RELATIVE SENSITIVITY OF SELECTED DISTRIBUTED LAG ESTIMATORS TO MEASUREMENT AND SPECIFICATION ERROR

Larry L. Bauer, S. Darrell Mundy, and Charles B. Sappington

The notion that one economic variable has an impact on another over several time periods is not uncommon to economic thought. Irving Fisher [5] was the first to introduce empirically the notion of distributed lags using different shapes of probability curves. Alt [2] continued the work, but made no *a priori* assumptions about the distribution of the lag as Fisher had done. Koyck [11] and Nerlove [13] returned to and extended the work of Fisher assuming *a priori* a monotonically declining distribution. More recently Jorgenson [10], Almon [1], and Johnson [8] have developed estimators which, like those of Alt, assume only a continuous lag function of unspecified shape. There is evidence, however, that the Jorgenson and Johnson techniques are extremely sensitive to errors of measurement and specification because in applied work, reasonable estimates of lag distributions occur infrequently [7, 12].

The purpose of this article is to estimate the relative sensitivity of four commonly used distributed lag estimation techniques. Sensitivity is measured by the ability of the technique to estimate known parameters when increasing amounts of error are added to some or all of the variables. In the real world, one particular technique usually is employed to estimate coefficients which then are judged by certain criteria, such as reasonable magnitudes and proper signs, as being acceptable. Possibly, a set of coefficients which is otherwise judged acceptable would include coefficients with magnitudes and present structures that are, if the true model were known, not very accurate. It would be useful to know which one of several possible techniques is most likely to give acceptable and accurate results.

PROCEDURE

Three sets of "true" or error-free time series data were generated, each with a lag distribution of different configuration. Four different techniques of estimating lag distributions were used on each set of "true" data. Random error was added to the dependent variable at 1, 3, 5, and 7 percent levels to simulate omitted variables or to otherwise represent errors in measurement. Finally, random error of 1, 3, 5, and 7 percent was used in both the dependent and the independent variables to simulate measurement and/or specification errors. For each of the four techniques, 27 runs, or nine per configuration, were performed: one for the no error, one for each of the four levels of error on the dependent variable only, and one for each of the four levels of error on both the dependent and independent variables. For one of the techniques each "run" involved 36 separate "subruns."

Configuration of Generated Lag Distributions

The three configurations of lag distribution in the generated data were monotonically declining, symmetric, and skewed. The monotonically declining distribution is represented by the geometric function generally associated with Koyck. A quadratic function was used to represent the symmetric or inverted "U" distribution. The skewed distribution was skewed to the right. These three configurations are typical of those most likely to be found in applied economic work.

Methods of Restricting Parameter Space of General Distributed Lag Function

Some serious problems arise at the outset in the empirical estimation of a distributed lag function. The distributed lag function in equation 1 cannot be estimated directly because there is an infinite number of regressors.

$$(1) \quad Y_t = \beta\{w_0X_t + w_1X_{t-1} + w_2X_{t-2} + \dots\} + U_t$$

$$w_i \geq 0 \text{ and } \sum_{i=0}^{\infty} w_i = 1$$

However, in empirical analysis a finite number of observations, and therefore degrees of freedom, are available and equation 1 cannot be

Larry L. Bauer is Associate Professor, Department of Agricultural Economics, Clemson University; S. Darrell Mundy is Associate Professor and Charles B. Sappington is Professor, Department of Agricultural Economics, University of Tennessee. The authors are listed alphabetically.

estimated in its present form. Also, multicollinearity most likely would occur among successive regressors. This equation therefore must be transformed, or reparameterized, to a working model which contains a finite number of variables.[1] in transforming the equation, one must make some reasonable yet restrictive assumptions about the pattern of the weights. The objective is to choose assumptions which allow the individual coefficients of the lag distribution to depend on a few parameters which would be relatively simple to estimate.

Three different methods of restricting the parameter space of a distributed lag function have been used in economic research: (1) the introduction of a finite number of lagged exogenous variables into the model, e.g., Almon and Johnson, (2) the introduction of lagged endogenous variables by using a reduced equation, e.g., Koyck, and occasionally (3) a method involving a combination of the first two e.g., Jorgenson.

