PREDICTING OUTPUT FROM SEEMINGLY UNRELATED AREA AND YIELD EQUATIONS

William Griffiths
University of New England

Graeme Thomson
La Trobe University, Bendigo

Tim Coelli
University of New England

Paper Contributed to the
Australian Agricultural Economics Society 39th Annual Conference,
University of Western Australia, Perth,
February 14-16, 1995.

PREDICTING OUTPUT FROM SEEMINGLY UNRELATED AREA AND YIELD EQUATIONS

William Griffiths, Graeme Thomson and Tim Coelli

(6-2-95)

1. Introduction

Predicting total production of a grain crop, for a given geological area in a given time period, is a frequently-faced problem in agricultural economics. One approach to this prediction problem is to carry out the following three steps: (1) Specify separate yield and area response equations. (2) From a time series of past observations, use econometric methods to estimate the parameters of each equation. (3) Use the estimated equations to make separate predictions of yield and area, and predict output as the product of predicted yield and predicted area. The argument for estimating separate area and yield equations rather than one single equation for output is usually based on the different decision processes and different variables that tend to underlie the area and yield equations. The area that is planted tends to depend upon price expectations (for the crop of interest as well as those of alternative crops), habit persistence (usually captured by a lagged dependent variable), input costs and rainfall at sowing time. Yield, on the other hand, depends upon climatic factors throughout the season and a variety of technology factors such as new plant varieties, new fertilizers and advances in crop rotation technologies. For further discussion of the specification of area response equations in Australian broad acre agriculture refer to Anderson (1974), Fisher (1975), Griffiths and Anderson (1978), Sanderson et al.

(1980) and Fisher and Munro (1983). For further discussion of the specification of yield equations refer to Guise (1969), Francisco and Guise (1988), Del Valle and Ray (1990) and Dillon and Anderson (1990). Two studies which consider the specification of both area and yield equations are Fisher (1978) and Coelli (1992).

One difficulty with modelling area and yield separately, and predicting output as the product of predicted yield and predicted area, is that it is not obvious that the simple product is an optimal predictor, and an appropriate expression for the standard error of the prediction error does not seem to be available in the literature. The object of this paper is to fill this gap. Assuming the area and yield equations comprise a two-equation system of seemingly unrelated regressions (Zellner, 1962), two predictors are suggested and the corresponding standard errors of their prediction errors are derived. The methodology is illustrated by predicting wheat output for the Corrigin Shire in Western Australia.

2. Model and Predictors

To explain previously generated observations on yield and area, assume we have the two equations

$$A = X_a\beta_a + U_a \tag{1}$$

$$Y = X_y\beta_y + U_y \tag{2}$$

where A and Y are T-dimensional vectors containing T past observations on area and yield, respectively; X_a and X_y are $(T \times K_a)$ and $(T \times K_y)$ matrices containing past observations on the explanatory variables that help describe movements in A and Y,

respectively; β_a and β_y are $(K_a \times 1)$ and $(K_y \times 1)$ area response and yield response coefficients, respectively; and U_a and U_y are $(T \times 1)$ normally distributed random vectors with zero means. Denoting the individual elements of U_a and U_y by u_{at} and u_{yt}, we represent the contemporaneous covariance matrix as

$$\Sigma = E\begin{bmatrix} u_{at}^2 & u_{at}u_{yt} \\ u_{at}u_{yt} & u_{yt}^2 \end{bmatrix} = \begin{bmatrix} \sigma_{aa} & \sigma_{ay} \\ \sigma_{ay} & \sigma_{yy} \end{bmatrix} \tag{3}$$

The subscript t in equation (3) denotes the t-th observation (t = 1,2,...,T). It is also assumed that the errors are uncorrelated over time in which case the joint covariance matrix for the complete error vector $(U_a' \; U_y')'$ is given by $\Sigma \otimes I_T$, where I_T is the T–dimensional identity matrix. Further, let

$$W = \begin{bmatrix} A \\ Y \end{bmatrix} \qquad X = \begin{bmatrix} X_a & 0 \\ 0 & X_y \end{bmatrix} \qquad \beta = \begin{bmatrix} \beta_a \\ \beta_y \end{bmatrix} \tag{4}$$

The model described by equations (1) through (4) is a two equation example of the standard seemingly unrelated regressions model introduced by Zellner (1962). The best linear unbiased estimator for β is the generalized least squares estimator

