FORECASTING MONTHLY SLAUGHTER COW PRICES WITH A SUBSET AUTOREGRESSIVE MODEL

Thomas H. Spreen, Richard E. Mayer, James R. Simpson, and James T. McClave

Cow-calf operations dominate cattle production in the Southeast. The 1978 January 1 U.S. cattle inventory report shows that breeding cows make up 45 percent of the total inventory in Region 4[1] compared with 33.3 percent nationally [3]. Further, in 1977 cow slaughter accounted for 64.8 percent of federally inspected cattle slaughter in Region 4 compared with 23.6 percent nationally [8].

Most analyses of the cattle industry tend to overlook the importance of slaughter cows to ranchers by focusing only on feeder and fed cattle prices [3, 5, 6, 8, 11, 14, 15, 16, 19]. Cow slaughter can be erratic and can result in drastic price fluctuation, part of which is seasonal. Cattle producers can both gain and lose from these erratic prices depending on when they choose to market their cull cows. Fall is historically the season of heaviest cow marketings, and Southeastern producers can cultivate temporary winter pastures and hold cull cows (as opposed to stocker calves) in anticipation of a price upswing. Thus, short-term forecasting of slaughter cow prices can assist cattle producers in the marketing decisions for cull cows.

A forecasting model of Florida cow prices is developed. A monthly model is estimated with data from the 1955-1975 period and post-sample analysis is performed for 1976 through June 1978.

METHOD

Methods for forecasting cattle prices differ considerably in approach and complexity. The major tool used in construction of these models is multiple regression. For example, the principal forecasting model of the livestock subsector used by the U.S. Department of Agriculture is a large-scale econometric model with 147 structural equations [17]. The major drawback of the structural approach, however, is stated succinctly by Johnson [7], who argues that a forecasting model is of little or no use if it projects values for some set of endogenous variables on the basis of a set of exogeneous variables which themselves must be projected. It may in fact be more difficult to forecast the exogenous variables than the endogeneous variables whose variation the exogeneous variables are purported to explain.

When the objective is to forecast, the time series analysis methods developed by Box-Jenkins [1] provide an alternative. Several authors have applied these techniques to selected data series—Leuthold et al. [9] to hog prices, Oliveira et al. [13] to lumber prices, and Oliveira et al. [14] to selected fed, feeder, and future cattle price series.

The Box-Jenkins procedure is to fit a model which is based exclusively on the past behavior of the data series of interest. The so-called ARMA model presupposes that a time series is composed of autoregressive (AR) and moving average (MA) components. A brief description of the Box-Jenkins procedure follows. A series Z_t is autoregressive (AR) of order p if

$$(1)\quad Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + U_t$$

where ϕ_i are parameters to be estimated, U_t is a series of uncorrelated random variables with mean 0 and variance σ_u^2 (commonly referred to as white noise), and Z_t is assumed to be stationary.[2]

The moving average (MA) model presupposes that the stationary series Z_t can be written as

$$(2)\quad Z_t = U_t - \theta_1 U_{t-1} + \ldots + \theta_q U_{t-q}$$

where θ_j $(j=1, \ldots, q)$ are parameters to be estimated and the U_t are defined as before.

Combining equations 1 and 2 gives the autoregressive — moving average (ARMA) model.

Thomas H. Spreen is Assistant Professor and James R. Simpson is Associate Professor, Food and Resource Economics Department, and James T. McClave is Associate Professor, Department of Statistics, the University of Florida. Richard E. Mayer is Assistant Manager, Swift & Co., Marshalltown, Iowa.

Florida Agricultural Experiment Station Journal Series No. 1751.

[1]Region 4 includes the following states: Kentucky, Tennessee, North Carolina, South Carolina, Georgia, Alabama, Mississippi, and Florida.

