NBER TECHNICAL PAPER SERIES

ESTIMATION OF POLYNOMIAL DISTRIBUTED LAGS AND LEADS WITH END POINT CONSTRAINTS

Donald W.K. Andrews

Ray C. Fair

Technical Working Paper No. 79

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1989

This paper is part of NBER's research program in Economic Fluctuations. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Technical Working Paper #79
October 1989

ESTIMATION OF POLYNOMIAL DISTRIBUTED LAGS AND LEADS WITH END POINT CONSTRAINTS

ABSTRACT

This paper considers the use of the polynomial distributed lag (PDL) technique when the lag length is estimated rather than fixed. We focus on the case where the degree of the polynomial is fixed, the polynomial is constrained to be zero at a certain lag length q, and q is estimated along with the other parameters. We extend the traditional PDL setup by allowing q to be real-valued rather than integer-valued, and we derive the asymptotic covariance matrix of all the parameter estimates, including the estimate of q. The paper also considers the estimation of distributed leads rather than lags, a case that can arise if expectations are assumed to be rational.

Donald W.K. Andrews
Cowles Foundation
Box 2125, Yale Station
New Haven, CT 06520

Ray C. Fair
Cowles Foundation
Box 2125, Yale Station
New Haven, CT 06520

I. Introduction

This paper considers the use of the polynomial distributed lag (PDL) technique of Almon (1965) when the lag length is estimated rather than fixed. We focus on the case where the degree of the polynomial is fixed, the polynomial is constrained to be zero at a certain lag length q, and q is estimated along with the other parameters. We extend the traditional PDL setup by allowing q to be real-valued rather than integer-valued. This extension plus a minor (and quite natural) modification of the PDL yields a regression function that is differentiable in q. Consequently, the model is simply a nonlinear regression model, and under standard assumptions the least squares estimate of q and various functions of q and the other parameters, such as the sum of the PDL coefficients, are consistent and asymptotically normal. Furthermore, if the errors are iid and normally distributed, these estimates are also asymptotically efficient. Estimates of their asymptotic variances and covariances are provided.

The paper also considers the estimation of distributed leads rather than lags. If expectations are rational and if the coefficients of the lead variables are assumed to lie on a polynomial, the PDL technique can be combined with Hansen's (1982) method of moments estimator to produce consistent and asymptotically normal estimates of all the parameters, including the lead length.

Considerable attention has been paid in the literature to the adverse effects of incorrectly specifying the lag length of PDLs, e.g., Schmidt and Waud (1973), Trevedi and Pagan (1979), Hendry, Pagan, and Sargan (1984) and references therein. For a fixed lag length the parameter estimates are usually inconsistent if the lag length is misspecified. For example, if the correct specification is for an explanatory variable to enter an equation only contemporaneously and if the lag length is specified to be greater than one, then the effect of the explanatory variable on the dependent variable will not be estimated consistently.

This misspecification problem does not arise if the lag length is estimated consistently. In consequence, several papers have considered estimating the lag length, e.g., Schmidt and Waud (1973) and Pagano and Hartley (1981). In each of these papers, however, no estimated standard error is obtained for the estimated lag length, and the estimated standard errors for the other parameter estimates are computed as though the estimated lag length is fixed. As has been recognized for some time -- see Schmidt (1973) and Frost (1975) -- such estimated standard errors understate the true variability of the parameter estimates. In contrast, this paper provides a standard error estimate for the lag length, and the estimated standard errors for the other parameter estimates take into account the estimation of the lag length.

