PREDICTION OF SHELL EGG PRICE–A SHORTRUN MODEL

Sujit K. Roy*

Variations in price occur frequently in the U. S. shell egg industry and, consequently, interest has been widespread among the traders and the producers with regard to such price variations. Quantitative models for the specific purpose of predicting shortrun egg price have so far been conspicuously absent. The basic objectives of this discussion are first, to present an econometric model to predict quarterly shell egg price and, second, to explain the casual factors which appear to affect price in the immediate future. Finally, alternative methods of estimating the predictive model have been compared with regard to their relative predictive ability.

Price at wholesale level is crucially important to the shell egg sector since the retail, as well as farm prices, are essentially determined with reference to the current wholesale price. This consititutes the reason for selecting the latter as the unit of investigation in the study. The wholesale egg market is characterized by base price quotations which originate in the central markets, such as Boston, New York City, Chicago, and Los Angeles. The New York City base price, as reported by the Urner Barry Publishing Company, is by far the most widely used series in the sector.

The total quantity of shell eggs available for final consumption at any given point in time is equal to aggregate production after allowing for the quantity of eggs broken for commercial use, the quantity of eggs used for hatching and the quantity of eggs in storage. Noncivilian purchases and net export of eggs are assumed to be only a negligible portion of total egg production. Since shell eggs are highly perishable, they must reach the final consumer market without considerable time lapse. Assuming a fairly stable demand structure in the shortrun, variation in shell egg price is, therefore, primarily due to variations in production and partially due to the shortrun fluctuations in hatching, breaking and storing activities. Under these specifications and assumptions, the following four functional relations were hypothesized to formulate the quarterly price prediction model:

$$\text{(i)} \ldots P_t = f(Y_t, H_t, B_t, S_t, P_{t-1}, D_1, D_2, D_3)$$

$$\text{(ii)} \ldots B_t = f(F_t, P_t, B_{t-1}, D_1, D_2, D_3)$$

$$\text{(iii)} \ldots Y_t = f(L_t, R_{t-1}, E_{t-1}, Y_{t-1}, D_1, D_2, D_3)$$

$$\text{(iv)} \ldots H_t = f(P_{t-1}/P_{t-5}, H_{t-1}, D_1, D_2, D_3)$$

where

P_{t-i} = simple average of the daily wholesale price per dozen of large extra fancy heavy grade eggs during the (t-i)-th quarter;[1]

Y_{t-i} = total production of eggs in millions during the (t-i)-th quarter;[2]

---

*Assistant professor of Agricultural Economics and Statistics, Texas Tech University.

This article is based on part of the results of a dissertation; (see Sujit K. Roy, "Econometric Models for Predicting Shortrun Egg Price," Ph.D. Thesis, The Pennsylvania State University, Department of Agricultural Economics and Rural Sociology, March 1969).

[1]Source: Urner Barry Publishing Company, Producer's Price-Current, New York, daily issues.

[2]Source: U. S. Department of Agriculture, Economic Research Service, *Selected Statistical Series for Poultry and Eggs through 1965*, Supplement to the "Poultry and Egg Situation," ERS 232, Revised May 1966, Tables 30, 31, 38, 40; U. S. Department of Agriculture, Economic Research Service, *Poultry and Egg Situation, Outlook Issues*, 1962 through 1968; also U. S. Department of Agriculture, Economic Research Service, *Poultry and Egg Situation*, five issues a year, April 1962 through April 1968.

H_{t-i} = total quantity of eggs used for hatching, in thousand cases, in the (t-i)-th quarter;[2]

B_{t-i} = total quantity of eggs broken commercially, in thousand cases, during the (t-i)-th quarter;[2]

S_{t-i} = shell eggs in cold storage, in thousand cases, on the first day of the (t-i)-th quarter;[2]

F_t = nonshell eggs in cold storage, in million pounds, on the first day of the t-th quarter;[2]

R_{t-i} = chicks placed for laying flock replacements, in millions, during the (t-i)-th quarter;[2]

E_{t-i} = eggs per 100 layers during the (t-i)-th period;[2]

L_t = number of layers, in millions, on farm on the first day of the t-th quarter;[2]

D_1 = 1 for 1st quarter (January through March), 0 otherwise;

D_2 = 1 for 2nd quarter (April through June), 0 otherwise;

D_3 = 1 for 3rd quarter (July through September), 0 otherwise;

i = 0, 1, 2, 3, ...

