CONFIDENCE INTERVALS FOR CORN PRICE AND UTILIZATION FORECASTS

By Lloyd D. Teigen and Thomas M. Bell[1]

In recent forecasting activity in ERS, researchers have used alternative scenarios as a means of bounding forecasting errors. One scenario is treated as the "most likely" occurrence. Two others present analysis of consequences when production is at its estimated upper and lower limits. Increasingly, we have been asked to assign probabilities of occurrence to each of these production ranges and to the corresponding ranges of all the other forecasted variables.

- An approximate variance of forecast is derived based on the structural coefficients and the variance around the structural equations. For the corn model, standard error of price was estimated to be $0.23 per bushel when neither acreage nor yield are known and $0.11 per bushel when production is known.
- Keywords: Corn, forecast, variance, confidence intervals.

In this article we estimate a variance of forecasted price and utilization levels for corn which can be used to estimate the probability associated with the ranges. We base the estimate on the variance around the structural equations of an econometric model of the corn sector used as part of the ERS forecasting process.

In particular, the forecast error for corn prices is 23 cents per bushel before planted acreage is known and 11 cents after the harvest. In applying these standard errors of estimate to official USDA forecasts, it must be assumed that the process of review and adjustment by commodity specialists does not increase variance.

The theory of forecasting variance is well known in the econometric literature. Goldberger and others (2) studied the variance of forecast when all exogenous variables are not random.[2] Feldstein (1) extended this to the situation in which the exogenous variables were subject to error. Schmidt (6) examined the variance of forecasts from a dynamic econometric model. The forecast variance estimated here uses a rather simple adjustment process to bypass the large-scale matrix computations required by the exact variance formulae.[3]

First, the variance estimator will be derived. After a brief statement of the empirical form of the structural model, the reduced-form variances will be evaluated. A set of point estimate forecasts for the 1978 crop year and their corresponding 95-percent confidence intervals will be presented, as will some implications for forecast analysis. The explicit structure of the econometric model of the corn sector used here appears as an appendix.

DERIVATION OF VARIANCE ESTIMATOR

This corn sector model consists of six equations which determine six endogenous variables. Food, feed, and export demand are functions of corn price. Ending commercial stocks are related to corn price and production. Price is determined by the identity which equates supply with utilization; production is a function of variables which are exogenous to the model. Each equation contains variables determined outside this system and also random disturbances (except for the identity).

Price and utilization levels for corn will be forecast by the model, given an estimate of the current corn supply. The variances of these forecasts are estimated from the variance of the corn supply and the variance of the random disturbance in each equation.

In this process we will assume that the coefficients in the equations are now parameters rather than estimates and that the intercept contains all exogenous variables and the disturbances in each equation. Finally, we assume that the disturbances are independent across equations.

In this formulation we can express the corn model as follows:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & b1 \\ 0 & 1 & 0 & 0 & 0 & b2 \\ 0 & 0 & 1 & 0 & 0 & b3 \\ 0 & 0 & 0 & 1 & -a & b4 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & -1 & -1 & -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} FD \\ FO \\ CX \\ CS \\ QC \\ PC \end{pmatrix} = \begin{pmatrix} U1 \\ U2 \\ U3 \\ U4 \\ U5 \\ U6 \end{pmatrix} \text{, or}$$

$Ay = U$, in matrix notation

The U_i terms contain both the predetermined variables and the disturbance terms in the equations. Sym-

[2] Italicized numbers in parentheses refer to items in citations at the end of this article.

[3] The exact variance of forecast in this six-equation system requires manipulation of matrices of dimension (6x84), (84x120), and (120x120), in addition to the (6x6) matrices used in this article.

bols for the six endogenous variables are defined in the appendix. The covariance matrix, S, for the U vector under the assumption of independence is a diagonal matrix whose nonzero elements are the squared structural standard errors of estimate. The structural variance of U_6 in the identity is assumed to be zero. The compact form for the solution of the model is:

$$y = A^{-1} = U. [a]$$

The covariance matrix for this solution is:

$$A^{-1}S(A^{-1})'$$

under the stated assumptions. Thus the expanded form for the variance of corn price is:

$$VAR(PC) = \left(\Sigma_1^4 \sigma_i^2 + (1+a)^2 \sigma_5^2\right)/(\Sigma b_i)^2$$

Similar expressions hold for the other variables. If the disturbances in the demand equations were positively correlated with the total supply, this estimate of price variance would be biased upward since the covariance terms would be subtracted from this equation. If this upward bias occurred, our variance estimates would err by predicting a larger variance for price and smaller variances for utilization than might otherwise be true.