II. Estimation of Distributed Lags

A Simple Example

It will be useful to begin with a simple example. Assume that the polynomial is linear and that there is one distributed lag variable:

$$(1)\quad Y_t = X_{1t}\beta + \int_0^q \alpha_{[j]} X_{2t-[j]}\,dj + u_t \;, \quad t = 1,\ldots,T$$

$$= X_{1t}\beta + \sum_{j=0}^{[q]-1} \alpha_j X_{2t-j} + (q-[q])\alpha_{[q]} X_{2t-[q]} + u_t \;, \quad t = 1,\ldots,T,$$

$$(2)\quad \alpha_j = \gamma_0 + \gamma_1 j \;, \quad j = 0,1,\ldots,[q],q \;,$$

$$(3)\quad \alpha_q = 0 \;,$$

where X_{1t} is a k-dimensional vector of explanatory variables other than X_{2t} and its lags, q is a real number greater than or equal to 1, and [q] is the integer part of q. Y_t and the X_{2t-j} are scalars. Equations (2) and (3) imply that

$$(4)\quad \alpha_j = -\gamma_1(q - j) \;.$$

Let

$$(5)\quad \theta = (\beta' \;\; \gamma_1 \;\; q)' \;, \quad X_t = (X_{1t}' \;\; X_{2t} \;\; X_{2t-1} \;\cdots\; X_{2t-[q]})' \;.$$

Given (4) and (5), equation (1) can be written as

$$(6)\quad Y_t = g(X_t,\theta) = X_{1t}\beta - \gamma_1\Big(\sum_{j=0}^{[q]-1}(q-j)X_{2t-j} + (q-[q])^2 X_{2t-[q]}\Big) + u_t$$

$$= X_{1t}\beta - \gamma_1 Q_{1t} + u_t \;.$$

An estimate of θ, denoted $\hat{\theta}$, can be obtained by minimizing the sum of squared residuals u'u, where $u' = (u_1, \ldots, u_T)$. One way this minimization can be done in practice is by searching over values of q. Given a value of q, Q_{1t} can be computed, and given Q_{1t}, equation (6) is linear in parameters and can thus be estimated by ordinary least squares. Thus, one can search over q by running least squares regressions to find the value that leads to the smallest overall sum of squared residuals. Alternatively, a gradient method can be used to compute the estimates, where the gradient is given in (7) below.

By writing the nonlinear regression function $g(X_t,\theta)$ in terms of an integral, as in (1), it is easy to see that it is a differentiable function of q and the other parameters. Thus, under standard conditions the nonlinear least squares estimator $\hat{\theta}$ is consistent and asymptotically normal (e.g., see Hansen (1982), Gallant (1987, Chs. 1,2), or Andrews and Fair (1988)). Note that q is identified only if $q \geq 1$, and it is an interior point of its parameter space, as is required for asymptotic normality, only if $q > 1$.

The estimation of the covariance matrix of $\hat{\theta}$ is straightforward. Let G be a Tx(k+2) matrix whose t-th row is

$$(7) \quad \frac{\partial}{\partial\theta'} g(X_t,\theta) = (X_{1t}' \,:\, -Q_{1t} \,:\, -\gamma_1(\sum_{j=0}^{[q]-1} X_{2t-j} + 2(q-[q])X_{2t-[q]})) .$$

An estimate of the covariance matrix of $\hat{\theta}$ is

$$(8) \quad \hat{V} = \hat{\sigma}^2(\hat{G}'\hat{G})^{-1} ,$$

where $\hat{\sigma}^2 = \hat{u}'\hat{u}/T$, $\hat{u}$ is the vector of estimated residuals from (6), and $\hat{G}$ is G evaluated at $\theta = \hat{\theta}$. The estimate $\hat{V}$ is appropriate when the errors $\{u_t: t\geq 1\}$ are independent, mean zero, variance σ^2 random variables conditional on $\{X_t: t\geq 1\}$. $\hat{V}$ is easy to compute in practice, since G is simply the matrix of regressors expanded by one column to include the derivative of $g(X_t,\theta)$ with respect to q.

In most PDL applications one is interested in the sum λ of the lag coefficients. In the present context λ is given by

$$(9) \quad \lambda = \int_0^q \alpha_{[j]} dj = -\gamma_1 \int_0^q (q-[j])dj = -\gamma_1 \left(\sum_{j=0}^{[q]-1} (q-j) + (q-[q])^2 \right) .$$