The first function representing the prediction equation for quarterly price consists of the four basic components of net supply of shell eggs. Current level of production (Y_t) and the quantity of shell eggs in storage at the end of the preceding period (S_t) would be expected to affect price (P_t) in the reverse direction. Variations in the quantity of eggs used for hatching (H_t) and the volume of eggs broken for commercial use (B_t) would presumably affect price directly.

The price function in the model contains three explanatory variables which relate to the current time period, i.e., Y_t, H_t and B_t. The variable S_t, however, may be considered as cold storage holding at the end of the preceding quarter. Thus, the content of the price function necessitated the formulation of functions to predict Y_t, H_t and B_t in order to make the price prediction model a 'closed' one.

Quantity of eggs broken during the current period (B_t) is presumably influenced by the demand for convenience food products. It was, however, assumed in the present study that the demand for such items is fairly stable in the shortrun and the variations in demand from one period to the next is more or less imperceptible. As the second function shows, the volume of eggs broken for commercial use (B_t) was hypothesized to be a function of the quantity of nonshell eggs at the end of the preceding period (F_t). For instance, if the quantity of nonshell eggs in storage is relatively large in the beginning of the period, less eggs may be expected to be broken during the quarter. Current level of shell egg price (P_t) as an explanatory variable may be expected to affect the dependent variable (B_t) in the opposite direction. When shell egg price is relatively low, an increased quantity of eggs may be moved through the breaking plants provided the demand situation in this alternative market is sufficiently strong.

Current level of production of shell eggs (Y_t), as hypothesized in the third function, is primarily dependent on the number of layers on farm at the beginning of the quarter (L_t), productivity of the layers or the average number of eggs per 100 layers (E_{t-1}), and the number of chicks placed for layer replacements during the preceding quarter (R_{t-1}). It was postulated that the latter variable (R_{t-1}) would affect production in the reverse direction because the preceding period's replacement chicks, which enter the laying flock during the current quarter, may lower the average productivity of the flock as young pullet's laying rate is generally lower than that of the matured ones.

An examination of the third function reveals that current production (Y_t) has been assumed to be independent of price during the same quarter (P_t). The producers may apparently adjust the level of current production in response to current price by altering the number of pullets entering the laying flock and/or the number of older layers to be removed from the farm. Such instantaneous adjustments within any given quarter, however, are quite difficult in practice. The number of pullets entering the laying flock during the current quarter would be determined at least four or five months ago when chicks were placed for laying flock replacements. The laying flock size cannot be enlarged instantaneously by adding additional pullets to the current flock simply because such pullets would not be available without prior plans. Furthermore, any significant number of older hens cannot be held back on the farm indefinitely because pullets attaining the laying age enter the flock and exert pressure on the existing plant capacity which cannot be increased in the shortrun. It is difficult for the producer to suddenly reduce or increase the number of hens to be sold since slaughter is usually finalized as a contract in advance. Thus, it is justifiable to assume that the producers' control over the adjustments in current production in response to variations in egg price during the same time period is insignificant.

The quantity of eggs used for hatching (H_t) is determined by the demand for broiler chicks and layer replacement chicks. The demand for broiler chicks, in its turn, is determined by the entire demand and supply structure of the broiler industry which, however, remains beyond the scope of the present study. The demand for layer replacement chicks, on the other hand, is dependent on the anticipations regarding the shell egg market in the future. It was assumed, as the fourth function shows, that the relative price index, P_{t-1}/P_{t-5}, would represent the current relative trend in shell egg price which may have considerable influence on the anticipations regarding the shell egg market.

The 'zero-one' variables, D_1, D_2, and D_3, were included in each of the functions to take into account part of the unexplained yet systematic quarterly variations in the dependent variables. Furthermore, each of the dependent variables in the above functions was assumed to be influenced by the lagged value of the same variable, i.e., P_{t-1}, B_{t-1}, Y_{t-1} and H_{t-1}. For example, if shell egg price during the preceding period remained at a relatively low level, current price may also tend to maintain a similar trend.

An examination of the structure of the model would reveal that the first two functions are a set of simultaneous equations. Two endogenous variables, P_t and B_t, appear in both functions. The other two dependent variables, Y_t and H_t, can be considered as exogenous variables for the subset of the first two functions. On the basis of these specifications, two-stage least squares (TSLS) estimates of the two over-identified simultaneous equations were derived. On the other hand, since the last two functions are exogenous in relation to the first two, ordinary least squares (OLS) estimates were obtained for the Y_t- and H_t- equations. The results, as presented below, reflect that in specific instances some of the initially hypothesized variables were eliminated in view of their statistical insignificance. The variables associated with the prime sign (') are in actual units, while the rest are expressed in logarithmic values of actual units. The estimates of the equations were based on 35 observations beginning with the first quarter of 1958.

