**Cowles Foundation Paper 171**

*Econometrica*, Vol. 29, No. 4 October (1961)

# THE ERROR OF FORECAST FOR MULTIVARIATE REGRESSION MODELS

BY JOHN W. HOOPER AND ARNOLD ZELLNER[1]

In this paper the problem of constructing a generalized error of forecast for a set of dependent variables in a multivariate regression model (which may be the set of reduced-form equations of an econometric model) is considered. The distribution of this multivariate error of forecast is derived and its use in the construction of probabilistic forecast regions is presented.

## 1. INTRODUCTION

IN THIS paper we present a generalized error of forecast for the set of dependent variables in a multivariate regression model. Since in many practical situations statisticians and econometricians are concerned with forecasting the future values of a group of dependent variables it seems desirable to be able to construct confidence regions and test hypotheses about the forecast of these variables.

The generalized error of forecast which we develop is a quite straightforward extension of the standard error of forecast used when making predictions with single equation regression models. We shall review this latter concept in Section 2. In Section 3 we derive an expression for the generalized error of forecast and show how it can be used to construct forecast regions and test hypotheses; an application of this statistic to a two-equation model is then made in Section 4. Finally, in Section 5 some further uses and limitations of the generalized error of forecast are discussed.

## 2. THE SINGLE EQUATION ERROR OF FORECAST[2]

The single equation regression model may be written as

$$y_n = \sum_{k=1}^{K} \beta_k X_{kn} + v_n \qquad (n = 1, 2, \ldots, N) \tag{2.1}$$

or in matrix notation as

$$y = \beta X + v \tag{2.2}$$

[1] We wish to express our appreciation for the many constructive comments which Professors T. C. Koopmans, A. T. James, and T. N. Srinivasan have made concerning both the substance and form of this paper. However, the final form and contents, including any errors which might remain, are the joint responsibility of the authors.

[2] The error of forecast for a single equation regression model appears to have a rather long history. It seems to have been first discussed by H. Working and H. Hotelling in [12]. For other references to this problem the reader is referred to [2, 181]. An excellent presentation of the error of forecast in matrix notation can be found in [11, 280–284].

where $y$ is a row vector of $N$ observations on the dependent variable, $X$ a $K \times N$ matrix of nonstochastic or fixed values taken by the $K$ independent variables ($K \leqslant N$), $\beta$ a row vector of $K$ unknown regression coefficients, and $v$ a random row vector consisting of $N$ independently distributed terms, each with zero expected value and a common finite variance $\sigma_v^2$. Upon estimating $\beta$ by the method of least squares, we obtain

$$y = bX + \bar{v} \tag{2.3}$$

where

$$b = yX'(XX')^{-1} = \beta + vX'(XX')^{-1} \tag{2.4}$$

and $\bar{v}$ is the vector whose components are the $N$ calculated residuals.

If we have a set of values available for the independent variables for the period in which the forecast is made, say $X_F$, (where $X_F$ is a $K \times 1$ vector) we can then predict a value of $y$, say $y_F^*$, by using the estimated regression relationship. We then obtain

$$y_F^* = bX_F \, . \tag{2.5}$$

The true value of $y$, say $y_F$, in this prediction period is

$$y_F = \beta X_F + v_F \tag{2.6}$$

where $v_F$ is the scalar value of the random term in the forecast period. When we subtract (2.6) from (2.5) we obtain as the error of forecast

$$y_F^* - y_F = (b-\beta)X_F - v_F \, . \tag{2.7}$$

The expected error is (using $E$ as the expected value operator)

$$E(y_F^* - y_F) = E[(b-\beta)X_F - v_F] = 0 \tag{2.8}$$

since $b$ is an unbiased estimator of $\beta$, $Ev_F = 0$, and $X_F$ is assumed to be nonstochastic.[3]

We notice that the forecast error in (2.7) may be divided into two parts. One source of error is due to the inaccuracies in estimating the regression coefficients and the other is due to the presence of the random term $v_F$. Thus the error of forecast is a function of two random variables, $b$ and $v_F$, and these variables are independently distributed since by assumption $v_F$ is independent of $v_n$ ($n = 1, \ldots, N$) and hence independent of $b$, which is but a linear function of $v_n$. If we now make the further assumption that the random terms, the $v_n$'s, are each normally distributed, then the forecast error is also normally distributed since it is just a linear combination of these normally distributed variables.