Description of Estimation Techniques

The four estimation techniques considered in this analysis are: (1) a least squares technique which directly estimates the conceptual model without reducing the parameter space (lagged exogenous in the statistical model), (2) the rational polynomial method (both lagged endogenous and exogenous variables), (3) the general polynomial method (lagged exogenous variables), and (4) the weighted average method (lagged exogenous variables). The last three methods all involve reparameterizing the conceptual model by reducing its parameter space. The first three do not impose *a priori* restrictions on the configuration of the lag distribution whereas the weighted average method allows for the selection of the "best" lag configuration from among many shapes and lengths of lags, all of which are reasonable on the basis of judgment or *a priori* knowledge.

Least squares method. The least squares technique applied to lag models is attributed to Tinbergen [14] and merely adds lagged variables sequentially until "the signs of the coefficients become erratic and cease to make sense."

Rational polynomial method. The rational polynomial technique was developed by Jorgenson [10, 6] and specifies that if in the relationship

$$Y_t = a + b_1 X_t + b_2 Z_t \tag{2}$$

a lagged response is hypothesized between X and Y, a lag distribution is imposed on X by

$$Y_t = a + b_1 \frac{N(L)}{D(L)} X_t + b_2 Z_t \tag{3}$$

where N(L) and D(L) are polynomials in the lag operator L and can be of any degree.[2] For example, if both are first-degree polynomials,

$$\frac{N(L)}{D(L)} = \frac{1 + cL}{1 - dL}$$

then

$$Y_t = a + b_1 \frac{1 + cL}{1 - dL} X_t + b_2 Z_t$$

and when both sides are multiplied by D(L), then

$$Y_t = a - ad + dY_{t-1} + b_1 X_t + cb_1 X_{t-1} + b_2 Z_t - db_2 Z_{t-1}. \tag{4}$$

This is the equation to be estimated; but because it is nonlinear in the parameters, an appropriate nonlinear regression technique must be used.

This technique estimates the parameter of the closed form of an infinite distribution where the closed form is the ratio of two polynomials, generally of third degree or less.

Estimation is accomplished by imposing ratios with different degrees of polynomial in the numerator and denominator and choosing one on the basis of reasonableness of coefficients and R^2.

General polynomial method. Whereas the rational polynomial techniques yield an infinite lag distribution, the general polynomial technique [8] estimates a distribution of predetermined length. With this technique, the weights of the lag distribution are represented by a polynomial, for example,

$$w(t) = a_0 + at_1 + a_2 t^2$$

where t is the lag time. In matrix notation this becomes:

$$\tilde{w} = \tilde{s}\,\tilde{a} \text{ or } \begin{bmatrix} w(0) \\ w(1) \\ w(2) \\ w(3) \\ w(4) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

where in this example the n by p matrix $\tilde{s}$ represents n = 5 or the current observation plus four lagged years and p = 3 is the number of parameters used to generate the lag

[1]The same problems would be likely to occur with a finite distribution, particularly if the hypothesized lag is relatively long.

[2]The lag operator L is a shorthand technique whereby: $LX_t = X_{t-1}$, $L^2X_t = X_{t-2}$.

distribution of length n. With a basic structure of

$$\tilde{y} = \tilde{x}\,\tilde{w} \tag{5}$$

where, for simplification, X is the only independent variable, and it has a lag distribution imposed on it by the vector $\tilde{w}$, estimation is carried out as follows.

Because

$$\tilde{w} = \tilde{s}\,\tilde{a}$$

the structure becomes

$$\tilde{y} = \tilde{x}\,\tilde{s}\,\tilde{a}. \tag{6}$$

Estimation can be carried out as

$$\tilde{y} = \tilde{v}\,\tilde{a} \tag{7}$$

where

$$\tilde{v} = \tilde{x}\,\tilde{s}.$$

In essence, the data are transformed with the matrix $\tilde{s}$ and reduced from n variables to p variables. The estimated a's are used to derive the n coefficients in the lag distribution.

