Queen's Economics Department Working Paper No. 3

A NOTE ON GAMMA DISTRIBUTED LAGS

Hiroki Tsurumi
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

10-1969

A NOTE ON GAMMA DISTRIBUTED LAGS

Hiroki Tsurumi

Discussion Paper No. 3

A NOTE ON GAMMA DISTRIBUTED LAGS*

Hiroki Tsurumi

1. Introduction

Solow [10] generalized Koyck's method for distributed lags to a family of J-shaped or unimodel lag distribution given by the Pascal distributions. Jorgenson [5] advances it to what he calls the general Pascal distribution. A distributed lag function is given by

$$(1) \quad \begin{aligned} y_t &= p_o x_t + p_1 x_{t-1} + p_2 x_{t-2} + \cdots \\ &= p_o x_t + p_1 L x_t + p_2 L^2 x_t + \ldots \\ &= [p_o + p_1 L + p_2 L^2 + \ldots] x_t = p\ (L) x_t \end{aligned}$$

where y_t and x_t are observed values of dependent and independent variables at time t and the coefficients are unknown paramenters, and L is the lag operator. Jorgenson defines the class of rational distributed functions as the sequence $\{ p_1, \ldots, p_j, \ldots \} \in l_1 = \{ p \mid \sum_{j=0}^{\infty} p_j = 1,\ p_j \geqq 0 \}$ and he represents p(L) by the quotient of two finite polynomials

$$(2) \quad p(L) = \frac{u(L)}{v(L)}$$

*The author wishes to thank Professor K. Kuga of Osaka University for his valuable comments on an earlier draft of the paper and for helping me on the discussion of the lag operators.

where $u(L) = u_o + u_1L + \ldots + u_mL^m$, and $v(L) = v_o + v_1L + \ldots + v_nL^n$, and $u(L)$ and $v(L)$ do not possess any common latent roots.

To analyse the representation of the infinite polynomial $p(L)$ by (2), let us first consider a lag operator $A = p(L)$ whose coefficients lie in a probabilistic space $\{p \mid \sum_{j=0}^{\infty} p_j = 1,\ p_j \geqq 0\}$, and

$$A\,\underset{\sim}{x}_t = [\ \sum_0^{\infty} p_j x_{t-j},\ \sum_0^{\infty} p_j x_{t-1-j},\ \ldots\ ] = \underset{\sim}{y}_t = [y_t, y_{t-1},\ \ldots]$$

and $\underset{\sim}{x}_t = (x_t, x_{t-1}, \ldots)$ lies in a space of bounded sequences of numbers, $X = \{\ |x_{t-j}| \leqq K_t$ for all j and $\Sigma K_t < +\infty\}$.(1) If we take the norm of operator A as $||A|| = \Sigma(p_j)^2$, then the operator A which maps X into X is a normed ring [Liusternik and Sobolev 8: pp. 90-91], and if we define the unit operator, $e = (1,0,0, \ldots,0, \ldots)$, then the theorem 2 on page 91 of Liusternik and Sobolev may be applied and $e + A$ will have an inverse operator $(e + A)^{-1}$.

Now let us suppose we have an operator space E, and $\underset{\sim}{y}_t \in X$, $\underset{\sim}{x}_t \in X$ and $(e + B)\,\underset{\sim}{y}_t = A\,\underset{\sim}{x}_t$ where A and B belong to the operator space E. Since $(e + B)\,\underset{\sim}{y}_t \in X$, and $A\underset{\sim}{x}_t \in X$, if $||B|| < 1$, then $\underset{\sim}{y}_t = (e + B)^{-1} A\,\underset{\sim}{x}_t$, and conversely if $\underset{\sim}{y}_t = (e + B)^{-1}A\,\underset{\sim}{x}_t$, then $(e + B)\,\underset{\sim}{y}_t = A\,\underset{\sim}{x}_t$.

---

(1) It is obvious that $\underset{\sim}{y}_t \in X$, since $\underset{\sim}{y}_{t-k} = [y_{t-k}, y_{t-k-1}, \ldots]$
$= [\Sigma p_j x_{t-j-k}, \Sigma p_j x_{t-j-k-1}, \ldots] \leqq [K_{t-k}, K_{t-k-1}, \ldots]$.