These variance estimates assume perfect knowledge of the system parameters. The usually applied variance of forecast definitions in the single equation case allow for variation due to the parameter estimates (for example, see 3 p. 375):

$$\sigma_F^2 = SEE^2 \left(1 + 1/n + (x-\bar{x})'(x'x)^{-1}(x-\bar{x})\right)$$

When there is only one independent variable, or when the X's are orthogonal in the multivariate case, the order of magnitude of this correction is approximately a factor of between $(n+k)/n$, and $(n+2k)/n$. Here k is the number of estimated parameters including intercept in the equation and n is sample size. As

---

[a] An expanded version of the reduced form of this system is the following:

$$FD, FO, CS = U_i - \frac{b_i}{\Sigma b_i}\left(\Sigma_1^4 U_i - (1+a)U_5\right), i = 1,2,3$$

$$CS = U_4 - a U_5 - \frac{b_i}{\Sigma b_i}\left(\Sigma_1^4 U_i - (1+a) U_5\right)$$

$$QC = U_5$$

$$PC = \left(\Sigma_1^4 U_i - (1+a) U_5\right)/\Sigma b_i$$

the X's depart from orthogonality, this factor increases with the correlation among the explanatory variables.

Thus, in a single-equation model, the standard error of forecast inflates the standard error of estimate by between $(1+k/2n)$ and $(1+k/n)$ to account for the variation due to coefficient estimation. Intuition suggests that an analog to this correction in the context of a system of estimates would define k as the average number of parameters per equation and n as the average sample size of each estimated equation.

ESTIMATED STANDARD ERRORS FOR THE ENDOGENOUS VARIABLES

Our empirical analysis used the estimated structural coefficients and standard errors of estimate to obtain the standard errors associated with the model solution. These specific estimates assume that the intercepts were known and the variance of each disturbance was that estimated with the structural equation. To generalize this approach, one could add the variance due to random exogenous variables in the system to the variance of the structural disturbance.

Table 1 presents the estimated standard errors of the endogenous variables in the corn model which correspond to three levels of production uncertainty. The standard errors presented in the first column come directly from the estimated equations. Those in the other three were derived from them. The first estimate of the reduced-form standard errors corresponds roughly to some point prior to the planting intentions report—at a time in which the standard error of the production estimate is 350 million bushels (an estimate used in the monthly Update of Food and Agriculture Outlook).[5] The second estimate is made after the planted acreage has been determined, and the only variation will be due to yield. The third estimate occurs after official production estimates have been made—so that the variance of production (from the standpoint of this model) is zero.

Thus, a March estimate of corn price for the next October-September year would have a standard error of about 23 cents per bushel, using a root mean square error of 350 million bushels for production. An October estimate of the upcoming crop year corn price would have a standard error of about 11 cents per bushel. Between these two estimates, the standard error of commercial exports would be reduced from 94 to 53 million bushels, and feed demand error would drop from 188 million bushels to 86 million bushels.

None of these absolute figures should be used to credit or discredit this model or the modeling system of which it is a part. However, they do indicate the inherent error tolerances in the estimated structure.

---

[5] ERS memorandum to Director of Economic Policy Analysis and Budget, regularly issued.

Table 1—Standard errors of corn price and utilization

Variable	Units	Structural standard error	Reduced from standard error		
			Preplanting	Midsummer	Postharvest
Price	Dol./bu.	≈0	.232	.148	.114
Production	Mil bu.		[1]350	[2]162.5	0
Feed demand	do.	148.2	188.1	115.7	85.69
Food demand	do.	13.39	13.98	13.53	13.41
Commercial exports	do.	42.67	93.99	64.16	53.10
Ending commercial stocks[3]	do.	83.95	125.1	87.52	73.90
Ending Government stocks	do.		427.3	227.5	25.7

[1] This corresponds to the root mean square error used in the monthly update of the Food and Agricultural Outlook. To estimate the production variance implied by the econometric model, square the yield times the standard error from the acreage equation and add the result to the square of acreage times the yield standard error. [2] All variation is due to yield uncertainty, since acreage is assumed known. [3] The price effect on next year's acreage was included with the price response in this equation, assuming soybean price is $6.40 per bushel.

