QUARTERLY PREDICTION MODELS
FOR LIVE HOG PRICES*

Richard J. Foote, Sujit K. Roy, and George Sadler

Pronounced short-run movements in production, marketings, slaughter and prices of live hogs create problems of uncertainty in the decision-making process of producers and other participants in the hog-pork industry. The objective of the present study was to develop quarterly prediction models for live hog prices, based on structural relations representing the price-determining forces in the sector.

The relative importance of the hog sector in the livestock industry is underscored by the fact that the annual cash income for the sector ranks second to beef. For instance, in 1970, hogs accounted for about 15 percent of the total cash receipts from all livestock products. About one-half of total hog production in 1970 was in the seven states of West North Central Division, with Iowa accounting for almost one-fourth of all hog production in the United States. This division, along with four contiguous and six southern states, accounted for about 90 percent of the total U.S. hog production [9]. Although the total consumption of pork has increased over past decades, pork has represented a smaller percentage of total meat consumption. The per capita consumption of pork has shown no appreciable increase since the middle fifties, while that of beef rose steadily during the period.

Hog prices tend to follow a seasonal pattern, directly related to the marketing of hogs, which in turn is related to time of farrowing and to feeding and breeding programs. Greatest concentration of farrowings is the spring pig-crop during March and April and during September for the fall pig-crop. Seasonal patterns for production of hogs tend to exhibit some degree of regularity, moreso in the summer months than during winter. Hog prices typically reach a high in the summer (July) when supplies are low, and a low in the fall (November) when they are relative large.

SPECIFICATION OF MODELS
AND STATISTICAL PROCEDURES

The quarterly simultaneous-equation models for predicting live hog prices consisted of the following three stochastic relationships and a market-clearing identity for pork.

Price-consumption relation for pork:

$$C_t = f(R_t, I_t, B_t, P_t) \tag{1}$$

Function relating live hog price and retail pork price:

$$H_t = f(R_t, W_t, T, Q_t) \tag{2}$$

Cold storage stocks of pork products relation:

$$S_{t+1} = f(C^*_{t+1} - Q^*_{t+1}, C^*_{t+2} - Q^*_{t+2}) \tag{3a}$$

or

$$S_{t+1} = f[R_t, \Delta I_t, I^*_{t+1}, S_t,$$

$$(F_{t-2} \times L_{t-2}), (F_{t-1} \times L_{t-1}),$$

$$(F_t \times L^*_t)] \tag{3b}$$

Market clearing identify for pork products:

$$C_t = Q_t + S_t - S_{t+1} \qquad (4)$$

Variables included in the preceding equations are defined as follows:

C = consumption of pork (civilian consumption) in the U.S. during the quarter, million pounds, derived from the identity, (4)

R = adjusted average weighted retail price of pork products for the quarter, cents per live hog equivalent pound. (The adjusted price series, used for statistical estimation, was developed by use of net yield conversion factors to express the series on a comparable basis with the live hog price series.)[1]

I = disposable personal income for the quarter at seasonally adjusted annual rates, billion dollars

B = consumption of beef during the quarter, million pounds

P = consumption of poultry; civilian consumption of turkeys and broilers, during the quarter, million pounds

H = price of live hogs at Omaha, Nebraska, 200-220 pounds, barrows and gilts, U.S. No. 1-2, cents per pound

W = average weekly earnings of production or nonsupervisory workers in meat packing plants during the quarter

T = time, where T=57 for 1957, T=58 for 1958, etc.

Q = production of pork million pounds, during the quarter; commercial pork production, 48 states of the U.S., excluding lard and rendered pork fat

S = stocks of cold storage pork first of quarter, million pounds; frozen and cured pork, cold storage holdings, 48 states of the U.S.

F = number of sows farrowed during the quarter, million head

L = average number of pigs saved per litter during the quarter

Variables with the asterisk refer to the expected value of the corresponding variables, and the subscripts, t, t-i, and t+i, (where i=1,2), refer to the current, lagged and succeeding quarters respectively.

Data for disposable income (I) were obtained from [10, 12, 13]; and data for average weekly earnings in meat packing plants (W) from [14]. Sources of data for all remaining variables were [7, 8, 9].

