the initial states. Notice that the correlations between actual and predicted values are very high and the mean absolute deviations and root mean squared errors are small. This suggests an excellent fit of the in-sample data. Figures 1 through 8 illustrate the out-of-sample forecasts. These graphs show the 232 post-sample observations, single period forecasts, and the residuals for each of the eight markets modeled. Notice that the residuals do not move systematically around zero, but rather they appear to be fairly random.

The model forecasts conducted through the 232 weeks of post-sample data were evaluated on root mean squared error and also subjected to Henriksson and Merton nonparametric tests (discussed below). The summary statistics for post sample performance are given in the right-most column of table 2. We observe no degradation of the in-sample results: the out-of-sample correlations and root mean squared errors remain good.

The Henriksson and Merton (HM) test provides an alternative to mean squared error forecast evaluation. The HM test ignores information about the sizes of the errors and simply tests whether or not the forecasts predict the correct direction of change. By eliminating magnitude information and using a nonparametric test it is possible to obtain a test statistic with a known

¹⁰ The critical value of these cointegration tests is also -3.12.

finite sample distribution under weak and plausible assumptions. The null hypothesis of the test is

$$H_0: Prob(y_{it} - y_{it-1} > 0 | \hat{y}_{it} - y_{it-1} > 0) + Prob(y_{it} - y_{it-1} \leq 0 | \hat{y}_{it} - y_{it-1} \leq 0) = 1 \quad (15)$$

for each of the $i = 1, 2, \dots, 8$ series in turn. To carry out the test, define

N_1 = the number of observations where prices actually rose,

N_2 = the number of observations where prices actually did not rise,

$N = N_1 + N_2$ = the total number of observations,

n_1 = the number of successful predictions given that prices rose,

n_2 = the number of unsuccessful predictions given that prices did not rise,

$n = n_1 + n_2$ = the number of times price rises were forecast.

Then the finite sample confidence level, c , is given by

$$c = 1 - \sum_{x=n_1}^{\min(N_1, n)} \binom{N_1}{x} \frac{\binom{N_2}{n-x}}{\binom{N}{n}}, \quad (16)$$

see Henriksson and Merton.

The values for the HM test were: $N_1 = 903$, $N = 1856$, $n_1 = 419$, and $n = 684$. These results lead to a confidence level of 1.0 that the forecasts are informative. We feel that this is an extraordinary result given the random walk natures of the series being modeled.

5 Arbitrage Price Indices for Live Cattle

The matrices D_i^* were calculated for all of the possible partitions of y where the CME futures is kept in y_1 . Notice that some weights may be negative. Negative weights suggest short positions in the corresponding market in the arbitrage portfolio. The arbitrage price index associated with an element of y_2 is constructed using the corresponding row of D^* . The price indices

were calculated for the 232 weeks of post sample data for all 35 possible partitions. Statistics for the difference between the current futures and the cash market indices are indicative of the relative strengths of the markets as compared to the futures which is the actual commodity being traded. In all cases, the two values tend to converge whenever the difference between them becomes relatively great. This demonstrates the cointegration and arbitrage forces at work in the marketplace. It also represents the market forces which tend to reduce basis variability when the market is performing efficiently. Thus, the constructed arbitrage price index might mitigate some of the adverse affects associated with futures contract delivery if implemented to replace the current system of arbitrary discount/premium/par schedules.

Graphs of the indices and discount/premiums are not included, but are available from the authors upon request. The estimated differentials between arbitrage prices and the futures price demonstrate the model's capacity to capture the characteristics of the markets. For example, without exception Joliet is estimated to trade at a discount to the futures. The mean absolute value of this discount is \$7.98 per cwt. This suggests a market with serious problems, and as it turns out, this market recently closed due to low volume. The low volume would suggest that disequilibrium prices might prevail, and subsequently, delivery as a percentage of total volume in Joliet should be high relative to other markets. In order to test this hypothesis we spoke with Edna Kucinski of the CME Delivery and Inspections Division. As an example, she quoted the following delivery statistics for the month of October 1987: Joliet 54, Peoria 55, Dodge City 46, and Sioux City 103 contracts. At first, this would appear to violate the hypothesis tendered above. However, one must consider the deliveries relative to the total volume of the markets. According to the USDA Market News Livestock, Meat, and Wool Reports for the month of October 1987 the volumes of cattle sold in these markets were: 2271, 9520, 30333, and 14991 head respectively. Thus, while the total number of contracts delivered was essentially the same

for Joliet, Peoria, and Dodge City, the delivery as a percentage of total volume was four times higher in Joliet than in Peoria, thirteen times higher than in Dodge City, and seven times higher than in Sioux City. This demonstrates that the par setting on deliveries at Joliet resulted in relatively more deliveries. Had the exchange set a reasonable discount then deliveries would have been discouraged and the price suppressing impacts of delivery in Joliet could have been avoided as well as the transactions costs associated with delivery.

The estimated discounts and premiums suggest that delivery at four of the cash markets during the post-sample period should have, on the average, been at a discount and three at par.¹¹ The average discounts for Sioux City, Amarillo, Joliet, and Peoria were 2.37, 2.06, 7.32, and 2.13, respectively. The average differential for the three markets deemed to settle around par were: -.21 for Dodge City, 1.06 for Omaha, and -1.66 for Greeley.

