COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 279

DISTRIBUTED LAGS, PREDICTION, AND SIGNAL EXTRACTION

David M. Grether

September 8, 1969

DISTRIBUTED LAGS, PREDICTION, AND SIGNAL EXTRACTION

by

David M. Grether*

Summary

A wide variety of economic models includes expectational variables among the list of variables determining behavior. In this paper it is shown that for a large class of time series, expectations about the future of observed series or about unobserved components of economic time series may lead to rational lag distributions. Specifically it is shown that rational distributed lags arise whenever the series (or the ρth differences of the series) have autoregressive, moving average representations and linear least squares forecasts are calculated. The orders of the lag distributions are also given. In Section 4 an example is given that shows the application of these results to an inventory adjustment model.

* The research described in this paper was carried out under grants from the National Science Foundation and from the Ford Foundation. I am indebted to Marc Nerlove and Joseph Stiglitz for helpful comments. They should not be held responsible for errors.

One of the difficulties with distributed lag models with a rational function of the lag operator is that the order of the operator is frequently not known *a priori*, though in some cases one may need this information in order to identify the structural parameters of interest. The results presented here show that the orders of the operators depend in a simple way upon the structure of the series being forecast or estimated, and, thus, the results should be useful in the formulation and estimation of distributed lag models.

Section 1. Background

It is commonly assumed that the behavior of economic agents depends upon expectational variables. Typical examples of these kinds of variables are forecasts of prices, orders placed etc., and normal levels of interest rates, permanent income and so on. Since the expectational variables are seldom directly observed, models which include variables of this type are often made empirically testable by specifying the way in which expectations are adjusted.

The simplest model of the type considered is

$$y_t = a_0 + a_1 x_t^* + u_t \tag{1.1}$$

where y_t is the level of some economic variable, x_t^* is an expectational variable, and u_t is a stochastic disturbance. A common assumption is

$$x_t^* - x_{t-1}^* = (1 - \beta)[x_{t-1} - x_{t-1}^*] . \tag{1.2}$$

Equations (1.1) and (1.2) give

(1.3) $y_t = \beta y_{t-1} + a_0(1 - \beta) + a_1(1 - \beta) x_{t-1} + u_t - \beta u_{t-1}$

The model (1.1 - 1.2) was used by Nerlove [5] to explain the acreage (y_t) that farmers alloted to a given crop x_t^* was the expected "normal" price for that crop. The basic idea was that farmers determined acreage allocation not on the basis of the price expected next year, but rather considered an average of expected prices over future periods. Equation (1.2) implies that

(1.4) $$x_t^* = (1 - \beta) \sum_{j=0}^{\infty} \beta^j x_{t-1-j}$$

so that x_t^* may be thought of as being a weighted average of all past values of the observed series x_t. More generally, x_t^* may be written as

(1.5) $$x_t = \sum_{j=0}^{\infty} b_j x_{t-j} = B(L)x_t \quad \text{where}$$

$$B(z) = \sum_{j=0}^{\infty} b_j z^j \quad \text{and} \quad L^j x_t = x_{t-j} \quad .$$

This representation is obviously intractable for estimation purposes, so one specifies a simpler form such as (1.2) before proceeding with estimation. While (1.2) is probably the most commonly used specification it is clear that a wide variety of other specifications could serve as well.

Suppose that

$$B(z) = P(z)/Q(z) \tag{1.6}$$

where $P(z)$ and $Q(z)$ are polynomials of degree p and q respectively. That is, $B(L)$ is a so-called rational lag distribution (Jorgenson [3]). The example given above corresponds to choosing $P(z) = 1 - \beta$ and $Q(z) = 1 - \beta z$. Assuming that $B(L)$ is a rational lag distribution allows one to transform (1.1) which in principle involves all the past values of x_t into an equation such as (1.3) which includes only a finite number of past x_t's and a finite number of lagged y_t's . Without further <u>a priori</u> specification there is no way to choose among a variety of lag distributions except on statistical grounds.

The following model also leads to (1.3) as the equation to be estimated (Muth [4], Nerlove [6])

$$y_t = a_0 + a_1 \hat{x}^*_{t,t-1} + a_t$$

$$(1.7) \quad x_t = x^*_t + \eta_t \quad E(\eta_t) = 0 \qquad E(\eta_t \eta_s) = \delta_{t,s}\sigma^2$$

$$x^*_t = x^*_{t-1} + \epsilon_t \quad E(\epsilon_t) = 0 \qquad E(\epsilon_t \epsilon_s) = \delta_{t,s}\lambda\sigma^2$$

$$E(\epsilon_t \eta_s) = 0$$

where $\delta_{t,s}$ is the Kronecker delta, and $\hat{x}^*_{t,t-1}$ is the linear minimum mean square error forecast of x^*_t made at time $t-1$ using past values of the observed series $\{x_t\}$. In this case β is given by

$$\beta = \frac{\lambda+2 - \sqrt{(\lambda + 2)^2 - 4}}{2} \tag{1.8}$$

Now λ is equal to $\sigma_\epsilon^2/\sigma_\eta^2$ so when λ is small any given observation on x_{t-1} gives little new information about x_t^* , and, therefore, expectations are revised only slightly. Conversely, when λ is large $1 - \beta$ is close to one and $\hat{x}^*_{t,t-1}$ is nearly equal to x_{t-1} . This type of behavior is clearly sensible, since when $\frac{\sigma_\epsilon^2}{\sigma_\eta^2}$ is big

most of the observed change in x_t will on the average be due to changes in x_t^* . Identifying $\hat{x}^*_{t,t-1}$ with expected normal prices makes sense since

$$\hat{x}^*_{t+j,\ t-1} = \hat{x}^*_{t,t-1} \quad j = 1, 2, 3, \ldots \tag{1.9}$$

Another model which differs only slightly from the one given in equations (1.7) also leads to the estimating equation (1.3).

$$\begin{aligned} & y_t = a_0 + a_1 \hat{x}_{t,t-1} + u_t \\ (1.10) \qquad & x_t - x_{t-1} = \epsilon_t - \beta\epsilon_{t-1} \qquad E(\epsilon_t) = 0 \quad |\beta| < 1 \\ & \qquad\qquad\qquad\qquad\qquad E(\epsilon_t\epsilon_s) = \delta_{t,s}\sigma^2 \end{aligned}$$

where $\hat{x}_{t,\ t-1}$ is the least squares forecast of x_t calculated at time $t-1$.

The models given by equations (1.7) and equations (1.10) lead to the same estimating equations as obtained from equations (1.1-1.2). However, they differ in one important aspect: the assumption about the lag distribution (1.2) has been replaced by two alternative assumptions, one, a behavioral assumption, is that expectations are formed according to a minimum mean square error criterion. The other assumption which could be tested using data on x_t specifies the stochastic structure of the observed series.