Consumption of pork in equation (1) is specified to depend on retail prices of pork and disposable personal income. Consumption of beef and poultry, two substitute products for pork, may be inversely related to pork consumption.

Live hog price is expressed as a function of retail pork price (2). The difference between the two can be, to some extent, explained by the processing cost, assumed to be represented partly by the wage of labor (W) in meat packing plants. Given the retail price of pork, an increase in labor costs would tend to reduce live hog prices. Pork production for the period (Q_t) is used in equation (2) as a proxy for the demand for marketing services. When Q_t increases, indicating increased pressure on existing limited marketing services, the increased distribution cost per unit may tend to lower live hog prices when the retail price of pork (R_t) is assumed constant. It may be assumed that available per-unit marketing services do not increase significantly in the short run.

The first equation representing the cold storage stocks relation, (3a), cannot be estimated directly, since it includes expectation variables such as the expected levels of consumption and production of pork one and two quarters ahead. However, the alternative version of the storage stock relation, (3b), which is derived from (3a), includes variables for which actual data are available from published sources. The derivation of (3b) from (3a), based on the approach used in [1], may be presented as follows:

Expected levels of production, Q^*_{t+1} and Q^*_{t+2} in (3a) are assumed to depend primarily on expected pig crops, marketable one and two quarters ahead. Since it takes approximately six months for pigs, from birth, to attain a 200-240 pound marketable weight, expected pig crops, and hence production (Q^*), may be assumed to be a function of lagged values of the number of sows farrowed (F) and the average litter size (L), i.e.,

$$Q^*_t = f(F_{t-2} \times L_{t-2})$$

$$Q^*_{t+1} = f(F_{t-1} \times L_{t-1})$$

$$Q^*_{t+2} = f(F_t \times L^*_t) \qquad (i)$$

---

[1] The live hog equivalent, in pounds, to one pound of retail pork products changed from 2.06 in 1957 to 1.97 in 1969 [11], and was estimated to have decreased over the period by an annual average rate of .004. The same rate of change was assumed to apply for years beyond 1969 in the present study. Reciprocals of the live hog equivalent in pounds to one pound of retail pork were used as net yield conversion factors. Retail pork prices were multiplied by the corresponding annual yield factors to obtain the adjusted retail price series.

Data on each of the independent variables including L_t^*, are published in [7]. The report contains a projected number of pigs per litter for the current quarter and the projected litter size is used in this model as L_t^*.

Expected consumption levels, C_{t+1}^* and C_{t+2}^*, equation (3a), may be assumed to be influenced by expected consumer incomes and expected prices of pork. Thus,

$$C_{t+1}^* = f(R_{t+1}^*, I_{t+1}^*)$$

and

$$C_{t+2}^* = f(R_{t+2}^*, I_{t+2}^*) \qquad \text{(ii)}$$

With respect to income expectations, it was assumed that projections regarding future income levels may be made on the basis of the most recent quarter to quarter change in income. Thus,

$$I_{t+1}^* = I_t + \Delta I_t \qquad \text{(iii)}$$

where

$$\Delta I_t = I_t - I_{t-1}$$

Expectations regarding prices (R*) were assumed to be based on Nerlove's specifications, which imply that price expectations are adjusted in proportion to the error made in the most recent past [5]. That is,
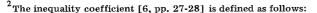
$$R_t^* - R_{t-1}^* = (R_{t-1} - R_{t-1}^*) \qquad \text{(iv-a)}$$

or

$$R_t^* = \beta R_{t-1} + (1-\beta) R_{t-1}^* \qquad \text{(iv-b)}$$

Based on the foregoing specifications on expectational variables for production, consumption, income and price and under some simplified assumptions, (3b) can be derived from the original relation, equation (3a). The derivation and assumptions are presented in the Appendix. The present treatment of the storage stocks relation differs significantly from those in earlier studies, such as Harlow [2] and Myers et. al. [4], with reference to the inclusion of expectational specifications.

The last equation of the model, (4), is the closing identity in the system. It equates consumption of

pork products with production of pork and net storage movement during the period.