It should also be noted that 20 different indices of the nearby futures price paired with a given spot market price could be constructed from equation (12) by placing the CME futures price in y_{2t} , the spot price in y_{1t} , and generating all of the potential partitions of y_t . The average of these indices could potentially be useful should cash settlement be chosen to replace physical delivery as the mechanism for arbitraging excessive basis deviations.

6 Conclusion

In this paper we have demonstrated a methodology for developing arbitrage price indices for futures settlement or cash settlement which are more efficient than simple indices constructed in an *ad hoc* fashion. The construction of these indices is based on the cointegration and spatial

¹¹ The costs of delivery (grading discounts, shrinkage, and additional transportation costs, etc.) are estimated to be approximately \$2.00 per cwt. Thus, if the average discount/premium was between \$±2, then that market was deemed to be on par on the average. This does not mean that the market should be arbitrarily traded that way; it is just a means of distinguishing the relative price differentials.

market arbitrage opportunities inherent in agricultural commodity markets. In general, ignoring these characteristics of the market in cash settlement or delivery would limit the futures contract in its ability to restore spatial market equilibrium and lead to undesirable and costly delivery.

An application to the live cattle market provided the empirical justification and demonstrated the construction of such indices. Aoki's state space model was used to develop the integrated multivariate model of spot and futures prices. Since the validity of the settlement indices depends critically on the accuracy of the underlying model, the model was carefully tested using 232 weeks of data not used in any way at all in model construction. Post sample mean squared errors associated with one week forecasts were small and the results of Henriksson and Merton nonparametric tests suggest that the forecasts are informative, validating the model as a basis for the proposed settlement indices.

The indices constructed have several advantages over less complicated alternatives. The new indices are time-varying, and the discount and premium structure is based on spatial market arbitrage relations. Thus, they should tend to improve the efficient operation of live cattle markets. The incentives for delivery in any given location will be negated because the index for that market will quickly adjust as a result of possible disequilibria. By averaging over arbitrage relations linking to a large number of markets the new indices reduce the potential for manipulation of the indices while maintaining the integrity of the marketplace. Physical delivery, with its large associated transactions costs, is an indication of failure of the market to incorporate all relevant events in the contract clearing mechanism. The proposed settlement indices are designed to reduce the incidence of delivery and thus the market overhead costs by incorporating temporal and regional effects in an optimal manner.

References

- Aoki, M. and A. Havenner, 1989, A Method for Approximate Representation of Vector-Valued Time Series and Its Relation to Two Alternatives, *Journal of Econometrics* 50, 89-95.
- Aoki, M. and A. Havenner, 1991, State Space Modeling of Multiple Time Series, *Econometric Reviews* 10, 1-59.
- Aoki, M., 1987, *State Space Modeling of Time Series* (Springer-Verlag, Berlin).
- Cerchi, M., 1989, State Space Methods in Asset Pricing, *Computers and Mathematics with Applications* 18 (no. 6/7), 581-590.
- Cerchi, M. and A. Havenner, 1988, Cointegration and Stock Prices: The Random Walk on Wall Street Revisited, *Journal of Economics Dynamics and Control* 12, 333-346.
- Engle, R. G. and C. W. J. Granger, 1987, Co-Integration and Error Correction: Representation, Estimation, and Testing, *Econometrica* 55, 251-276.
- Goodwin, B. and T. Schroeder, 1991, Cointegration Tests and Spatial Price Linkages in Regional Cattle Markets, *American Journal of Agricultural Economics* 73, 452-464.
- Granger, C. W. J., 1983, Co-integrated Variables and Error-Correcting Models, Economics Department Working Paper 83-13, University of California, San Diego.
- Granger, C. W. J., 1986, Developments in the Study of Cointegrated Economic Variables, *Oxford Bulletin of Economics and Statistics* 48, 213-28.
- Havenner, A. and M. Aoki, 1988, An Instrumental Variable Interpretation of Linear Systems Theory, *Journal of Economic Dynamics and Control* 12, 49-54.
- Henriksson, R. D. and R. C. Merton, 1981, On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills, *Journal of Business* 54, 513-553.
- Ross, S.A., 1976, The Arbitrage Theory of Capital Asset Pricing, *Journal of Economic Theory* 13, 341-360.