The least squares estimate of λ, $-\hat{\gamma}_1 \left(\sum_{j=0}^{[\hat{q}]-1} (\hat{q}-j) + (\hat{q}-[\hat{q}])^2 \right)$, has asymptotic variance

$$(10) \quad \sigma^2(\hat{\lambda}) = (\partial\lambda/\partial\gamma_1 \quad \partial\lambda/\partial q) V_2 (\partial\lambda/\partial\gamma_1 \quad \partial\lambda/\partial q)' \quad ,$$

where V_2 is the 2x2 covariance matrix of $(\hat{\gamma}_1 \quad \hat{q})'$, i.e., the lower right 2x2 block of the covariance matrix of $\hat{\theta}$, and

$$(11) \quad \partial\lambda/\partial\gamma_1 = -q^2 - [q]^2/2 + q[q] - [q]/2 \quad ,$$

$$(12) \quad \partial\lambda/\partial q = -\gamma_1(2q-[q]) \quad .$$

$\sigma^2(\hat{\lambda})$ can be estimated using the lower right 2x2 block of $\hat{V}$ in (8) and evaluating (11) and (12) at $q = \hat{q}$ and $\gamma_1 = \hat{\gamma}_1$.

We now consider various extensions of model (1).

Endogenous Explanatory Variables

If X_{2t} or some of the variables in X_{1t} are endogenous and if a matrix Z of first stage regressors is available, equation (1) can be estimated by two stage least squares (2SLS). $\hat{\theta}$ is obtained by minimizing $u'Z(Z'Z)^{-1}Z'u$, and the estimated covariance matrix is

$$(13) \quad \hat{V} = \hat{\sigma}^2(\hat{G}'Z(Z'Z)^{-1}Z'\hat{G})^{-1} \quad .$$

Again, $\hat{\theta}$ can be computed by searching over values of q. Given q, the problem is a standard 2SLS estimation problem. Alternatively, a gradient method can be used.

Quadratic Polynomials

If the polynomial is quadratic:

(14) $$\alpha_j = \gamma_0 + \gamma_1 j + \gamma_2 j^2 \quad , \; j = 0,1,\ldots,[q],q \; ,$$

(15) $$\alpha_q = 0 \; ,$$

and so

(16) $$\alpha_j = -\gamma_1(q-j) - \gamma_2(q^2-j^2) \; .$$

In this case, $q \geq 2$ is needed for identification, θ contains an extra element γ_2, i.e., $\theta' = (\beta' \; \gamma_1 \; \gamma_2 \; q)$, and equation (6) becomes

(17) $$Y_t = X_{1t}\beta - \gamma_1 Q_{1t} - \gamma_2 Q_{2t} + u_t \; ,$$

where Q_{1t} is as in (6) and

(18) $$Q_{2t} = \sum_{j=0}^{[q]-1} (q^2-j^2)X_{2t-j} + (q^2-[q]^2)(q-[q])X_{2t-[q]} \; .$$

Equation (7) becomes

(19) $$\frac{\partial}{\partial\theta'} g(X_t,\theta) = (X'_{1t} : -Q_{1t} : -Q_{2t} : \\ -\gamma_1\{\sum_{j=0}^{[q]-1} X_{2t-j} + 2(q-[q])X_{2t-[q]}\} \\ -\gamma_2\{2q\sum_{j=0}^{[q]-1} X_{2t-j} + (3q+[q])(q-[q])X_{2t-[q]}\}) \; .$$

Multiple Distributed Lag Variables

If model (1) contains a second distributed lag variable, say X_{3t-j}, two cases need to be considered, one in which the lag lengths for X_2 and X_3 are the same and the other in which they are not. If they are the same, the new term in (1) is $\int_0^q \eta_{[j]} X_{3t-[j]} dj$, where (assuming a linear polynomial) $\eta_j = \delta_0 + \delta_1 j$, $j = 0,1,\ldots,[q],q$, and $\delta_q = 0$. θ now contains an extra element δ_1, i.e., $\theta' = (\beta' \ \gamma_1 \ \delta_1 \ q)$, and equation (6) becomes

$$(20) \quad Y_t = X_{1t} - \gamma_1 Q_{1t} - \delta_1 R_{1t} + u_t ,$$

where

$$(21) \quad R_{1t} = \sum_{j=0}^{[q]-1} (q-j) X_{3t-j} + (q-[q])^2 X_{3t-[q]} .$$

