Simultaneous Prediction of Price Magnitude and Direction

by

L. R. Cloman, J. E. Epperson, N. R. French, and T. T. Fu

SIMULTANEOUS PREDICTION OF PRICE MAGNITUDE AND DIRECTION

L.R. Cloman, J.E. Epperson, N.R. French and T.T. Fu*

Accurate price forecasts are an essential tool for agricultural decisionmakers. In today's complex economic environment the quality of forecasts becomes even more crucial to the survival of agricultural firms. In this paper techniques are examined with respect to the potential for improving forecasts relative to traditional economic forecasting methods.

In earlier research Menkhaus and Adams consider the problem faced by feeder calf producers in whether to sell Spring calves in the Fall or to hold the animals until the following Spring. They emphasize that price movements are composed of magnitude and direction components and argue that price forecasts can be improved by considering both components. They present a forecasting approach in which the direction of price movements are first predicted using discriminant analysis and then these predictions are incorporated into the prediction of price magnitude.

This paper extends the idea of the potential importance of considering both direction and magnitude in making accurate price forecasts. However, rather than estimating direction and magnitude separately, a model is presented which incorporates the simultaneous nature of changes in price magnitude and direction.

The methodology for this paper follows that used by Menkhaus and Adams and compares it to a simultaneous Probit-OLS process. The analysis is presented as follows: Discriminant analysis is used to predict the direction of movement between Fall feeder calf prices and Spring yearling prices. The discriminant analysis approach is then compared to regression analysis with respect to accuracy in predicting price direction. The usefulness of incorporating the direction of price movement as a variable in a model to predict price magnitude is examined. Finally, the discriminant and regression analyses are compared with a two equation model that simultaneously determines direction and magnitude of price movements. One equation in the system forecasts price magnitude, the other determines price direction. A two-stage Probit-OLS estimation method is employed.

Models

The following discriminant model was specified as in Menkhaus and Adams work and predicts the direction of movement between October_t feeder calf prices and March_{t+1} yearling prices in September_t:

*Graduate research assistant, professor, graduate research assistant, and post doctoral associate, respectively, Division of Agricultural Economics, University of Georgia.

(1) $PREDIR_{t+1} = f(SEPTYL_t, MSDIR_t, SEPTSL_t, SEPTC_t, PERCENT_t)$

where

$PREDIR_{t+1}$ = discriminant score showing direction of price movement (up or down) between October feeder calf prices in year t and March yearling prices in year t+1;
$SEPTYL_t$ = September yearling price (deflated) in year t (cwt.);
$MSDIR_t$ = direction of price movement between March yearling prices and September yearling prices in year t where 1 designates upward price movements and 0 signals downward price movements;
$SEPTSL_t$ = September slaughter steer prices (deflated) in year t (cwt);
$SEPTC_t$ = September corn price (deflated) in year t (cents per bu.);
$PERCENT_t$ = percent of January 1 cattle and calf inventory slaughtered in year t.

A regression model was then estimated using the same independent variables as in the discriminant model:

(2) $PDIFF_{t+1} = f(SEPTYL_t, MSDIR_t, SEPTSL_t, SEPTC_t, PERCENT_t)$

where PDIFF is the March yearling price in year t+1 (deflated) minus the October feeder calf price in year t (deflated). The other variables are as previously defined.

Next we estimated regression models in which a variable indicating the direction of price movement was included as an exogenous variable. The model was first estimated using the actual directional movements (DIR), and then re-estimated using the predicted directional movements (PREDIR) obtained through discriminant analysis:

(3) $PDIFF_{t+1} = f(SEPTYL_t, MSDIR_t, SEPTSL_t, SEPTC_t, PERCENT_t, DIR_{t+1})$

(3a) $PDIFF_{t+1} = f(SEPTYL_t, MSDIR_t, SEPTSL_t, SEPTC_t, PERCENT_t, PREDIR_{t+1})$

where DIR is the actual direction of price movement between the March yearling price in year t+1 (deflated) minus the October feeder calf price in year t (deflated), such that 1 indicates up and 0 down, and PREDIR is the discriminant predicted direction of price movement between $March_{t+1}$ yearling price and $October_t$ yearling price (1 indicates up and 0 down). The remaining variables are as previously defined.