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_a \\ \hat{\beta}_y \end{pmatrix} = \left[X'(\Sigma^{-1} \otimes I_T)X\right]^{-1} X'(\Sigma^{-1} \otimes I_T)y \tag{5}$$

with covariance matrix

$$(6) \qquad V(\hat{\beta}) = \left[X'(\Sigma^{-1} \otimes I_T)X\right]^{-1} = \begin{bmatrix} \sigma^{aa}X_a'X_a & \sigma^{ay}X_a'X_y \\ \sigma^{ay}X_y'X_a & \sigma^{yy}X_y'X_y \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} V_{aa} & V_{ay} \\ V_{ay}' & V_{yy} \end{bmatrix}$$

In equation (6) the symbols σ^{aa}, σ^{yy} and σ^{ay} denote the elements of Σ^{-1}; the submatrices V_{aa}, V_{ay} and V_{yy} have been introduced because they appear in subsequent expressions for predictors and their standard errors. In practice the elements in the contemporaneous error covariance matrix Σ are unknown and are estimated using least squares residuals. Details of this procedure and other information about the seemingly unrelated regressions model can be found in any standard econometrics text; see, for example, Judge et al. (1988, Ch.11).

Given values for the explanatory variables in the next period, denoted by the $(K_a \times 1)$ and $(K_y \times 1)$ vectors x_a and x_y, respectively, the problem is to predict next period's output, which is the product of next period's area a, and next period's yield y, with the latter two quantities being given by

$$(7) \qquad a = x_a'\beta_a + u_a$$

$$(8) \qquad y = x_y'\beta_y + u_y$$

The random errors u_a and u_y are assumed to be a joint drawing from the bivariate normal distribution $N(0, \Sigma)$, consistent with the data generating process for the sample observations.

Thus, we are attempting to predict

(9) $$q = a.y$$

$$= \left(x_a'\beta_a + u_a\right)\left(x_y'\beta_y + u_y\right)$$

$$= \left(x_a'\beta_a\right)\left(x_y'\beta_y\right) + x_a'\beta_a u_y + x_y'\beta_y u_a + u_a u_y$$

The choice of predictor and the variance of the prediction error depend on the level of recognition of parameter uncertainty. Three cases involving differing degrees of recognition of parameter uncertainty can be identified. The three cases are: (i) assume all parameters (β_a, β_y and Σ) are known, (ii) recognize that β_a and β_y are unknown but assume Σ is known; and (iii) recognize that β_a, β_y and Σ are all unknown.

The usual approach under case (i) is to: (1) Assume all parameters (β_a, β_y and Σ) are known. (2) Derive the relevant predictor and the variance of the prediction error under this assumption. (3) In the expressions derived under (2), replace the unknown parameters by estimates. This is the approach typically adopted in time-series analysis when autoregressive and/or moving-average models are used for forecasting. See, for example, Judge et al. (1988, p.705-713). Under case (ii) we recognize that β_a and β_y (but not the elements in Σ) are unknown before a predictor is chosen, and derive the variance of the prediction error under this assumption. In this case only Σ and its elements are replaced by estimates to make the predictor operational. This is the approach typically taken for best linear unbiased prediction in generalized least squares models. See, for example, Judge et al. (1988, p.343-346). Under case (iii) where we recognize uncertainty in both $\left(\beta_a, \beta_y\right)$ and Σ the finite sample properties of the predictors appear intractable. We therefore do not consider

this third case in this paper. The properties of the predictors of cases (i) and (ii) are discussed below.

2.1 Case (i): All Parameters Assumed Known

When all parameters are assumed known, the minimum variance predictor for q is the expectation of equation (9) that is given by

$$\hat{q}_1 = E(q) = \left(x_a'\beta_a\right)\left(x_y'\beta_y\right) + \sigma_{ay} \tag{10}$$

