USE OF FORECASTS IN DECISIONMAKING: THE CASE OF STOCKER CATTLE IN FLORIDA

Thomas H. Spreen and Carlos A. Arnade

Abstract

The decision to overwinter feeder cattle hinges directly on the forecast of spring cattle prices. An analysis of price forecasts from several alternative models is presented. The models are evaluated using both the traditional mean square criterion and alternative criteria. The alternative criteria evaluate the profitability of the decision implemented based upon the forecast.

Key words: forecasting, beef cattle, feeder cattle.

Cattle production is dispersed throughout the Southeast with a typical operation producing weaned calves in the fall. In Florida, weaned calves weigh approximately 400 pounds and generally are not placed in feedlots. The calves are grown out on high roughage diets to approximately 600 pounds. This intermediate phase is called stocker cattle production or "backgrounding" (Florida Department of Agriculture and Consumer Services). In the Southeast, temporary winter pastures of rye, ryegrass, or cereal grains provide roughage supplemented by small amounts of corn. Nearly all winter pastures use cool-season annual grasses which cannot survive summer heat and must be replanted each fall.

An important decision confronts the cattle producer each fall. The producer can immediately sell weaned feeder calves or retain the calves and initiate stocker cattle production. If a cattle producer chooses to produce stocker cattle, a major proportion of the cost is incurred at the beginning of the production process. Estimates indicate that nearly 90 percent of the cost of stocker cattle production is incurred in establishment of pasture and foregone revenues from retention of calves (Arnade). These are sunk costs and can be distinguished from fixed costs by noting sunk costs are not incurred unless production is initiated.

The purpose of this paper is to analyze the decision to produce stocker cattle. An economic decision model is developed. Statistical loss functions are reviewed and alternative loss functions are proposed for evaluation of price forecasts used in making economic decisions. Alternative statistical models are estimated for stocker steer prices. The forecasts from these models are evaluated using both traditional statistical loss functions and decision-related criteria.

A CONCEPTUAL FRAMEWORK

In this section, a conceptual model of the decision to produce stocker cattle is presented. Consider a firm which produces a single output Y from a set of inputs $X' = (X_1, \ldots, X_n)$ according to the production relation:

$$(1) \qquad Y = f(X_1, \ldots, X_n).$$

Suppose the firm operates in competitive markets and thus takes input and output prices as exogenous so that its profit function is:

$$(2) \qquad \pi = PY - RX,$$

where P is the price of Y and $R = (r_1, \ldots, r_n)$ is the vector of prices of X.

Assume that output is deterministic; that is, if a producer commits input levels $X^0 = (X^0_1, \ldots, X^0_n)$, then output Y^0 is realized where:

$$(3) \qquad Y^0 = f(X^0).$$

Next, introduce time into the production process; i.e., suppose that if production is initiated in period t, that the output Y cannot be marketed until period t+1. Furthermore, assume that all inputs are committed in period t such that all costs are "sunk." Then, the expected profit function is:

$$(4) \qquad E_t(\pi_{t+1}) = E_t(P_{t+1})Y_{t+1} - R_tX_t,$$

where $E_t(\pi_{t+1})$ is the expectation of profit to be realized in period t+1 formed in period t and $E_t(P_{t+1})$ is expected product price in period t+1 formed in period t.

Thomas H. Spreen is Associate Professor in the Food and Resource Economics Department at the University of Florida and Carlos A. Arnade is a Graduate Research Assistant in the Department of Agricultural Economics at the University of Maryland.

Florida Agricultural Experiment Station Journal Series No. 5579.

Conventional economic theory suggests that the producer reacts to $E_t(\pi_{t+1})$ in making the production decision. Assuming all resources are divisible, one could hypothesize that:

$$(5) \qquad Y_{t+1} = G[E_t(\pi_{t+1})],$$

where Y_{t+1} is planned output in period $t+1$ and $\frac{dY_{t+1}}{dE_t(\pi_{t+1})} > 0$.