Finally, equations (4) and (5) were estimated simultaneously using a two-stage Probit-OLS estimation procedure:

(4) $PDIFF_{t+1} = f(SEPTYL_t, MSDIR_t, SEPTSL_t, PERCENT_t, DIR_{t+1})$

(5) $DIR_{t+1} = f(SEPTYL_t, MSDIR_t, SEPTSL_t, SEPTC_t, PERCENT_t)$

where DIR in equation (5) is a dichotomous variable representing the difference in the $March_{t+1}$ yearling price and $October_t$ price (1 indicates a price rise between $October_t$ and $March_{t+1}$ and 0 signals a price decrease), DIR in equation (4) is an index to be defined later (Maddala, p. 245), and the remaining variables are as previously specified.

Identification of the two-equation system required that equation (4) have one less exogenous variable than equation (5) (Maddala). Therefore, based on the predictive power of the model SEPTC was omitted from equation (4).

The methodology for estimating equations (4) and (5) follows Heckman's two-stage simultaneous estimation procedure (Maddala, p. 244) and can best be described using:

(6) $Y_1 = \gamma_1 Y_2^* + \beta_1' X_1 + u_1$

(7) $Y_2^* = \beta_2' X_2 + u_2$

where

Y_1 = difference in $March_{t+1}$ and $October_t$ yearling prices;

Y_2^* = a variable observed only as a dichotomous variable such that if $Y_2 > 0$, then $Y_2^* = 1$, otherwise $Y_2^* = 0$ (i.e. 1 indicates prices rose between $October_t$ and $March_{t+1}$ and 0 signals a price decline);

X_1 = vector of explanatory variables for Y_1;

X_2 = vector of explanatory variables for Y_2^*;

γ_1 = scalar coefficient;

β_1, β_2 = vectors of coefficients;

u_1, u_2 = random error terms with bivariate normal distribution.

The reduced form equations for this system are:

(8) $Y_1 = \Pi_1 X + v_1$

(9) $Y_2^* = \Pi_2 X + v_2$

where

X = combined exogenous variables in X_1 and X_2;

Π_1, Π_2 = vectors of reduced form coefficients;

v_1, v_2 = random error terms with bivariate normal distribution.

Equation (7) forecasts the direction of price movement. This predicted directional variable is endogenous to the price magnitude equation (6).

Heckman's two-stage estimation process provides consistent estimates for equations (6) and (7). It should be noted that in this particular model the reduced form equation (9) and the structural form equation (7) are the same, thus $\hat{\beta}_2 = \hat{\Pi}_2$.[1]

The estimation process begins by obtaining an estimate of β_2. However, since Y_2^* is observed only as a dichotomous variable, we can only estimate β_2/σ_2, where $\sigma_2^2 = var(v_2)$. Hence, equation (7) is estimated by probit Maximum Likelihood (ML) in the form

(10) $$Y_2^{**} = \frac{Y_2^*}{\sigma_2} = \frac{\beta_2}{\sigma_2} X_2 + \frac{u_2}{\sigma_2} .$$

Equation (6) is now written as

$$(11) \quad Y_1 = \gamma_1 \sigma_2 Y_2^{**} + \beta_1' X_1 + u_1.$$

In the second step of the estimation process, equation (11) is estimated by OLS after the predicted index value of Y_2^{**}, $\beta_2/\sigma_2 X_2$, is substituted into the equation (Maddala, p. 245). Therefore, the estimable parameters in this model are $\gamma_1 \sigma_2$, β_1, β_2/σ_2, σ_1, and σ_{12}/σ_2.

The asymptotic covariance matrix for equation (10) is that provided by probit ML estimation of the equation. However, the asymptotic covariance matrix for equation (11) must be derived using a procedure similar to the one used by Amemiya for a simultaneous tobit model.