In model (1.7) the variable being forecast was unobservable, while in (1.10) the expectational variable was the forecasted value of the observed series; yet in both cases the calculation of the expectational variable by least squares led to a form of a rational distributed lag. In the next section it will be shown that these results generalize to a very wide class of stochastic specifications.

Section 2. Prediction

In what follows it is assumed that the stochastic processes are mean zero covariance stationary processes of the moving average, auto-regressive type. That is,

$$E(x_t) = 0$$

$$E(x_t \, x_{t-k}) = \gamma(k) \qquad (2.1)$$

$$E(x_t^2) = \gamma(0) < \infty$$

The results presented below can be extended to processes whose ρ^{th} differences satisfy (2.1), but for ease of exposition equations (2.1) will be assumed to hold throughout.[1]

The covariance generating function of $\{x_t\}$ is defined as

(2.2) $$g_{xx}(z) = \sum_{-\infty}^{\infty} \gamma(k) z^k$$

If $H(z) = \sum_{-\infty}^{\infty} h_i z^i$ is the Laurent expansion of a function which converges in an annulus containing the unit circle, then

(2.3)
$$[H(z)]_+ = \sum_{0}^{\infty} h_i z^i$$
$$[H(z)]_- = \sum_{-\infty}^{-1} h_i z^i$$
$$[H(z)]_k = \sum_{k}^{\infty} h_i z^i \qquad \text{also converge.}$$

Any covariance stationary time series may be represented as (Wold [10])

(2.4) $x_t = \eta_t + \zeta_t$ where $\{\eta_t\}$ and $\{\zeta_t\}$ are mutually uncorrelated and

$$\eta_t = \sum_{j=0}^{\infty} \rho_j e^{i\lambda_j t}$$

$$\zeta_t = \sum_{i=0}^{\infty} b_i \epsilon_{t-i} \qquad E(\epsilon_t) = 0$$

$$E(\epsilon_t \epsilon_s) = \delta_{t,s} \sigma^2$$

$$\Sigma b_j^2 < \infty$$

$\{\eta_t\}$ is called the linearly deterministic part of the series and $\{\zeta_t\}$ is called the non-deterministic part. Throughout this paper it is assumed that all processes are purely non-deterministic, that is, they are representable as a one-sided moving average with white noise inputs:

(2.5) $$x_t = \sum_{i=0}^{\infty} b_i \varepsilon_{t-i} = B(L)\varepsilon_t \qquad \text{where} \quad B(z) = \sum_{i=0}^{\infty} b_i z^i$$

and L is the lag operator. For any such process $\{x_t\}$ it can be shown (Whittle [9]) that the least squares forecast of x_{t+v} calculated at time t is given by

(2.6) $$\hat{x}_{t+v,t} = \gamma(L)x_t \qquad \text{where}$$

$$\gamma(z) = \frac{1}{B(z)} \left[\frac{B(z)}{z^v} \right]_+ .$$

where $B(z)$ is the canomical factorization of the covariance generating function i.e.

(2.7) $g_{xx}(z) = \sigma^2 B(z)B(z^{-1})$, and all the roots of $B(z)$ lie outside the unit circle. By convention σ^2 is chosen so that $b_0 \equiv 1$.

The following theorem due to Whittle [9, p. 93] will be used often in the analysis below:

Theorem: Let $Q(\ )$ be a function of z analytic in $\delta < |z| < \delta^{-1}$ and θ a number such that $|\theta| < 1$. Then $R(z) = (1-\theta z)^{\rho} \left[\frac{Q(z)}{(1-\theta z)^{\rho}}\right]_+$

$= [Q(z)]_+ + \pi_{\rho}(z)$ where

$$\pi_{\rho}(z) = \sum_{i=0}^{\rho-1} \frac{d^i[Q(z)]}{dz^i} - \Big|_{z=\theta^{-1}} \frac{(z-\theta^{-1})^i}{i!} \qquad \text{2.}$$

The second example in Section 1 showed that the adaptive expectations model could be rationalized as a least squares forecast of a certain type of time series. The generalization of that example is contained in the following.

Theorem: Let $N(\ )$ and $D(\ )$ be polynomials of degreee n and d respectively, and let $x_t = B(L)\ \epsilon_t = \frac{N(L)}{D(L)}\ \epsilon_t$ be a covariance stationary time series. Then the least squares forecast of x_{t+v} made at time t is given by $\hat{x}_{t+v,\ t} = \gamma(L)x_t$, where $\gamma(z) = \frac{M(z)}{N(z)}$ and $M(\ )$ is a polynomial of degree $\text{Max}\ [n - v,\ d - 1,\ 0]$.

Proof:

$$\gamma(z) = \frac{1}{B(z)} \left[\frac{B(z)}{z^v}\right]_+ \tag{2.8}$$

$$= \frac{D(z)}{N(z)} \left[\frac{N(z)z^{-v}}{D(z)}\right]_+$$

$$= \frac{D(z)}{N(z)} \frac{[N(z)z^{-v}]_+}{D(z)} + \left[\frac{[N(z)z^{-v}]_-}{D(z)}\right]_+$$

$$= \frac{[N(z)z^{-v}]_+}{N(z)} + \frac{D(z)}{N(z)} \left[\frac{[N(z)z^{-v}]_-}{D(z)}\right]_+$$

Assume that all the roots of $D(\ )$ are distinct (this is not necessary as will become clear, but merely facilitates the exposition) and denote them by λ_i , $i = 1, 2, \ldots, d$. Expanding $\frac{1}{D(\ )}$ by partial fractions yields

$$(2.9) \qquad \gamma(z) = \frac{[N(z)z^{-v}]_+}{N(z)} + \frac{D(z)}{N(z)} \sum_{i=1}^{d} \left[\frac{\delta_i [N(z)z^{-v}]_-}{1 - \lambda_i z}\right]_+$$

$$= \frac{[N(z)z^{-v}]_+}{N(z)} + \frac{D(z)}{N(z)} \sum_{i=1}^{d} \delta_i \left[\frac{[N(z)z^{-v}]_-}{1 - \lambda_i z}\right]_+$$

where

$$\frac{1}{D(z)} = \sum_{i=1}^{d} \frac{\delta_i}{1 - \lambda_i z} .$$

Applying Whittle's theorem to each term in the summation gives

$$(2.9) \qquad \delta_i \left[\frac{[N(z)z^{-v}]_-}{1 - \lambda_i z}\right]_+ = \delta_i \frac{[N(z)z^{-v}]_-}{1 - \lambda_i z} \Big|_{z = \lambda_i^{-1}}$$

So on combining terms

$$(2.10) \qquad \gamma(z) = \frac{[N(z)z^{-v}]_+}{N(z)} + \frac{E(z)}{N(z)} = \frac{M(z)}{N(z)}$$

Now $E(\ )$ is a polynomial of order $e = d-1$, so $m = \max[n-v, d-1, 0]$. In general $\gamma(\ )$ is a rational function whose denominator is the numerator of $B(\ )$ [i.e. the moving average part] and the numerator of which is of order $\max[n-v, d-1]$. The presence of multiple roots in $D(\ )$ does not alter the conclusion in any way. If λ_i is a p^{th} order root, then instead of $\left[\frac{\delta_i[N(z)z^{-v}]_-}{1-\lambda_i z}\right]_+$ in (2.9) one would have $\sum_{j=1}^{p}\left[\frac{\delta_i^j[N(z)z^{-v}]_-}{(1-\lambda_i z)^j}\right]_+$. The theorm of Whittle can now be applied to each term in this latter sum and on combining terms this will give

$$\frac{S_i(z)}{(1-\lambda_i z)^p} \quad \text{where the order of } S_i(\ ) \text{ is } p-1 .$$

In either case the statement about the order of the polynomials in $\gamma(\ )$ is still true.