Endogenous variables simultaneously determined within the system are R_t, H_t, S_{t+1} and C_t. Each of the three stochastic relations is over-identified, and subsequently (1), (2) and (3b) were estimated by using the three-stage least squares method. This procedure yields consistent estimators, which are generally more efficient than corresponding two-stage least squares estimators [15, 3]. Alternative estimates for the model for each quarter were obtained on the basis of per capita or total data for quantity and income variables. Furthermore, both two-stage and three-stage least squares estimates were evaluated. The models, estimated separately for each quarter, were based on 15 years' data beginning with 1957. The estimated model presented here for each quarter was preferred to the alternative estimates in terms of relative accuracy of predicting the endogenous variables, especially the live hog price, H_t. The predictive performance of each model was evaluated by using Theil's inequality coefficients (U_2) for each endogenous variable.[2]

ESTIMATED STRUCTURAL MODELS AND EVALUATION

The three-stage least squares estimates of the three stochastic structural relations for each model are presented below. The variables with prime (') signs were expressed in per capita terms, otherwise the variables were used as defined earlier. The value within parentheses below the coefficient is the ratio of the coefficient to its standard error.

First Quarter:

$$C_t' = \underset{(17.36)}{17.70} \underset{(-1.501)}{-0.1117B_t'} \underset{(8.143)}{+3.537I_t'}$$

$$\underset{(-9.665)}{-0.2837R_t} \qquad \text{(1)}$$

$$H_t = \underset{(0.7651)}{4.506} \underset{(-2.392)}{-0.7341Q_t'} \underset{(-2.073)}{-0.0551W_t}$$

$$\underset{(0.6651)}{+0.0758T} \underset{(10.78)}{+0.9120R_t} \qquad \text{(2)}$$

---

[2]The inequality coefficient [6, pp. 27-28] is defined as follows:

$$U_2 = \sqrt{\Sigma(\Delta P_{it} - \Delta A_{it})^2 \div \Sigma \Delta A_{it}^2}$$

where $\Delta P_{it} = P_{it} - P_{it-1}$, and $\Delta A_{it} = A_{it} - A_{it-1}$, and P_{it} and A_{it} are the predicted and actual levels, respectively, of the ith variable during the tth period. In the case of perfect predictions, that is, $P_{it} = A_{it}$ for all t periods, $U_2 = 0$. On the otherhand, $U_2 = 1$ for naive "no-change" extrapolative forecasts.

$$S'_{t+1} = 4.954 + 7.160\Delta I'_t + 0.5377 I^{*'}_{t+1}$$
$$(1.458) \quad (0.9287) \quad (0.9055)$$

$$+0.5626 S'_t - 1.620(F_{t-2} \times L_{t-2})$$
$$(1.996) \quad\quad (-0.1053)$$

$$-0.1088 R_t \quad\quad\quad (3)$$
$$(-1.229)$$

Second Quarter:

$$C_t = 3506.0 - 0.03564 B_t - 0.06662 P_t$$
$$(7.681) \quad (-3.047) \quad (-.2000)$$

$$+4.160 I_t - 68.47 R_t$$
$$(3.927) \quad (-4.543) \quad\quad\quad (1)$$

$$H_t = -14.86 - .00074 Q_t - .1461 W_t$$
$$(-2.613) \quad (-.3176) \quad (-2.840)$$

$$+.3767T + .9953 R_t$$
$$(3.071) \quad (8.900) \quad\quad\quad (2)$$

$$S_{t+1} = -711.3 + .07953 I^{*}_{t+1} + .7124 S_t$$
$$(-1.967) \quad (.2736) \quad\quad (.3937)$$

$$+35.89(F_{t-2} \times L_{t-2}) \quad\quad (3b)$$
$$(1.416)$$

Third Quarter:

$$C_t = 4426.0 - .4506 B_t - .7525 P_t$$
$$(10.81) \quad (-3.248) \quad (-2.531)$$

$$+7.902 I_t - 55.08 R_t$$
$$(7.465) \quad (-5.376) \quad\quad\quad (1)$$

$$H_t = -38.94 - .2258 W_t + .9873T$$
$$(-2.844) \quad -3.587) \quad (3.022)$$

$$+.7798 R_t \quad\quad\quad (2)$$
$$(5.251)$$

$$S_{t+1} = 211.5 + 1.033 I^{*}_{t+1} + .1522 S_t$$
$$(.7698) \quad (3.240) \quad\quad (1.085)$$

$$+14.99(F_{t-2} \times L_{t-2}) - 15.28 R_t \quad (3b)$$
$$(1.510) \quad\quad\quad (-1.584)$$