This predictor is an unbiased predictor in the sense that the expectation of its prediction error is zero. That is, $E(\hat{q}_1 - q) = 0$. It is made operational by replacing β_a, β_y and σ_{ay} by their estimates $\hat{\beta}_a, \hat{\beta}_y$ and $\hat{\sigma}_{ay}$. Compared to what might be termed the naive predictor $\left(x_a'\beta_a\right)\left(x_y'\beta_y\right)$, note the existence of the additional term σ_{ay}. A positive correlation between the errors implies that, on average, their product will be positive, and conversely for negatively correlated errors. In the Appendix we show that variance of the prediction error is given by

$$E\left[\left(\hat{q}_1 - q\right)^2\right] = \left(x_a'\beta_a\right)^2\sigma_{yy} + \left(x_y'\beta_y\right)^2\sigma_{aa} + 2\left(x_a'\beta_a\right)\left(x_y'\beta_y\right)\sigma_{ay} + \sigma_{aa}\sigma_{yy} + \sigma_{ay}^2 \tag{11}$$

The square root of this quantity is the standard error of the prediction error that can be used in conjunction with $\hat{q}_1$ to form a confidence interval for future output. Since, in

practice, β_a, β_y and the elements in Σ are replaced by consistent estimates, such a confidence interval will be a large sample approximate one.

2.2 Case (ii): Only Σ Assumed Known

The expression for the prediction error variance in equation (11) recognizes uncertainty about the values of the future errors u_a and u_y but it does not recognize the sampling error that occurs in the estimation of β_a, β_y and Σ. To recognize the uncertainty in β_a and β_y it is natural to suggest the predictor

$$q^* = \left(x_a'\hat{\beta}_a\right)\left(x_y'\hat{\beta}_y\right) + \sigma_{ay} \tag{12}$$

However, this predictor is biased because

$$E\left[\left(x_a'\hat{\beta}_a\right)\left(x_y'\hat{\beta}_y\right)\right] \neq \left(x_a'\beta_a\right)\left(x_y'\beta_y\right). \tag{13}$$

In the Appendix we show that

$$E\left[\left(x_a'\hat{\beta}_a\right)\left(x_y'\hat{\beta}_y\right)\right] = \left(x_a'\beta_a\right)\left(x_y'\beta_y\right) + x_a'V_{ay}x_y \tag{14}$$

Consequently, a predictor $\hat{q}_2$ that recognizes uncertainty in the estimation of β_a and β_y, and that is unbiased in the sense that the expectation of its prediction error is zero, $E(\hat{q}_2 - q) = 0$, is

$$\hat{q}_2 = \left(x_a'\hat{\beta}_a\right)\left(x_y'\hat{\beta}_y\right) - x_a'V_{ay}x_y + \sigma_{ay} \tag{15}$$

Furthermore, from the Appendix we see that the variance of its prediction error is

$$(16)\quad E\left[(\hat{q}_2 - q)^2\right] = (x_a'\beta_a)^2 x_y' V_{yy} x_y + (x_y'\beta_y)^2 x_a' V_{aa} x_a$$
$$+ 2(x_a'\beta_a)(x_y'\beta_y) x_a' V_{ay} x_y + (x_a' V_{ay} x_y)^2 + x_a' V_{aa} x_a x_y' V_{yy} x_y + (x_a'\beta_a)^2 \sigma_{yy}$$
$$+ (x_y'\beta_y)^2 \sigma_{aa} + 2(x_a'\beta_a)(x_y'\beta_y)\sigma_{ay} + \sigma_{aa}\sigma_{yy} + \sigma_{ay}^2$$

The first five terms in this expression represent the added uncertainty associated with estimation of β_a and β_y; the last five terms are identical to $E\left[(\hat{q}_1 - q)^2\right]$. For computational purposes a partial matrix algebra representation of (16) might be convenient. It can be shown that

$$(17)\quad E\left[(\hat{q}_2 - q)^2\right] = z'\left[\Sigma + x'\left[X'(\Sigma^{-1} \otimes I)X\right]^{-1} x\right] z + x_a' V_{aa} x_a x_y' V_{yy} x_y$$
$$+ (x_a' V_{ay} x_y)^2 + \sigma_{aa}\sigma_{yy} + \sigma_{ay}^2$$

where

$$(18)\quad \underset{2\times 1}{z} = \begin{pmatrix} x_y'\beta_y \\ x_a'\beta_a \end{pmatrix} \qquad \underset{(K_a+K_y)\times 2}{X} = \begin{pmatrix} x_a & 0 \\ 0 & x_y \end{pmatrix}$$

After replacing unknown parameters with their estimates, the predictor $\hat{q}_2$ and the corresponding standard error of prediction error, calculated as the square root of (16) or (17), can be used to construct a large sample approximate confidence interval for next period's output. Since $\hat{q}_2$ and its standard error reflect uncertainty in the estimation of β_a and β_y, we would expect this predictor to lead to a better approximation in finite samples than $\hat{q}_1$ and its standard error.