Resources are not all divisible, however, since there are cases in which the producer's decision is discrete, i.e., to produce or not produce. In this case, the producer reacts to $E_t(\pi_{t+1})$ according to:

$$(6a) \qquad Y_{t+1} = 0 \text{ if } E_t(\pi_{t+1}) < 0,$$

$$(6b) \qquad Y_{t+1} = \bar{Y}_{t+1} \text{ if } E_t(\pi_{t+1}) \geq 0,$$

where $\bar{Y}_{t+1}$ is the output level if production is initiated. If production is initiated, actual profit is:

$$(7) \qquad \pi_{t+1} = P_{t+1}\bar{Y}_{t+1} - R_tX_t.$$

The only stochastic variable on the right-hand-side of (7) is P_{t+1}. Let P^*_{t+1} be the value of P_{t+1} which makes π_{t+1} equal to 0. Then P^*_{t+1} represents a "trigger price" because from (6a−b), $E_t(P_{t+1}) \geq P^*_{t+1}$ will cause the producer to initiate production, while $E_t(P_{t+1}) < P^*_{t+1}$ will not. The term "breakeven price" is also used for P^*_{t+1}.

LOSS FUNCTIONS

A typical statistical model for price forecasting relates the variable of interest, in this case P_t, to some set of explanatory variables (Z_t). The functional relationship can be written as:

$$(8) \qquad P_t = g(Z_t, \theta),$$

where θ is a set of parameters which may be estimated using some statistical procedure.

Mean Square Error

Statistical estimation requires the specification of some criterion on which alternative estimates can be evaluated. Only in this way can a "best" or "optimal" estimate be determined. Nearly all techniques used in the estimation of forecasting models use mean square error as the criterion. Let $\hat{\theta}$ be the estimate of θ using the mean square error criterion, then:

$$(9) \quad \hat{\theta} = \min_{\theta} \left[\frac{1}{T} \sum_{t=1}^{T} [P_t - g(Z_t, \theta)]\right]^2.$$

The loss function is the sum of the squared "errors" (averaged over the sample). The error is the difference between observed P_t and predicted P_t, $\hat{P}_t$, where:

$$(10) \qquad \hat{P}_t = g(Z_t, \hat{\theta}).$$

Mean square error is intuitively appealing as a loss function since the general notion is to find that value of θ, $\hat{\theta}$, such that $\hat{P}_t = g(Z_t, \hat{\theta})$ is as close to P_t as possible. A small "error" results in a smaller penalty than a large "error."

Other loss functions are plausible such as mean absolute error or the minimax criterion which is to minimize the maximum error of any single observation over the sample. The production decision problem presented here suggests at least two other possible criteria on which a statistical model used to forecast P_{t+1} could be evaluated.

Decision-Related Loss Functions

In a discrete decision framework (e.g. (6a-b)), an effective forecasting model is one that can correctly predict whether $P_{t+1} \geq P^*_{t+1}$ or $P_{t+1} < P^*_{t+1}$. Accuracy of the forecast in the MSE sense is not important except that a small MSE forecast should more correctly predict the relationship between P_{t+1} and P^*_{t+1}. It is plausible, however, that a model with a higher MSE will more accurately predict whether P_{t+1} is less than or greater than P^*_{t+1}. If P_{t+1} is slightly less than P^*_{t+1}, an inaccurate prediction far below P_{t+1} would be more useful than one close to but above P^*_{t+1}.

An alternative criterion is to define $\hat{D}_t$ to represent realized net returns when the decision implied by the forecast is implemented in period t. Define D^*_t to be the net returns from the correct decision, and $D_t = D^*_t - \hat{D}_t$. D_t is zero if the decision implemented is optimal in the sense that D^*_t is realized and D_t positive indicates that a "wrong" decision was made. Let

$$(11) \qquad D = \sum_{t=1}^{T} D_t.$$

A larger D indicates a less useful model than a smaller D.