A 1978 FORECAST AND ITS CONFIDENCE INTERVAL

We now illustrate the use of these standard errors of forecast. The reduced-form solution of the model will be presented and graphically illustrated using basic data from an update to the Food and Agricultural Outlook. This forecast represents exogenous information that was available May 1, 1977. Because we focus only on the changes in forecast variance through the year, the exogenous variables were fixed throughout the analysis. Thus the point estimates will not change among the alternatives.

Table 2 presents the food, feed, exports, and stock demand equations as functions only of the endogenous price and production variables. The effects of the exogenous variables were collapsed into the intercepts, using the values given in the table. The price slope of the demand for commercial stocks was adjusted to reflect the price response of planted acres. Total demand is the sum of the individual demand equations plus policy exports of 200 million bushels.

Table 2—Exogenous variables and simplified demand equations

Independent variable	Intercept	Production slope	Price slope	Exogenous variable and value assumed					
Food use	448.36	0	-19.649	YPD 688					
Feed use	5,601.57	0	-827.454	PM 10.0	PL 1.87	LO 1.029			
Commercial exports	2,276.53	0	-383.195	PS 6.40	SEEC 16,801	SX 21,221	AUX 102,500	PLX_{t-1} 1.00	XIX 1.72
Commercial stocks	1,201.06	.1049	[1]-317.73	AP 83.9	GS 0				
Total demand	9,727.51	0	-1,548.03	GX 200	GS 0				
Total supply	7,068	0	0	[2]QC 6,219	CS_{t-1} 849				

[1] This coefficient reflects the response of next year's acreage to current corn price, assuming the soybean price is $6.40 per bushel. [2] The acreage harvested is 71.9 million with 86.5 bushels per harvested acre.

Equating total demand with production plus beginning stocks, we can solve for the market clearing price. In this case, the equilibrium price would be $1.72 per bushel. However, the loan rate for corn is $1.75, so that either Government purchases for inventory or additional policy exports are needed to raise the corn price to $1.75. A 46.5-million bushel Government purchase would raise the price the necessary 3 cents. The variance of the Government stock estimate is the sum of the variance of production and the variance of each demand estimate. From table 1 the preplanting standard error of Government stocks is calculated as 427 million bushels, so that the estimated purchase is not statistically different from zero.

The confidence interval for each endogenous variable is calculated from its standard error by multiplying it by the appropriate value of the t-statistic inflated by the $(1+k/n)$ factor which accounts for parameter variation. The k/n correction factor allows the predetermined variables to be up to two standard deviations from their sample means. The result is a conservative estimate of the probability that the interval will cover the actual value. To learn what variability of price and utilization is due to the variables we collapsed into the intercepts, sensitivity analysis must be performed. Upper and lower bounds must be set for the exogenous variables and the system solved again to obtain a new set of price and utilization levels. The maximum should be treated as one scenario and the minimum, another.

The 95-percent confidence interval for the price forecast, as an example, is $P \pm S_p t.05 \, (1+k/n)$. With an average of 5 parameters, 20 observations, and 16 degrees of freedom, the preharvest confidence interval is $p \pm .59$. Since the equilibrium price level ($1.72) is below the support price, the lower bound of the confidence interval would be at the support price ($1.75) and the upper bound would be $2.34. The 95-percent confidence intervals for the other variables are presented in table 3. The point estimates which represent the "most likely" occurrences are also shown. In addition, confidence intervals based on the midsummer and postharvest variance estimates are included in the table. These intervals are illustrated in figures 1 and 2.

Figure 1 shows the preplanting 95-percent confidence intervals for supply, demand, and price. Figure 2 contains the midsummer and postharvest confidence bands. The outer limits of the shaded bands in this figure define the midsummer confidence interval and the inside of the shaded bands defines the 95-percent confidence intervals when harvest is known. Since the same exogenous data are used in each case, the point estimates of price and utilization are the same. By overlaying the two figures, one can trace the effects of increasing information on forecast precision. The information which reduces the variance of production narrows the confidence intervals for price and consumption.