The covariance matrix of the two-stage estimates of equation (11) can be estimated by

$$(12) \quad \mathrm{Var}(\hat{\alpha}_1) = c(H'X'XH)^{-1} + (\gamma_1 \sigma_2)^2 (H'X'XH)^{-1} H'X'XV_0X'XH(H'X'XH)^{-1}$$

where

$\alpha_1' = (\gamma_1 \sigma_2, \beta_1')$;
$H = (\beta_2, J_1)$;
J_1 = matrix consisting of 1's and 0's so that $XJ_1 = X_1$;
$c = \sigma_1 - 2 \gamma_1 \sigma_{12}$;
V_0 = covariance matrix of the probit ML estimate of β_2 and

$$\mathrm{Cov}(v_1, v_2) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}.$$

Data

The models were estimated over the period 1926 to 1970. The years 1971 through 1981 were reserved for post sample prediction testing. Cattle statistics were obtained from selected issues of the U.S.D.A. publications _Livestock and Meat Statistics_ and _Livestock and Meat Situation_. Corn prices were obtained from the _Commodity Year Book_. Consumer price index data are from the Department of Commerce _Survey of Current Business_.

To calculate the values of the variables (PDIFF) and (DIR) the weighted average of all weights and grades of March feeder calf prices at Kansas City in year t+1 and the October price of good and choice feeder steer calves at Kansas City in year t were compared. September yearling prices per cwt. (deflated), SEPTYL, are the weighted average of all weights and grades of September feeder steers at Kansas City. The movement in yearling prices (MSDIR) were calculated using March and September yearling steer prices at Kansas City in year t. The variable (SEPTC) is the September 15 average price (deflated) of corn in cents per bu. received by U.S. farmers. The deflated September slaughter price (SEPTSL) represents the average cost per 100 lbs. of sales out of first hands for choice slaughter steers at Chicago for 1926 through 1949. Over the period 1950 to 1981 this price represents the price of choice slaughter steers at Omaha, 900-1100 lbs. The inventory used to calculate PERCENT is the January 1 cattle and calf inventory.

Slaughter used in calculation of PERCENT represents the sum of cattle and calf commercial slaughter.

Results

Estimation results of the discriminant, regression, and simultaneous Probit-OLS models are delineated in Table 1. The primary test, however, of any forecasting model is its predictive power. The forecasting ability of these models are evaluated according to how well price magnitude and direction were predicted for the in-sample period of 1926 to 1970 and for the post sample period of 1971 to 1981.

For the in-sample period the discriminant function, equation (1), correctly classified 22 of the 25 downward price movements and 13 of the 20 upward price movements for an overall 78 percent correct classification. This compares to an 80 percent correct classification obtained by Menkhaus and Adams.[2]

Because the dependent variable of the regression equation (2) was specified as a difference, it too can indicate price direction in sign. Both studies found that the regression model correctly indicated 73 percent of the price movements. Price magnitude was evaluated using Theil's U_2 coefficient. For the in-sample period 1926 to 1970, the regression equation (2) without a turning point variable yielded a U_2 of 0.59 (Menkhaus and Adams' U_2 was 0.51) which indicated that the model was better than a naive no-change model in predicting price magnitude. A value of 1 for the U_2 represents equivalency with a naive no-change extrapolation model.

The potential usefulness of incorporating the direction of price movement as a variable in a prediction model was evaluated using equation (3). In this case, the actual directional movements were included exogenously as a dummy variable. As expected, the accuracy of directional indications provided by the regression (3) improved to 87 percent (Menkhaus and Adams, 96 percent). The addition of the exogenous binary directional variable reduced Theil's U_2 to 0.41 (Menkhaus and Adams' U_2 was 0.36).

As noted above, the discriminant function correctly classified 78 percent of the directional movements. When these predicted price direction movements are used as an exogenous variable in the regression model (3a), the predicted price difference indicates price direction correctly 76 percent of the time. This compares to 73 percent in the regression model (2) without an exogenous directional variable. Theil's U_2 for this regression (3a) was 0.58 compared to a 0.59 for regression model (2).

Our alternative two-equation simultaneous system produced the following results. The Probit equation (5) correctly identified 78 percent of the price movements, the regression equation (4) correctly identified 73 percent of the directional movements in price, and the U_2 statistic was 0.59.