MODEL I

(I-i): Two-stage Least Squares Estimates

$$P_t = \underset{(1.3874)}{8.5804} - \underset{(0.0209)}{0.0745} D_1 - \underset{(0.0221)}{0.0804} D_2 -$$

$$\underset{(0.3993)}{2.3048} Y_t + \underset{(0.1533)}{0.7080} H_t - \underset{(0.0414)}{0.0782} (B_t/B_{t-1}) +$$

$$\underset{(0.0967)}{0.1919} P_{t-1}$$

$$R^2 = 0.87,\ F_{(6,28)} = 29.03,\ SEE^3 = 0.0263$$

(I-ii): Two-stage Least Squares Estimates

$$B_t = \underset{(0.1739)}{4.3525} + \underset{(0.0379)}{0.2209} D_2 + \underset{(0.0352)}{0.1414} D_3 -$$

$$\underset{(0.0920)}{0.4841} F_t - \underset{(0.1977)}{1.0573} (P_t/P_{t-4})$$

$$R^2 = 0.85,\ F_{(4,30)} = 40.62,\ SEE = 0.0787$$

(I-iii): Ordinary Least Squares Estimates

$$Y'_t = \underset{(18018.0144)}{-96825.8316} + \underset{(249.3628)}{420.9953} D_1 +$$

$$\underset{(234.1354)}{1076.4857} D_2 + \underset{(0.1976)}{0.3458} Y'_{t-1} + \underset{(0.8791)}{1.7438} E'_{t-1}$$

$$- \underset{(4.2017)}{8.3967} R'_{t-1} + \underset{(7090.8521)}{39603.9389} L_t$$

$$R^2 = 0.91,\ F_{(6,28)} = 45.11,\ SEE = 309.6752$$

(I-iv): Ordinary Least Squares Estimates

$$H_t = \underset{(0.3224)}{0.7606} + \underset{(0.0207)}{0.0349} D_1 + \underset{(0.0160)}{0.0297} D_2 +$$

$$\underset{(0.0958)}{0.7719} H_{t-1} + \underset{(0.0934)}{0.6702} (H_{t-4}/H_{t-5})$$

$$R^2 = 0.89,\ F_{(4,30)} = 60.36,\ SEE = 0.0287$$

The numbers in parentheses immediately below the regression coefficients are the respective standard errors.

An important aspect of the estimate of the P_t-equation is the omission of S_t because of its statistical insignificance. It may be suggested that the quantity of shell eggs in storage on a particular day, i.e., the first day of the quarter, is only a negligible portion of the total market supply of eggs during the three-month period and, hence, has little impact on quarterly shell egg price. The signs related to all regression coefficients, excepting that associated with (B_t/B_{t-1}), confirm the hypotheses regarding the respective casual relations. The estimated coefficient of

[3]SEE is the abbreviation for standard error of estimate.

(B_t/B_{t-1}) is negative which is contrary to the expected sign. It may be observed, however, that the coefficient of this variable is not statistically as well-determined as the others. Most of the coefficients were tested to be significantly different from zero at the 1 or 5 percent level of significance. The fourth equation (I-iv) is admittedly rather "naive" in its structure, yet in terms of R^2 and other statistical tests it appears adequate as a prediction equation.

Alternative estimates of the first two equations, i.e., P_t- and B_t- equations, were obtained by utilizing ordinary least squares method, although the method is statistically inappropriate in deriving estimates of simultaneous equation model. However, such methods sometimes may yield predictions which are perhaps as accurate as others derived through more appropriate techniques.[4] The ordinary least squares estimates of the P_t- and B_t- equations are presented here under Model II. The estimates of Y_t- and H_t-equations in Model II are, however, the same as in Model I.