Fourth Quarter:

$$C_t = 3191.0 - .01073 B_t - .03416 P_t$$
$$(16.58) \quad (-.1392) \quad (-.1720)$$

$$-4.522 I_t - 70.64 R_t \quad\quad (1)$$
$$(8.502) \quad (-10.17)$$

$$H_t = -7.658 - .01037 Q_t + .6796T$$
$$(-1.405) \quad (-7.064) \quad (3.134)$$

$$+.5239 R_t \quad\quad\quad (2)$$
$$(3.505)$$

$$S_{t+1} = 3725.0 + 5.638\Delta I_t + 5.594 I^{*}_{t+1}$$
$$(1.108) \quad (0.5382) \quad (1.182)$$

$$-1.582 S_t + 16.62(F_{t-2} \times L_{t-2})$$
$$(.7952) \quad\quad (.7911)$$

$$-26.46(F_{t-1} \times L_{t-1}) - 148.2 R_t \quad (3b)$$
$$(-.5329) \quad\quad\quad (-1.313)$$

It was observed during the initial phase of estimation that coefficients of certain relatively minor variables for some quarters were associated with inconsistent signs and large standard errors. Those variables were omitted from the final estimated models, assuming that the related coefficients were insignificant. Results of the consumption-price relation, equation (1), for the first quarter, indicate that consumption of poultry, P_t, as a substitute product was excluded because of an inconsistent sign. However, for all other quarters, both beef and poultry consumption entered the consumption-price relation with the expected negative signs. The coefficients of retail price of pork and income were associated with relatively low standard errors. With regard to equation (2), ratios of the coefficients to respective standard errors for most variables appeared to be relatively high. Two variables, W_t or Q_t, were excluded from the equations for the last two quarters because of inconsistent signs of coefficients and relatively large standard errors.

Estimated equations for storage stocks, (3b), indicate that the current retail pork price influences end-of-period storage stocks (S_{t+1}) in the negative direction, while income (I^{*}_{t+1}) or the change in income (ΔI_t) affects S_{t+1} directly. Other variables in the equation [such as ($F_{t-2} \times L_{t-2}$) and S_t] were allowed to enter the equations regardless of signs,

since there was no firm *a priori* knowledge or expectation regarding the direction of effects. The ratios of coefficients to respective standard errors in (3b) were in most cases relatively low.

Values of the four endogenous variables were calculated from the reduced form equations, which were derived by solving the three estimated stochastic relations and the identity, (4), for each quarter. The inequality (U_2) coefficients, computed from the calculated and actual values of the endogenous variables for the period of fit, are presented in Table 1. The U_2 coefficients for all endogenous variables were less than 1 for the selected models. The estimates of live hog prices generally involved smaller errors than those for retail prices for the first three quarters. Errors for retail pork prices in the fourth quarter were smaller than those for other quarters. Estimates of consumption of pork, C_t, and storage stocks, S_{t+1}, appear to be reasonably accurate. The largest U_2 coefficient for these two quantity variables was less than 0.4.

Price predictions were generated from the structural models for three years (1972-73-74) beyond the period of fit to examine the models' performance during a period of unusual market situations. Predictions and actual levels of live hog prices for the period are presented in Table 2. Predicted prices for four out of twelve quarters involved errors of 5 cents or more. In some cases, however, although the error magnitude was large, the direction of change was predicted correctly. Given the abnormal market conditions of the period, the models seemed to be sensitive enough to capture the price variations for most quarters.

CONCLUDING REMARKS

The results of the 4-equation quarterly models, estimated by three-stage least squares, indicated that most of the expected relations among the major variables were "stable" over the period of fit. A significant feature of the models was the storage blocks relation, based on expectational specifications on certain variables. As a methodological note, it may be observed that the predictive accuracy of the structural models appeared to improve when the first round ordinary least squares equations included, instead of all exogenous variables in the system, only a subset of such variables selected on the basis of consistency of signs of the coefficients. Furthermore, similar monthly structural models in another phase of the research failed to produce prediction results as accurate as those obtained from the present quarterly models. All four quarterly models seemed to have performed satisfactorily with regard to prediction of live hog prices. Predictions for three years beyond the period of fit also appeared to be reasonable, in spite of unusual market conditions of the period.