Equation (7) becomes

$$(22) \quad \frac{\partial}{\partial \theta'} g(X_t, \theta) = (X'_{1t} : -Q_{1t} : -R_{1t} : \\ -\gamma_1 (\sum_{j=0}^{[q]-1} X_{2t-j} + 2(q-[q]) X_{2t-[q]}) \\ -\delta_1 (\sum_{j=0}^{[q]-1} X_{3t-j} + 2(q-[q]) X_{3t-[q]})) .$$

If the lag lengths are not equal, the new term in (1) is $\int_0^r \eta_{[j]} X_{3t-[j]} dj$, where (assuming a linear polynomial) $\eta_j = \delta_0 + \delta_1 j$, $j = 0,1,\ldots,[r],r$, and $\delta_r = 0$. θ now contains two extra elements δ_1 and r, i.e., $\theta' = (\beta' \ \gamma_1 \ \delta_1 \ q \ r)$. Equation (6) becomes equation (20) except that r replaces q in the definition of R_{1t} given in (21). Equation (7) becomes

$$(23)\quad \frac{\partial}{\partial\theta'} g(X_t,\theta) = (X'_{1t} : -Q_{1t} : -R_{1t} : -\gamma_1\{\sum_{j=0}^{[q]-1} X_{2t-j} + 2(q-[q])X_{2t-[q]}\} : -\delta_1\{\sum_{j=0}^{[r]-1} X_{3t-j} + 2(r-[r])X_{3t-[r]}\}) .$$

With two lag lengths the computational burden of searching becomes more burdensome, and a gradient method is likely to be much faster.

The extension to models with quadratic polynomials and more than two lagged variables is straightforward. In addition, the extension is straightforward to models with a PDL on both the dependent variable and various independent variables, as in the class of autoregressive distributed lag models considered in Hendry, Pagan, and Sargan (1984).

Nonlinearity

Finally, equation (1) -- and thus $g(X_t,\theta)$ in (6) -- can be nonlinear in parameters other than just q. Given q, the minimization of u'u need not be an ordinary least squares problem, and the derivatives of $g(X_t,\theta)$ with respect to θ can be more involved than those in (7). This means, among other things, that the case in which u_t is n-th order autoregressive can be handled easily. Equation (1) can be quasi differenced using the autoregressive parameters in order to eliminate the autoregressive part of the error, and the autoregressive coefficients can be incorporated into θ. This merely converts the problem into one in which $g(X_t,\theta)$ is more nonlinear in parameters than otherwise.

Estimation and Testing of the Degree of the Polynomial

Thus far we have considered the case where the degree of the polynomial is fixed. It is possible, however, to estimate both the lag length and the degree of the polynomial and to test the adequacy of a specified polynomial degree. With the lag length q treated as a real-valued parameter to be estimated, a sequence of models with PDLs of increasing degrees is a sequence of nested nonlinear regression models. Therefore, any of a number of standard consistent model selection procedures can be applied to estimate the polynomial degree. For example, one can use a downward sequential t- or F-testing procedure, as in Pagano and Hartley (1981), or one can use Akaike's information criterion, Schwartz's criterion, Mallow's C_p criterion, cross validation, or generalized cross-validation, etc. With a consistent model selection procedure, the asymptotic variances given above are still valid (because the correct model is selected with probability that goes to one as T goes to infinity), but the accuracy of the asymptotic approximation is likely to suffer.

The adequacy of a given choice of polynomial degree can be tested using an asymptotic t- or F-test as in Pagano and Hartley (1981). A RESET or RASET specification test can be used to test whether the degree of the polynomial is correct and whether the PDL restriction itself is appropriate -- see Harper (1977).

This completes the discussion of distributed lags. The cases considered in this section can be easily extended and combined, and in each case it is straightforward to treat the lag length or lengths as parameters to be estimated and to estimate their standard errors.