We may note that even if $(e + B)$ and A are finite polynomials, $(e + B)^{-1}A$ will be infinite since $(e + B)^{-1} = e - B + B^2 - \ldots$ is infinite.

Koyck lag is obtained by putting $A = \{1-\lambda, 0,\ldots,0,\ldots\}$, $B = \{0,-\lambda,0,\ldots,0,\ldots\}$, and Solow's Pascal distributed lag is obtained by putting $A = \{(1-\lambda)^r,0,\ldots,0,\ldots\}$ and $B = \{0,-\binom{r}{1}\lambda, \binom{r}{2}\lambda^2,\ldots\}$. Jorgenson's representation (2) corresponds to $(e + B) = \{1, v_1,\ldots,v_n\}$ and $A = \{u_o,u_1,\ldots,u_m\}$.

Koyck and Solow generates the distributed lag coefficients $\{p_k\}$ by *a priori* specifying the distributions with one or two paramenters, whereas Jorgenson tries to approximate $\{p_k\}$ by two finite polynomials whose coefficients are given by $\{v_k\}$ and $\{u_k\}$. Consequently Jorgenson's positions are twofold: (i) the infinite series $\{p_k\}$ may be approximated to a desired degree of accuracy by two finite polynomials and (ii) it is better to have fewer number of unknown coefficients $\{v_k\}$ and $\{u_k\}$ to be estimated. The degrees of polynomials, m and n are given *a priori*. However, it seems difficult to satisfy the two criteria (i) and (ii) above simultaneously: as long as one seeks the position (i), the degrees of polynomials, m, and n may get larger and larger.(2)

---

(2) In the approximation theorem 2 [5:p.142] Jorgenson states $\{v_k\}$ and $\{u_k\}$ may be taken to correspond to the probability distributions of a non-negative, integer-valued, random variable. However, unless one constrains v_k and u_k to be non-negative *a priori*, it seems difficult to estimate them to be non-negative.

Furthermore, if one uses any estimation procedures which do not impose any constraints on the unknown parameters, then one will face two difficulties: (1) there is no guarantee that $\{v_k\}$ or $\{u_k\}$ are non-negative, and (2) the transformation from $y_t = u(L)/v(L)x_t + u_t$ to $v(L)y_t$ $u(L)x_t + v(L)u_t$ will generate autocorrelation.(3)

In summary one may be obliged to choose one of the two approaches to the distributed lags: one is to follow the Koyck-Solow line by specifying a priori the shape of distributed lags, and the other is to follow Jorgenson's approach by choosing the two finite polynomials of degrees n and m with a hope to attain a desired level of accuracy in approximating $\{p_k\}$.

Since the latter approach tends to involve three problems, i.e. (i) the choice of n and m, (ii) estimation procedures which guarantee nonnegative coefficients, and (iii) autocorrelation, it may be better to specify the shape of distributed lags with one or two parameters.

In this note we propose to represent the unknown parameters of the distributed lags $\{p_k\}$ by a family of J-shaped or unimodal lag distributions given by the gamma function, and once we obtain the estimates of the parameter(s) of the gamma distribution by an appropriate estimation

---

(3) These two points, especially the second, may be surmounted by using a technique given in [4]. The first point is also applicable to the Almon distributed lags.

procedure, we approximate the distribution to a discrete distribution. As discussed in the following sections, the estimation of gamma distributed lags is simplier than that of Solow's Pascal distributed lags.

2. Gamma Distributed Lags

Let the sequence of the unknown coefficients in equation (1) follow the gamma distribution

$$(3)\quad p_k = \frac{1}{\Gamma(s)}\, k^{s-1}\, e^{-k} \qquad k \geqq 0,\ s > 0$$

where $\frac{1}{\Gamma(s)}\int_0^\infty k^{s-1}e^{-k}dk = 1$, and $p_o = 0$ for $s \nless 1$.

Then as illustrated in Figure 1, for different values of s, $\{p_k\}$ has the unimodal distributions. The maximum value of p_k is reached at $k = s - 1$.