7 Tables and Figures

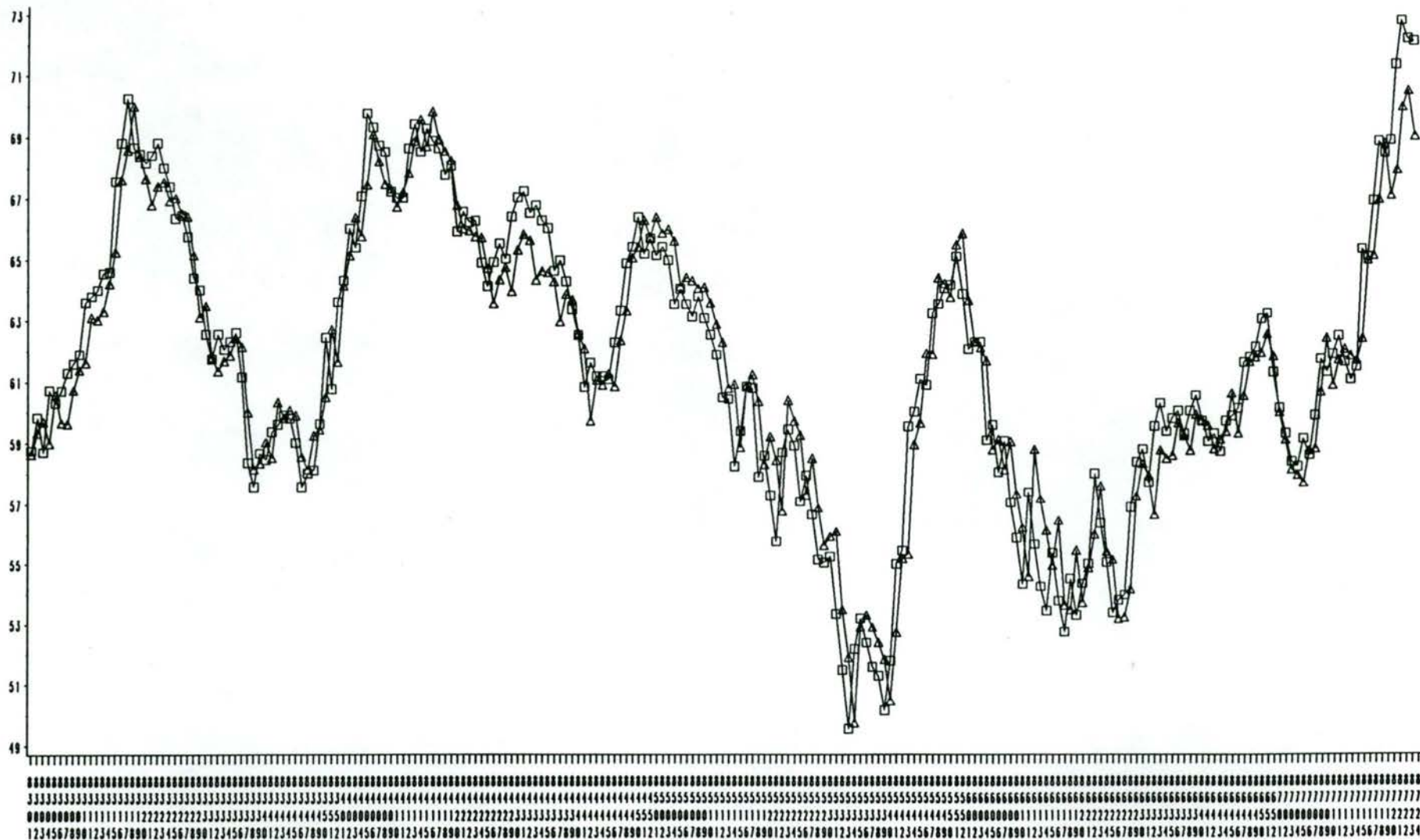
Table 1: <i>Coefficient Estimates</i>			
A^*			
0.9893	- 0.0256	- 0.0219	- 0.0332
0.0	0.6815	0.0274	- 0.1194
0.0	- 0.1607	0.7295	0.0571
0.0	- 0.2987	- 0.0224	0.5030
B^{**}			
0.0047	- 0.1846	- 0.3466	- 0.0717
- 0.0147	0.2092	- 0.1800	- 0.0634
0.0017	- 0.3759	0.2227	- 0.2612
- 0.0121	0.2454	- 0.1471	0.0711
- 0.0485	- 0.0616	0.1107	0.1046
- 0.0101	- 0.2197	0.0708	- 0.0281
0.0086	0.1990	- 0.1427	0.2106
- 0.0303	0.2517	0.4049	0.0719
C^*			
-11.9128	- 0.1583	- 0.3069	0.4894
-12.3892	0.5909	- 0.3792	0.1250
-11.6381	- 0.6512	0.2908	0.3607
-12.2591	0.5883	- 0.2648	0.2912
-11.8430	- 0.1423	- 0.1815	0.4296
-11.6531	- 0.4809	0.2767	0.6612
-12.0363	0.6091	- 0.4088	0.4275
-11.8364	0.8246	1.0634	0.2254
Order: Sioux City, Amarillo, Joliet, Dodge City, Omaha, Peoria, Greeley, and Futures. The state initial conditions were .9834, .6435, 1.1372, and .3456.			

Table 2: <i>Summary Statistics in Levels</i>		
<i>Statistic</i>	<i>In Sample</i>	<i>Out of Sample</i>
<i>Sioux City Steers (\$/cwt)</i>		
Mean:	52.57	61.61
Root mean squared error:	1.19	1.35
Average error:	0.027	0.329
Mean absolute deviation:	0.904	1.073
Simple correlation:	.990	.924
<i>Amarillo Steers (\$/cwt)</i>		
Mean:	53.36	63.12
Root mean squared error:	1.23	1.15
Average error:	0.030	0.036
Mean absolute deviation:	0.938	0.902
Simple correlation:	.990	.939
<i>Joliet Steers Price (\$/cwt)</i>		
Mean:	52.46	61.28
Root mean squared error:	1.34	1.46
Average error:	0.025	0.672
Mean absolute deviation:	0.858	1.144
Simple correlation:	.991	.917
<i>Dodge City Steers (\$/cwt)</i>		
Mean:	52.93	63.03
Root mean squared error:	1.25	1.44
Average error:	0.028	0.665
Mean absolute deviation:	0.938	1.135
Simple correlation:	.990	.922