The most crucial test of these models, however, is how well they predict in the post sample period from 1971 to 1981. These validation results are presented in Table 2 and with the exception of slight differences in the

Table 1. Estimation Results for the Discriminant, Regression, and Two-Stage Probit-OLS Models, 1926-1970

Variable	Discriminant Coefficients		Regression Coefficients			Two-Stage Probit-OLS Coefficients	
	Standardized	Unstandardized	W/O Direction Variable	W/Actual Direction Variable	W/Discriminant Predicted Direction Variable	Price Magnitude Equation	Price Direction Equation
	(1)[a]		(2)	(3)	(3a)	(4)	(5)
Constant	---	-9.717904	-8.4693 (-2.317)[b]	-0.7440 (-0.261)	-9.9481 (-2.456)	7.4890 (6.035)	-10.0140 (-3.068)
SEPTYLt	-0.92884	-0.13338	-0.5229 (-4.494)	-0.3930 (-4.652)	-0.5002 (-4.178)	-0.3838 (-0.030)	-0.0876 (-1.068)
MSDIRt	0.27620	0.64397	1.9212 (2.301)	1.2939 (2.170)	1.8648 (2.219)	0.9144 (4.082)	0.6337 (1.172)
SEPTSLt	0.34396	0.04843	0.1773 (1.364)	0.1301 (1.417)	0.1566 (1.181)	0.1798 (0.134)	-0.0016 (-0.018)
SEPTCt	0.44589	0.00853	0.0161 (1.868)	0.0078 (1.259)	0.0212 (2.022)		0.0102 (1.421)
PERCENTt	0.82305	27.32489	31.0920 (2.925)	4.4758 (0.522)	33.9120 (3.039)	-13.4403 (-82.754)	28.0320 (2.945)
DIRt+1				3.9160 (6.370)		1.5886 (0.042)	
PREDIRt+1					-0.8918 (-0.860)		
			R^2 = 0.63	R^2 = 0.82	R^2 = 0.64	R^2 = 0.64	Log-Likelihood = -19.58
			F = 13.628	F = 29.65	F = 11.40	F = 13.63	
			U2 = 0.59	U2 = 0.41	U2 = 0.58	U2 = 0.59	

[a]Corresponding equation numbers.

[b]t-statistics.

Table 2. Predictions Using Discriminant, Regression, and Two-Stage Probit OLS Models, 1971-1981

Year	Discriminant		Regression				Two-Stage Probit-OLS	
	Actual	Predicted (1)[a]	Actual (cwt)	Predicted W/O Direction Variable (cwt) (2)	Predicted W/Actual Direction Variable (cwt) (3)	Predicted W/Discriminant Predicted Direction Variable (cwt) (3a)	Predicted Price Magnitude (cwt) (4)	Predicted Price Direction (5)
1971	0	0	-6.05	-4.67429	-4.95366	-4.63905	-4.67477	.06550
1972	0	0	-3.45	-5.09336	-5.19838	-5.05342	-5.09386	.04360
1973	0	0	-1.51	-5.83253	-2.04455	-5.38883	-5.83306	.09180
1974	0	0	-15.12	-8.75409	-7.71551	-7.99072	-8.75462	.00570
1975	0	0	-3.35	-1.34633	-2.16461	-1.11462	-1.34679	.11700
1976	1	1	0.41	0.62670	2.47138	-0.11210*	0.62618	.51600
1977	0	0	-2.52	-1.88337	-3.35122	-1.81991	-1.88383	.37070
1978	1	1	1.79	-0.09832*	1.48790	-0.85132*	-0.09880*	.73240
1979	1	0*	0.97	-3.97560*	-1.34360*	-3.66787*	-3.97614*	.38970**
1980	0	0	-13.81	-9.17867	-8.18951	-8.85351	-9.17921	.00520
1981	0	0	-10.14	-6.77767	-6.18550	-6.73097	-6.77818	.00870
				U2 = 0.46	U2 = 0.45	U2 = 0.49	U2 = 0.46	

*Indicates an incorrect classification.

**Indicates an incorrect classification using a 50-50 criterion.

[a]Numbers in parentheses are the corresponding equation numbers.