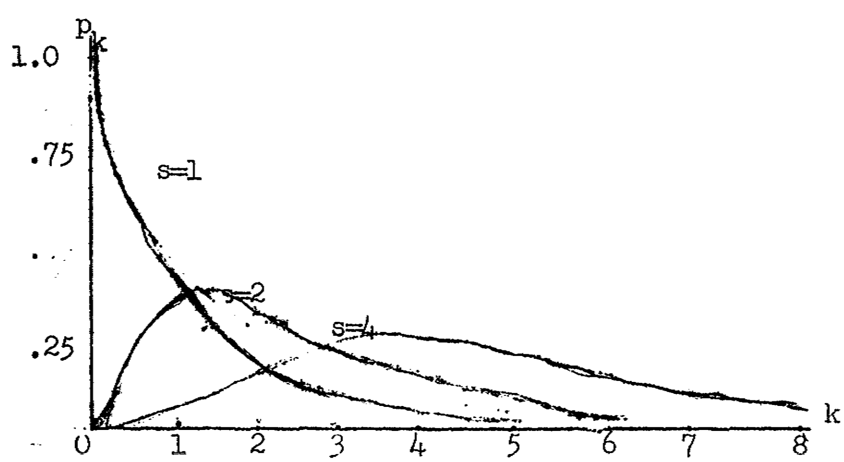

Figure 1 Gamma Distribution

If p_k is given as in equation (3), equation (1) becomes

$$(1)'\quad y_t = \frac{1}{\Gamma(s)}\sum_{k=0}^{\infty} k^{s-1}e^{-k}L^k\, x_t = p(L)\, x_t.$$

One may approximate the distribution (3) in a discrete version, since the values of k, in discrete time period analysis, are integer values.

$$(4)\quad \frac{1}{\Gamma(s)} k^{s-1} e^{-k}\, dk \sim \frac{1}{G} \sum_{k=0}^{\infty} k^{s-1} e^{-k}$$

where $G = \sum_{k=0}^{\infty} k^{s-1}e^{-k}$.

As given in Figure 1, p_o is zero for $s \neq 1$, and thus in this case equation (1)' becomes

$$y_t = p_1 x_{t-1} + p_2 x_{t-2} + \cdots .$$

If one is certain that p_o is not zero, then the gamma distribution (3) may be changed to

$$(5)\qquad p_{k-1} = \frac{1}{G} k^{s-1}e^{-k} \qquad k \geq 1,\ s > 0.$$

3. Estimation of the Gamma Distributed Lags

The estimation of the gamma distributed lags (4) or (5) involves the problem of nonlinearity in the parameter, and thus an appropriate nonlinear estimation method should be used. In our example we use the modified nonlinear least squares method of Marquardt [9]. As an illustration of the gamma distributed lags we will estimate the investment function of the total manufacturing industry in Canada using annual data, and we use Jorgenson's investment function given in [6]:

$$(6)\qquad I_t = \mu(L)[K_t^E - K_{t-1}^E] + \delta K_{t-1} + u_t$$

where I_t = gross investment at time t

K_t^E = expected capital stock at time t

K_{t-1} = net capital stock at the beginning of time t

δ = depreciation rate

$\mu(L)$ = distributed lag function

u_t = disturbance term.

We represent the expected capital stock at time t by

$$(7) \qquad K_t^E = \beta \frac{V_t}{P_t}$$

where V_t and P_t are respectively the current value of output measured in terms of value added in total manufacturing and price deflator for the total manufacturing output. (4)

Since we use annual data and since investment includes some items which will be completed within the same period (i.e. capital item charged to the operating expenses), let us use the gamma distributed lag structure of equation (5). If we substitute (7) into (6), we will obtain

---

(4) Jorgenson uses the user cost, but we used P_t as the price deflator because data to derive the user cost were not available.

The sources of data are as follows: Fixed capital series were taken from Fixed Capital Flows and Stocks Manufacturing 1926-1960, Dominion Bureau of Statistics, Catalogue No. 13-523 August 1966. Total manufacturing fixed capital series in constant 1957 dollars are given in p.A-6 for 1926-1960. From 1961 to 1967, the figures were provided by Mr. C. Braithwaite of D.B.S.