Table 2 (continued): <i>Summary Statistics in Levels</i>		
<i>Statistic</i>	<i>In Sample</i>	<i>Out of Sample</i>
<i>Omaha Steers Price (\$/cwt)</i>		
Mean:	52.54	61.21
Root mean squared error:	1.19	1.04
Average error:	0.027	0.016
Mean absolute deviation:	0.897	0.814
Simple correlation:	.990	.947
<i>Peoria Steers Price (\$/cwt)</i>		
Mean:	52.34	60.25
Root mean squared error:	1.22	1.14
Average error:	0.024	-0.230
Mean absolute deviation:	0.923	0.922
Simple correlation:	.989	.946
<i>Greeley Steers Price (\$/cwt)</i>		
Mean:	53.05	62.68
Root mean squared error:	1.24	1.36
Average error:	0.028	0.368
Mean absolute deviation:	0.948	1.063
Simple correlation:	.990	.920
<i>Live Cattle Futures (\$/cwt)</i>		
Mean:	53.28	62.70
Root mean squared error:	1.56	1.62
Average error:	0.022	0.477
Mean absolute deviation:	1.216	1.226
Simple correlation:	.983	.879

Figure 1: Actual and Post-Sample Forecasts of Sioux City Choice Steer Prices
Weekly Data, Jan 1983 - May 1987

Dollars/cwt

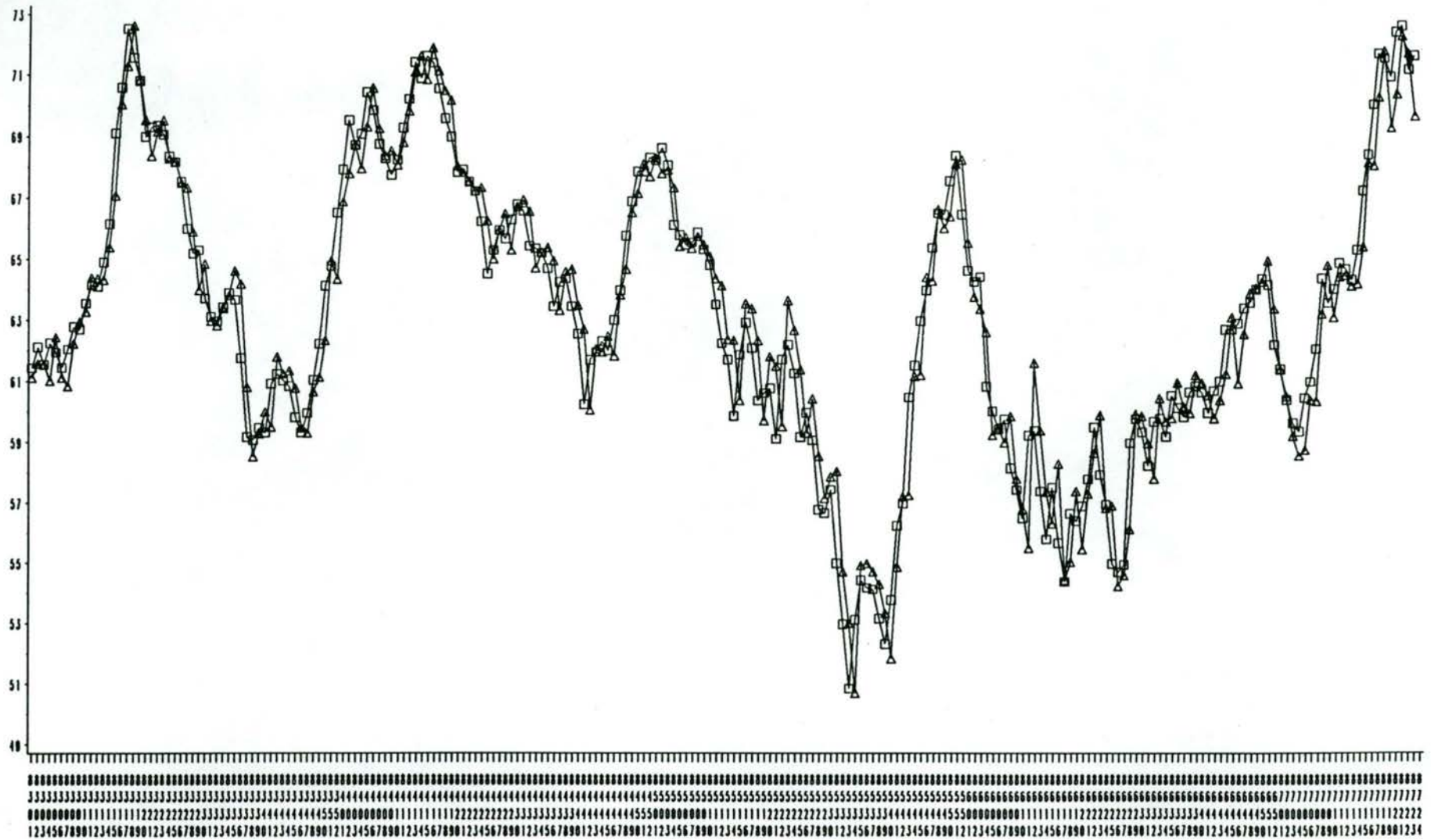


Year / Week

LEGEND ■■■ ACTUAL ▲▲▲ FORECAST

Figure 2: Actual and Post-Sample Forecasts of Amarillo Choice Steer Prices
Weekly Data, Jan 1983 - May 1987