2.3 Uncorrelated Errors

If it happens that the error for the yield equation is uncorrelated with the error for the area equation, then the expressions for the predictors and the variances of the prediction errors simplify considerably. For the case of known parameters we have

(19) $$\hat{q}_1 = (x_a'\beta_a)(x_y'\beta_y)$$

(20) $$E\left[(\hat{q}_1 - q)^2\right] = (x_a'\beta_a)^2\sigma_{yy} + (x_y'\beta_y)^2\sigma_{aa} + \sigma_{aa}\sigma_{yy}$$

Recognizing uncertainty in the estimation of β_a and β_y yields

(21) $$\hat{q}_2 = (x_a'\hat{\beta}_a)(x_y'\hat{\beta}_y)$$

Note that, in this case, the naive predictor is the natural one and that $\hat{\beta}_a$ and $\hat{\beta}_y$ are the ordinary least squares estimators $\hat{\beta}_a = (X_a'X_a)^{-1}X_a'A$ and $\hat{\beta}_y = (X_y'X_y)^{-1}X_y'Y$. Also, using the fact that $V_{aa} = \sigma_{aa}(X_a'X_a)^{-1}$ and $V_{yy} = \sigma_{yy}(X_y'X_y)^{-1}$, the variance of the prediction error is given by

(22) $$E\left[(\hat{q}_2 - q)^2\right] = \sigma_{yy}(x_a'\beta_a)^2\left[1 + x_y'(X_y'X_y)^{-1}x_y\right]$$
$$+ \sigma_{aa}(x_y'\beta_y)^2\left[1 + x_a'(X_a'X_a)^{-1}x_a\right]$$
$$+ \sigma_{aa}\sigma_{yy}\left[1 + x_a'(X_a'X_a)^{-1}x_a x_y'(X_y'X_y)^{-1}x_y\right]$$

Whether one opts for the uncorrelated error version in equations (21) and (22) or the correlated error version in (15) and (16) will depend on whether there are likely to be common omitted factors that influence both yield and area. For those who prefer to base their decision on an hypothesis test, the Lagrange multiplier test suggested by Breusch and Pagan (1980) can be employed. Under the null hypothesis $H_0: \sigma_{ay} = 0$, the statistic $\lambda = T\hat{\sigma}_{ay}^2 / \hat{\sigma}_{aa}\hat{\sigma}_{yy}$ has an approximate chi-square distribution with one degree of freedom. This statistic, as well as the likelihood ratio test statistic value for the same hypothesis are routinely printed by the computer soft ware SHAZAM (1993).

3. An Example

To illustrate the methodology we used data collected from the Corrigin Shire in Western Australia. This Shire is located approximately 200 km east of Perth. It is a typical wheat-sheep broadacre farming area with predominantly winter rainfall and an average annual rainfall of 365 mm. Equations for wheat yield and for area sown to wheat were specified. The explanatory variables used in the area equation were

1. lagged area (A_{t-1}): This variable is typically included in supply response functions of this nature to reflect partial adjustment towards a desired area, the partial adjustment being attributable to inability to make short-run changes to fixed input levels.
2. lagged yield (Y_{t-1}): The yield achieved in the previous year has been included to pick up the 'catch up' effect. It appears to be a widely held belief that following a poor year farmers tend to plant more wheat. Some argue that the farmers wish to replace grain reserves run down during the poor year. Others argue that farmers

wish to catch up on lost income. Some also argue that during a poor year the applied fertiliser is not fully utilised, hence another wheat crop is put in the next year to use up the unused nutrients. It is also argued that seed bed preparation after a poor year is much cheaper due to smaller weed populations, hence encouraging an increase in plantings.

3. lagged price (P_{t-1}): We assume that, prior to planting, the wheat grower bases his/her estimate of the likely wheat price at harvest time largely on the price received for the previous crop. Based on past studies, and on what he sees as the likely behaviour of farmers, Coelli (1992) argues for this choice in preference to more complex adaptive or rational expectations structures.