The output figures were obtained from census value added by manufacturing given in M.C. Urquhart (ed.) Historical Statistics of Canada, Cambridge University Press, 1965, Q-11 in p. 463 and in Canada Year Book, various issues.

The price index, P_t, was taken from "the general wholesale price indexes," Canadian Statistical Review, various issues. The series based on 1935-39=100 were converted to a 1957=100 basis.

$$(8) \qquad I_t = \frac{\beta}{G} \sum_{k=1}^{\infty} k^{s-1}e^{-k}x_{t-k} + \delta K_{t-1} + u_t$$

where $x_t = \frac{V_{t-k+1}}{P_{t-k+1}} - \frac{V_{t-k}}{P_{t-k}}$.

The first term of the right hand side of equation (8) is an infinite series, and for estimation it is necessary to limit the sum to some finite number. Following Klein [7] and Dhrymes [3], we truncate the infinite sum at m and rewrite (8) in the alternate form

$$(8)' \qquad I_t = \beta\eta + \frac{\beta}{G} \sum_{k=1}^{m} k^{s-1}e^{-k}x_{t-k} + \delta K_{t-1} + u_t$$

where $\eta = \frac{1}{G} \sum_{k=m+1}^{\infty} k^{s-1}e^{-k}M$, and $M = \max_{t,k}(x_{t-k})$.(5)

The choice of m is not a trivial matter, since the estimate of s may depend upon m. As a possible criterion, one may turn to the

---

(5) The Klein-Dhrymes technique gives $= \lambda^t \sum_{j=0}^{\infty} \lambda^j x_{-j}$ which is independent of t, and this is possible due to the exponentiality of the Koyck lag. If we follow their line more closely, then we could re-formulate (8) as follows:

$$I_t = \frac{\beta}{G} \sum_{k=0}^{t-1} k^{s-1}e^{-k}x_{t-k} + \frac{\beta}{G} \sum_{k=t}^{\infty} k^{s-1}e^{-k}x_{t-k} + \delta K_{t-1} + u_t$$

$$= \frac{\beta}{G} \sum_{k=0}^{t-1} k^{s-1}e^{-k}x_{t-k} + \frac{\beta}{G} \sum_{k=0}^{\infty} (t+k)^{s-1}e^{-(t+k)}x_{-k} + \delta K_{t-1} + u_t$$

$$= \frac{\beta}{G} \sum_{k=0}^{t-1} k^{s-1}e^{-k}x_{t-k} + \frac{\beta}{G} t^{s-1}e^{-t} \sum_{k=0}^{\infty} (1+ \frac{k}{t})^{s-1}e^{-k}x_{-k} + \delta K_{t-1} + u_t$$

Since $(1 + \frac{k}{t})^{s-1} \lesseqgtr (1+k)^{s-1}$, we may approximate

$$I_t = \frac{\beta}{G} \sum k^{s-1}e^{-k}x_{t-k} + \frac{\beta}{G} t^{s-1}e^{-t} \eta_o + \delta K_{t-1} + u_t$$

where $\eta_o = \sum_{k=0}^{\infty} (1+k)^{s-1}e^{-k}x_{-k}$.

This approach, compared to the one presented in equation (8)', has an estimation errors which are built in when we approximated (1+k/t) by (1+k), and they do not disappear. Whereas equation (8)' guarantees that the estimation errors disappear as $m \to \infty$.

principle of the nonlinear least squares which is to minimize the sum of squares, $Q = \sum_{t=1}^{T} (y_t - f_t)^2$, where y_t and f_t are respectively dependent variable and nonlinear function whose coefficients are to be estimated. Since the estimate of the variance of u_t, $\hat{\sigma}_u^2$, is given by $Q/(T-k)$, k being the number of parameters to be estimated, we may choose m such that $\hat{\sigma}_u^2$ is a minimum.(6)