Weighted average method. The fourth technique used is the weighted average approach. With this method, the weights of a predetermined distribution are used to calculate a single variable which is essentially a weighted average of all the past data. As an example, the single variable calculated for one of the lag functions used in this study is

$$X_t^{(1)} = .60X_t + .24X_{t-1} + .096X_{t-2} + .0384X_{t-3} + .0154X_{t-4} + .0061X_{t-5} + .0025X_{t-6} + .0010X_{t-7} + .0004X_{t-8} + .0002X_{t-9}. \tag{8}$$

Other single variables, $X_t^{(2)}$ and $X_t^{(3)}$, were calculated for a steeper and less steep monotonically declining function of the same length lag. Then $X_t^{(4)}$, $X_t^{(5)}$ and $X_t^{(6)}$ were calculated for a steep, middle, and less steep monotonically declining function with a lag of 12 periods; three more with a eight-year lag, and the last three for a six-year lag or a total of 12 montonically declining functions. Twelve single variables were similarly calculated for the skewed configuration as well as 12 for the geometric distribution. Then the $X_t^{(i)}$'s (i = 1, 2 ... 36) were used, in turn, in the estimating equation

$$Y_t = a + b_1X_t^{(i)} + b_2Z_t \tag{9}$$

to determine whether the method would allow for the selection of the one proper shape and length of lag from among the 36.

Data

The "true" error-free model for this experiment is:

$$Y = a + b_zZ + b_0X_t + b_1X_{t-1} + ... + b_9X_{t-9}. \tag{10}$$

The coefficients, b_0 through b_9, are positive; that is, the "true" length of the lag is 10 time periods for each of the three distributions. The variable Z is a trend variable with the values 1, 2, 3, ..., 30. In the error-free case, the dependent variable for each observation was generated from the equation using the known or "true" coefficients for each independent variable.[3] The values of the constant terms were selected such that the means of Y were equal for the three sets of data. The true overall regression coefficient or the total effect of the X variable is the sum of b_0 through b_9 and the weights of the lag distribution are determined by dividing each of the b's by this sum.

A random normal table was used to add error, so the population from which a sample was drawn would have E(u) = 0 and E(uu') = $\sigma^2 I$ which are assumed in all four estimating techniques. Any small sample would not be likely to have these same attributes, but knowing these population characteristics is an advantage over having to assume them.[4]

Measurement error could reasonably be expected to be proportional to the magnitude of the variable involved. For this reason, the errors were specified as a percentage of the particular observation to which the drawing applied. This process would be expected to introduce some degree of heteroscedasticity into the relationships; but in real world data, heteroscedasticity is probably as common as homoscedasticity. Thus, the drawing for the errors was changed to percentages. For example, if a drawing were −2.000 for the 1 percent error case, that particular observation on that particular variable was 98 percent of the value of the true value. The drawings were conducted for 30 observations on Z and 30 for each

[3]To add a realistic amount of multicollinearity, data on X are actually annual hog prices from 1942 through 1971.

[4]Two of the anonymous reviewers suggested that the mean of the sample error term should be reported. The mean is not reported because, in applied empirical work, it is never known; additionally, small sample properties are not within the scope of this article.

of the three dependent variables. For the lagged variable, X, it was necessary to have data lagged in excess of 10 periods, the true length, for the general polynomial and weighted average techniques. The data used in this analysis allowed for 19 lagged observations necessitating 49 drawings for the lagged variable. This resulted in a total of 169 drawings for each level of error. Separate drawings were made for each level of error.

Criteria of Satisfactory Results

For the three estimating techniques that place no *a priori* restrictions on the shape of the lag distribution, the usual criteria for evaluating results in empirical application are: (1) reasonableness as an explanation of the phenomenon in question, i.e., signs of coefficients are as expected, (2) a smooth and continuous lag distribution with weights that are all between zero and one and which sum to one, and (3) the magnitude of the coefficient of determination and the overall F ratio.

The requirements for the weighted average technique differ to the extent that *a priori* restrictions, based on the judgment of the researcher, are placed on the configuration of the lag distribution, guaranteeing that criterion 2 is met. However, this criterion is replaced by that of choosing from among the subruns (36 in this study) the largest t-statistic on the coefficient of the weighted average variable subject to the presence of the hypothesized signs for it and the other independent variable.