To illustrate the various procedures and the general nature of the calculations the following example will be treated in some detail.

$$(2.11)\quad x_t = \frac{(1-\alpha L)(1-\beta L)}{(1-\delta L)(1-\lambda L)}\epsilon_t \qquad \{\epsilon_t^2\} \text{ white noise}$$

$$E(\epsilon_t^2) = \sigma^2$$

The generating function for a v-step predictor is

$$(2.12)\qquad \gamma(z) = \frac{1}{B(z)}\left[\frac{B(z)}{z^v}\right]_+$$

$$= \frac{(1-\delta z)(1-\lambda z)}{(1-\alpha z)(1-\beta z)}\left[\frac{(1-\alpha z)(1-\beta z)}{(1-\delta z)(1-\lambda z)z^v}\right]_+$$

The first step in obtaining the predictor is to calculate the value of the expression under the $[\ \ ]_+$ operator. Now

$$(2.13)\qquad \frac{1}{(1-\delta z)(1-\lambda z)} = \frac{1}{\delta-\lambda}\left[\frac{\delta}{1-\delta z} - \frac{\lambda}{1-\lambda z}\right]$$

So

$$(2.14)\qquad \left[\frac{(1-\alpha z)(1-\beta z)}{(1-\delta z)(1-\lambda z)z^v}\right]_+ = \frac{1}{\delta-\lambda}\left[\frac{\delta(1-\alpha z)(1-\beta z)z^v}{1-\delta z} - \frac{\lambda(1-\alpha z)(1-\beta z)z^{-v}}{1-\lambda z}\right]_+$$

$$= \frac{\delta}{\delta-\lambda}\left[\frac{(1-\alpha z)(1-\beta z)z^{-v}}{(1-\delta z)}\right]_+$$

$$- \frac{\lambda}{\delta-\lambda}\left[\frac{(1-\alpha z)(1-\beta z)z^{-v}}{(1-\lambda z)}\right]_+$$

For $v = 1$

$$(2.15)\qquad \left[\frac{(1-\alpha z)(1-\beta z)z^{-1}}{1-\pi z}\right]_+ = \frac{-(\alpha+\beta)+\alpha\beta z}{1-\pi z} + \frac{\pi}{1-\pi z}$$

$$= \frac{\pi-(\alpha+\beta)+\alpha\beta z}{1-\pi z}$$

from Whittle's theorem. Here $(1 - \alpha z)(1 - \beta z)z^{-1}$ plays the role of $Q(z)$ in the statement of the theorem. Using the above relation and combining terms gives

$$\left[\frac{(1 - \alpha z)(1 - \beta z)}{(1 - \delta z)(1 - \lambda z)z}\right]_+ = \frac{(\delta + \lambda) - (\alpha + \beta) + (\alpha\beta - \lambda\delta)z}{(1 - \delta z)(1 - \lambda z)} \tag{2.16}$$

For $v = 2$

$$\left[\frac{(1 - \alpha z)(1 - \beta z)z^{-2}}{(1 - \pi z)}\right]_+ = \frac{\pi^2 - \alpha\beta - (\alpha + \beta)\pi}{1 - \pi z} \tag{2.17}$$

which gives

$$\left[\frac{(1 - \alpha z)(1 - \beta z)}{(1 - \delta z)(1 - \lambda z)z^2}\right]_+ = \frac{(\delta + \lambda)^2 - (\alpha + \beta)(\lambda + \delta) - (\alpha\beta - \lambda\delta) - \lambda\delta(\lambda + \delta - (\alpha + \beta)z}{(1 - \delta z)(1 - \lambda z)} \tag{2.18}$$

In general for v greater than two

$$\left[\frac{(1 - \alpha z)(1 - \beta z)}{(1 - \delta z)(1 - \lambda z)z^v}\right]_+ = \frac{C_v - (\alpha + \beta)C_{v-1} + \alpha\beta C_{v-2} - \lambda\delta[C_{v-1} - (\alpha + \beta)C_{v-2} + \alpha\beta C_{v-3}]z}{(1 - \delta z)(1 - \lambda z)} \tag{2.19}$$

where the sequence $\{C_i\}$ is defined by

$$\sum_{i=0}^{\infty} C_i z^i = \sum_{j=0}^{\infty} \delta^j z^j \sum_{k=0}^{\infty} \lambda^k z^k \tag{2.20}$$

If the pair (α, β) and the pair (δ, λ) are the same, then all expressions vanish. If this were the case then the process $\{x_t\}$ would simply be white noise and its forecast values would be zero. The generating functions for the forecasts are

$$\gamma_1(z) = \frac{(\lambda + \delta) - (\alpha + \beta) - (\delta\lambda - \alpha\beta)z}{(1 - \alpha z)(1 - \beta z)}$$

(2.21)

$$\gamma_2(z) = \frac{[(\delta + \lambda)^2 - (\lambda + \delta)(\alpha + \beta) - (\alpha\beta - \lambda\delta)] - \lambda\delta(\lambda + \delta - (\alpha + \beta))z}{(1 - \alpha z)(1 - \beta z)}$$

and in general

$$\gamma_v(z) = \frac{C_v - (\alpha + \beta)C_{v-1} + \alpha\beta C_{v-2} - \lambda\delta[C_{v-1} - (\alpha + \beta)C_{v-2} + \alpha\beta C_{v-3}]z}{(1 - \alpha z)(1 - \beta z)}$$

As is always the case except when $\{x_t\}$ is an autoregression, the generating functions are infinite series. However, if $\hat{x}_{t+v,\ t}$ appears as an explanatory variable in an equation the form of the lag distribution is such that the equation can be transformed into an equation containing two lagged values of the dependent variable and the current and lagged values of $\{x_t\}$. Note that the transformation will, in general, introduce serial correlation among the disturbances of the equation to be estimated.