[2]Technically, all roots of $m^p - \sum_{j=1}^{p} m^{p-j}\phi_j = 0$ are inside the unit circle.

$$(3)\quad Z_t - \phi_1 Z_{t-1} - \ldots - \phi_p Z_{t-p} = U_t - \theta_1 U_{t-1} - \ldots - \theta_q U_{t-q}$$

To more completely write the ARMA model, one uses the backshift operator B, where $B^n Z_t = Z_{t-n}$. *Substitution into equation 3 gives*

$$(4)\quad (1\text{-}B\phi_1 - B^2\phi_2 - \ldots - B^p\phi_p)Z_t = (1\text{-}B\theta_1 - \ldots - B^q\theta_q)U_t.$$

The general ARMA model in equation 4 applies to any stationary autoregressive-moving average process of orders p and q (autoregressive and moving average portions of the model, respectively). Stationarity of the series is a necessary assumption. Often, however, the series will not be stationary, but instead will contain a trend component. In such a case, the series must be transformed in some manner that will render the series stationary so the stationarity assumption will not be violated. One commonly used procedure is differencing of the series. Using the backshift operator notation adopted in equation 4, one would, for example, write the first differenced series as $\Delta Z_t = (1\text{-}B)Z_t$ where $\Delta Z_t = (Z_t - Z_{t-1})$.

A model which includes the differencing necessary to render the series stationary is the ARIMA model (integrated autoregressive-moving average model).

The complete ARIMA (p, d, q) model, where p is the order of the autoregressive component, q is the order of the moving average component, and d is the degree of consecutive differencing required to achieve stationarity, is given by

$$(5)\quad (1\text{-}B\phi_1 - B^2\phi_2 - \ldots - B^p\phi_p)\,(1\text{-}B)^d Z_t = (1\text{-}B\theta_1 - \ldots - B^q\theta_q)U_t.$$

The first step in constructing an ARIMA model is model identification, i.e., one must determine whether the time series at hand is AR, MA, ARMA, or ARIMA, and the order of the process. This step entails examination of the sample autocorrelations and partial autocorrelations of the series of interest. For specific details of this procedure, see Nelson [12]. The procedure is somewhat subjective and often more than one tentative model is "identified."

One explanation for the plausibility of several ARIMA models for the same time series is given by Wold's Theorem [4]. Essentially, this theorem states that a stationary process may be approximated arbitrarily close by *both* a finite order AR *and* a finite order MA model. The order of the approximating model may have to be very large, and in these cases an ARIMA model is often proposed. The ARIMA model has the advantage that few terms are needed to describe a wide variety of time series processes, whereas AR or MA models may need many terms to describe the process adequately. ARMA models, however, have the distinct disadvantage of being difficult to estimate and use in forecasting.

McClave [10] has developed a procedure which approximates stationary processes by making use of subset AR models. In a subset model, some of the coefficients are constrained to be zero, allowing one to fit a model with higher order lags while estimating few coefficients. For example, the subset AR model

$$Z_1 - \phi_1 Z_{t-1} - \phi_3 Z_{t-3} - \phi_{12} Z_{t-12} = U_t$$

might be used for a monthly series having seasonal components, but only three parameters need be estimated. The subset model can be readily identified, estimated, and used in forecasting while maintaining a parsimonious representation.

To determine the best subset model, McClave uses a likelihood ratio test: the Max χ^2 technique. The user first selects a value K such that the maximum lag in the model is at most K. The procedure next searches and determines the best model of order 1, 2, 3 up to order K. This is accomplished, say for order q, by considering all possible combinations of the K lags taken q at time to find the combination which minimizes the residual variance. Let σ_q^2 be the residual variance associated with the best subset model of order q. The σ_q^2's are then sequentially tested by the statistic

$$(6)\quad M_{q+1} = (N\text{-}q\text{-}1)\ \frac{\sigma_q^2 - \sigma_{q+1}^2}{\sigma_{q+1}^2}$$

where N is the sample size and σ_{q+1}^2 is the residual variance associated with the best model of order q+1. As pointed out by McClave, the asymptotic distribution of M_{q+1} is not known (except for q=0), but a conservative test for the hypothesis that σ_{q+1}^2 is not significantly smaller than σ_q^2 is reject if

$$(7)\quad M_{q+1} > \overline{\chi^2}(K-q), \alpha$$

where $\overline{\chi^2}(K-q),\alpha$ is the upper α level corresponding to the maximum order statistic in a sequence of K−q χ^2 random variables each with one degree of freedom. The results of many simulations [5] indicate that the procedure selects the correct model with high probability. Furthermore, the procedure tends to select parsimonious representations, a desirable characteristic of ARMA models.