[3] If $X_F$ is stochastic, as would be the case if a forecast of the independent variables for the prediction period were used, it would be necessary to assume that $X_F$ and $b-\beta$ are distributed independently for $E(y_F^* - y_F)$ to be equal to zero.

The variance of the forecast error is given by[4]

$$\sigma_F^2 = E(y_F^* - y_F)^2 = Ev_F^2 + X_F'[E(b-\beta)'(b-\beta)]X_F \tag{2.9}$$

which can be written as

$$\sigma_F^2 = \sigma_v^2[1 + X_F'(XX')^{-1}X_F] . \tag{2.10}$$

The positive square root of (2.10) is the population standard error of forecast. As an estimator of $\sigma_F^2$ we use the unbiased estimator

$$\hat{\sigma}_F^2 = \hat{\sigma}_v^2[1 + X_F'(XX')^{-1}X_F] \tag{2.11}$$

where

$$\hat{\sigma}_v^2 = \frac{yy' - b(XX')b'}{N-K} . \tag{2.12}$$

We can now consider the statistic

$$t = \left(\frac{y_F^* - y_F}{\sigma_F}\right) \Big/ \left(\frac{\hat{\sigma}_F}{\sigma_F}\right) . \tag{2.13}$$

This is the familiar $t$ statistic with $N-K$ degrees of freedom since, under the assumption that the random terms are normally distributed, $y_F^* - y_F$ is normally distributed with expected value zero and variance $\sigma_F^2$, $(N-K)\hat{\sigma}_F^2/\sigma_F^2$ is distributed as a $\chi^2$ variable with $N-K$ degrees of freedom, and $\hat{\sigma}_F$ and $y_F^* - y_F$ are independently distributed.[5] We can then make probability statements about $t$ in the form

$$\Pr(|t| > t_\alpha) = \alpha \tag{2.14}$$

when testing hypotheses about $y_F$, where $t_\alpha$ is the value of the $t$ statistic at the $\alpha$ level of significance, or in the form

$$\Pr(y_F^* - \hat{\sigma}_F t_\alpha < y_F < y_F^* + \hat{\sigma}_F t_\alpha) = 1 - \alpha \tag{2.15}$$

for the purpose of determining a confidence interval for $y_F$.[6]

[4] The variance of the forecast error as given by (2.9) is valid regardless of whether the random terms are normally distributed. The assumption of normality is needed in order to determine the distribution of the statistic defined in (2.13).

[5] Cf. [**10**, 206, 298].

[6] This is not a confidence interval in the classical sense because $y_F$ is a random variable and not a population parameter. The interval in (2.15) is a similar beta-expectation tolerance interval and it can be interpreted as meaning that there is a $1-\alpha$ level of confidence of one future observation ($y_F$) falling in the interval $y_F \pm \hat{\sigma}_F t_\alpha$. The similar beta-expectation tolerance interval should not be confused with the beta-content tolerance interval with confidence level $1-\alpha$. The beta-content tolerance interval is constructed so that there is a probability $1-\alpha$ that beta per cent of future

## 3. THE GENERALIZED ERROR OF FORECAST

We now consider the following multivariate regression model

$$y_{in} = \sum_{k=1}^{K} \pi_{ik} X_{kn} + v_{in} \qquad (i = 1,2,\ldots,G;\, n = 1,2,\ldots,N) \tag{3.1}$$

which can be written in matrix notation as

$$Y = \Pi\, X + V\,. \tag{3.2}$$

$Y$ represents a $G \times N$ matrix of the $N$ sets of sample values taken by the $G$ jointly dependent variables, $\Pi$ a $G \times K$ matrix of unknown regression coefficients, $X$ a $K \times N$ nonstochastic matrix[7] of values taken by the independent variables, and $V$ a $G \times N$ matrix of random terms. It is also assumed that the $N$ column vectors of $V$ are independent random drawings from a $G$-dimensional normal population such that each column has a zero expected value and the columns have a common variance-covariance matrix, $\Sigma_{vv}$.[8] These conditions on $V$ may be summarized as

$$Ev_{in} = 0 \qquad (i = 1,\ldots,G;\, n = 1,\ldots,N) \tag{3.3}$$

and

$$Ev_{in}v_{jn'} = \begin{cases} 0, & n \neq n'\,, \\ \sigma_{ij}, & n = n'\,, \end{cases} \qquad \text{for all } i \text{ and } j\,. \tag{3.4}$$

Thus interdependence between the random terms in different equations in the same time period is allowed but not between random terms of different time periods.