Dollars/cwt



Year / Week

LEGEND ■■■ ACTUAL ▲▲▲ FORECAST

Figure 3: Actual and Post-Sample Forecasts of Joliet Choice Steer Prices
Weekly Data, Jan 1983 - May 1987

Dollars/cwt

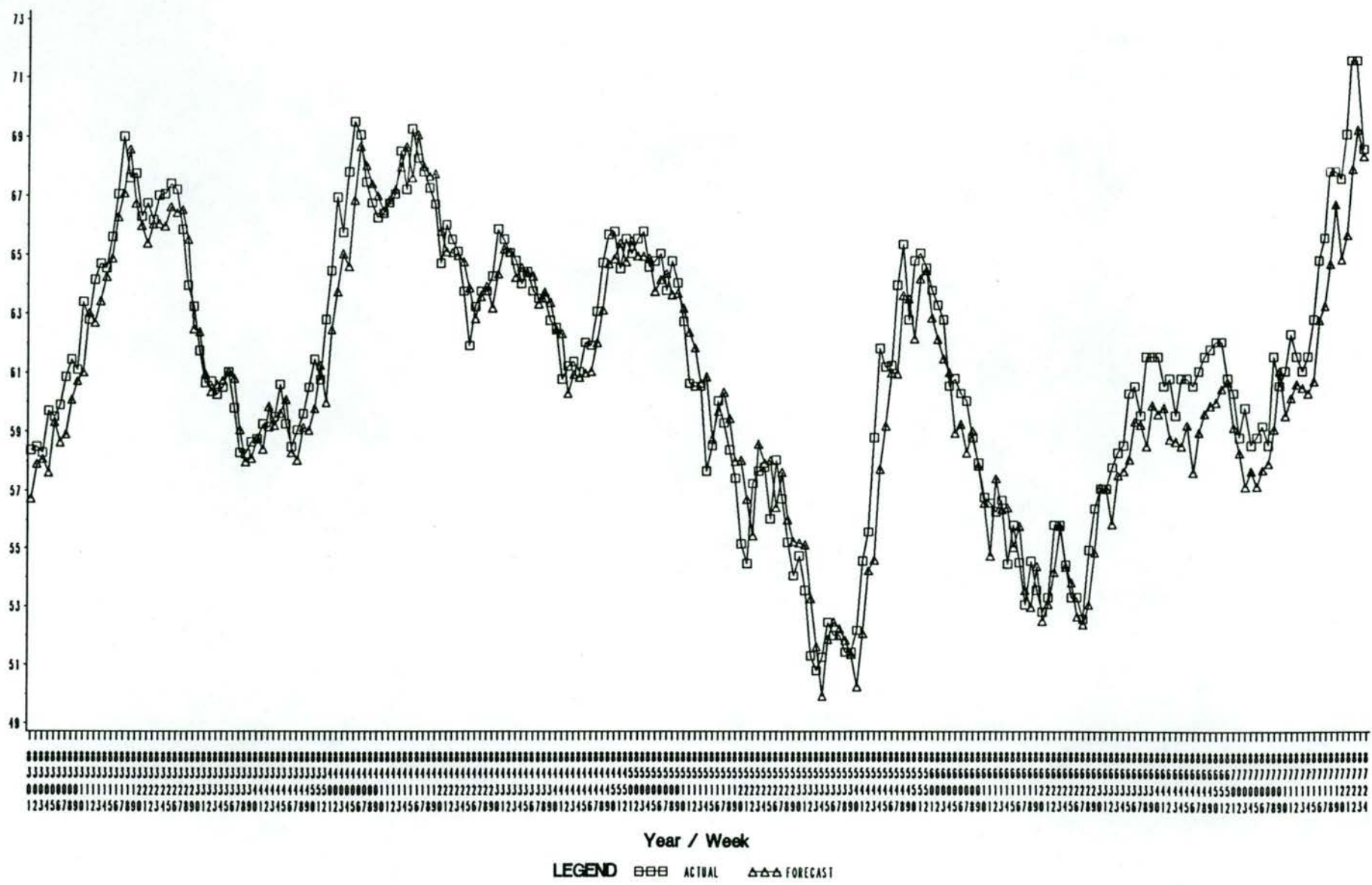
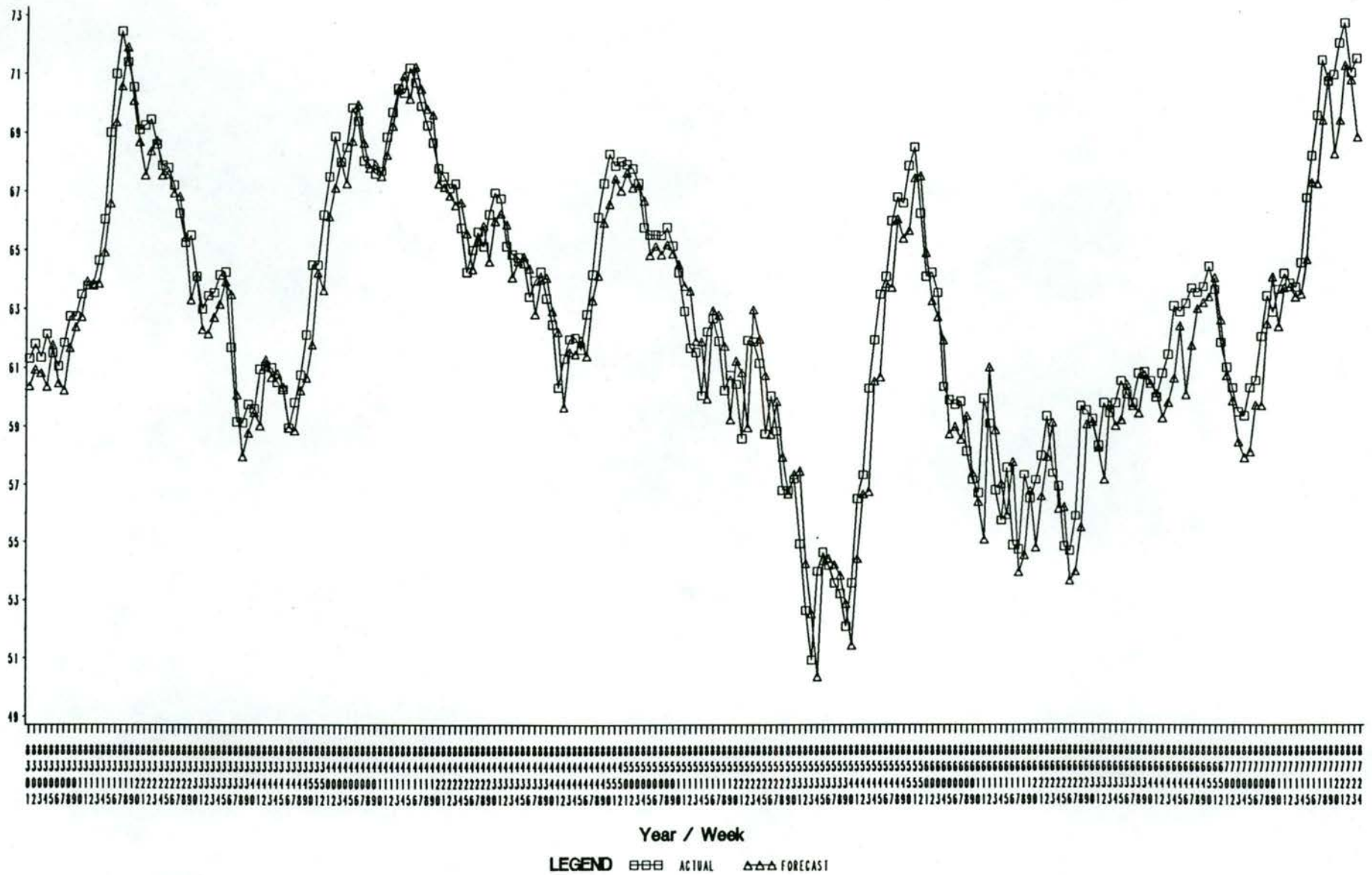