MODEL ESTIMATION

McClave's procedure was applied to a time series consisting of average monthly prices of

Florida slaughter cows, averaged over all grades. The data covered the 1955-1975 period. An average of all grades was used because grading procedures have changed, and averaging over all slaughter cow categories permitted construction of a time series of sufficient length for effective use of the Max χ^2 procedure.

Initial analyses indicated that first order differencing was necessary to remove a trend and a twelfth order difference was used to remove a strong seasonal effect. Thus, the series

$$Y_t = (1\text{-}B^{12})(1\text{-}B)Z_t$$

was deemed stationary and was analyzed by the Max χ^2 procedure.

The results of the Max χ^2 are presented in Table 1. As can be seen from the significance levels in the table, the Max χ^2 procedure selected a model of order five with lags of 1, 6, 12, 24, and 36 as the "optimal" model. The "best" model of order 6 did not result in a significant reduction in residual variance compared with that achieved in the "best" order 5 model.

TABLE 1. RESULTS OF THE MAX χ^2 PROCEDURE

Model	Lags Which Yield Minimum Residual Variance	Residual Variance	Max χ^2	Significance Level
0	None	0.9212		
1	12	0.7372	58.1	0.0
2	6, 12	0.6555	28.9	0.00
3	6, 12, 24	0.6138	15.7	0.003
4	6, 12, 24, 36	0.5764	14.9	0.005
5[a]	1, 6, 12, 24, 36	0.5472	12.2	0.020
6	1, 4, 6, 12, 24, 36	0.5300	7.4	0.247

[a]Optimal model according to the specified Max χ^2 test criterion; significance level = .05.

Estimation of the parameters of the model of order 5 by the Yule-Walker equation [4] yields

$$(8)\quad \underset{(.058)}{(1 + .186B} + \underset{(.053)}{.213B^6} - \underset{(.053)}{.650B^{12}} - \underset{(.065)}{.341B^{24}} - \underset{(.0685)}{.215B^{36})}(1\text{-}B^{12})(1\text{-}B)Z_t = U_t$$

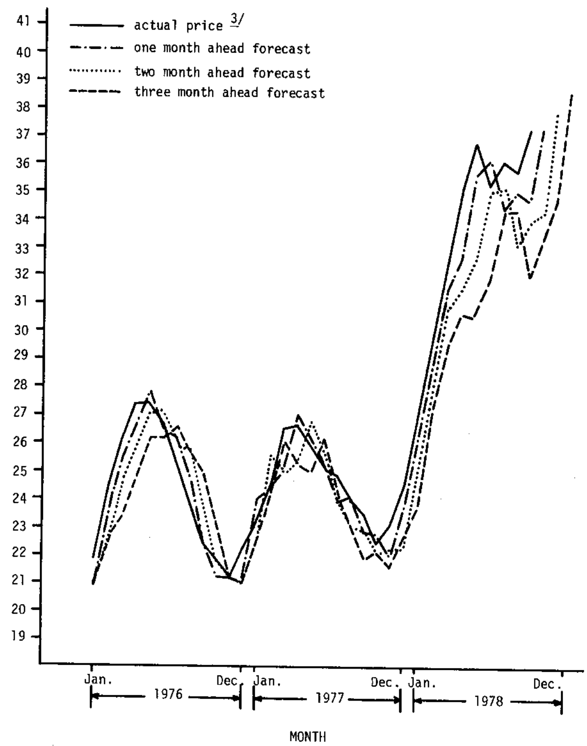
where the numbers in parentheses are the estimated standard errors.