Section 3. Signal Extraction

Recently Nerlove [6] considered the derivation of distributed lags in situations similar to the first example given in section 1. Specifically he assumed that economic agents reacted to estimates of of unobserved components of economic time series. Nerlove treated several examples and derived the lag distributions. In this section it is shown that Nerlove's results generalize to all covariance stationary processes of the mixed moving average, auto-regressive type. In addition an example is given which leads to a lag distribution of the Pascal type.

Assume that $\{x_t, y_t\}$ are jointly covariance stationary non-deterministic time series. The cross covariance generating function is defined by

(3.1) $$g_{yx}(z) = \sum_{k=-\infty}^{\infty} \Gamma_{yx}(k) z^k$$

where $$\Gamma_{yx}(k) = E(y_t x_{t-k})$$

If $g_{xx}(z) = \sigma^2 B(z) B(z^{-1})$, then the least squares estimate of y_t is given by

(3.2) $$\hat{y}_{t,t} = \gamma(L) x_t \quad \text{where}$$

$$\gamma(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{g_{yx}(z)}{B(z^{-1})} \right]_+$$ 3.

The case to be considered here is

(3.3) $$x_t = y_t + \eta_t$$

where x_t is the observed series, y_t is the signal, and η_t is the noise. Assuming that

(3.4) $$E(y_t \eta_{t-k}) = 0 \quad k = 0, \pm 1, \ldots$$

equation (3.2) reduces to

$$\hat{y}_{t+v,\,t} = \gamma(L)x_t$$

(3.5) $$\gamma(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{g_{yy}(z)}{B(z^{-1})z^v}\right]_+$$

Theorem: Let $y_t = \frac{P(L)}{Q(L)} \epsilon_t$, $\eta_t = \frac{R(L)}{S(L)} \zeta_t$

and let $x_t = y_t + \eta_t$. $\{\epsilon_t\}$, $\{\zeta_t\}$ are mutually uncorrelated white noise sequences, and $P(\ )$, $Q(\ )$, $R(\ )$, and $S(\ )$ are polynomials of degree p , q , r , and s respectively. The least squares estimate of y_{t+v} made at time t is given by

$$\hat{y}_{t+v,\,t} = \gamma(L)x_t$$

(3.6)

$$\gamma(z) = \frac{S(z)N(z)}{T(z)}$$

where $\sigma^2 T(z)T(z^{-1}) = \sigma_\epsilon^2 P(z)S(z)P(z^{-1})S(z^{-1}) + \sigma_\zeta^2 R(z)Q(z)R(z^{-1})Q(z^{-1})$.

σ^2 being chosen that $t_0 = 1$. The order of the polynomial $N(\ )$ is $n = \max\{P - v , q - 1, 0\}$.

Proof:

First note that

$$g_{yy}(z) = \sigma_\epsilon^2 \frac{P(z)P(z^{-1})}{Q(z)Q(z^{-1})} \quad \text{and}$$

$$g_{xx}(z) = \frac{\sigma^2 T(z)T(z^{-1})}{Q(z)S(z)Q(z^{-1})S(z^{-1})} \quad .$$

Thus from (3.5)

$$(3.7) \qquad \gamma(z) = \frac{Q(z)S(z)}{\sigma^2 T(z)} \left[\frac{\sigma_\epsilon^2 P(z)P(z^{-1})S(z^{-1})Q(z^{-1})}{Q(z)Q(z^{-1})T(z^{-1})z^v} \right]_+$$

$$= \frac{Q(z)S(z)}{\sigma^2 T(z)} \left[\frac{\sigma_\epsilon^2 P(z)P(z^{-1})S(z^{-1})}{T(z^{-1})z^v Q(z)} \right]_+$$

The expression under the $[\ ]_+$ operator can be evaluated using Whittle's theorem. To see this imagine expanding $1/Q(z)$ by partial fractions and applying Whittle's theorem to each term in the resulting sum. In this case, $\sigma_\epsilon \frac{P(z)P(z^{-1})S(z^{-1})}{T(z^{-1})z^v}$ plays the role of $Q(z)$ in the statement of the theorem. These calculations give

$$(3.8) \quad \left[\frac{\sigma_\epsilon^2 P(z)P(z^{-1})S(z^{-1})}{T(z^{-1})Q(z)z^v} \right]_+ = \frac{\left[\frac{\sigma^2 P(z)P(z^{-1})S(z^{-1})}{T(z^{-1})z^v} \right]_+}{Q(z)} + \frac{V(z)}{Q(z)}$$

$$= \frac{N(z)}{Q(z)}$$

where the order of $V(\ )$ is $q - 1$ and the order of $N(\ )$ is $\max\,[p - v\ ,\ q - 1\ ,\ 0]$.

As was the case with predictions, the result allows for recursive calculation of $\hat{y}_{t+v,t}$ and for the derivation of distributed lags in models using $\hat{y}_{t+v,t}$ as an exogenous variable. Again the orders of the lag operators are simple functions of the prediction period and of the orders of the operators in the two components.

Example.

Suppose one wishes to predict

(3.9) $$y_t = \frac{\epsilon_t}{(1 - \theta L)(1 - \rho L)}$$

where y_t is observed with an error, i.e. one observes

(3.10) $$x_t = y_t + \eta_t$$

$$= \frac{\epsilon_t}{(1 - \theta L)(1 - \rho L)} + \eta_t$$

$\{\epsilon_t\}$, $\{\eta_t\}$ white noise and uncorrelated with $E(\epsilon_t^2) = \sigma^2$

$E(\eta_t^2) = \lambda\sigma^2$

Then

(3.11) $$x_t = \frac{\epsilon_t + (1 - \theta L)(1 - \rho L)\eta_t}{(1 - \theta L)(1 - \rho L)}$$

so

(3.12) $$g_{xx}(z) = \frac{\sigma^2[1 + \lambda(1 - \theta z)(1 - \rho z)(1 - \theta z^{-1})(1 - \rho z^{-1})]}{(1 - \theta z)(1 - \rho z)(1 - \theta z^{-1})(1 - \rho z^{-1})}$$

$$= \frac{\sigma_1^2[(1 - \alpha z)(1 - \beta z)(1 - \alpha z^{-1})(1 - \beta z^{-1})]}{(1 - \theta z)(1 - \rho z)}$$

where α , β are the roots of $1 + \lambda(1 - \theta z)(1 - \theta z^{-1})(1 - \rho z)(1 - \rho z^{-1})$ that lie inside the unit circle. To obtain $\hat{y}_{t+v,t}$

$$(3.13)\quad \gamma(z) = \frac{(1-\theta z)(1-\rho z)}{\sigma_1^2(1-\alpha z)(1-\beta z)}\left[\frac{\sigma^2(1-\theta z^{-1})(1-\rho z^{-1})}{(1-\theta z)(1-\theta z^{-1})(1-\rho z)(1-\rho z^{-1})(1-\alpha z^{-1})(1-\beta z^{-1})z^v}\right]_+$$