4. lagged input prices (PI_{t-1}): An index of input prices lagged by one period was included with the expectation that a rise in the price of inputs would have a negative influence upon the area planted. As the wheat crops are generally planted in May or June of one financial year and harvested in November or December of the following financial year, the use of lagged input prices appeared appropriate.

5. quota dummy (D_t): A dummy variable which takes the value 1 in the years 1970 and 1971 and 0 elsewhere was introduced to capture the likely negative effect of quota restrictions that were in place during those years (see Fisher 1975).

6. trend (T_t): A linear time trend was included in the area equation to attempt to proxy factors such as the release of land for agriculture and technological change which may have a systematic effect upon the area planted to wheat in a shire. A quadratic effect was initially considered but the squared term was omitted because of its statistical insignificance.

7. rainfall at sowing time (RS_t): Breaking rains are likely to be an important variable influencing farmers' decisions to plant wheat. Rainfall for the 3-month period April, May and June has been included, along with its squared term (RS_t^2), the squared term being introduced to allow for the possibility of diminishing returns to breaking rains. Rainfall was expressed as a ratio, relative to the average rainfall over the sample period.

To explain average wheat yield in Corrigin Shire, we used monthly rainfalls and a linear trend variable to reflect technological change. The distribution of rainfall is important since rainfall during the germination, growing and flowering periods is necessary, and is likely to have differing effects on yield. For this reason monthly rainfalls, from May through to October ($R5_t, R6_t, \ldots, R10_t$) as well as their squares ($R5_t^2, R6_t^2, \ldots, R10_t^2$), were included. The estimated coefficients of the August rainfall variables were observed to be very small relative to the estimates for the other months, highly insignificant, and of the incorrect sign. An examination of the rainfall pattern over the sample period for June and July indicated good reliable rainfall, suggesting soil moisture was unlikely to be a limiting factor in August. The August rainfall variables were therefore dropped from the yield equation and the system of equations re-estimated.

The data set consisted of 39 observations for the period 1950-88. Details of the data sources are described in Coelli (1992). Briefly, observations on yields and areas planted were taken from various publications of the Australian Bureau of Statistics. Rainfall figures were those recorded at the Corrigin Post Office. Wheat price and the

general input price index were taken from the Commodity Statistical Bulletin published by the Australian Bureau of Agricultural and Resource Economics.

The area and yield equations were initially estimated using 37 observations, the first observation being dropped to allow for lagged variables and the last observation being dropped to permit a comparison of actual and predicted values. The equations were estimated as a two-equation seemingly unrelated regression model allowing for contemporaneous correlation between the errors that may result from common omitted influences. The estimated equations, with standard errors in parentheses, are:

$$\begin{aligned}\hat{A}_t = {} & 9.48 + \underset{(8.03)}{} \\ & \underset{(0.088)}{0.631\ A_{t-1}} - \underset{(3.0)}{14.3\ Y_{t-1}} + \underset{(0.048)}{0.183\ P_{t-1}} \\ & - \underset{(0.049)}{0.239\ PI_{t-1}} - \underset{(3.2)}{12.1\ D_t} + \underset{(0.156)}{0.994\ T_t} \\ & + \underset{(11.9)}{16.6\ RS_t} - \underset{(5.36)}{7.30\ RS_t^2} \\ & R^2 = 0.954\end{aligned}$$

$$\begin{aligned}\hat{Y}_t = {} & -0.186 + \underset{(0.323)}{} \\ & \underset{(0.0033)}{0.0156\ T_t} + \underset{(0.209)}{0.434\ R5_t} - \underset{(0.090)}{0.133\ R5_t^2} \\ & + \underset{(0.283)}{0.527\ R6_t} - \underset{(0.113)}{0.196\ R6_t^2} + \underset{(0.255)}{0.360\ R7_t} - \underset{(0.107)}{0.208\ R7_t^2} \\ & + \underset{(0.199)}{0.427\ R9_t} - \underset{(0.0693)}{0.0597\ R9_t^2} + \underset{(0.166)}{0.153\ R10_t} - \underset{(0.0600)}{0.0705\ R10_t^2} \\ & R^2 = 0.680\end{aligned}$$