26

Table 3—Ninety-five percent confidence intervals for variables in the corn model

	Lower bound	Point estimate	Upper bound
Preplanting:			
Production	5329	6219	7109
Price	1.75	1.75	2.34
Feed demand	3675	4153	4632
Food demand	378	414	450
Commercial exports	1367	1606	1845
Commercial stocks	327	645	963
Midsummer:			
Production	5806	6219	6632
Price	1.75	1.75	2.13
Feed demand	3859	4153	4448
Food demand	380	414	448
Commercial exports	1443	1606	1769
Commercial stocks	422	425	868
Postharvest:			
Production	6219	6219	6219
Price	1.75	1.75	2.04
Feed demand	3935	4153	4372
Food demand	380	414	448
Commercial exports	1471	1606	1741
Commercial stocks	457	425	833

IMPLICATIONS

Recent forecasting work in ERS has typically derived point estimates of the consequences of three alternative scenarios. Increasingly, probability statements have been requested for each scenario. To the extent that these scenarios estimate the effects of events outside the modeling systems, this variance framework will not help the analysts. But to the extent that the scenarios simply bracket the most likely occurrences, these standard errors could be used to provide an interval estimate of a single most likely scenario, together with the likelihood that the interval contains the observations in the period of forecast. Thus, one interval estimate, requiring little more time to prepare than a single point estimate, would provide more information than the point estimates for three separate scenarios. Analysts would gain more time to evaluate events outside the modeling system and to interpret their analysis to policymakers and the general public.

APPENDIX: STRUCTURE OF THE CORN MODEL

This model of the corn sector was originally developed by Womack (7). It consists of structural demand equations for four components of utilization, a recursive production submodel, and price determination
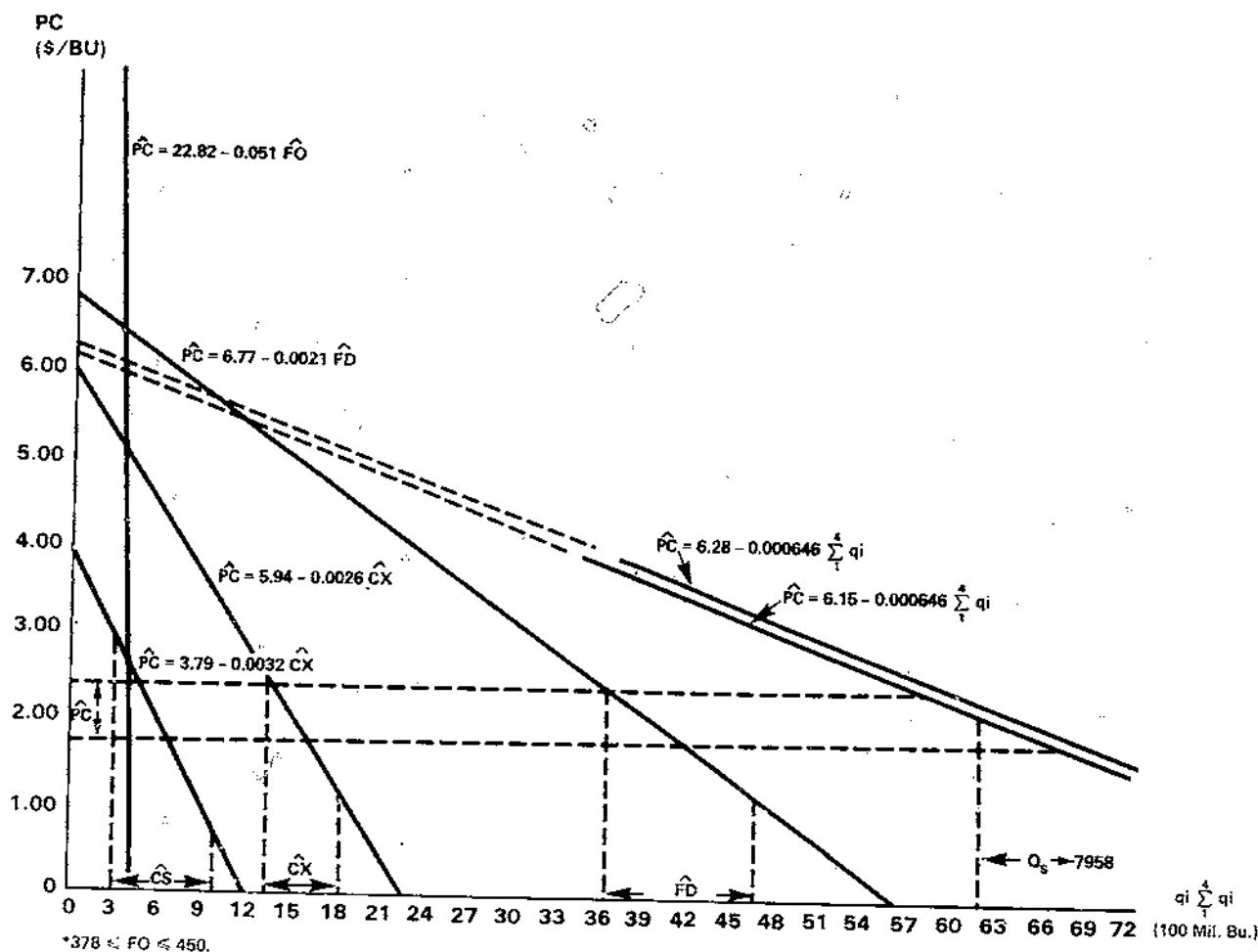
FIGURE 1. 95 PERCENT PRE-PLANTING CONFIDENCE
LIMITS FOR CORN