$$= \frac{\sigma^2}{\sigma_1^2}\frac{(1-\theta z)(1-\rho z)}{(1-\alpha z)(1-\beta z)}\left[\frac{1}{(1-\theta z)(1-\rho z)(1-\alpha z^{-1})(1-\beta z^{-1})z^v}\right]_+$$

$$= \frac{\sigma^2}{\sigma_1^2}\frac{(1-\theta z)\ (1-\rho z)}{(1-\alpha z)(1-\beta z)(\theta-\rho)}\left[\frac{\theta}{(1-\theta z)(1-\alpha z^{-1})(1-\beta z^{-1})z^v} - \frac{\rho}{(1-\rho z)(1-\alpha z^{-1})(1-\beta z^{-1})z^v}\right]_+$$

$$= \frac{\sigma^2}{\sigma_1^2}\ \frac{(1-\theta z)(1-\rho z)}{(1-\alpha z)(1-\beta z)(\theta-\rho)}\left\{\left[\frac{\theta}{(1-\theta z)(1-\alpha z^{-1})(1-\beta z^{-1})z^v}\right]_+ - \left[\frac{\rho}{(1-\rho z)(1-\alpha z^{-1})(1-\beta z^{-1})z^v}\right]_+\right\}$$

Letting $\dfrac{z^{-v}}{(1-\alpha z^{-1})(1-\beta z^{-1})}$ play the role of $Q(z)$ in Whittle's theorem and noting that for v positive $[Q(z)]_+$ vanishes:

$$(3.14) \qquad \gamma(z) = \frac{\sigma^2}{\sigma_1^2} \frac{(1-\rho z)\theta^{v+1}}{(1-\alpha z)(1-\beta z)(\theta-\rho)(1-\alpha\theta)(1-\beta\theta)}$$

$$- \frac{\sigma^2}{\sigma_1^2} \frac{(1-\theta z)\rho^{v+1}}{(1-\alpha z)(1-\beta z)(\theta-\rho)(1-\alpha\rho)(1-\beta\rho)}$$

This somewhat unlovely expression is clearly of the form

$$(3.15) \qquad \gamma(z) = \frac{\delta + kz}{(1 - \alpha z)(1 - \beta z)}$$

Now δ and k can be calculated by combining terms as in the earlier example. There is, however, another approach available which for numerical calculations should be relatively simple. If $z = \theta^{-1}$ the the second term in (3.14) is zero and

$$(3.16) \quad \gamma(\theta^{-1}) = \frac{\sigma^2}{\sigma_1^2} \frac{(1-\rho\theta^{-1})\theta^{v+1}}{(1-\alpha\theta^{-1})(1-\alpha\theta)(1-\beta\theta^{-1})(1-\beta\theta)(\theta-\rho)}$$

$$= \frac{\sigma^2}{\sigma_1^2} \frac{\theta^{v}}{(1-\alpha\theta)(1-\alpha\theta^{-1})(1-\beta\theta)(1-\beta\theta^{-1})}$$

$$= \theta^{v}$$

as

$$\sigma^2[1 + \lambda(1-\theta z)(1-\theta z^{-1})(1-\rho z)(1-\rho z^{-1})] = \sigma_1^2[(1-\alpha z)(1-\alpha z^{-1})(1-\beta z)(1-\beta z^{-1})]$$

so

$$(3.17) \qquad \frac{\sigma^2}{\sigma_1^2} = (1-\alpha\theta)(1-\alpha\theta^{-1})(1-\beta\theta)(1-\beta\theta^{-1})$$

$$= (1-\alpha\rho)(1-\alpha\rho^{-1})(1-\beta\rho)(1-\beta\rho^{-1})$$

By a similar argument using the second equation in (3.17) it is easily seen that

$$\gamma(\rho^{-1}) = \rho^{v} \tag{3.18}$$

These two restrictions are sufficient to determine δ and k. The form of $\gamma(\ )$ as determined yields the following recursive scheme

$$\hat{y}_{t+v,t} - (\alpha + \beta)\hat{y}_{t+v-1,t-1} + \alpha\beta\hat{y}_{t+v-2,\ t-2} = \delta x_t + kx_{t-1} \tag{3.19}$$

Nerlove [6] considered examples which produced distributed lag patterns which closely resembled the so-called Pascal distributed lag (i.e. a convolution of several Koyck-Nerlove lag distributions all with the same parameter). No example in which the lag distribution was exactly of this type was given, however. Using the results presented above examples can be easily generated. For instance,

$$x_t = \frac{P(L)}{Q(L)}\epsilon_t + \frac{R(L)}{S(L)}\eta_t \qquad E(\epsilon_t^2) = B^2$$

$$= (1-\beta L)^2\epsilon_t + (1-\alpha L)(1-\gamma L)\eta_t \qquad E(\eta_t^2) = A^2F \quad F > 0$$

where

$$(1-\alpha z)(1-\alpha z^{-1}) + \frac{2B}{A}(1-\beta z)(1-\beta z^{-1}) = F(1-\gamma z)(1-\gamma z^{-1}) .$$

Then

$$\sigma^2T(z)T(z^{-1}) = B^2(1-\beta z)^2(1-\beta z^{-1})^2 + A^2F(1-\alpha z)(1-\alpha z^{-1})(1-\gamma z)(1-\gamma z^{-1})$$

$$= B^2(1-\beta z)^2(1-\beta z^{-1})^2 + A^2(1-\alpha z)(1-\alpha z^{-1})((1-\alpha z)(1-\alpha z^{-1}) + \frac{2B}{A}(1-\beta z)(1-\beta z^{-1}))$$

$$= B^2(1-\beta z)^2(1-\beta z^{-1})^2 + 2AB(1-\alpha z)(1-\alpha z^{-1})(1-\beta z)(1-\beta z^{-1})$$

$$+ A^2(1-\alpha z)^2(1-\alpha z^{-1})^2$$

$$= [B(1-\beta z)(1-\beta z^{-1}) + A(1-\alpha z)(1-\alpha z^{-1})]^2$$

$$= (D(1-\delta z)(1-\delta z^{-1}))^2$$

$$= D^2(1-\delta z)^2(1-\delta z^{-1})^2$$

So $T(z) = (1-\delta z)^2$

From (3.6)

$$\gamma(z) = \frac{S(z)N(z)}{T(z)} \qquad N = \max\,\{P-v,\ Q-1,\ 0\}$$

In this example

$$S(z) \equiv 1$$

$$T(z) = (1-\delta z)^2$$

$$P = 2$$

$$Q = 0$$

So for $v = 2$ the solution is a second order Pascal type distributed lag.

Section 4. Other Applications.

The results presented in Section 2 and 3 can be extended to a wider class of problems. In general if the exogenous variables have rational spectral densities (are of the moving average, auto-regressive type) or if their p^{th} differences have rational spectra, then maximizing (minimizing) a quadratic objective function leads to a rational distributed lag. Further, estimating the stochastic properties of the exogenous variables provides information about the lag distribution and may aid in the identification of the structural parameters. To illustrate these points an example of a inventory adjustment, production smoothing model will be considered.