III. Estimation of Distributed Leads

Suppose that X_{2t-j} in (1) is replaced by X^e_{2t+j}, where the latter is the expected value of X_{2t+j} and all expectations are assumed to be formed at the end of period t-1, before information for period t is available. Let the expectation error for X^e_{2t+j} be

$$(24) \quad {}_{t-1}\epsilon_{t+j} = X_{2t+j} - X^e_{2t+j}, \; j = 0,1,\ldots,[q] .$$

Equation (1) in this case is

$$(25) \quad Y_t = X_{1t}\beta + \int_0^q \alpha_{[j]} X^e_{2t+[j]} dj + u_t, \quad t = 1,\ldots,T,$$

$$= X_{1t}\beta + \sum_{j=0}^{[q]-1} \alpha_j X_{2t+j} + (q-[q])\alpha_{[q]} X_{2t+[q]} + v_t ,$$

where

$$(26) \quad v_t = \int_0^q \alpha_{[j]} \; {}_{t-1}\epsilon_{t+[j]} dj + u_t .$$

Given (2) and (3), a new equation (6) can be derived:

$$(27) \quad Y_t = g(X_t,\theta) = X_{1t}\beta - \gamma_1\{\sum_{j=0}^{[q]-1} (q-j) X_{2t+j} + (q-[q])^2 X_{2t+[q]}\} + v_t$$

$$= X_{1t}\beta - \gamma_1 Q_{1t} + v_t ,$$

where X_t now denotes $(X'_{1t}\; X_{2t}\; X_{2t+1} \ldots X_{2t+[q]})'$.

Consider first 2SLS estimation of (27). Let Z_t be a vector of first stage regressors. A necessary condition for consistency is that Z_t and v_t be uncorrelated. This will be true if both u_t and the ${}_{t-1}\epsilon_{t+j}$ are mean zero and uncorrelated with Z_t. The assumption that u_t is mean zero and uncorrelated with Z_t is the usual 2SLS assumption. The assumption that the

${}_{t-1}\epsilon_{t+j}$ are mean zero and uncorrelated with Z_t is the rational expectations assumption. If expectations are formed rationally and if the variables in Z_t are used (perhaps along with others) in forming the expectations of the X_{2t+j}, then Z_t and the ${}_{t-1}\epsilon_{t+j}$ are uncorrelated. Therefore, given this assumption (and the other standard assumptions that are necessary for consistency), the 2SLS estimator of θ is consistent. It minimizes $v'Z(Z'Z)^{-1}Z'v$.

A problem with the 2SLS estimator in this context is that it ignores the m-dependent property of v_t. Because of the ${}_{t-1}\epsilon_{t+j}$, v_t will in general be m-dependent with $m = [q]-1$ if q is not an integer and $m = [q]-2$ if q is an integer. The 2SLS estimates are consistent, but the standard formula for their covariance matrix in (8) is incorrect and the estimates are not efficient within the class of limited information estimators. Hansen's (1982) method of moments estimator takes account of the m-dependent character of v_t. It is based on minimizing $v'ZM^{-1}Z'v$, where M is some consistent estimate of $\lim T^{-1}E(Z'vv'Z)$. In order to estimate M one needs an estimate of v_t in (27), such as the 2SLS estimate $\hat{v}_t$.

A general way of computing M is as follows. Let $f_t = \hat{v}_t Z_t$. Let $R_j = T^{-1}\Sigma_{t=j+1}^{T} f_t f_{t-j}'$, $j = 0, 1, \ldots, m$. M is then $(R_0 + R_1 + R_1' + \ldots + R_m + R_m')$. In many cases computing M in this way does not yield a positive definite matrix, and something else must be done. Hansen (1982), Cumby, Huizinga, and Obstfeld (1983), and Andrews (1988), among others, discuss the computation of M based on an estimate of the spectral density matrix of $Z_t'v_t$ evaluated at frequency zero. A third approach is to compute M under the following homoskedasticity assumption:

$$(28) \qquad E[v_t v_s | Z_t, Z_{t-1}, \ldots] = E[v_t v_s] \quad \text{for} \quad t \geq s ,$$

which says that the contemporaneous and serial correlations in v do not depend on Z. This assumption is implied by the assumption that $E[v_t Z_s] = 0$ for $t \geqq s$ if normality is also assumed. Under this assumption M can be computed as follows. Let $a_j = T^{-1}\Sigma_{t=j+1}^{T} \hat{v}_t \hat{v}_{t-j}$ and $B_j = T^{-1}\Sigma_{t=j+1}^{T} Z_t Z'_{t-j}$, $j = 0, 1, \ldots, m$. M is then $(a_0 B_0 + a_1 B_1 + a_1 B'_1 + \ldots + a_m B_m + a_m B'_m)$.