In this study, estimation was carried out as though the true coefficients were not known. The foregoing criteria were applied to choose the best or most reasonable lag distribution in the same manner as would be done in an actual research situation. The problem then was to decide whether the estimated distributions were satisfactory approximations of the true distribution. This determination had to be made subjectively by comparing the coefficients and the plots of the estimated lag distribution with those of the true distribution.

RESULTS

In empirical economic analysis there is essentially always error of both specification and of measurement in the data. The amount of error necessary to cause an estimation technique to yield unsatisfactory results may not always be relevant. Such is the case in this analysis; however, the relative sensitivity of various estimating techniques to error is of interest here.

The results indicate that whether a particular level of error was in the dependent variable only or in both the dependent and independent variables made no discernible difference in results for the estimators with one exception.[5] The small sample size, 30, may have been at fault in that case but a sample size of 30 is not considered unusually small in applied economic research.

Least squares method. Estimation with ordinary least squares gave perfect estimates of the shape and length of all three configurations in the error-free model. With 1 percent error added, however, the results failed to meet the required criteria of continuousness, positive lag weights, and reasonable results. As would be expected, the same was true for the 3, 5, and 7 percent error levels.

General polynomial method. The general polynomial technique estimated the symmetric distribution perfectly (within rounding error) with the error-free data. However, the technique failed to estimate the monotonically declining and skewed distributions with any reasonable degree of accuracy for the specifications chosen—polynomial degrees of zero, first, second, third, and fourth degrees with nine lagged years of data on X.[6] The results in both cases failed to meet all criteria. A possible reason for the inaccurate estimation of the lag distribution and the other coefficients is misspecification caused by not employing polynomial forms greater than fourth degree. An indication of this possibility occurred with the monotonically declining distribution; there were fewer negative weights as the degree of the polynomial was increased from zero through fourth degree. Also, the fact that the estimates of the weights more nearly approached the true values implies that polynomials of degree higher than four might have improved accuracy.[7] However, no discernible systematic relationship between the degree of polynomial and the number of negative lag weights was found when the skewed distribution was estimated.

Introduction of error into the data in-

[5]Some discernible difference was present with the general polynomial technique. Incorporating error in X as well as Y did compound the usual problems of missed turning points, incorrect magnitudes and signs of the weights, and an increased number of negative weights.

[6]Only the number of lagged years corresponding to the true lag was tried with the general polynomial form. Relatively high sensitivity to errors in specification with respect to functional form (degree of the polynomial) and measurement precluded in this study the consideration of specification error caused by incorrect length of lag.

[7]A popular assumption in using the general polynomial technique is that most economic phenomena can be explained by using relatively simple polynomials of fourth degree or smaller. For example, Chen et al. [4, p. 78] point this out in their use of a low-order general polynomial lag model.

creased the frequency of the problems of missed turning points and incorrect magnitudes and signs on the lag weights. The technique performed reasonably well in estimating the parameters of the symmetric model at all levels of error on X and Y, correctly estimating the signs and turning points but incorrectly estimating the magnitude of the weights. When error was introduced to the data for the monotonically declining and skewed distributions, the results were judged to be unsatisfactory regardless of level of error.

Rational polynomial method. The rational polynomial technique gave perfect results on the error-free data when the true distribution being estimated was either monotonically declining or skewed. When the true shape of the lag was symmetric, however, the Jorgenson method yielded results which were continuous, positive, and reasonable but failed to approximate the configuration of the true lag distribution. Progressing to a third degree polynomial did not improve the estimates.[8]

With the addition of error the rational polynomial technique produced estimates of the monotonically declining and skewed distributions that were continuous, positive, and reasonable; however, the length and shape of the lag distribution were estimated inaccurately as error increased. With 1 and 3 percent error in the variables the results were judged as acceptable; with 5 and 7 percent error the results were not acceptable, especially with respect to the constant term and the coefficient of the nonlagged variable.

Weighted average method. The estimation technique using several (36 in this analysis) predetermined distributions to collapse the variable on which the lagged distribution was imposed perfectly estimated all three of the no-error situations. As error was added, however, the acceptability of the results declined with respect to the estimated value of the constant term and the coefficient of the nonlagged variable. The results were very similar to those for the rational polynomial technique except that the weighted average method successfully estimated the symmetric distribution.