Figure 4: Actual and Post-Sample Forecasts of Dodge City Choice Steer Prices
Weekly Data, Jan 1983 - May 1987

Dollars/cwt



Weekly Data, Jan 1983 - May 1987

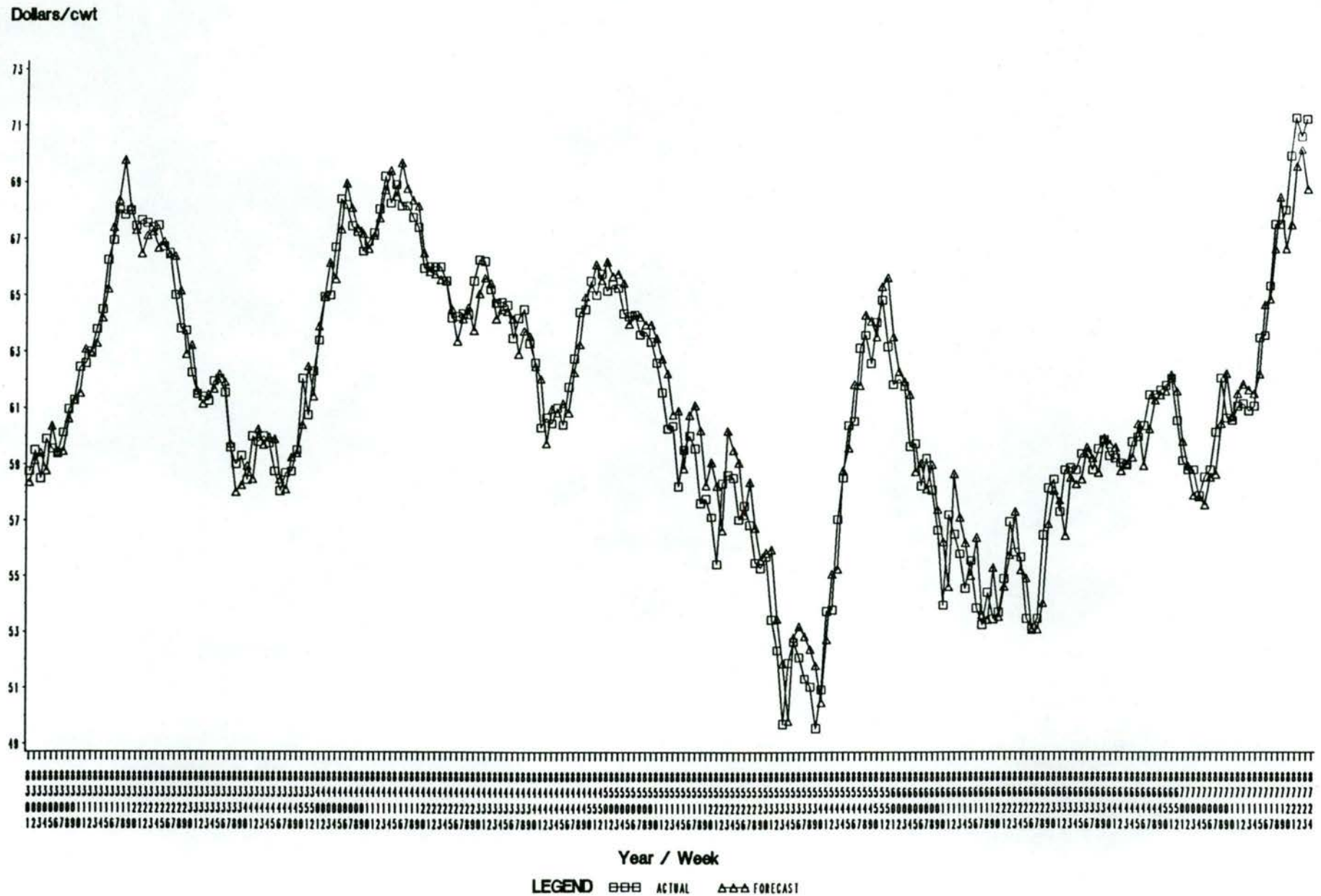


Figure 6: Actual and Post-Sample Forecasts of Peoria Choice Steer Prices
Weekly Data, Jan 1983 - May 1987