The complete estimation procedure in the case of polynomial distributed leads can now be summarized. 1) Estimate β, γ_1, and q in (27) by 2SLS, which minimizes $v'Z(Z'Z)^{-1}Z'v$. This requires searching over values of q or using a gradient method. 2) Given these estimates, compute $\hat{v}_t$ from (27). Then compute M in one of the above ways. 3) Estimate β, γ_1, and q in (27) by minimizing $v'ZM^{-1}Z'v$. This again requires searching over values of q or using a gradient method. These are the final parameter estimates. The estimated covariance matrix of these estimates is

$$(29) \qquad \hat{V} = T(\hat{G}'ZM^{-1}Z\hat{G})^{-1} ,$$

where the elements of G are as in (7) except that X_{2t+j} replaces X_{2t-j} and $X_{2t+[q]}$ replaces $X_{2t-[q]}$.

The various extensions discussed in Section II can also be applied here. The modifications needed for the case of leads rather than lags are slight, and they will not be discussed further.

IV. Monte Carlo Results

It is of interest to see how good an approximation the asymptotic standard errors of the estimates of q and λ are in finite samples. We consider three examples in this section, which are taken from equations estimated in Fair (1989). The equations are price equations for fairly disaggregate commodities, with distributed lag or lead values of an aggregate price variable added to pick up aggregate expectational effects on individual price setting behavior. The data are monthly. The first two equations were estimated for the period August 1958 - April 1989, a total of 359 observations, and the third equation was estimated for the period July 1958 - August 1983, a total of 302 observations. The first equation includes 20 explanatory variables plus lags of the aggregate price variable. The second equation has three fewer explanatory variables, and the third equation has one fewer. Eleven of the explanatory variables in each equation are seasonal dummy variables. The polynomial was taken to be linear, and the aggregate price variable was the only variable to which the polynomial lag distribution was applied. Given the lag length q, each equation is linear in parameters.

The Monte Carlo experiments were run as follows. Each equation was estimated first using the historical data. The estimated parameters were used as the true parameters for purposes of the Monte Carlo experiments. The error term in the equation was assumed to be normal with mean zero and variance $\hat{\sigma}^2$, where the latter is the estimated variance of the equation. For each repetition a new data set was generated by drawing error terms from this distribution and using these error terms plus the estimated parameters to compute new values of the dependent variable. The equation was then reestimated using the new data, and the parameter estimates were recorded (including the estimate of q). One thousand repetitions were made for each equation. The largest value of q allowed was 72. (The smallest value of q allowed was 1.)

The means and variances of the parameter estimates over the 1000 repetitions were calculated. Of particular interest is the comparison of the Monte Carlo and asymptotic estimates of the standard errors of $\hat{\lambda}$ and $\hat{q}$. For the first equation the estimates of λ and q from the historical data were .1256 and 27.34, with asymptotic standard errors of .0307 and 6.46. The average estimates of λ and q from the 1000 repetitions were .1259 and 27.00, with Monte Carlo standard errors of .0305 and 7.72. The biases in the estimates are thus fairly small, as are the differences in the standard errors. The range of the estimates of q over the 1000 repetitions was 1 to 72. There were 6 occurrences of an estimate of 1 and 2 occurrences of an estimate of 72.

The results were similar for the second equation. The estimates of λ and q from the historical data were .1184 and 12.56, with asymptotic standard errors of .0286 and 5.72. The average estimates from the 1000 repetitions were .1209 and 12.53, with Monte Carlo standard errors of .0270 and 5.82. The range of the estimates of q was 1 to 35.78. There were 30 occurrences of an estimate of 1. The estimates of q are in units of months, and so the estimated standard errors of the lag length for these two cases are around 6 or 7 months.