By applying the method of least squares to (3.2) we obtain

$$Y = PX + \hat{V} \tag{3.5}$$

where $P$, the matrix of estimated regression coefficients is[9]

---

observations will be included in such intervals. For a complete discussion of these various types of tolerance intervals the reader is referred to [**5**]. An interesting earlier interpretation of the standard error of forecast is given in [**3**]. A discussion of beta-content tolerance intervals as applied to single equation regression models is given in [**8**, 255–258] and [**13**]. We are indebted to Professor L. R. Klein for calling to our attention the problem of defining a forecast interval for random variables.

[7] We also make the usual assumption that the rank of $X$ is $K$ which ensures that none of the $X$'s can be expressed as a linear combination of the others and that $K \leqslant N$.

[8] It should be noticed that this model is completely analogous to what is referred to as the reduced form equations in econometrics. Under this interpretation the $y$'s are the set of endogenous variables, the $\pi$'s the reduced-form regression coefficients, the $x$'s the exogenous variables, and the $v$'s the reduced-form disturbances. The reader is referred to [**9**, 113–121] for precise definitions of these terms. It should be noticed that we are here excluding reduced form equations in which some of the $x$'s are lagged endogenous variables.

[9] The expression for $P$ in (3.6) also shows that $P$ is an unbiased estimator of $\Pi$, since when taking expected values of both sides we have $E[VX'(XX')^{-1}] = 0$ so $EP = \Pi$.

$$P = YX'(XX')^{-1} = \Pi + VX'(XX')^{-1} \tag{3.6}$$

and $\hat{V}$ is the matrix of calculated residuals.

If we now desire to predict simultaneously[10] the values of the $G$ dependent variables using the estimated regression coefficients and a set of values for the independent variables, $X_F$, the forecast would be

$$Y_F^* = PX_F \tag{3.7}$$

where $Y_F^*$ is a column vector of forecast values and $X_F$ is a column vector of nonstochastic known values for the independent variables. The observed value of $Y$ in the prediction period, say $Y_F$, is given by

$$Y_F = \Pi\, X_F + V_F \tag{3.8}$$

where $V_F$ is the column vector of values assumed by the random terms in the forecast period. The error of forecast is

$$Y_F^* - Y_F = (P - \Pi)X_F - V_F\,. \tag{3.9}$$

The expected error is

$$E(Y_F^* - Y_F) = E[(P - \Pi)X_F - V_F] = 0 \tag{3.10}$$

since $P$ is an unbiased estimator of $\Pi$, $EV_F = 0$, and $X_F$ is nonstochastic and thus independent of $P$. The variance-covariance matrix of the forecast error is given by

$$\Sigma_{FF} = E[(Y_F^* - Y_F)(Y_F^* - Y_F)'] \tag{3.11}$$

$$= E[(P - \Pi)X_F X_F'(P - \Pi)'] - E(P - \Pi)X_F V_F' - EV_F X_F'(P - \Pi)' + EV_F V_F'\,.$$

The two middle terms in (3.11) are equal to zero since $V_F$ and $P$ are independent and $EV_F = 0$.

To evaluate the first term in (3.11) we let $(P - \Pi)$ be a $G \times K$ matrix $A$ with a typical element $[a_{ij}]$ and $X_F X_F'$ be a $K \times K$ matrix $M_F$ with typical element $[m_{kk'}]$. We find that

[10] We are here concerned with finding a statistic that will enable us to make probability statements about the set of dependent variables and not each variable considered separately. The latter has been done under the assumption that the coefficients of the structural equations were estimated by the method of full maximum likelihood and also for the resulting reduced form equations. Cf. [2].