Assume that a firm's cost in period t are given by

$$(4.1)\quad C_t = \lambda_1(\hat{P}_t - \hat{P}_{t-1})^2 + \lambda_2(\hat{I}_t - a - \alpha\hat{S}_t)^2 + \lambda_3\hat{P}_t + \lambda_1\hat{I}_t$$

where

$\hat{P}_t$ = production in period t

$\hat{S}_t$ = sales in period t

$\hat{I}_t$ = inventories at the end of the t^{th} period

$$\hat{P}_t = \hat{S}_t + \hat{I}_t - \hat{I}_{t-1} \qquad \underline{4.}|$$

At time $t-1$ the firm is assumed to choose the level of production for period t in order to minimize

$$(4.2)\qquad V = E\{\sum_{v=0}^{\infty} \rho^v C_{t+v}\} \qquad 0 \leq \rho \leq 1$$

Given the accounting identies between production sales, and the change in inventories it sufficies to determine either production or inventories. While formally it makes little difference which quantity is set it seems more natural to view the firm as determining the level of production directly rather than the level of inventories. Partly this is because determining production levels seems the more basic decision, but also the choice of which variable to determine implies different reasons for the holding of inventories. The point is that the decision as to the level of production or inventories for period t must be made before the t^{th} period, based upon sales expected in period t and in subsequent periods. In fact, these expectations will not be fulfilled exactly so that all plans will not be carried out. If the firm sets the level of production then forecasting errors will show up in differences between actual end of period inventories and the planned level of inventories. On the other hand, if inventories are treated as being determined, then forecasting errors will cause unintended fluctuations in the level of production. This latter case seems implausible in that one of the reasons for holding inventories is presumably to protect against errors in forecasting sales, and, also, production plans are more likely to be somewhat inflexible. Furthermore, if the forecast error made in the t^{th} period partially determines the t^{th} period's production, it is no longer so clear why one assumes that production plans for period t must be made prior to knowing period t's sales. For this example it is assumed that production is the control variable, though one could as easily take it to be inventories.

Let

$$\hat{S}_t = \bar{S} + S_t \qquad E(S_t) = 0 \qquad \bar{S} \text{ a constant}$$

$$(4.3)\quad \hat{I}_t = \bar{I} + I_t$$

$$\hat{P}_t = \bar{P} + P_t \qquad \bar{P} = \bar{S}$$

Then (4.2) may be written as

$$(4.4)\quad V = E\{\sum_{v=0}^{\infty} \rho^{v}(\lambda_1(P_{t+v} - P_{t-1+v})^2 + \lambda_2(I_{t+v} - \alpha S_{t+v})^2\}$$

$$+ \sum_{v=0}^{\infty} \rho^{v}(\lambda_3\bar{P} + \lambda_4\bar{I} + \lambda_2(\bar{I} - a - \alpha\bar{S}))$$

The second term is easily minimized, so the problem reduces to

$$\min E\{\sum_{v=0}^{\infty} \rho(\lambda(P_{t+v} - P_{t+v-1})^2 + (1 - \lambda)(I_{t+v} - \alpha S_{t+v})^2\}$$

where $P_t = S_t + I_t - I_{t-1}$ and all series have mean zero. Substituting

$$(4.5)\qquad I_t = \frac{P_t - S_t}{1 - L}$$

into (4.4) the problem becomes

$$\min_{\text{wrt. } \{P_{t+v}\}} E\{\sum_{v=0}^{\infty} \rho^{v}(\lambda(P_{t+v} - P_{t+v-1})^2 + (1 - \lambda)\left(\frac{P_{t+v} - S_{t+v}}{1 - L} - \alpha S_{t+v}\right)^2\}$$

$$0 \leq \lambda \leq 1$$

The meaning of the expression (4.5) is simply that inventories at any time are the sum of past discrepancies between production and sales plus any initial stock.[5]

Differentiating the objective function with respect to P_{t+v} gives the following first order conditions:

$$(4.6)\quad \frac{(\lambda(1-L)^2(1-\rho L^{-1})^2 + 1-\lambda)P_{t+v} - (1-\lambda)(1+\alpha(1-L)\hat{S}_{t+v,t-1})}{(1 - L)(1 - \rho L^{-1})} = 0$$

$v = 0, 1, 2, \ldots$

In (4.6) S_{t+v} has been replaced by its certainty equivalent $\hat{S}_{t+v,t-1}$ (see Theil [7])

Now

$$(4.7)\quad \lambda(1-z)^2(1-\rho z^{-1})^2 + (1-\lambda) = \lambda z^{-2}(z-\beta)(z-\bar{\beta})(z-\frac{\rho}{\beta})(z-\frac{\rho}{\bar{\beta}})$$

$$= \lambda|\beta|^2(z-\frac{1}{\beta}z)(1-\frac{1}{\bar{\beta}}z)(1-\frac{\rho}{\beta}z^{-1})(1-\frac{\rho}{\bar{\beta}}z^{-1})$$

$$= R(z)Q(z^{-1})$$

where

$$R(z) = \sqrt{\lambda}|\beta|(1-\frac{1}{\beta}z)(1-\frac{1}{\bar{\beta}}z)$$

$$Q(z^{-1}) = \sqrt{\lambda}|\beta|(1-\frac{\rho}{\beta}z^{-1})(1-\frac{\rho}{\bar{\beta}}z^{-1})$$

By inspection it is seen that if β is a root of (4.7), then $\frac{\rho}{\beta}$ is also a root. Next it will be shown that there is always a root β with $|\beta| > 1$.

If z_0 is a root of (4.7) then $(\lambda \neq 0)$

$$(4.8)\quad (1-z_0)(1-\rho z_0^{-1}) = \pm i \sqrt{\frac{1-\lambda}{\lambda}}$$

$$\text{or } z = \frac{\pm\ 1 + \rho - i\sqrt{\frac{1-\lambda}{\lambda}} \pm \sqrt{(1 + \rho - i\sqrt{\frac{1-\lambda}{\lambda}})^2 - 4\rho}}{2}$$

From (4.7) it is easily seen that for any λ in (0, 1) there must be at least one root outside the unit circle if $\rho = 0$ or if $\rho = 1$.

Further, it is clear that for any value of ρ in (0, 1) there can be no roots on the unit circle. But for any fixed λ the roots are bounded continuous functions of ρ, so there must be at least one root outside the unit circle, i.e., $|\beta| > 1$.[6]

The preceding argument implies that (4.6) may be written as:

$$(4.9) \quad \frac{R(L)Q(L^{-1})P_{t+v} - (1-\lambda)(1+\alpha(1-L))\hat{S}_{t+v,t-1}}{(1-L)(1-\rho L^{-1})} = \qquad v = 0, 1, 2, \ldots$$

where the roots of $R(z)$ lie outside the unit circle, and the roots of $Q(z)$ lie inside the unit circle.