The third equation is one in which the asymptotic standard error of $\hat{q}$ is larger than $\hat{q}$. The estimates of λ and q from the historical data were .1860 and 11.09, with asymptotic standard errors of .0442 and 14.01. The average estimates from the 1000 repetitions were .1925 and 14.29, with Monte Carlo standard errors of .0429 and 13.12. The range of the estimates of q was 1 to 72. There were 118 occurrences of an estimate of 1 and 8 occurrences of an estimate of 72. In this case the Monte Carlo estimate of q is noticeably larger than the estimate from the historical data (14.29 versus 11.09). Also, the Monte Carlo standard error of $\hat{q}$ is smaller than the asymptotic standard error of $\hat{q}$ (13.12 versus 14.01). These two results are explained by the fact that the distribution of $\hat{q}$ is truncated at 1. In this third case, where the standard error of $\hat{q}$ is large relative to $\hat{q}$ itself, there is considerable probability mass at 1. This leads $\hat{q}$ to be biased upward. It also leads to the Monte Carlo standard error being smaller than the asymptotic standard error. The asymptotic formula thus overstates the standard error of $\hat{q}$ for large values of the standard error.

The results for the third equation are fairly typical of the overall results in Fair (1989), where q is generally not estimated with precision. In many of the estimated equations the standard error of $\hat{q}$ is at least as large as $\hat{q}$. This imprecision does not, however, carry over to the estimates of λ. The latter generally are fairly precise. In other words, the data seem fairly good at tacking down the sum of the lag coefficients, but not the lag length itself.

V. Conclusion

Since it is quite rare that lag and lead lengths are known with certainty, the ability to estimate them and the standard errors of their estimates should prove useful in practice. The estimated standard errors should help one in deciding how much confidence to place on the overall estimated lag or lead distributions.

REFERENCES

Almon, Shirley, "The Distributed Lag Between Capital Appropriations and Expenditures," Econometrica, 33 (January 1965), 178-196.

Andrews, Donald W.K., "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," Cowles Foundation Discussion Paper No. 877R, 1988.

Andrews, Donald W.K., and Ray C. Fair, "Inference in Econometric Models with Structural Change," Review of Economic Studies, (October 1988), 615-640.

Cumby, Robert E., John Huizinga, and Maurice Obstfeld, "Two-Step Two-Stage Least Squares Estimation in Models with Rational Expectations," Journal of Econometrics, 21 (August 1983), 333-355.

Fair, Ray C., "Inflationary Expectations and Price Setting Behavior," mimeo, August 1989.

Frost, Peter A., "Some Properties of the Almon Lag Technique When One Searches for Degree of Polynomial and Lag," Journal of the American Statistical Association, 70 (September 1975), 606-612.

Gallant, A.R., Nonlinear Statistical Models. New York: Wiley, 1987.

Hansen, Lars, "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica, 50 (July 1982), 1029-1054.

Harper, Charles P., "Testing for the Existence of a Lagged Relationship within Almon's Method, "Review of Economics and Statistics, 59 (May 1977), 204-210.

Hendry, David F., Adrian Pagan, and J. Dennis Sargan, "Dynamic Specification," Chapter 18 in Handbook of Econometrics, Volume 2, ed. by Z. Griliches and M.D. Intriligator, New York: North-Holland, 1984.

Pagano, Marcello, and Michael J. Hartley, "On Fitting Distributed Lag Models Subject to Polynomial Restrictions," Journal of Econometrics, 16 (June 1981), 171-198.

Schmidt, Peter, "On the Difference Between the Conditional and Unconditional Asymptotic Distributions of Estimates in Distributed Lag Models with Integer-valued Parameters," Econometrica, 41 (January 1973), 165-169.

Schmidt, Peter, and Roger N. Waud, "The Almon Lag Technique and the Monetary Versus Fiscal Policy Debate," Journal of the American Statistical Association, 68 (March 1973), 11-19.

Trevedi, Pravin, and Adrian Pagan, "Polynomial Distributed Lags: A Unified Treatment," Economic Studies Quarterly, 30 (1979), 37-49.