An alternative statistic is to let:

$I_t = 1$ if $P_t \geq P^*_t$ and $\hat{P}_t \geq P^*_t$
or if $P_t < P^*_t$ and $\hat{P}_t < P^*_t$, and
$= 0$, otherwise;

that is, $I_t = 1$ if the forecast, $\hat{P}_t$, is "right" in that a correct decision is made and $I_t = 0$ if the forecast is "wrong." Then,

$$(12) \quad I = \frac{1}{T} \sum_{t=1}^{T} I_t$$

measures the proportion of correct forecasts. An I close to one indicates a useful model.[1,2]

Use of Logit Models

Given a time series of observed prices P_t and breakeven prices P^*_{t+1}, one could calculate a time series of the profitability of a particular enterprise. In the framework of equation (6a−b), an enterprise initiated in period t is profitable if P_{t+1} exceeds P^*_{t+1}. Define

$$R_{t+1} = 1 \text{ if } P_{t+1} \geq P^*_{t+1} \text{ and}$$

$$= 0, \text{ otherwise.}$$

Instead of focusing effort on forecasting the quantitative time series P_t, one could forecast the qualitative time series R_t. Thus, the model

$$(13) \qquad R_t = h(Z_t, \Gamma)$$

is the forecasting equation, where Γ is a vector of parameters to be estimated. Since R_t is qualitative, an ordinary least squares procedure is not appropriate. The method of logit analysis is appropriate in this case. Using a maximum likelihood estimator, a value of Γ is determined, call it Γ_{MLE}, such that $h(Z_t, \Gamma_{MLE})$ is most likely to have generated the pattern of zeros and ones observed in R_t. This is equivalent to maximization of the statistic I (eq. 12). See Judge, et al. (pp. 521-525) for the details of logit estimation.

EMPIRICAL ANALYSIS

An empirical illustration of the implications of alternative loss functions in the estimation of forecasting models is presented. The decision to sell weaned cattle or retain the calves and initiate stocker cattle production in North Florida is analyzed. Most producers in this area wean calves in the fall. If stocker production is undertaken, cool season pastures of rye, ryegrass, or some other small grain are cultivated. The animals are generally kept until late spring. In this analysis, assume stocker production is initiated in the fourth quarter with calves weighing 400 pounds and ends in the second quarter. Using cost studies by Gunter and Westberry, the feeding cost of backgrounding feeder calves on rye and ryegrass pasture supplemented with small amounts of corn (or an equivalent amount of hay in those years in which it was cheaper) is estimated over the 1960 to 1981 period. Fourth quarter Kansas City 400-500 pound medium frame No. 1 (MF1) steer prices (USDA, *Agricultural Prices*) are used to represent the cost of the 400-pound feeder calf. Second quarter 600-700 pound Kansas City MF1 steer prices (USDA, *Agricultural Prices*) are used for the output price.[3]

Estimation of Forecasting Models

The empirical analysis includes estimation and evaluation of five different forecasting models.[4] Several models are analyzed to examine the relationship between the mean square error and the I and D statistics associated with various forecasting models. If models with a small mean square error also have large I statistics (I close to 1) and small D statistics (D close to zero), then mean square error is a good guide for selection of models whose forecasts are used in decisionmaking. On the other hand, if there is little relationship between mean square error and the I and D statistics across models, mean square error is not a useful criterion for model selection whose forecasts directly influence decision making. Four models forecast the quantitative series of quarterly Kansas City 600-700 pound feeder prices, while the other model employs the logit procedure to predict the profitability of overwintering feeder cattle in North Florida. All models are estimated over the 1960 to 1975 period, and *ex ante* forecasts are derived over the 1976 to 1981 period.

Two deterministic models are analyzed: the no-change model and the trend model. The no-change model says that forecasted price is equal to current price; i.e.,

$$(14) \qquad P^*_{t+1} = P_t.$$

Forecasts from this model are referred to as "naive forecasts."

The trend model is based on the notion that current price trends will persist. Mathematically,

$$(15) \qquad P^*_{t+1} = P_t + (P_t - P_{t-1}).$$

---

[1] Consider a model with I close to zero, which indicates the model is usually "wrong." This model could be useful if the decisionmaker followed a rule to do the opposite of what the model recommends. The model, however, is wholly inadequate as a predictor of the time series of interest and thus it is questionable if it consistently would provide a "wrong" forecast.

[2] The statistical properties of I and D are not treated in this paper.

[3] Choice grade prices are used before the change in the feeder cattle grading system was made. Kansas City prices are used because prior to 1970 Florida feeder calf price time series are incomplete.