from the supply-utilization identity. The endogenous variables are corn prices, production, and levels of feed, food, export and stock demand. The production model was patterned after Houck and others (3, 4), and it consists of an equation for planted acreage and an equation for yields. The demand equations describe feed use, food use, commercial exports, and ending commercial stocks.

Many questions of model specification, estimation techniques, sample periods, and so on, could be addressed to this model. Since our purpose was to estimate forecast variance from the given model, we did not consider such questions. For answers to them, see (7, 3, 4).

Feed demand in the model is a function of corn price, soybean meal price, and indexes of livestock prices and livestock output. The sample period is 1950-72.

$$FD = -827.454 \ PC + 57.975 \ PM + 1004.86 \ PL$$
$$(-3.01) \qquad (1.50) \qquad (3.21)$$

$$+ \ 3490.36 \ LO \ -448.852$$
$$(7.43) \qquad (-0.68)$$

$$R^2 = .935 \qquad SEE = 148.2 \ mil. \ bu. \qquad DW = 1.304$$

Food demand is a function of corn price and real personal disposable income. This equation was estimated using 2SLS, so no R^2 is presented. The sample period is 1948-72.

$$FO = -19.649 \ PC + .4732 \ YPD + 122.2786$$
$$(-1.28) \qquad (19.90)$$

$$SEE = 13.39 \ mil. \ bu. \qquad DW = 0.49$$

27

PC
($/BU)

$\hat{PC} = 22.82 - 0.051\ \hat{FO}$

$\hat{PC} = 6.77 - 0.0021\ \hat{FD}$

$\hat{PC} = 6.28 - 0.000646\ \overset{4}{\underset{1}{\Sigma}}\ qi$

$\hat{PC} = 6.15 - 0.000646\ \overset{4}{\underset{1}{\Sigma}}\ qi$

$\hat{PC} = 5.94 - 0.0026\ \hat{CX}$

$\hat{PC} = 3.79 - 0.0032\ CS$

$\hat{PC}$

$\overline{q}_s$

7481

$\hat{CS}$ $\hat{CX}$ $\hat{FD}$

$qi\ \overset{4}{\underset{1}{\Sigma}}\ qi$
(100 Mil. Bu.)

·380 · FO · 448.