In the area equation all estimated coefficients have the correct signs and, with the exception of the rainfall variables, are significantly different from zero at a 5%

significance level. Although we have not been able to obtain precise estimates of the effects of the rainfall variables, they are retained because of their obvious importance. A similar remark can be made about the various rainfall variables in the yield equation. Few are significantly different from zero, but all are obviously important, and they do have the correct signs. Evidence on the existence of contemporaneous correlation between the errors was not conclusive. The Lagrange multiplier test statistic value of 3.07 was less than the 5% critical value of 3.84 from the $\chi^2_{(1)}$ distribution, but the likelihood ratio test value of 7.09 was greater. We retained the assumption of contemporaneously correlated errors for our calculations of the predictions and their standard errors. To check for autocorrelated errors we estimated each equation separately and computed the values of the Durbin-Watson statistic, and Durbin's h-statistic. There was no evidence of autocorrelation.

The estimated area and yield equations were used to predict area and yield for the next period and to compute values for the various predictors of output and their standard errors. Also, the equations were re-estimated an additional 4 times, omitting the last 5, 4, 3 and 2 observations, respectively. In each case one-step ahead predictions were made.

The various results appear in Table 1. There are a number of observations we can make. First, the naive predictor $\left(x_a'\hat{\beta}_a\right)\left(x_y'\hat{\beta}_y\right)$, the predictor that does not recognize any parameter uncertainty $\hat{q}_1$, and the predictor that recognizes coefficient uncertainty $\hat{q}_2$, all give essentially the same predictions. The values for $\hat{q}_1$ are slightly higher than those for $\left(x_a'\hat{\beta}_a\right)\left(x_y'\hat{\beta}_y\right)$, reflecting a positive value for the error correlation

Table 1

One-Step Ahead Predictions for Output With
Corresponding Standard Errors and Confidence Intervals

	Year				
	1984	1985	1986	1987	1988
Actual Output	175 8	80 3	134 8	128.7	156.9
$\left(x_a'\hat{\beta}_a\right)\left(x_y'\hat{\beta}_y\right)$	144 0	114.0	126.3	124.2	119.5
$\hat{q}_1$	144 6	114.5	126.8	124.5	119.8
$se(\hat{q}_1 - q)$	23 2	20 4	19.7	19 8	17.4
95% CI	(99.1,190.1)	(74 5,154.5)	(88.2,165 4)	(85.7,163.3)	(85.7,153.9)
$\hat{q}_2$	144.5	114.4	126.7	124.5	119.8
$se(\hat{q}_2 - q)$	25.8	23.4	23.2	22.6	19.6
95% CI	(93.9,195.1)	(68 5,160.3)	(81.2,172.2)	(80 2,168.8)	(81.4,158.2)

σ_{ay}, and the values for $\hat{q}_2$ are very slightly less than those for $\hat{q}_1$. In general, we would expect the predictors to give similar values when the estimated equations are good fits with high R^2's. Under these circumstances that part of the prediction attributable to the systematic components $\left(x_a'\hat{\beta}_a\right)$ and $\left(x_y'\hat{\beta}_y\right)$ will be large relative to the covariance between the errors. The predictors are likely to yield different predictions when the equations are poor fits, and the contemporaneous error correlation is high.

Although the new predictors we have derived do not, in this particular case, yield results very different from the so-called naive predictions, it is important to assess the reliability of predictions, and the expressions for standard errors that we have derived

are useful for this purpose. Values for these standard errors, and the 95% prediction confidence intervals derived from them, also appear in Table 1. With the exception of the interval derived from $\hat{q}_1$ in 1988, each interval contains the corresponding actual output for that year. Thus, although some of the predictions miss the mark rather badly, if a proper assessment of the reliability of the predictions is given, the realizations of output should not generate surprise. Note that the standard errors for $\hat{q}_2$ are only slightly greater than those for $\hat{q}_1$, indicating that most of the prediction uncertainty originates from equation error uncertainty not coefficient uncertainty.

Summary and Conclusions

Although the problem of predicting output from separate yield and area equations is a common one, issues relating to choice of an appropriate predictor, and the standard error of the prediction error, seem to have been neglected in the literature. We have attempted to fill this void within the context of the general seemingly unrelated regressions model. Results for the case where the errors of the yield and area equations are uncorrelated emerge as a special case. In our empirical example the correlated errors had little bearing on the predictions, but it was clear that assessment of the reliability of the predictions, through computation of appropriate standard errors, was important.