TABLE 1. INEQUALITY (U_2) COEFFICIENTS FOR THE FOUR ENDOGENOUS VARIABLES OF THE QUARTERLY LIVE HOG PRICE MODELS

Quarter	Price of live hogs, H_t	Retail Price of Pork, R_t	Consumption of Pork, C_t	End-of-quarter storage stocks, S_{t+1}
1st [a]	0.2242	0.8337	0.2475	0.3173
2nd	0.4400	0.8479	0.3237	0.3920
3rd	0.2487	0.8419	0.1588	0.2253
4th	0.3914	0.2026	0.2476	0.2797

[a]Consumption and storage stocks, C_t and S_{t+1}, were in per capita terms for the first quarter model.

TABLE 2. PREDICTED AND ACTUAL QUARTERLY LIVE HOG PRICES, CENTS/POUNDS FOR YEARS BEYOND THE PERIOD OF FIT

Quarter	1972 $\hat{H}_t$	1972 H_t	1973 $\hat{H}_t$	1973 H_t	1974 $\hat{H}_t$	1974 H_t
1st	24.31	25.64	34.48	35.96	34.09	39.62
2nd	27.98	22.74	35.57	37.68	31.56	30.12
3rd	32.36	29.42	48.70	48.93	41.77	37.71
4th	29.38	29.82	34.36	42.97	34.49	38.87

Derivation of Storage Stocks Relation, (3b), from (3a)
Based on Expectational Specifications

Four sets of expectational variables were specified as follows:

$$Q_t^* = f(F_{t-2} \times L_{t-1}), \quad Q_{t+1}^* = f(F_{t-1} \times L_{t-1}) \quad \text{and} \quad Q_{t+2}^* = f(F_t \times L_t^*) \tag{i}$$

$$C_{t+1}^* = f(R_{t+1}^*, I_{t+1}^*) \quad \text{and} \quad C_{t+2}^* = f(R_{t+2}^*, I_{t+1}^*) \tag{ii}$$

$$I_{t+1}^* = I_t + \Delta I_t, \text{ where } \Delta I_t = I_t - I_{t-1} \tag{iii}$$

$$R_t^* - R_{t-1}^* = \beta(R_{t-1} - R_{t-1}^*) \tag{iv-a}$$

or

$$R_t^* = \beta R_{t-1} + (1-\beta) R_{t-1}^* \tag{iv-b}$$

Substituting C_{t+1}^* and Q_{t+1}^* (i=1,2) by corresponding variables indicated in (i) and (ii), equation (3a) may be rewritten as follows:

$$S_{t+1} = f(R_{t+1}^*, I_{t+1}^*, F_{t-1} \times L_{t-1}, R_{t+2}^*, I_{t+2}^*, F_t \times L_t^*) \tag{v}$$

Since I_{t+1}^* and I_{t+2}^* would be highly correlated in (v), I_{t+2}^* was omitted and ΔI_t was introduced into the equation for estimation purposes. Also, R_{t+1}^* and R_{t+2}^* are expected to have high multicollinearity. Hence, eliminating R_{t+2}^* from (v), the equation is presented in linear form as follows:

$$S_{t+1} = a + b_1 R_{t+1}^* + b_2 I_{t+1}^* + b_3 \Delta I_t + b_4 (F_{t-1} \times L_{t-1}) + b_5 (F_t \times L_t^*) \tag{vi}$$