Now

$$(4.10) \quad \frac{R(L)Q(L^{-1})P_{t+v} - (1-\lambda)(1+\alpha(1-L))\hat{S}_{t+v,t-1}}{(1-L)(1-\rho L^{-1})} =$$

$$\frac{R(L)Q(L^{-1}) - (1-\lambda)(1+\alpha(1-L))S_{t,t-1}}{1-L} +$$

$$\rho \left[\frac{R(z)Q(L^{-1}) - (1-\lambda)(1+\alpha(1-L)\hat{S}_{t+v+1,\ t-1}}{(1-L)(1-\rho L^{-1})} \right] .$$

But the second term is zero so one may write

$$(4.11) \quad \frac{R(L)P_t}{1-L} = (1-\lambda)(1+\alpha(1-L))\tilde{S}_{t,t-1}$$

where $\tilde{S}_{t+v,t-1} = \dfrac{\hat{S}_{t+v,t-1}}{Q(L^{-1})}$ is a weighted average of all sales expected in future periods.

Now

$$\frac{(1+\alpha(1-L))\hat{S}_{t,t-1}}{1-L} = \text{const.} + \alpha\tilde{S}_{t,t-1} + \hat{\tilde{S}}_{t,t-1} + \hat{\tilde{S}}_{t,t} + \hat{\tilde{S}}_{t-1,t} + \ldots$$

(4.12)

$$\frac{1+\alpha(1-L)\hat{S}_{t-1,t-2}}{1-L} = \text{const.} + \alpha\hat{\tilde{S}}_{t-1,t-2} + \hat{\tilde{S}}_{t-1,t-2} + \hat{\tilde{S}}_{t-2,t-2} + \ldots$$

so

$$(4.13)\quad R(L)P_t = (1-\lambda)\alpha(\tilde{S}_{t,t-1} - \tilde{S}_{t-1,t-2}) + (1-\lambda\left[\tilde{S}_{t,t-1} + \sum_{j=0}^{\infty}(\tilde{S}_{t-1-j,t-1} - \tilde{S}_{t-1-j,t-2})\right]$$

The last steps in the derivation of (4.13) assumed that the firm has been running optimally in previous periods. The purpose of examining the solution of the problem is to illustrate the derivation of distributed lag models consistent with optimizing behavior. When the econometrician estimates the parameters of such a model he will typically have to use data on firms (or more likely industries) over a relatively short period of time. In general there will be no reason to assume that the optimization began with the first observation the investigator has; rather it seems more natural to assume a long (but unobserved) history of optimal running and that the initial conditions are sufficiently far in the past so that their influence on the observed behavior is negligible.

To gain some understanding of the preceding equation consider the case in which changing the level of production is costless. If $\lambda = 0$, $R(L)Q(L^{-1}) = 1$ so $\tilde{S}_{t+v,t-1} = \hat{S}_{t+v,t-1}$ and the solution is

(4.14) $$P_t = \alpha(\hat{S}_{t,t-1} - \hat{S}_{t-1,t-2}) + \hat{S}_{t,t-1} + (S_{t-1} - \hat{S}_{t-1,t-2})$$

In this case production is simply set equal to expected sales plus the expected change in the desired level of inventories plus a correction factor for the error in forecasting the sales of the current period.

If one assumes that $S_t = B(L)\,\epsilon_t$; that is, S_t is a non-deterministic covariance stationary time series and that forecasts are calculated according to (2.6) i.e., least squares forecasts, then

$$R(L)P_t = \gamma(L)S_{t-1}$$

(4.15) where

$$\gamma(z) = \lim_{\delta \nearrow 1} \frac{1-\delta z}{B(z)} \left[\frac{B(z)(1-\lambda)(1+\alpha(1-\delta z))}{(1-\delta z)Q(z^{-1})z}\right]_+$$

Proof:

$$(4.16)\qquad \frac{1-\delta z}{B(z)}\left[\frac{B(z)(1-\lambda)(1+\alpha(1-\delta z))}{(1-\delta z)Q(z^{-1})z}\right]_+ =$$

$$\frac{(1-\lambda)\alpha(1-\delta z)}{B(z)}\left[\frac{B(z)}{Q(z^{-1})z}\right]_+ + \frac{(1-\delta z)}{B(z)}\left[\frac{B(z)}{(1-\delta z)Q(z^{-1})z}\right]_+$$

Now $\frac{1}{B(z)}\left[\frac{B(z)}{Q(z^{-1})z}\right]$ is the generating function of the operator that estimates $\frac{S_t}{Q(L^{-1})}$. Thus as δ goes to one the first term will give $(1-\lambda)\ \alpha(\tilde{S}_{t,t-1} - \tilde{S}_{t-1,t-2})$. Applying Whittle's theorem to the second term gives

$$(4.17)\qquad \frac{(1-\lambda)(1-\delta z)}{B(z)}\left[\frac{B(z)}{(1-\delta z)Q(z^{-1})z}\right]_+ = \frac{(1-\lambda)}{B(z)}\left[\frac{B(z)}{Q(z^{-1})z}\right]_+ +$$

$$\frac{(1-\lambda)\left(\left[\frac{B(z)}{Q(z^{-1})z}\right]_-\right)\Big|_{z=\frac{1}{\delta}}}{B(z)}$$

The first term clearly gives $(1-\lambda)\ \tilde{S}_{t,t-1}$ so all that remains to be shown is that the second term leads to $\Sigma(\tilde{S}_{t-j-1,t-1} - \tilde{S}_{t-j,t-2})$.

Now

(4.18) $$\hat{S}_{t-j,t} = \sum_{k=0}^{\infty} k^q \hat{S}_{t-j-1+k,t-1} \quad \text{where}$$

$$\frac{1}{Q(z^{-1})} = \sum_{k=0}^{\infty} k^q z^{-k}$$

and

(4.19) $$\frac{1}{B(z)}\left[\frac{B(z)}{z^{k-j}}\right]_+ - \frac{z}{B(z)}\left[\frac{B(z)}{z^{k-j+1}}\right]_+ = \begin{cases} 0 & k < j \\ \dfrac{b_{k-j}}{B(z)} & k \geq j \end{cases}$$

So the generating function for

(4.20) $$\sum_{j=0}^{\infty} \hat{S}_{t-j-1,t-1} - \hat{S}_{t-j-1,t-2} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} k\, q(\hat{S}_{t-j+k-1,t-1} - \hat{S}_{t-j+k-1,t-2})$$

is given by

(4.21) $$\frac{1}{B(z)} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} {}_{k+j}q b_k = \frac{\left[\dfrac{B(z)}{Q(z^{-1})z}\right]_-}{B(z)} \Bigg|_{z = 1}$$

Q.E.D.