Although the results that are derived are exact finite sample results, they lose their exact finite sample applicability when unknown parameters are replaced with estimates. Apart from the use of large sample approximations, there does not seem to be any easy solution to this problem within a sampling theory framework. From a Bayesian perspective, however, estimation of the predictive probability density

function for output, and its mean and variance, does not present a problem. Research in this direction is in progress.

Finally, it is worth noting that the methodology introduced in this paper has wider applicability than is suggested by the empirical example. It can be used not only for predicting the product of area and yield, but also for predicting the product of any two dependent variables within a regression framework, whether or not they can be classified as "seemingly unrelated". For example, for predicting the total quantity of sawntimber used for dwellings, the Australian Bureau of Agricultural and Resource Economics (1989) estimates separate equations to explain the quantity of timber used per dwelling, and the number of new dwellings. They do not seem to have considered methodology like ours for the construction of their predictions and associated standard errors.

Appendix

To economize on repetitive symbols, let

$$z_a = x_a'\beta_a \qquad z_y = x_y'\beta_y$$

$$\hat{z}_a = x_a'\hat{\beta}_a \qquad \hat{z}_y = x_y'\hat{\beta}_y$$

Our first task is to derive the prediction–error variance given in equation (11). Working in this direction, we have

(A1)
$$\left(\hat{q}_1 - q\right)^2 = \left(z_a u_y + z_y u_a + u_a u_y - \sigma_{ay}\right)^2$$

$$= z_a^2 u_y^2 + z_y^2 u_a^2 + \left(u_a u_y - \sigma_{ay}\right)^2 + 2z_a z_y u_a u_y + 2z_a u_a u_y^2 - 2z_a \sigma_{ay} u_y + 2z_y u_a^2 u_y - 2z_y \sigma_{ay} u_a$$

Using the fact that all third moments for the bivariate normal distribution are zero, the expectation of this quantity is

(A2)
$$E\left[\left(\hat{q}_1 - q\right)^2\right] = z_a^2 \sigma_{yy} + z_y^2 \sigma_{aa} + \mathrm{var}\left(u_a u_y\right) + 2z_a z_y \sigma_{ay}$$

From equation (6) in Bohrnstedt and Goldberger (1969),

(A3)
$$\mathrm{var}\left(u_a u_y\right) = \sigma_{aa}\sigma_{yy} + \sigma_{ay}^2$$

and hence

$$\text{(A4)} \qquad E\left[\left(\hat{q}_1 - q\right)^2\right] = z_a^2\sigma_{yy} + z_y^2\sigma_{aa} + 2z_a z_y\sigma_{ay} + \sigma_{aa}\sigma_{yy} + \sigma_{ay}^2$$

which agrees with equation (11).

Moving to the predictor which recognizes uncertainty in the estimation of β_a and β_y, we wish to evaluate

$$\text{(A5)} \qquad E\left(\hat{z}_a\hat{z}_y\right) = E\left(x_a'\hat{\beta}_a\hat{\beta}_y'x_y\right)$$

$$= E\left[\left(x_a' \quad 0'\right)\hat{\beta}\hat{\beta}'\begin{pmatrix}0\\ x_y\end{pmatrix}\right]$$

Now,

$$\text{(A6)} \qquad E\left[\hat{\beta}\hat{\beta}'\right] = \mathrm{Var}(\hat{\beta}) + E(\hat{\beta})E(\hat{\beta})'$$

$$= \left[X'\left(\Sigma^{-1}\otimes I\right)X\right]^{-1} + \beta\beta'$$

where $\mathrm{Var}(\cdot)$ denotes the covariance matrix. Substituting (A6) into (A5) gives

$$\text{(A7)} \qquad E(\hat{z}_a\hat{z}_y) = \left(x_a' \quad 0'\right)\left[X'\left(\Sigma^{-1}\otimes I\right)X\right]^{-1}\begin{pmatrix}0\\ x_y\end{pmatrix} + \left(x_a' \quad 0'\right)\beta\beta'\begin{pmatrix}0\\ x_y\end{pmatrix}$$

$$= x_a'V_{ay}x_y + z_a z_y$$

which is the result in equation (14).