Denotes reduction in error due to acreage uncertainty
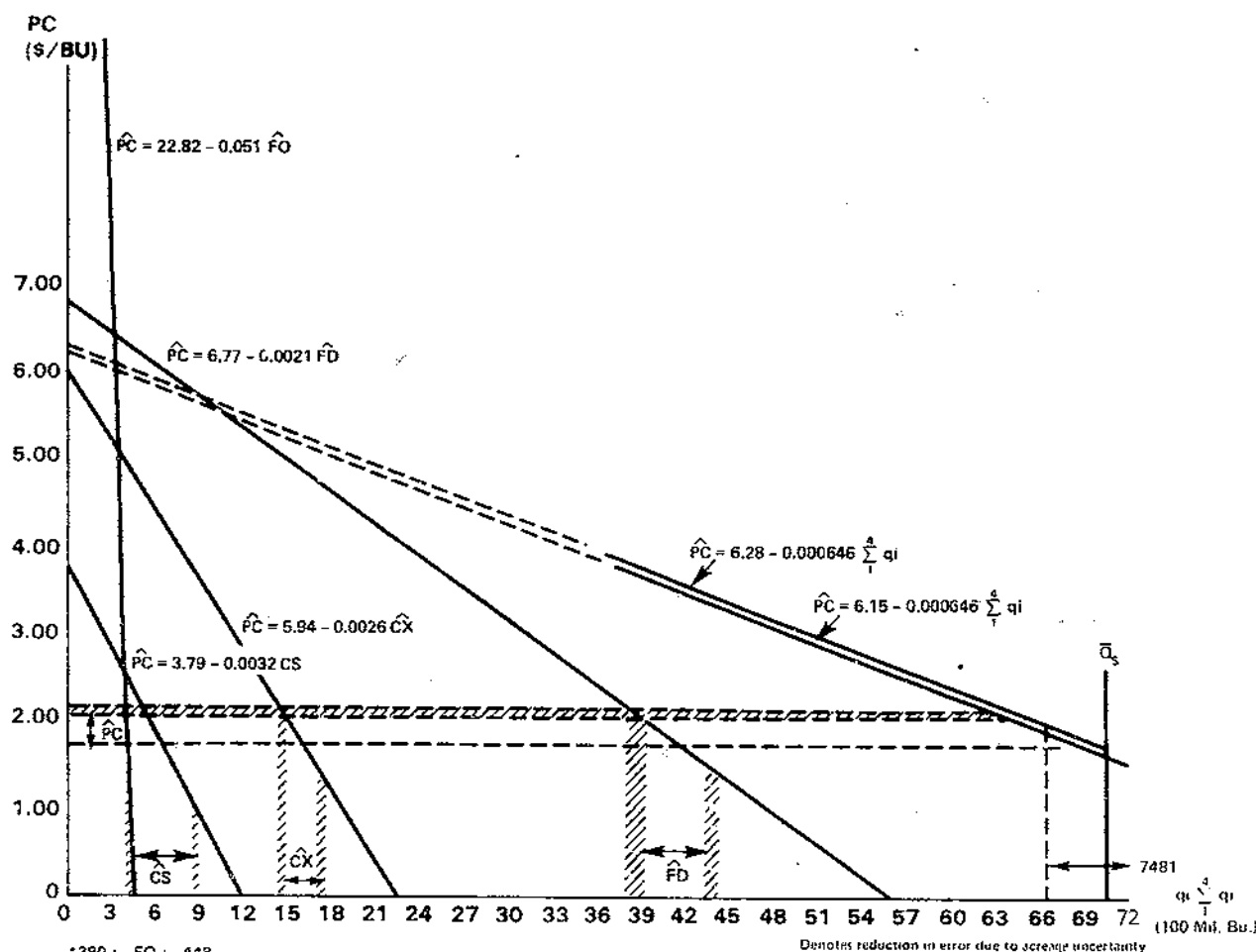
FIGURE 2. 95 PERCENT MIDSUMMER AND POST-HARBEST
CONFIDENCE LIMITS FOR CORN

USDA

NEG ERS 2900-77 (9)

Commercial exports are a function of U.S. corn price, U.S. soybean price, and production plus beginning stocks in EEC and exporting countries; and animal units, livestock price, and personal consumption expenditures in importing countries. The sample period is 1956-73.

$$CX = -383.195\ PC + 294.363\ PS - .0356\ SEEC$$
$$\quad\ (-3.13)\qquad\ \ (5.50)\qquad\quad (-2.43$$

$$-.0279SX + .020AUX$$
$$(-3.74)\quad (2.12)$$

$$+\ 1255.98\ PLX_{t-1} + 610.236XIX - 2772.79$$
$$(2.16)\qquad\qquad (1.20)$$

$$R^2 = .98\qquad SEE = 42.67\ \text{mil. bu.}$$

Ending commercial stocks are related to corn price, this year's production, acres planted for next year's har-

vest, and ending Government stocks.[1] The sample period is 1949-73.