POST-SAMPLE EVALUATION

Data for the years 1976-1978 were not used in model development and thus were available for analysis of the model's performance during that period. The procedure was to use the model to forecast 12 months ahead successively using December 1975, January 1976, February 1976, etc. as the origin data. Specifically, the initial forecast used information available in December 1975 to provide monthly forecasts of expected prices for the period January 1976 through December 1976. The next forecast used information available in January 1976 to provide monthly forecasts of expected prices for the period February 1976 through January 1977. The last forecast used information available in December 1977 to provide monthly forecasts of price expectations for the period January 1978 through December 1978. The result was a series of 24 forecasts of 12 months each.

Actual prices and 1, 2, and 3-month-ahead forecasts are plotted in Figure 1. Visual inspection shows that the model performed well for 1976 and 1977.

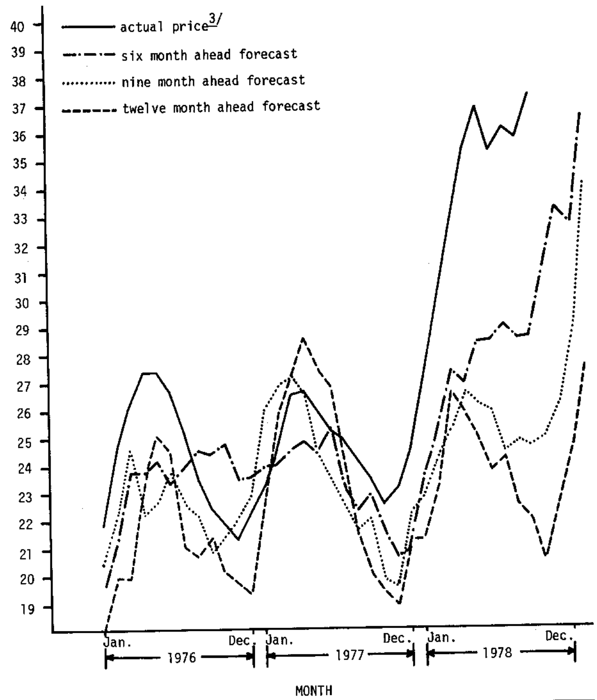
FIGURE 1. COMPARISON OF ACTUAL AND 1, 2, AND 3-MONTH-AHEAD FORECASTS OF FLORIDA SLAUGHTER COW PRICE



Actual prices and 6, 9, and 12-month-ahead forecasts are plotted in Figure 2. The model's accuracy diminishes as the forecast horizon is extended. Mean absolute errors for forecast horizons of 1 through 12 months computed for the 1976 through 1977 period are listed in Table 2. The data available for the period beginning January 1978 are not directly comparable[3] with those used in model development

[3]Data up to and including December 1977 represent a weighted average of Florida auction prices *and* estimates of direct marketing prices averaged across all grades. Data beginning January 1978 represent only a weighted average of Florida auction prices averaged across all grades. The Florida Crop and Livestock Reporting Service compiles auction prices monthly but then adjusts these figures annually by their estimates of direct marketing volumes and prices. Thus, the data after December 1977 are to be considered preliminary.

FIGURE 2. COMPARISON OF ACTUAL AND 6, 9, AND 12-MONTH-AHEAD FORECASTS OF FLORIDA SLAUGHTER COW PRICE



and post-sample evaluation and thus mean absolute forecast errors were computed for the period from January 1976 through December 1977 only. The table illustrates that the model serves as a good predictor of prices particularly in the short-term but, as expected, forecast accuracy diminishes as forecast horizon is extended. Even at 12 months ahead, however, the mean absolute error is less than 10 percent. No other forecasting models for Florida cow prices are available for comparison.

SUMMARY AND CONCLUDING COMMENTS

A short-term forecasting model of Florida slaughter cow prices is formulated and estimated, and post-sample analysis demonstrates its effectiveness as a forecasting tool. The model is very accurate in giving 1, 2, and 3-month-ahead projections, but less satisfactory in more distant forecasts.