(3.12) $E[(P-\Pi)X_FX_F'(P-\Pi')] = EAM_FA' =$

$$\begin{bmatrix} \sum_{k'=1}^{K}\sum_{k=1}^{K}\sigma_{k'k}^{(1)(1)} m_{kk'} & \ldots & \sum_{k'=1}^{K}\sum_{k=1}^{K}\sigma_{k'k}^{(1)(G)} m_{kk'} \\ \vdots & & \vdots \\ \sum_{k'=1}^{K}\sum_{k=1}^{K}\sigma_{k'k}^{(G)(1)} m_{kk'} & \ldots & \sum_{k'=1}^{K}\sum_{k=1}^{K}\sigma_{k'k}^{(G)(G)} m_{kk'} \end{bmatrix}$$

where $\sigma_{k'k}^{(i)(j)}$ is the element in the $k$th row and $k$th column of the variance-covariance matrix of the $i$th and $j$th rows of $P$. The covariance matrix between $P_i$ and $P_j$, two rows of $P$, is

$$(3.13)\quad E(P_i-\Pi_i)'(P_j-\Pi_j) = (XX')^{-1}X(EV_i'V_j)X'(XX')^{-1} = (XX')^{-1}X\sigma_{ij}I_NX'(XX')^{-1} = \sigma_{ij}(XX')^{-1}$$

where $V_i$ is the $i$th row of $V$ as defined in (3.2) and $V_j$ the $j$th row. This means that the row vector of $GK$ components, $(P_1, P_2,\ldots,P_G)$ is normally distributed (since the columns of $V$ are normally distributed) with an expected value $(\Pi_1, \Pi_2,\ldots,\Pi_G)$, and with a variance-covariance matrix, $\Sigma_{pp}$, where[11]

$$(3.14)\qquad \Sigma_{pp} = \begin{bmatrix} \sigma_{11}(XX')^{-1} & \ldots & \sigma_{1G}(XX')^{-1} \\ \vdots & & \vdots \\ \sigma_{G1}(XX')^{-1} & \ldots & \sigma_{GG}(XX')^{-1} \end{bmatrix}$$

and $\Sigma_{pp}$ is of order $GK \times GK$.

From (3.12) and (3.14), letting $(XX')^{-1}_{kk'}$ be a typical element of $(XX')^{-1}$, we have for a typical element $[\sum_{k,k'}\sigma_{gg'}(XX')^{-1}_{kk'}m_{kk'}]$ of $E(AM_FA')$ that

$$(3.15)\qquad [\sum_{k,k'}\sigma_{gg'}(XX')^{-1}_{kk'}m_{kk'}] = \sigma_{gg'}\sum_{k,k'}(XX')^{-1}_{kk'}m_{kk'} = \sigma_{gg'}\,q$$

where

$$(3.16)\qquad q = X_F'(XX')^{-1}X_F$$

is a scalar. So we find that

$$(3.17)\qquad E(AM_FA') = q\Sigma_{vv}\,.$$

From (3.11) and (3.17), we obtain for the variance-covariance matrix of $Y_F^*-Y_F$, $\Sigma_{FF}$, that

$$(3.18)\qquad \Sigma_{FF} = q\Sigma_{vv} + \Sigma_{v_Fv_F} = (1+q)\Sigma_{vv}$$

since $\Sigma_{v_Fv_F} = \Sigma_{vv}$.

[11] Cf. [1, 182] for an alternative derivation of these results. $\Sigma_{pp}$ is recognized as the Kronecker product of the matrices $\Sigma_{vv}$ and $(XX')^{-1}$, i.e., $\Sigma_{pp} = \Sigma_{vv} \otimes (XX')^{-1}$.

We shall not in general know $\Sigma_{vv}$ so we estimate it by the unbiased estimator[12] $S_{vv}$, where

$$S_{vv} = (YY' - PXX'P')/(N-K)\,. \tag{3.19}$$

Thus from (3.18) and (3.19) we find that the estimated variance-covariance matrix of the error of forecast, $S_{FF}$, is

$$S_{FF} = (1+q)S_{vv}\,. \tag{3.20}$$

For the purpose of making probability statements about the forecast we can use Hotelling's $T^2$ statistic,[13] where

$$T^2 = (Y_F^* - Y_F)'S_{FF}^{-1}(Y_F^* - Y_F)\,. \tag{3.21}$$

Now it has been shown[14] that the distribution of

$$\frac{(N-K-G+1)T^2}{(N-K)G} \tag{3.22}$$

is the $F$ distribution with $G$ and $N-K-G+1$ degrees of freedom.

We can now, with the use of (3.22), construct forecast regions or test hypotheses about the forecast. The procedure is to choose a level of significance, say $\alpha$, and then find the corresponding value of $F$, say $F_\alpha$, in the tables of the $F$ distribution. Then we have the result that the set of points for which the inequality

$$\left(\frac{N-K-G+1}{(N-K)G}\right) T^2 \leqslant F_\alpha \tag{3.23}$$

[12] Cf. [**1**, 183] for a proof that this is the unbiased estimator of $\Sigma_{vv}$.