We have experimented with m = 4,5,6,7,8,9, and as shown in Table 1, at m=6, the estimate of the variance of u_t, $\hat{\sigma}_u^2$, reaches a minimum within this range of m. As a set of initial values for the estimation of equation (8)' we chose $\hat{\eta}_o = 1.0$, $\hat{\beta}_o = .16$, $\hat{s}_o = 2.0$, and $\hat{\delta}_o = .0685$. The choice of $\hat{\delta}_o$ was from the average depreciation value of capital stock series, whereas $\hat{\eta}_o$, $\hat{\beta}_o$, and $\hat{s}_o$ were arbitrarily chosen.(7) In Table 1 the figures in parentheses were the estimated standard errors of the coefficients, $\bar{R}^2$ is the coefficient of determination adjusted for degrees of freedom, var is the estimated variance of the disturbance term, and DW is the Durbin-Watson test statistics.

---

(6) The range of m to be examined will be limited by the number of data available at hand, and the minimum if attained will be a local one. In our experiment the sample period without counting the lags is from 1949 to 1967 for all values of m.

(7) Since the nonlinear least squares method only guarantees a local minimum, we tested whether converged values of the parameters, especially the estimate of s, may differ by the choice of initial value of s. For this purpose we chose $s_o = 1.5$ and $s_o = 4.0$, while holding the initial values of other parameters unchanged. m was set at 6. The converged values of the estimated coefficients were the same up to four decimal places.

Table 1: Changes in the Upper Sum Number m and Estimated Results

m	$\hat{\eta}$	$\hat{\beta}$	$\hat{s}$	$\hat{\delta}$	$\bar{R}^2$	var	DW
4	736.4415 (418.5039)	.7710 (.1801)	2.8243 (.4548	.0277 (.0211)	.82	27165.5	1.37
5	748.2283 (363.9757)	.8216 (.1671)	2.9991 (.4190)	.0204 (.0200)	.84	23173.9	1.40
6	780.2656 (350.9799	.8486 (.1577)	3.0295 (.3619)	.0142 (.0199)	.85	21167.2	1.41
7	795.3340 (364.9233)	.8461 (.1617)	2.9901 (.3645)	.0128 (.0207)	.85	21727.1	1.46
8	797.6076 (368.2027)	.8489 (.1639)	2.9901 (.3629)	.0120 (.0210)	.85	21849.1	1.44
9	797.0016 (368.5688)	.8522 (.1653)	2.9882 (.3605)	.0115 (.0211)	.85	21853.2	1.44

The lag structure estimated as in Table 1 for m = 6 is presented in Figure 2. The peak of the lag response occurs at k=2.0295, or during the previous period because of the way lag is set up by equation (5).

Following the criterion of choosing m above, we chose $m = 6$. The estimated results tend to be stable over the change in the value m. The Durbin-Watson test statistics indicate no first-order autocorrelation at .01 level.(8)

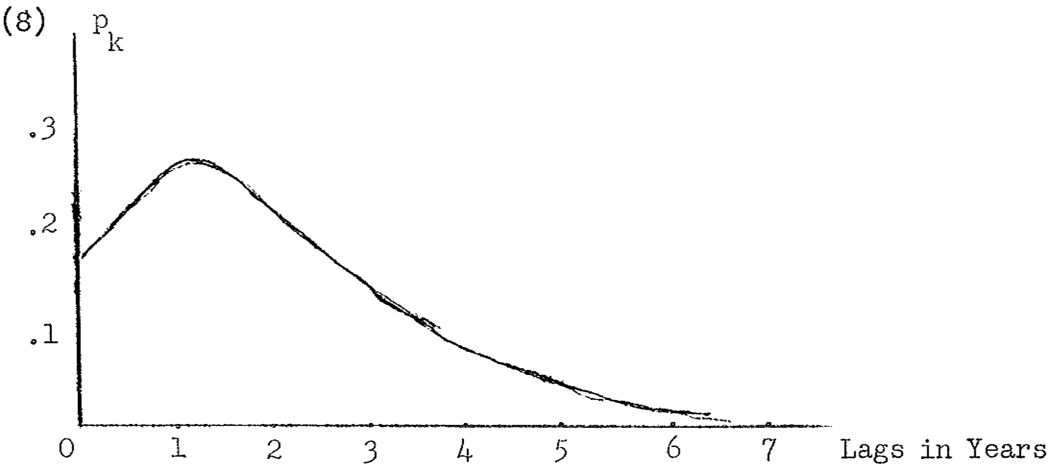

Figure 2: The Time Form of the Distributed Lags

4. Conclusions

The gamma distributed lag presented in the previous section seems to work fairly well and compared to the Pascal distributed lag, the former involves only one unknown parameter to be estimated, s, which does not have to be an integer, and thus the estimation procedure is easier than that of the Pascal distributed lag.