Finally, to derive the prediction-error variance in equation (16),

$$(A8) \qquad \left(\hat{q}_2 - q\right)^2 = \left[\hat{z}_a \hat{z}_y - \left(z_a z_y + x_a' V_{ay} x_y\right) - z_a u_y - z_y u_a - \left(u_a u_y - \sigma_{ay}\right)\right]^2$$

Retaining only those terms with nonzero expectations, we have

$$(A9) \qquad E\left[\left(\hat{q}_2 - q\right)^2\right] = Var\left(\hat{z}_a \hat{z}_y\right) + z_a^2 \sigma_{yy} + z_y^2 \sigma_{aa} + 2 z_a z_y \sigma_{ay} + Var\left(u_a u_y\right)$$

Again utilizing equation (6) in Bohrnstedt and Goldberger (1969) yields

$$\begin{aligned}(A10) \quad E\left[\left(\hat{q}_2 - q\right)^2\right] &= z_a^2 x_y' V_{yy} x_y + z_y^2 x_a' V_{aa} x_a + 2 z_a z_y x_a' V_{ay} x_y \\ &+ x_a' V_{aa} x_a x_y' V_{yy} x_y + \left(x_a' V_{ay} x_y\right)^2 + z_a^2 \sigma_{yy} + z_y^2 \sigma_{aa} \\ &+ 2 z_a z_y \sigma_{ay} + \sigma_{aa} \sigma_{yy} + \sigma_{ay}^2\end{aligned}$$

This is the result in equation (16).

References

Anderson, K. (1974), 'Distributed Lags and Barley Acreage Response Analysis', Australian Journal of Agricultural Economics 18, 119–132.

Australian Bureau of Agricultural and Resource Economics (1989), 'Long Term Consumption Projections for Forest Products', Discussion Paper No.89.6, Canberra.

Bohrnstedt, G.W. and Goldberger, A.S. (1969), 'On the Exact Covariance of Products of Random Variables', Journal of the American Statistical Association 64, 1439–1442.

Breusch, T.S. and Pagan, A.R. (1980), 'The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics', Review of Economic Studies 47, 239–254.

Coelli, T.J. (1992), 'Forecasting Wheat Production Using Shire Level Data', Paper contributed to the 36th Annual Conference of the Australian Agricultural Economics Society, ANU, Canberra.

Del Valle, M. and Ray, D.E. (1990), 'Explaining Oklahoma Wheat Yields Using Monthly Precipitation and Temperature Averages', Current Farm Economics, Oklahoma Agricultural Experiment Station, 29–35.

Dillon, J.L. and Anderson, J.R. (1990), The Analysis of Response in Crop and Livestock Production, 3rd Edition, Pergamon, UK.

Fisher, B.S. (1975), 'Supply Response in the Wheat–Belt of South-Eastern Australia: The Impact of Delivery Quotas on Wheat Plantings', Australian Journal of Agricultural Economics, 19, 81–93.

Fisher, B.S. (1978), 'An Analysis of Storage Policies for the Australian Wheat Industry', Research Report No. 7, Department of Agricultural Economics, University of Sydney.

Fisher, B.S. and Munro, R.G. (1983), 'Supply Response in the Australian Extensive Livestock and Cropping Industries: A Study of Intentions and Expectations', Australian Journal of Agricultural Economics, 27, 1-11.

Francisco, E.M. and Guise, J.W.B. (1988), 'A Note on Establishing Yield–Rainfall Relationships', Agricultural and Forest Meteorology 42, 75–81.

Griffiths, W.E. and Anderson, J.R. (1978), 'Specification of Agricultural Supply Functions – Empirical Evidence on Wheat in Southern NSW', Australian Journal of Agricultural Economics 22, 115–128

Guise, J.W.B. (1969), 'Factors Associated With Variations in the Aggregate Average Yield of New Zealand Wheat (1918-1967)', American Journal of Agricultural Economics 51, 866–881.

Judge, G.G., Hill, R.C., Griffiths, W.E., Lütkepohl, H. and Lee, T.–C. (1988), Introduction to the Theory and Practice of Econometrics, 2nd edition, John Wiley and Sons, New York.

Sanderson, B.A., Quilkey, J.J. and Freebairn, J.W. (1980), 'Supply Response of Australian Wheatgrowers', Australian Journal of Agricultural Economics, 24, 129–140.

SHAZAM (1993), User's Reference Manual 7.0, McGraw–Hill, New York.

Zellner, A. (1962), 'An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias', Journal of the American Statistical Association 57, 348–368.