Summary of acceptable results

With error added to the data, both the least squares and general polynomial techniques seemed unable to produce satisfactory results. Therefore, the results of the other two techniques are compared and analyzed. The results from the rational polynomial and the weighted average techniques for the monotonically declining and the skewed distributions with 3 and 5 percent error added both to the dependent and independent variables are shown in Tables 1 and 2. Also presented for purposes of

TABLE 1. TRUE PARAMETERS OF MONOTONICALLY DECLINING CONFIGURATION, RATIONAL POLYNOMIAL AND WEIGHTED AVERAGE ESTIMATES OF PARAMETERS, 3% AND 5% ERROR ADDED TO DEPENDENT AND INDEPENDENT VARIABLES

		3% error		5% error	
Variable	True	Rational polynomial	Weighted average	Rational polynomial	Weighted average
Constant	10	11.6002	11.07	28.1176	22.40
w_0	.60	.6211	.60	.6615	.5949
w_1	.24	.2353	.24	.1949	.2387
w_2	.096	.0892	.096	.0945	.0978
w_3	.0384	.0338	.0384	.0284	.0372
w_4	.0154	.0128	.0154	.0136	.0157
w_5	.0061	.0048	.0061	.0041	.0059
w_6	.0025	.0018	.0025	.0020	.0039
w_7	.0010	.0007	.0010	.0006	.0020
w_8	.0004	.0003	.0004	.0003	.0010
w_9	.0002	.0001	.0002	.0001	.0010
w_{10}	.0000				.0010
b_z	.5	.5116	.39	.5564	.56
overall b	6	5.8639	6	4.7955	5.11

TABLE 2. TRUE PARAMETERS OF SKEWED CONFIGURATION, RATIONAL POLYNOMIAL AND WEIGHTED AVERAGE ESTIMATES OF PARAMETERS, 3% AND 5% ERROR ADDED TO DEPENDENT AND INDEPENDENT VARIABLES

		3% error		5% error	
Variable	True	Rational polynomial	Weighted average	Rational polynomial	Weighted average
Constant	12.79	24.5356	18.47	42.1505	21.96
w_0	.25	.2345	.2486	.3086	.2067
w_1	.2775	.3128	.2777	.3647	.2120
w_2	.2054	.1776	.2051	.1724	.2155
w_3	.1260	.1084	.1270	.0815	.1784
w_4	.0705	.0657	.0708	.0385	.0901
w_5	.0368	.0398	.0363	.0182	.0530
w_6	.0182	.0241	.0181	.0086	.0283
w_7	.0090	.0146	.0091	.0041	.0124
w_8	.0045	.0088	.0054	.0019	.0035
w_9	.0013	.0053	.0018	.0009	

[8]According to Griliches [6, p. 27], "In practice one will not be interested in a...polynomial of higher order than 2 or 3."

Table 2 Continued

Variable	True	3% error Rational polynomial	3% error Weighted average	5% error Rational polynomial	5% error Weighted average
w_{10}		.0032		.0004	
w_{11}		.0019		.0002	
w_{12}		.0011		.0001	
w_{13}		.0007			
w_{14}		.0004			
w_{15}		.0002			
w_{16}		.0001			
b_z	.5	.6477	.7	.6119	.32
overall b	6	5.2228	5.51	4.2348	5.66

comparison are the true coefficients. It should be emphasized that acceptable results were judged by subjective comparison of the estimated coefficients with the true and, in this study, known coefficients. In the tables the 3 percent error results were judged to be acceptable and the 5 percent error results were judged as unacceptable. The 1 percent error situation for the two techniques is not presented because the true coefficients were closely duplicated. As would be expected, the 7 percent error results were even more divergent from the true parameters than were the estimates for the 5 percent case. The results in the 5 and 7 percent error cases met all the criteria for satisfactory results that usually would be available in applied research. However, the results were judged unacceptable because of the wide divergency from the true parameters which are never known in the real world. The results in the table indicated to the authors that, on the basis of statistical observations alone, there is very little difference between the two techniques.