U_2 statistics are the same as those of Menkhaus and Adams. The discriminant function (1) properly classified 10 out of the 11 cases. The regression (2) without a directional variable correctly indicated 9 of 11 directional moves in price and yielded a U_2 of 0.46. When the actual direction movements were included in the regression equation (3) via a binary variable, price direction was correctly indicated in 10 of 11 cases and the U_2 improved to 0.45. However, when the predicted discriminant directional movements were included exogenously in the model (3a), price direction was correctly identified in only 8 of the cases and the U_2 increased to 0.49. The probit equation (5) of the simultaneous system yielded correct directional indications in 10 of 11 cases. The price magnitude difference equation (4) also correctly identified 9 of the 11 directionals. The U_2 statistic for this equation was 0.46. An interesting result was the fact that the simultaneous system price magnitude difference forecasts were essentially identical to the forecasts of the regression equation (2) without a directional variable.

Conclusions

The purpose of this paper was to demonstrate a two-stage Probit-OLS method which may be useful in improving forecasting results. This method was compared with the methodology and estimation results developed by Menkhaus and Adams.

Tomek and Robinson have pointed out that price has two dimensions, the first is magnitude and the second is direction. It follows that the ability to forecast accurately hinges largely on an ability to reflect both dimensions of price in a forecasting approach.

Menkhaus and Adams used discriminant analysis to predict directional movements and included these predictions in a price forecasting equation in an effort to enhance the accuracy of the price forecast. The two-stage simultaneous method developed in this paper was used to capture the simultaneous nature of changes in price magnitude and direction.

The idea of forecasting price direction and incorporating it into a price magnitude forecast is both theoretically and intuitively appealing. When the actual direction of movement in price was used as a binary variable in the regression model, the price magnitude difference forecast was improved which supports the idea that both components of price must be considered. However, in the in-sample period when the predicted discriminant directionals were used in the regression equation, the U_2 statistic was 0.58 which was slightly better than the U_2 of 0.59 obtained from the regression equation without a turning point variable but worse than the U_2 of 0.41 obtained from the regression with the actual directional variable.

In comparison, the two-stage Probit-OLS procedure yielded a U_2 of 0.59 during the in-sample period which was the same for the regression without a turning point variable but was slightly worse than for the regression with predicted discriminant directionals, and worse than the regression with the actual directional variable.

In the post sample period the regression with actual turning points included as a variable, again performed the best with a U_2 of 0.45. The regression equation without a directional variable was slightly worse with a U_2 of 0.46 which again was identical to the Probit-OLS procedure result. The equation with the discriminant predicted directional variable performed the worst during the 1971-1981 period.

These results from the different models suggest that refinement of the directional equations may improve the forecasting results. In this particular application the Probit-OLS estimation procedure did not improve forecasts. However, these results would not likely hold for all economic applications and consequently the two-stage Probit-OLS method warrants consideration as a potentially useful forecasting method.

FOOTNOTES

[1]This is because Y_1 does not appear in equation (7).
[2]The difference in results between the two studies may be due to a slight difference in calculation of the PERCENT variable. Menkhaus and Adams used an average of cattle and calf slaughter for the first nine months of each year and multiplied this average by twelve whereas this study used the actual cattle and calf slaughter for each year.

References

Amemiya, T. "The Estimation of a Simultaneous Equation Tobit Model." International Economic Review 20 (1979):169-81.

Commodity Research Bureau, Inc. Commodity Year Book, New York. Selected Years.

Heckman, James J. "Dummy Endogenous Variables in a Simultaneous Equation System." Econometrica 46 (1978):931-959.

Maddala, G.S. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge University Press. 1983.

Menkhaus, D.J. and R.M. Adams. "Forecasting Price Movements: An Application of Discriminant Analysis." Western Journal of Agricultural Economics. 6 (1981):229-238.

Tomek, W.G. and K.L. Robinson. Agricultural Product Prices. Cornell University Press. 1972.

U.S. Department of Agriculture. Agricultural Prices. Economics, Statistics and Cooperatives Service. Washington, D.C., selected issues.

U.S. Department of Agriculture. Livestock and Meat Situation. Economics, Statistics and Cooperatives Service, Washington, D.C., selected issues.

U.S. Department of Agriculture. Livestock and Meat Statistics. Stat. Bul. No. 522, Ag. Marketing Service, Washington, D.C. selected issues.

U.S. Department of Commerce. Survey of Current Business, selected issues.