MODEL II

(II-i): Ordinary Least Square Estimates

$$P_t = \underset{(1.4108)}{8.8578} - \underset{(0.0207)}{0.0820}\,D_1 - \underset{(0.0212)}{0.0911}\,D_2 -$$

$$\underset{(0.4081)}{2.3643}\,Y_t + \underset{(0.1576)}{0.7121}\,H_t - \underset{(0.0366)}{0.0501}\,(B_t/B_{t-1})$$

$$+ \underset{(0.0978)}{0.1683}\,P_{t-1}$$

$$R^2 = 0.86,\ F_{(6,28)} = 27.18,\ \text{SEE } 0.0271$$

(II-ii): Ordinary Lease Squares Estimates

$$B_t = \underset{(0.1859)}{4.3545} + \underset{(0.0405)}{0.2227}\,D_2 + \underset{(0.0376)}{0.1436}\,D_3 -$$

$$\underset{(0.0983)}{0.4851}\,F_t - \underset{(0.2003)}{0.9270}\,(P_t/P_{t-4})$$

$$R^2 = 0.83,\ F_{(4,30)} = 34.64,\ \text{SEE } 0.0842$$

A comparison of the OLS and TSLS estimates of the P_t-equation would disclose that the constant term and the coefficients of D_1, D_2, Y_t and H_t are over-estimated by the former. On the other hand, the parameters related to (B_t/B_{t-1}) and P_{t-1} are under-estimated in the OLS model. The differences in the magnitude of the corresponding OLS and TSLS coefficients in the P_t- and B_t- equations are, however, not very substantial. The TSLS equations, i.e., (II-i) and (II-ii), appear to have a slight edge over the corresponding OLS equations in terms of R^2 and the standard error of estimate.

The alternative price prediction models were compared on the basis of (i) the relative accuracy of prediction of *direction* of change and (ii) the *magnitude* of prediction errors. These two criteria supplement each other in testing the accuracy of prediction. When two models are compared, it is sometimes possible that one of them is inferior to the other in predicting direction of change even though it is relatively better in terms of the average of the absolute deviations of the predicted estimates from the actual values. Observed prices of 16 quarters, beginning with the first quarter of 1964, were compared with the corresponding prices predicted by the two models. With regard to prediction of direction of change, each of the models correctly specified the direction for 13 quarters. More specifically, two of the actual positive changes were predicted by both models as negative changes and one of the realized negative changes was specified by each model as a positive change. Incidentally, each of the models incorrectly predicted the direction of change for the same observations.

The closeness of the magnitude of the predicted estimates to that of the observed values is the other major criterion of predictive accuracy. One suggested method involves the comparison of U-coefficient[5] which is defined as follows:

$$U = \sqrt{\Sigma u_{it}^2} \div \sqrt{\Sigma A_{it}^2}$$

where A_{it} is the observed value of the i-th variable to be predicted and u_{it} is the deviation of the predicted value from the corresponding A_{it} value. The U-coefficients of Model I and Model II were 0.02046 and 0.02619, respectively. Although none of the models can lay claim to superiority over the other in terms of prediction of direction of change, the difference between the U-coefficients offers a legitimate ground to favor the TSLS model (i.e., Model-II).

The foregoing prediction model was formulated also for the purpose of analysis of the structural relationships. The estimated structural parameters

[4]In fact, the ordinary least squares estimates of a monthly shell egg price prediction model formulated elsewhere by the author yielded predictions superior to those of the comparable two-stage least squares model. The monthly prediction model, however, was in part a simultaneous equation system; see Roy, op. cit. pp. 105-110.

[5]H. Theil, *Applied Economic Forecasting*, pp. 26-29, North-Holland Publishing Company, Amsterdam, 1966.

were found to be generally consistent with the expected causal relationships among the variables. It is hoped that the postulated model would provide not only a predictive device, but also a somewhat deeper insight into the causes of quarterly variations in shell egg price.

Lack of data often leads to the omission of important variables and, hence, becomes a major limitation of econometric models. Several relevant variables could not be included in the present model because appropriate time series data were not available. For instance, prices of eggs used for hatching and prices paid for eggs broken commercially might have improved the H_t- and B_t-equations, respectively. Similarly, data on culling rate would have been useful in improving the predictive accuracy of the equation representing shell egg production. The structure of the model would become more comprehensive and complete if the specified data were available. Efforts should be made in developing time series data on such variables.

Future work may be directed toward the use of alternative estimation techniques, such as the three-stage least squares. The relative performance of the predictive models estimated by the two-stage and three-stage least squares would be of considerable interest to the econometricians.

The predictive accuracy of the present model, as that of any other econometric study using time series data, is susceptible to the changes in the structural relations over time. Thus, an extended forecast too far beyond the sample time period may yield unreliable and poor results. A regular updating of the data and re-estimation of the model would be highly desirable since the shell egg sector has been undergoing rapid structural changes in recent years.