$$CS = 288.364\ PC + .1049\ QC - 8.138\ AP$$
$$\quad\ (-3.79)\qquad\ (3.92)\qquad (-2.18)$$

$$-.1375\ GS + 1231.47$$
$$(-2.67)\quad (4.00)$$

$$R^2 = .825\qquad SEE = 83.95\ \text{mil. bu.}\qquad DW = 1.935$$

Planted acreage for corn is a function of the maximum of the effective support or farm price of corn relative to soybean price, the deficiency payments for corn, the effective support price of soybeans, a sorghum acre-

---

[1] This equation is specified slightly different from Womack's (7).

28

age variable, time, and a dummy variable. The sample period is 1950-74.

$$AP_{t-1} = 23.002\ PC^*/PS - 43.627\ DPC - 5.907\ PFS$$
$$\phantom{AP_{t-1} = }(6.06)(-7.13)(-2.46)$$

$$-.287\ APS^* - .311\ (T-48)$$
$$(-2.02)(-2.65)$$

$$+\ 7.153\ D661 + 88.825$$
$$(6.06)$$

$$R^2 = .965\qquad DW = 1.532\qquad SEE = 1.72\ mil.\ acres$$

Yield per harvested acre is a function of the fertilizer/corn price ratio, weather in the current and preceding year, two dummy variables, and a logarithmic time trend.[2] The sample period was 1951-71. This equation was published by Houck and Gallagher (4).

$$YLD_{t-1} = -.5101\ PF/PC + 4019W_{t-1} + .3525W$$
$$(6.32)(3.09)\phantom{W_{t-1}}(2.83)$$

---

[2] Currently, within the Forecast Support Group of ERS, researchers are developing yield estimates based on weather information available at different months of the year. Results will allow more finely graded changes in confidence bands than those derived.

$$+\ 7.675\ D61 - 15.28D70 + 38.093\ LN(T)$$
$$(3.67)(5.24)(12.72)$$

$$-\ 66.31$$
$$(5.34)$$

$$\overline{R}^2 = .980\qquad DW = 2.12\qquad SEE = 2.472\ bu.$$

Because yield and planted acres are calendar year variables and the time subscripts refer to the crop years, the production for the current crop year equals the product of acres planted and yields in the previous crop year.

$$QC = AP_{t-1}\ *\ YLD_{t-1}$$

The system is closed by the supply-utilization identity which equates production plus beginning stocks to feed, food, export, and stock demand, plus Government stocks and policy exports.

$$QC + CS_{t-1} + GS_{t-1} = FD + FO + CS$$

$$+\ CS + GS + GX$$

A complete rationalization of the specification and a description of the precise content of each variable are found in (7).

CITATIONS

(1) Feldstein, Martin. "The Error of Forecast in Econometric Models When the Forecast Period Exogeneous Variables are Stochastic." *Econometrica*, 39: 55-60, Jan. 1971.

(2) Goldberger, A. S., A. L. Nagar, and H. S. Odeh. "The Covariance Matrices of Reduced Form Coefficients and Forecasts for a Structural Econometric Model." *Econometrica*, 29:556-573, Oct. 1961.

(3) Houck, J. P., M. E. Abel, M. E. Ryan, P. W. Gallagher, R. G. Hoffman, and J. B. Penn. *Analyzing the Impact of Government Programs On Crop Acreage.* Econ. Res. Serv., U.S. Dept. Agr., Tech. Bul. 1548, Aug. 1976.

(4) Houck, J. P. and P. W. Gallagher. "The Price Responsiveness of U.S. Corn Yields." *Am. J. Agr. Econ.* 58:731-734.

(5) Kmenta, Jan. *Elements of Econometrics.* MacMillan, and Co., New York, 1971.

(6) Schmidt, Peter. "The Asymptotic Distribution of Forecasts in the Dynamic Simulation of an Econometric Model." *Econometrica*, 42: 303-309, Mar. 1974.

(7) Womack, Abner W. *The U.S. Demand for Corn, Sorghum, Oats and Barley: An Econometric Analysis.* Univ. Minn., Dept. Agr. and Applied Econ., Econ. Rpt. 76-5, St. Paul, Aug. 1976.