[13] Cf. [**7**]. Notice that we could also use the statistic

$$(1+q)T^2 = (Y_F^* - Y_{F'})S_{vv}^{-1}(Y_F^* - Y_F)$$

which has the same distribution as $T^2$ except for the multiplicative constant, $1+q$. In computing a forecast region this expression is easier to use as it eliminates the necessity of dividing the elements of $S_{vv}^{-1}$ by $1+q$.

[14] Cf. [**1**, 105–106]. That this theorem applies in our case may be seen by considering the following: (i) $Y_F^* - Y_F$ is normally distributed with zero expected value and variance-covariance matrix $\Sigma_{FF}$, and (ii) $(N-K)S_{FF}$ is distributed independently of $Y_F^* - Y_F$ as the sum of $N-K$ vector products, i.e., as $\Sigma_{n=1}^{N-K} s_n s'_n$ (where $s_n$ is a $G$-element random column vector), where these vectors are independent of each other and each has a multinormal distribution with zero expected value and variance-covariance matrix $\Sigma_{FF}$.

holds forms the area of the forecast region.[15] This region may be interpreted in the same way as the forecast interval for a single equation.[16] Specifically, if repeated samples are taken, holding $X$ and $X_F$ fixed, then $1-\alpha$ per cent of the time a region as in (3.23) will cover the true values of the forecasted variables, i.e., $Y_F$.

It is interesting to determine those values of the independent variables used for the forecast, i.e., $X_F$, which would minimize the value of $q = X'_F (XX')^{-1}X_F$. As can be seen from (3.18), minimizing $q$ will minimize the generalized variance[17] of the error of forecast. This is the multivariate analogue to the minimization of the variance of the error of forecast in single equation models.

To take account of the constant term in each equation, we let $X_{1t}$ be a dummy variable, i.e., $X_{1t} \equiv 1$ for all $t$. Then taking our observations as deviations from the sample means so that $z_{kt} = x_{kt} - \bar{x}_k$ and $z_{kF} = x_{kF} - \bar{x}_k$ for $k = 1, \ldots, K$ and $t = 1, \ldots, N$ we can write

$$1 + q = \left(1 + \frac{1}{N} + Z'_F(ZZ')^{-1}Z_F\right). \tag{3.24}$$

Since $(ZZ')$ is at least a positive semi-definite matrix this expression is minimized when $Z_F$ is the null vector, i.e., when $X_{kF} = \bar{X}_k$. Thus choosing the sample means of the exogenous variables as the values of the exogenous variables for the prediction will minimize the generalized variance of the error of forecast.[18] For the latter we obtain

$$|\Sigma_{FF}|_{\min} = \left(1 + \frac{1}{N}\right)^G |\Sigma_{vv}|. \tag{3.31}$$

## 4. CONSTRUCTION OF FORECAST REGIONS: AN EXAMPLE

For the purpose of illustrating the construction of a forecast region for a set of equations we shall use the familiar model of Haavelmo.[19] This model

[15] It should be noticed that this is a direct generalization of the forecast region for a single equation model with $K$ independent variables. In this case $G = 1$, so we obtain, substituting $t^2$ for $T^2$, the $t^2$ distribution, which is, as is well known, the $F$ distribution with 1 and $N-K$ degrees of freedom.

[16] Cf. p. 546 above.

[17] The generalized variance of a multivariate distribution is defined as the determinant of the variance-covariance matrix. Therefore, the generalized variance of the error of forecast is $|\Sigma_{FF}| = |(1+q)\Sigma_{vv}| = (1+q)^G|\Sigma_{vv}|$. This is a minimum when $q$ is a minimum since $|\Sigma_{vv}| > 0$ because $\Sigma_{vv}$ is a positive definite matrix and the determinant of a positive matrix is always positive.

[18] We are indebted to one of the referees for suggesting this alternative and simpler derivation of this point.