[4] These models are selected from a set of seventeen models analyzed by Arnade.

The forecasted price is equal to current price plus the difference between current price and the price observed in the previous period. If the price in period t is greater than the price in period $t-1$, then forecasted price in period $t+1$ is greater than price in period t. Forecasts from this model are called extrapolative forecasts, because the forecast is an extrapolation of the last two observed prices. Within sample mean square error (MSE) and the I statistic for the no-change model and the trend model are shown in Table 1.

The third model analyzed is a Box-Jenkins time series model. Quarterly data over the 1960-1975 period yields 64 observations. The data are analyzed using standard Box-Jenkins procedures (Nelson, pp. 69-142). Analyses indicated that the appropriate autoregressive-integrated-moving average (ARIMA) model is a first-differenced, first-order autoregressive model. Using the backshift operator B ($BP_t = P_{t-1}$), the estimated model is:

$$(16) \qquad (1 - .324B)\,(1 - B)P_t = U_t,$$

where U_t is the random disturbance. The Box-Pierce test was applied to the estimated residuals of equation (16) (Box and Pierce). The hypothesis that the residuals are a white noise process could not be rejected. Within sample mean square error, the I statistic and the estimated Box-Pierce statistic for the ARIMA model are shown in Table 1.

A different philosophical approach is taken in the formulation of price forecasting models based upon the structure of the market for feeder cattle. For forecasting purposes, a structural econometric model is written in reduced form. In reduced form, however, it still may be difficult to use the model for forecasting if several of the explanatory variables are current exogenous variables. The use of lagged endogenous and exogenous variables can ameliorate this difficulty.

TABLE 1. ESTIMATED PRICE FORECASTING MODELS, ASSOCIATED MEAN SQUARE AND I STATISTICS FOR QUARTERLY KANSAS CITY MF1 600-700 POUND FEEDER CATTLE PRICES OVER THE PERIOD 1960-1975

Model	MSE [a]	I [b]
No change: .. $\hat{P}_t = P_{t-1}$	7.15	.625
Trend: ..... $\hat{P}_t = P_{t-1} + (P_{t-1} - P_{t-2})$	9.80	.733
ARIMA: ..... $(1 - .312B)\,(1 - B)P_t = U_t$ [c]	6.65	.733

[a] Mean square error.
[b] Percentage of correct predictions relative to the estimated breakeven price.
[c] Estimated Box-Pierce Q statistic with 22 degrees of freedom is 22.46. The 95th percentage point from a χ^2 statistic with 22 degrees of freedom is 33.9 (for details see Box and Pierce).

In the particular problem of this study, only second quarter feeder cattle prices are of interest. A reduced-form equation is:

$$(17) \quad P_s = \alpha_0 + \alpha_1 X_{1,s-1} + \alpha_2 X_{2,s} + \alpha_3 X_{3,s-1} + U_s\,,$$

where:

P_s = second quarter 600-700 pound Kansas City MF1 steer prices in year s,

$X_{1,s-1}$ = index of the ranch costs in year $s-1$ (calculated on a per head basis, see Arnade, p. 53),

$X_{2,s}$ = marketings from feedlots in the seventeen largest cattle feeding states in year s, and

$X_{3,s-1}$ = average fourth quarter U.S. corn prices in year $s-1$.

U_s is a random disturbance and α_i, $i=0, \ldots, 3$ are parameters to be estimated. $X_{1,s-1}$ is an index of prices paid by farmers for hay, fertilizer, etc. and should reflect the cost of producing calves, and hence acts as a supply shifter for feeder cattle. Current marketings from feedlots ($x_{2,s}$) and the price of corn ($X_{3,s-1}$) should capture the factors that influence feedlots and hence act as demand shifters. $X_{1,s-1}$ and $X_{3,s-1}$ are known in the fourth quarter when predicting the second quarter price next year, while $X_{2,s}$, fed marketings, is not known and must be forecasted. The parameters of equation (17) are estimated using ordinary least squares. The estimated parameters, mean square error (MSE), and the I statistic for the model are shown in Table 2. The model hereafter is referred to as the least squares model.