CONCLUSIONS

As error was increased, the results were consistently more divergent from the known parameters. Though results from repeated sampling of the random error at each level of error would be necessary to specify firm conclusions, the consistency of the results does indicate the following rather firm hypotheses which apply at least to small samples.

1. The least squares technique is not satisfactory for the estimation of distributed lags if one assumes that there is error of either specification and/or measurement.
2. The general polynomial technique, at least up through a fourth degree polynomial, is unsatisfactory for the estimation of a distributed lag model under essentially all conditions.
3. The rational polynomial technique will not accurately estimate a symmetric distribution, perhaps because the technique assumes, or is biased toward, an infinite lag. This method works as well as any if the true lag does have a long "tail."
4. In this analysis, the weighted average technique worked as well as the rational polynomial technique and better than the other two methods in all the tested situations.
5. If the only interest is in the "best" estimates of the true parameters, this experiment does not indicate a clear choice between the rational polynomial and the weighted average methods based on the accuracy with which the two estimated the parameters.
6. Because the rational polynomial technique gives positive, continuous, and reasonable results with 7 percent error added to both the dependent and independent variables in the monotonically declining and skewed distributions, the unresolved question arises of how much error is required before these criteria are not met.
7. If acceptable estimates of a distributed lag are of high priority, the weighted average method would be the selected technique. If, in addition, simplicity were desirable, the weighted average method would again be the selected technique. There is a problem of searching for the true distribution and also a problem of when to stop searching.
8. One characteristic of the weighted average technique of possible methodological interest is that the economist is hypothesizing the shape of the lag configuration before statistical estimation rather than "letting the data determine the economic model" as the other techniques allegedly do. An important consideration in light of this characteristic is that the "costs" of making prior assumptions seemingly approached zero with the data sets in this study. An implication is that distributed lag models based on the generation of economic hypotheses prior to statistical estimation perform equally as well empirically as models based on *ex post* statistical results. Hence, methodologically, a rational economic researcher would be most likely to choose the weighted average technique for applied econometric research over the other approaches considered in this study if sufficient data were available to compute the weighted average.

REFERENCES

[1] Almon, S. "The Distributed Lag Between Capital Appropriations and Expenditures," *Econometrica.* Volume 22, 1965, pp. 178-196.

[2] Alt, F. F. "Distributed Lags," *Econometrica,* Volume 10, 1942, pp. 113-128.

[3] Bauer, L. L. "The Effect of Technology on the Farm Labor Market," *American Journal of Agricultural Economics,* Volume 51, 1969, pp. 605-618.

[4] Chen, D., R. Courtney, and A. Schmitz. "A Polynomial Lag Formulation of Milk Production Response," *American Journal of Agricultural Economics,* Volume 54, 1972, pp. 77-83.

[5] Fisher, I. "Our Unstable Dollar and the So-called Business Cycle," *Journal of American Statistics Association,* Volume 20, 1925, pp. 179-202.

[6] Griliches, Z. "Distributed Lags: A Survey." *Econometrica,* Volume 35, 1967, pp. 16-49.

[7] Hileman, A. E. "Time Lags in the Impact of Public Investment in Water Resources: The Tennessee Valley Region: 1936-1968," unpublished Ph.D. dissertation, University of Tennessee, 1971.

[8] Johnson, T. "A Note on Distributed Lags in Polynomial Form," paper presented to an econometric work group at North Carolina State University, 1969.

[9] Johnston, J. *Econometric Methods,* 2nd ed. New York: McGraw-Hill Book Company, Inc., 1960.

[10] Jorgenson, D. W. "Rational Distributed Lag Functions," *Econometrica,* Volume 34, 1966, pp. 135-149.

[11] Koyck, L. M. *Distributed Lags and Investment Analysis.* Amsterdam: North-Holland Publishing Company, 1954.

[12] Mundy, S. D. *Technology and the Farm Labor Market,* unpublished Ph.D. dissertation, University of Tennessee, 1972.

[13] Nerlove, M. *Distributed Lags and Demand Analysis,* USDA Handbook No. 141, 1958.

[14] Tinbergen, J. "Long-Term Trade Elasticities," *Metroeconomica,* Volume 1, 1949, pp. 174-185.