[19] Cf. [**6**, 83–91] for a complete description of this model.

may be written as a set of two regression or reduced form equations. We have

$$c_t = \pi_{11}X_{1t} + \pi_{12}X_{2t} + v_{1t} \qquad (t = 1,\ldots,M) \tag{4.1}$$
$$y_t = \pi_{21}X_{1t} + \pi_{22}X_{2t} + v_{2t}$$

where the jointly dependent variables are consumers' expenditures ($c_t$) and disposable income ($y_t$). The independent variables are $X_{1t} \equiv 1$ and gross investment $X_{2t}$. The random terms are $v_{1t}$ and $v_{2t}$. The regression coefficients as estimated by least squares are

$$p_{11} = 298.554\,, \qquad p_{12} = 1.499\,, \tag{4.2}$$
$$p_{21} = 285.787\,, \qquad p_{22} = 2.105\,,$$

and the inverse of the estimated variance-covariance matrix of the random terms is

$$S_{vv}^{-1} = \begin{bmatrix} .02507 & -.02570 \\ -.02570 & .03125 \end{bmatrix}. \tag{4.3}$$

The inverse of the moment matrix of the independent variables is

$$(XX')^{-1} = \begin{bmatrix} .7329396 & -.00701333 \\ -.00701333 & .00007498 \end{bmatrix}. \tag{4.4}$$

For the values of the independent variables in the forecast period we shall choose[20] $X_{1F} = 1$ and $X_{2F} = 100$. We then obtain the equation

$$1 + q = 1.08007\,. \tag{4.5}$$

Now from (3.22) we know that $(N-K-G+1)T^2/(N-K)G$ has the $F$ distribution, $F_{G,\,N-K-G+1}$. In our example $N = 13$, $K = 2$, and $G = 2$. So $F = 4.10$ at the 5 per cent level of significance. The corresponding value of $(1+q)T^2$ is

$$(1+q)T^2_{5\%} = \frac{(4.10)(1.08007)(11)(2)}{10} = 9.742\,. \tag{4.6}$$

For the forecast values of the dependent variables we have

$$c_F^* = 448.454\,, \tag{4.7}$$
$$y_F^* = 496.287\,.$$

Using this and (4.3), (4.4), and (4.6), we obtain

$$9.742 = .02507(c_F^* - c_F)^2 - .05139(c_F^* - c_F)(y_F^* - y_F) + .03125(y_F^* - y_F)^2\,. \tag{4.8}$$

[20] The value of $X_{2F}$ is close to the mean of this variable in the sample period, i.e., $\bar{X} = 93.5385$.

This represents an ellipse in the parameter space of $c_F$ and $y_F$, where the center of the ellipse is $(\overset{*}{c}_F, \overset{*}{y}_F)$. The area covered by the ellipse is the forecast region for $c_F$ and $y_F$, the true values of the dependent variables in the forecast period. This is ellipse $A$ in Figure 1. Ellipse $B$ is the larger forecast region that results from using $X_{1F} = 1$ and $X_{2F} = 200$ as the values of the independent variables in the forecast period at the same level of significance.

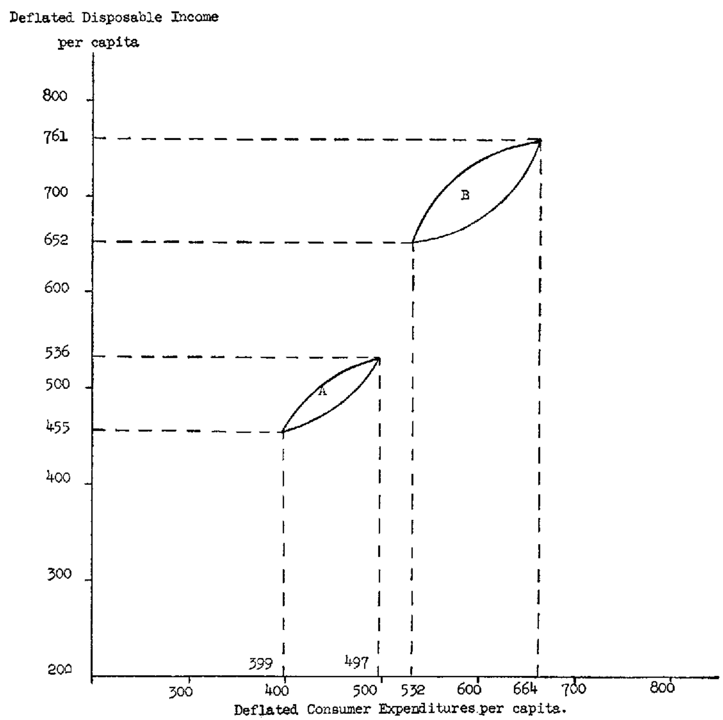


FIGURE 1.—Joint forecast regions for $C_F$ and $Y_F$; 5% level of significance.