The logit model (18) uses the same regressors as in equation (17), and is of the form:

$$(18) \quad R_s = \beta_0 + \beta_1 X_{1,s-1} + \beta_2 X_{2,s} + \beta_3 X_{3,s-1} + V_s\,,$$

where:

R_s = 1 if the actual second quarter 600-700 pound Kansas City MF1 steer price exceeds the breakeven price in year s, and = 0 otherwise.

β_i, $i=0, \ldots, 3$ are parameters to be estimated, V_s is a random disturbance and all other variables are as defined. The estimated parameters and the I statistic for the logit model are shown in Table 2. Mean square error is not applicable for the logit model.

TABLE 2. ESTIMATED LEAST SQUARES AND LOGIT MODELS, ASSOCIATED MEAN SQUARE AND I STATISTIC FOR SECOND QUARTER KANSAS CITY, MF1 600-700 POUND FEEDER CATTLE PRICES FOR THE PERIOD 1960-1975

Model	Intercept	X_1[a]	X_2[b]	X_3[c]	MSE	I
Least squares	−.182	1.024	−.569	−3.073	5.92	.625
	(3.32)[d]	(.109)	(.22)	(1.35)		
Logit	8.36	.179	−.49	−4.55	[e]	.75
	(5.64)	(.125)	(.03)	(4.22)		

[a] Estimated cost of producing a feeder calf in the previous year. For details see Arnade (p. 53).
[b] Annual fed cattle marketings in the seventeen largest cattle feeding states (thousand head) (USDA, Livestock and Meat Statistics).
[c] Average fourth quarter U.S. corn price (USDA, Agricultural Prices).
[d] The figures in parentheses are the estimated standard errors of the estimated parameters.
[e] Mean square error is not applicable for the logit model.

The parameters of the least squares model were estimated by minimizing mean square error while the parameters of the logit model were estimated in a manner equivalent to maximizing the I statistic. Comparison of the estimated parameters of the two models demonstrates that minimizing mean square error and maximizing the I statistic are not equivalent.

Computation of Forecasts

To better assess the forecasting properties of the five models, post sample evaluation is performed over the 6-year period 1976 to 1981. Predicted prices from each model as well as the breakeven price and actual price are shown in Table 3.

The ARIMA, least squares, and logit models are all updated. For example, to generate the 1977 forecast, the model is re-estimated using data up through 1976. All updating is done by simply augmenting the sample. After each update, the residuals of the ARIMA model are checked via the Box-Pierce statistic for white noise. The tests show that the (1,1,0) model is adequate to reduce the series to white noise over the 1960 to 1981 period (although the estimate of the first order autoregressive coefficient varies slightly).

The least squares and logit models are updated in a similar fashion. The parameter estimates of these models are also stable as the sample is augmented.

To generate forecasts from both the least squares and logit models, fed marketings (X_2) must be forecasted. An ARIMA model is estimated for this series and used to provide forecasted values of fed marketings.

TABLE 3. PREDICTED, ACTUAL, AND BREAKEVEN SECOND QUARTER KANSAS CITY MF1 600-700 POUND FEEDER CATTLE PRICES OVER THE PERIOD 1976-1981

Model	1976	1977	1978	1979	1980	1981
No change	38.05[a]	36.90	41.25	67.44	83.30	75.53
Trend	43.71	33.00	39.75	75.02	81.77	72.08
ARIMA	39.32	35.49	40.70	69.03	81.80	74.32
Least squares	40.20	45.84	63.13	84.93	80.97	77.86
Logit	.012[b]	.485	.825	.820	.920	.415
Breakeven	39.32	44.14	60.19	84.09	79.57	78.64
Actual	43.89	41.10	58.00	86.74	70.43	70.65

[a] Price per hundredweight.
[b] The probability that the actual price exceeds the breakeven price.

Forecast Evaluation

The performance of the five models in *ex ante* forecasting is summarized in Table 4. Mean square error (MSE), proportion of correct forecasts (I), and profit deviation (D) are computed for each model.