Dollars/cwt

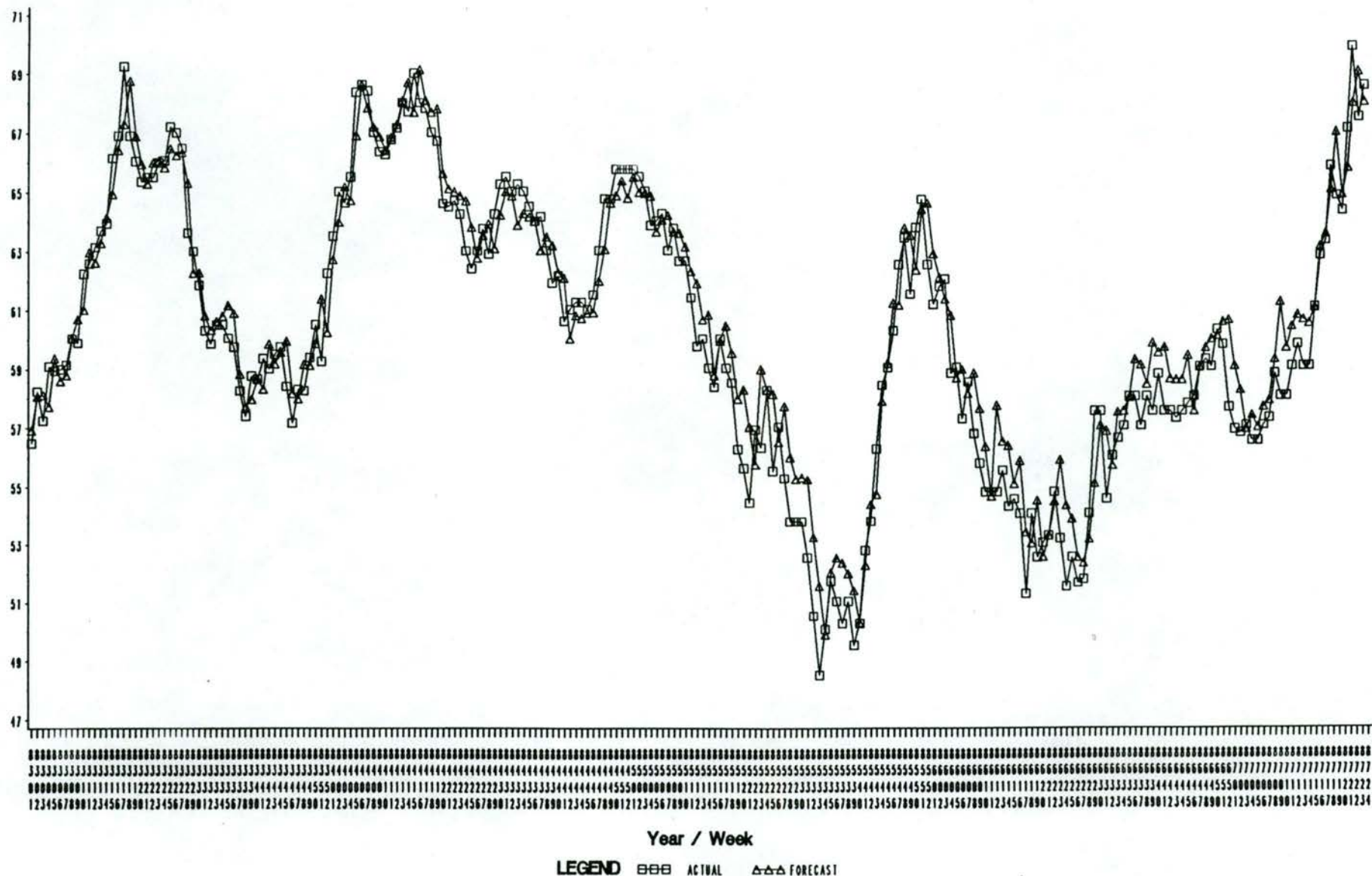
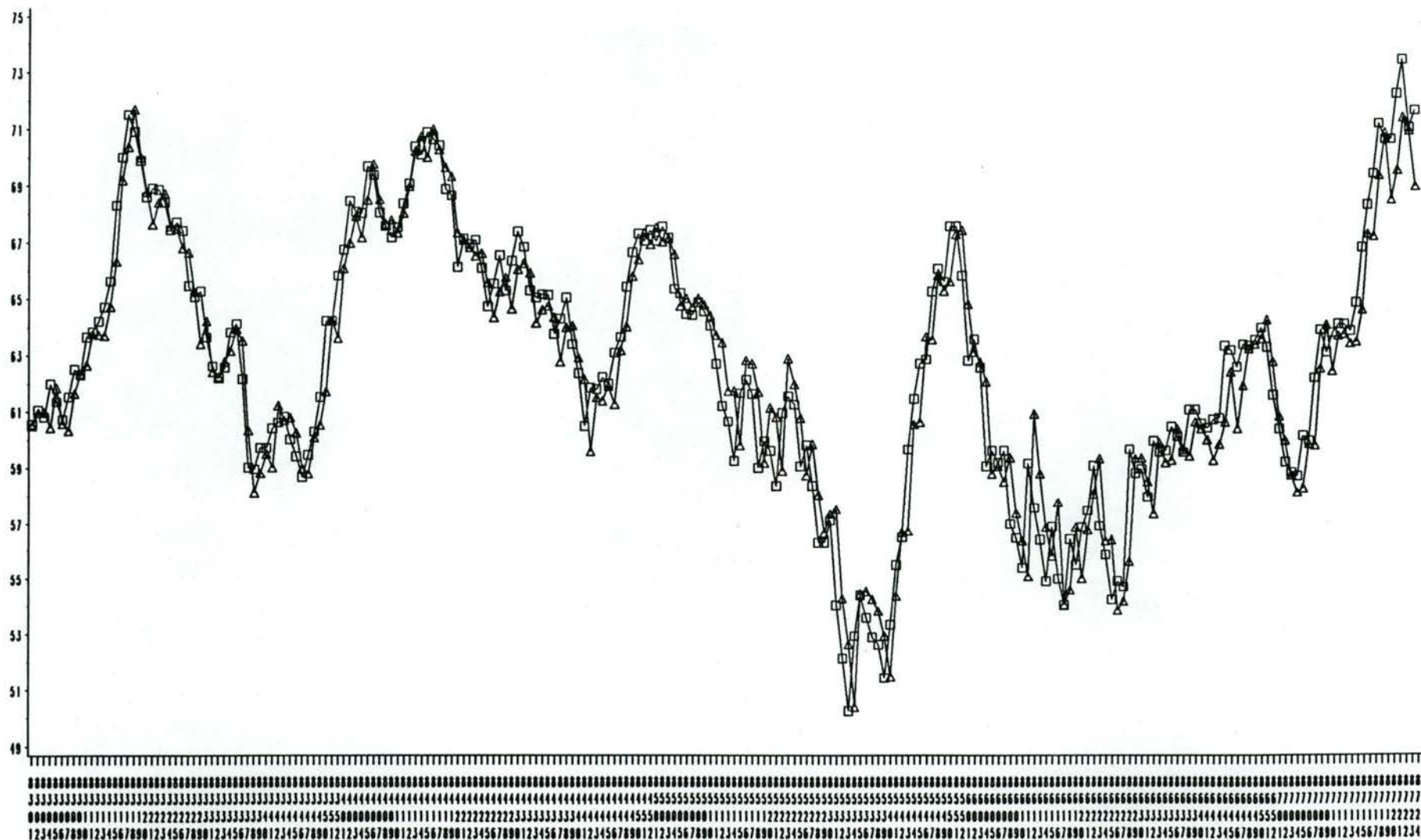


Figure 7: Actual and Post-Sample Forecasts of Greely Choice Steer Prices
Weekly Data, Jan 1983 - May 1987

Dollars/cwt



Year / Week

LEGEND ■■■ ACTUAL ▲▲▲ FORECAST

Weekly Data, Jan 1983 - May 1987