## 5. CONCLUSIONS

Our analysis has yielded an expression for the error of forecast for interdependent regression equations and the distribution of the resulting statistic has been given. In an illustrative calculation this result has been employed to construct forecast regions. Calculation of such regions is extremely important in appraising the forecasts provided by systems of equations. Since point forecasts, unaccompanied by a forecast region, may on occasion be seriously misleading, we recommend that those who forecast take the additional trouble needed to construct forecast regions.

Further, our work provides what is necessary to test hypotheses about forecasts. Thus, for example, it is possible to test the hypothesis that a forecast from a reduced form system is not significantly different from a "judgment forecast."

Finally, we wish to point out several limitations associated with our result. Probability statements made about forecast regions constructed as we indicated above are valid providing the vector of exogenous variables employed in making the forecast is fixed, that is, nonstochastic. Given this condition, one can state, as is usually done, that the constructed region will cover the true values of the endogenous variables in the forecast period a certain proportion of the time in repeated trials. However, in an actual set of repeated forecasts it may be impossible to hold fixed all the exogenous variables and in this situation the interpretation given above for the forecast region does not apply.[21] A similar situation prevails when some of the exogenous variables have to be forecast; that is, the vector $X_F$ can no longer be regarded as nonstochastic. Further, it would be highly desirable to extend the present analysis to establish the properties of forecasts from interdependent reduced form equations that have lagged endogenous variables in the set of predetermined variables and that are estimated utilizing "small" samples of data.

*Yale University and University of Wisconsin*

## REFERENCES

[1] ANDERSON, T. W.: *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley and Sons, Inc., 1958.

[2] BROWN, T. M.: "Standard Errors of Forecast of a Complete Econometric Model," *Econometrica*, Vol. 22, April, 1954, pp. 178–192.

[3] EISENHART, C.: "The Interpretation of Certain Regression Methods and Their Use in Biological and Industrial Research," *Annals of Mathematical Statistics*, Vol. 10, June, 1939, pp. 162–186.

[4] FISHER, R. A.: *Statistical Methods and Scientific Inference*, Edinburgh and London: Oliver and Boyd, 1956.

[5] FRASER, D. A. S., AND I. GUTTMAN: "Tolerance Regions," *Annals of Mathematical Statistics*, Vol. 27, March, 1956, pp. 162–179.

[6] HAAVELMO, T.: "Methods of Measuring the Marginal Propensity to Consume," Chapter IV in Wm. C. Hood and T. C. Koopmans, *Studies in Econometric Method*, New York: John Wiley and Sons, Inc., 1953, pp. 75–91. Reprinted from *Journal of the American Statistical Association*, Vol. 42, March, 1947, pp. 105–122.

[7] HOTELLING, H.: "The Generalization of Student's Ratio," *Annals of Mathematical Statistics*, Vol. 2, August, 1931, pp. 360–378.

[8] KLEIN, L. R.: *Econometrics*, Evanston, Illinois: Row, Peterson, and Company, 1953.

[9] KOOPMANS, T. C., AND WM. C. HOOD: "The Estimation of Simultaneous Linear Economic Relationships," Chaper VI in Wm. C. Hood and T. C. Koopmans,

[21] R. A. Fisher discusses this problem in a different context. Cf. [4, 83–85].

*Studies in Econometric Method*, New York: John Wiley and Sons, Inc., 1953, pp. 112–199.

[10] MOOD, A. M.: *Introduction to the Theory of Statistics*, New York: McGraw-Hill Book Company, Inc., 1950.

[11] STONE, R.: *Measurement of Consumers' Expenditure and Behavior in the United Kingdom, 1920–1938*, Cambridge: The University Press, 1954.

[12] WORKING, H., AND H. HOTELLING: "Applications of the Theory of Error to the Interpretation of Trends," *Journal of the American Statistical Association*, Vol. 24 (Papers and Proceedings), March, 1929, pp. 73–85.

[13] WALLIS, W. A.: "Tolerance Intervals for Linear Regression," *Proceedings of the Second Berkeley Symposium on Mathermatical Statistics and Probability*, edited by J. Neyman, Berkeley: The University of California Press, 1951, pp. 43–51.