The least squares model is the best model both within sample and post sample MSE. The trend model has the highest within sample MSE. On the other hand, its post sample MSE is second best, although its post sample MSE differs little from that of the no change and ARIMA models.

In post sample evaluation, all models were correct 50 percent of the time except the trend model which had a two-thirds correct rate. The trend model also gave the best value of D, with the least squares model yielding the second lowest D.

The results suggest that the minimum MSE model is not the most useful model in terms of implementation of the correct decision. The statistics I and D measure approximately the same phenomenon, as the same model yielded the best I and D values.

TABLE 4. SUMMARY OF FORECAST PERFORMANCE FOR SECOND QUARTER KANSAS CITY MF1 600-700 POUND FEEDER CATTLE PRICES OVER THE PERIOD 1976-1981

Model	MSE	I	D
No change	12.21	.5	95.56
Trend	10.71	.67	69.42
ARIMA	11.60	.5	95.56
Least squares	6.17	.5	84.49
Logit	[a]	.5	92.74

[a] Not applicable.

CONCLUDING REMARKS

Statistical forecasting models are usually evaluated on the basis of a criterion in which large discrepancies between predicted and observed values are penalized greater than small discrepancies. Mean square error, mean absolute error, and Thiel's U statistic are examples of such loss functions. Forecasts, however, may be used as input to decisionmaking. In this context, a more appropriate forecast evaluation would be to analyze the outcome of the decision implemented.

In this paper, a model is developed for the decision to implement stocker cattle production. The decision is assumed to hinge on the

relationship between fixed, known production costs and the predicted price of stocker cattle. The model is specified for the decision to overwinter feeder cattle in North Florida.

Five forecasting models are presented. The analysis demonstrates that a statistical criterion, mean square error, and alternative criteria, which are based on the outcome of the decision implemented, are not equivalent. A relatively simple trend model, which exhibited a high within-sample mean square error, proved to be a better forecast tool, based on the outcome of the stocker decision, than more statistically sophisticated models.

The evaluation is performed over 1976 to 1981, a period characterized by wide price swings with high prices more than double low prices over the period. The performance of all five models is marginal, which is consistent with other recent studies dealing with forecasting and cattle markets (Brandt, Harris). Thus, the implication of the study that the trend model is a superior forecasting model is limited by peculiarities of the feeder cattle market in the post sample forecasting period.

REFERENCES

Arnade, Carlos A. "An Evaluation of Alternative Price Forecasting Mechanisms in Relation to the Florida Stocker Decision." M.S. Thesis, University of Florida. 1982.

Box, G.E.P., and G.M. Jenkins. *Time Series Analysis Forecasting and Control,* 2nd ed. San Francisco: Holden-Day. 1976.

Box, G.E.P., and D.A. Pierce. "Distributions of Residual Autocorrelations in ARIMA Time Series Models." *Journal of the American Statistical Association* 64:1209. 1970.

Brandt, Jon A. "Statistical and Economic Performance Evaluations of Alternative Forecasting Approaches for Livestock Prices." Selected paper presented at the American Agricultural Economics Association meetings, W. Lafayette, Indiana, July 1983.

Florida Department of Agriculture and Consumer Services. *Beef Cattle in Florida.* Bulletin No. 28. Gainesville, Florida 1977.

Gunter, D., and G. Westberry. "The Cost and Returns of Grazing Stocker Calves on Annual Pastures in North Florida." Food and Resource Economics Department, University of Florida, Gainesville. 1979.

Harris, Kim S. "Model Selection Among Alternative Steer Prices Forecasting Techniques." Selected paper presented at the American Agricultural Economics Association meetings, W. Lafayette, Indiana, July 1983.

Judge, G., R. Hill, W. Griffiths, H. Lutkepohl, and T.C. Lee. *Introduction to the Theory and Practice of Econometrics.* New York: John Wiley and Sons. 1982.

Nelson, C.R. *Applied Time Series Analysis for Managerial Forecasting.* San Francisco: Holden-Day. 1973.

U.S. Department of Agriculture. *Agricultural Prices.* Washington, D.C. 1960-1982.

U.S. Department of Agriculture. *Livestock and Meat Statistics.* Washington, D.C. 1960-1982.