Let all additive terms, except $b_1 R_{t+1}^*$, be represented by Z_{t+1}. Thus, (vi) is abbreviated to (vii):

$$S_{t+1} = b_1 R_{t+1}^* + Z_{t+1} \tag{vii}$$

Lagging (vii) by one time period,

$$S_t = b_1 R_t^* + Z_t \tag{viii}$$

or

$$R_t^* = \frac{1}{b_1}(S_t - Z_t) \tag{ix}$$

and also,

$$R_{t-1}^* = \frac{1}{b_1}(S_{t-1} - Z_{t-1}) \tag{x}$$

Recalling (iv-b) and substituting R_{t-1}^* by (x) in the equation, the following equation is obtained:

$$R_t^* = \beta R_{t-1} + (1-\beta)\frac{1}{b_1}(S_{t-1} - Z_{t-1}) \tag{xi}$$

Using simplified coefficients, (xi) is rewritten:

$$R_t^* = B_1 R_{t-1} + B_2 S_{t-1} - B_3 Z_{t-1} \qquad \text{(xii)}$$

Using (xii) for the succeeding period,

$$R_{t+1}^* = B_1 R_t + B_2 S_t - B_3 Z_t \qquad \text{(xiii)}$$

Substituting (xiii) in (vii),

$$S_{t+1} = b_1 (B_1 R_t + B_2 S_t - B_3 Z_t) + Z_{t+1} \qquad \text{(xiv)}$$

The equation may be presented more explicity and in a general form as follows:

$$S_{t+1} = f(R_t, S_t, I_t^*, \Delta I_{t-1}, F_{t-2} \times L_{t-2}, F_{t-1} \times L_{t-1}^*, I_{t+1}^*, \Delta I_t, F_{t-1} \times L_{t-1}, F_t \times L_t^*) \qquad \text{(xv)}$$

The foregoing equation was simplified to obtain equation (3b) of the basic model by excluding some variables (e.g., I_t^*, ΔI_{t-1} and $F_{t-1} \times L_{t-1}^*$) which were believed to be highly correlated or to be adequately represented by other variables in the equation.

REFERENCES

[1] Foote, Richard J., Robert R. Williams and John A. Craven. *Quarterly Shorter-Term Price Forecasting Models Relating to Cash and Futures Quotations for Pork Bellies*, Economic Research Service, USDA, Technical Bulletin No. 1482, July 1973.

[2] Harlow, Arthur A. *Factors Affecting the Price and Supply of Hogs*, Agricultural Marketing Service, USDA, Technical Bulletin No. 1274, 1962.

[3] Mandansky, A. "On the Efficiency of Three-Stage Least-Squares Estimation," *Econometrica*, Volume 32, No. 1-2, January-April 1964, pp. 51-66.

[4] Myers, L. H., J. Havlicek, Jr. and P. L. Henderson. *Short-Term Price Structure of the Hog-Pork Sector of the United States*, Research Bulletin No. 855, Indiana Agricultural Experiment Station, Purdue University, Lafayette, Indiana, February 1970.

[5] Nerlove, Marc. *The Dynamics of Supply: Estimation of Farmers' Response to Price*, Baltimore: The John Hopkins Press, 1958.

[6] Theil, H. *Applied Economic Forecasting*, Amsterdam: North-Holland Publishing Co., 1966.

[7] U.S. Department of Agriculture, *Hogs and Pigs*, Statistical Reporting Service.

[8] ———. *Livestock and Meat Situation*, U.S. Department of Agriculture, Economic Research Service.

[9] ———. *Livestock and Meat Statistics*, Supplement to Statistical Bulletin No. 333, U.S. Department of Agriculture, Economic Research Service, June 1971.

[10] ———. *Poultry and Egg Situation*, U.S. Department of Agriculture, Economic Research Service.

[11] ———. *Price Spreads for Beef and Pork, Revised Series 1949-1969*, U.S. Department of Agriculture, Economic Research Service, May 1970, p. 24.

[12] ———. *Selected Statistical Series for Poultry and Eggs Through 1968, Revised January 1970*, U.S. Department of Agriculture, Economic Research Service, ERS 232.

[13] U.S. Department of Commerce. *Survey of Current Business*, Bureau of Economic Analysis.

[14] U.S. Department of Labor Statistics. *Employment and Earning Statistics of the United States*, Tables C-2, C-5, C-6, C-7, line 2011, Washington, D.C.

[15] Zellner, A. and H. Theil. "Three-State Least Squares: Simultaneous Estimation of Simultaneous Equations," *Econometrica*, Volume 30, pp. 54-78.