The model provides a valuable decision-making aid to Florida cattlemen who must determine when to market slaughter cows. As shown in Figure 1, the model forecasted a strong price increase in early 1978 which did in fact occur. It would have been very profitable for cattlemen to over-winter cull cows to take advantage of the price rise.

TABLE 2. MEAN ABSOLUTE FORECAST ERRORS BY FORECAST HORIZON (months ahead)

Forecast Horizon (months)	Mean Absolute Forecast Error
1	0.61
2	1.01
3	1.25
4	1.49
5	1.69
6	1.94
7	2.01
8	2.01
9	2.07
10	2.28
11	2.49
12	2.61

REFERENCES

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

[2] Florida Department of Agriculture. *Florida Agricultural Statistics-Livestock Summary,* Florida Crop and Livestock Reporting Service, 1977.

[3] Franzmann, John R. and Rodney L. Walker. "Trend Models of Feeder, Slaughter, and Wholesale Beef Cattle Prices," *American Journal of Agricultural Economics,* Volume 54, 1972, pp. 507-512.

[4] Fuller, Wayne A. *Introduction to Statistical Time Series.* New York: John Wiley and Sons, Inc., 1976.

[5] Handke, Paul and Gene Futrell. "Forecasting Guidelines for Choice Steer and Heifer Prices," *Economic Report Series No. 12*, Department of Economics, Iowa State University, January 1978.

[6] James, John B. "Factors Affecting Price Differences of Cattle in the Southwest," Ph.D. dissertation, Texas A&M University, 1970.

[7] Johnson, Stanley R. "Agricultural Sector Models and Their Interface with the General Economy: Discussion," *American Journal of Agricultural Economics*, Volume 59, 1977, pp. 133-136.

[8] Jones, Lonnie L. "Short-Term Forecasting Models for Average Monthly Prices of Wholesale Beef," unpublished Ph.D. dissertation, The Ohio State University, 1967.

[9] Leuthold, R.M., J.A. MacCormick, A. Schmitz, and D.C. Watts. "Forecasting Daily Hog Prices and Quantities: A Study of Alternative Forecasting Techniques," *Journal of the American Statistical Association*, Volume 65, 1970, pp. 90-107.

[10] McClave, James T. "Estimating the Order of Autoregression Models: The Max χ^2 Method," *Journal of the American Statistical Association*, Volume 73, 1978, pp. 122-128.

[11] McIlroy, Walter. "Application of the Econometric Model, Distributed Lag, to Forecasting Slaughter Cattle Price," unpublished Ph.D. dissertation, Oklahoma State University, 1973.

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

[13] Oliveira, R.A., A.J. Buongiorno, and A.M. Kniotck. "Time Series Forecasting Models of Lumber Cash, Futures, and Basis Prices," *Forest Science*, Volume 23, 1977, pp. 268-280.

[14] Oliveira, R.A., C.W. O'Connor, and G.W. Smith. "Short-Run Forecasting Models of Beef Prices," paper presented at the Western Agricultural Economics Association Annual Meeting, Bozeman, Montana, July 1978.

[15] Rohdy, D. Dale, G.H. Hoffman, and A.G. Madsen. "Short-Run Price Forecasting Model for Fed Cattle," Colorado Agricultural Experiment Station Technical Bulletin No. 108, December 1969.

[16] Teigen, Lloyd. "Costs, Loss and Forecasting Error: An Evaluation of Models for Beef Prices," unpublished Ph.D. dissertation, Michigan State University, 1973.

[17] U.S. Department of Agriculture, CED/USDA, *Baseline*, November 1977.

[18] U.S. Department of Agriculture, *Livestock and Meat Statistics*, 1977.

[19] Whyte, Charles D. "Dynamic Short-Term Estimating and Forecasting Models for Wholesale Beef Price," unpublished Ph.D. dissertation, The Ohio State University, 1969.