---

(8) If the first order autocorrelation exists, we may transform equation (8)' into

$$I_t^* = \alpha_o + \frac{\beta}{G} \sum_{k=1}^{m} k^{s-1} e^{-k} x_{t-k}^* + \delta K_{t-1}^* + e_t$$

where $I_t^* = I_t - \rho I_{t-1}$, $\alpha_o = \beta\eta_o(1-\rho)$, $x_{t-k}^* = x_{t-k} - \rho x_{t-k-1}^*$, $K_t^* = K_t - \rho K_{t-1}$, and $e_t = u_t - \rho u_{t-1}$. This equation may be estimated by the nonlinear least squares method.

One may advance the gamma distribution to a more general form than equation (4). We may put

$$p_k = \frac{1}{G} k^{s-1} e^{-\alpha k} \qquad \alpha > 0$$

where $G = \Sigma\, k^{s-1} e^{-\alpha k}$, and p_k attains maximum at $k = \frac{s-1}{\alpha}$. Koyck lag may be represented by putting $s = 1$. We tested this general form for $m = 6$, and the results were

$$\hat{\eta} = \underset{(563.3444)}{902.8861}, \quad \hat{\beta} = \underset{(.2821)}{.7951}, \quad \hat{s} = \underset{(1.9057)}{2.8497}, \quad \hat{\alpha} = \underset{(.9453)}{.9912}$$

$$\hat{\delta} = \underset{(.0230)}{.01127}, \qquad \bar{R}^2 = .83, \text{ var} = 23667.2, \quad DW = 1.35$$

In general the generalized form tended to have a strong multicollinearity between the first derivative variable associated with α and that with s. The estimated simple correlation between α and s was .98.

Equation (8)' is moderately nonlinear and the nonlinear least squares method seems to be applied without any difficulty. The consistent properties of nonlinear least squares under certain regularity conditions were given in Hartley and Booker [2]. If there is a simultaneous determination of investment, I_t, and output, V_t, then we may apply a technique such as given in [1].

By truncating the infinite series as we did in equation (8)' we may incur some errors on the part of the estimators, and the magnitude of the error will depend on the size of m as well as on the value of s. When m gets large relative to s, the maximum order of errors will tend to be zero. This may be illustrated as follows:

The gamma distributed lag model without truncation is given by

$$(9) \qquad I_t^* = \frac{\beta}{G} \sum_{k=1}^{\infty} k^{s-1} e^{-k} x_{t-k} + \delta K_{t-1} + u_t$$

whereas the truncated version is

$$(10) \qquad I_t = \beta\eta + \frac{\beta}{G} \sum_{k=1}^{m} k^{s-1} e^{-k} x_{t-k} + \delta K_{t-1} + u_t$$

where $\eta = \frac{1}{G} \sum_{k=m+1}^{\infty} k^{s-1} e^{-k} M$. The difference between (9) and (10) in estimating parameters arises from the first order partials associated with β and s. For equation (9) they are

$$\frac{\partial I_t^*}{\partial \beta} = \frac{1}{G} \sum_{k=1}^{\infty} k^{s-1} e^{-k} x_{t-k}$$

$$\frac{\partial I_t^*}{\partial s} = \beta \sum_{k=1}^{\infty} \frac{\partial\left(\frac{1}{G} k^{s-1} e^{-k} x_{t-k}\right)}{\partial s} .$$

whereas for equation (10) they are

$$\frac{\partial I_t}{\partial \beta} = \eta + \frac{1}{G} \sum_{k=1}^{m} k^{s-1} e^{-k} x_{t-k}$$

$$\frac{\partial I_t}{\partial s} = \beta \sum_{k=1}^{m} \frac{\partial\left(\frac{1}{G} k^{s-1} e^{-k} x_{t-k}\right)}{\partial s} .$$

Comparing $\frac{\partial I_t}{\partial \beta}$ with $\frac{\partial I_t^*}{\partial s}$, we find that

$$\frac{\partial I_t}{\partial \beta} = \epsilon_1 M + \frac{1}{G} \sum_{k=1}^{m} k^{s-1} e^{-k} x_{t-k}$$

where $\epsilon_1 = \frac{1}{G} \sum^{\infty} k^{s-1} e^{-k}$. It is clear that as $m \to \infty$, $\epsilon_1 \to 0$, and $\frac{\partial I_t}{\partial \beta} \to \frac{\partial I_t^*}{\partial \beta}$

And for $\frac{\partial I_t^*}{\partial s}$ and $\frac{\partial I_t}{\partial s}$ we find that

$$\frac{\partial I_t^*}{\partial s} = \frac{\beta}{G} \sum_{k=m+1}^{\infty} k^{s-1} e^{-k} x_{t-k} + \frac{\beta}{G} \sum_{k=1}^{m} k^{s-1} e^{-k} x_{t-k}$$

$$= \Delta + \frac{\partial I_t}{\partial s}$$

where $\Delta = \sum_{k=m+1}^{\infty} \frac{\partial(\frac{1}{G} k^{s-1} e^{-k} x_{t-k})}{\partial s}$ is the neglected term due to truncation, and it is a source of errors caused by using equation (10) instead of (9).

The maximum order of this source of errors will be given as follows; since as argued in section 1 x_{t-k} is a bounded sequence such that $|x_{t-k}| \lneqq K_t$ for all k, Δ becomes

$$|\Delta| = \left| \sum_{k=m+1}^{\infty} \frac{\partial(\frac{1}{G} \; k^{s-1} e^{-k} x_{t-k})}{\partial s} \right|$$

$$\lneqq \sum_{k=m+1}^{\infty} \frac{\partial(\frac{1}{G} k^{s-1} e^{-k})}{\partial s} K_t = \epsilon_2 K_t$$

where $\epsilon_2 = \sum_{k=m+1}^{\infty} \frac{\partial(\frac{1}{G} k^{s-1} e^{-k})}{\partial s}$. The maximum order of the source of errors, ϵK_t, will tend to zero as $m \to \infty \Rightarrow \epsilon \to 0$. How fast ϵ_2 as well as ϵ_1 tends towards zero depends on the value of s.

Queen's University.

References

1. Eisenpress, H. and J. Greenstadt: "The Estimation of Nonlinear Econometric Systems," Econometrica Vol. 34, 1966, pp. 851-861.

2. Hartley, H.O. and A. Booker, "Nonlinear Least Squares Estimation," Annals of Mathematical Statistics, Vol. 36, 1965, pp. 638-650.

3. Dhrymes, P.J., "Efficient Estimation of Distributed Lags with Autocorrelated Errors," International Economic Review, Vol. 10, pp. 47-67.

4. Dhrymes, P.J., L.R. Klein and K. Steiglitz, "Estimation of Distributed Lags," Mimeographed, 1968.

5. Jorgenson, D.W., "Rational Distributed Lag Functions," Econometrica, Vol. 32, 1966, pp. 135-149.

6. __________, "Anticipations and Investment Behavior," The Brookings Quarterly Econometric Model of The United States, Rand McNally and Company, 1965, pp. 35-92.

7. Klein, L.R., "The Estimation of Distributed Lags," Econometrica, Vol. 26, 1958, pp. 553-565.

8. Liusternik, L. and V. Sobelev, Elements of Functional Analysis, Frederick Ungar, New York, 1965.

9. Marquardt, D.W., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," SIAM, Applied Mathematics, Vol. 11, 1963, pp. 431-441.

10. Solow, R.M., "On a Family of Lag Distributions," Econometrica, Vol. 28, 1960, pp. 393-406.