Suppose that S_t is a moving average, auto-regressive process e.g. let

(4.22) $$B(z) = \frac{C(z)}{D(z)}$$

where C( ) and D( ) are polynomials of order c and d respectively. Then

$$(4.23)\qquad \frac{\gamma(z)}{1-\lambda} = \underset{\delta \nearrow 1}{\text{limit}}\ \frac{(1-\delta z)D(z)}{C(z)} \left[\frac{C(z)(1+\alpha(1-\delta z))}{D(z)Q(z^{-1})z(1-\delta z)}\right]_+$$

Let

$$(4.24)\qquad S(z) = \frac{C(z)(1+\alpha(1-\delta z))}{Q(z^{-1})z}$$

using the same type of argument as was used in Sections 2 and 3, it follows that

$$(4.25)\qquad \frac{\gamma(z)}{1-\lambda} = \frac{[S(z)]_+}{C(z)} + \frac{T(z)}{C(z)}$$ where $T(z)$ is of order d .

$[S(z)]_+$ is obviously of order c so

$$(4.26)\qquad \gamma(z) = (1-\lambda)\frac{V(z)}{C(z)}$$ where $V(z)$ is of order $\max\,[c, d, 0]$.

This gives

(4.27) $R(L)C(L)P_t = V(L)S_{t-1}$ a rational lag distribution. Thus as is the case with the expectational models treated above, adjustment models of this type reduce to forms suitable for éstimation.

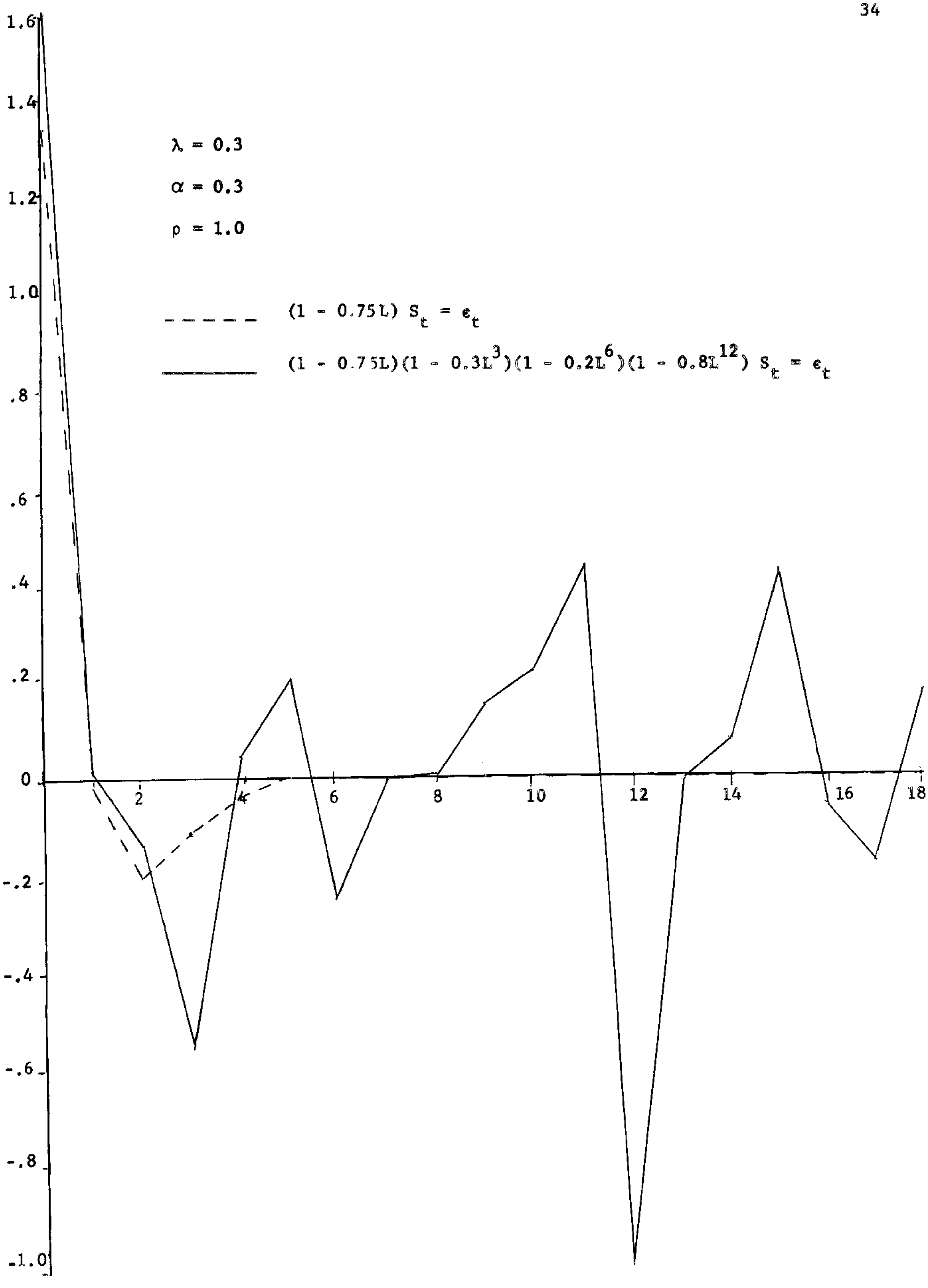
Certain features of the distribution that differ from lag patterns used in practice should be noted. First, except for the possibility of fortuitous cancelation, the distributed lag is <u>never</u> of the geometric type so commonly used. If, for example, the sales series is auto-regressive then (4.27) becomes

$$(4.28)\qquad P_t = \frac{V(L)}{R(L)}\, S_{t-1}$$

where the order of $V(\ )$ is the order of the auto-regression and $R(\ )$ is of order two. Introducing a moving average part into the structure of the sales series increases the order of the operator in the denominator. In dealing with distributed lags it is common to assume that the weights sum to one, are all positive, and that all the roots of the polynomial in the denominator are real. The example given here satisfies only the first of these conditions. Thus the calculation of mean lags would be meaningless.

The lag distribution depends not only upon the parameters of the cost function but also upon the nature of the sales series. Two firms with identical cost functions may appear to adjust at different rates because of the differences in the sales series they face. Without some knowledge of the way in which observed adjustment patterns depend on the sales series, one could not tell whether or not the firms had the same cost of adjustment. Simply looking at the first few terms in the lag distribution could be seriously misleading.

Figure 1 illustrates this point dramatically. The graph shows the first eighteen terms in the lag distributions for two firms with λ equal to 0.3, α equal to 0.3, and ρ one. Note that when ρ is one the firm is minimizing average cost over the future rather than discounting future costs. In both cases the sales series is assumed to be an auto-regression so that the coefficients of lagged production (which depends only on λ in this case) are the same, though as noted above these coefficients would also change if


1.6
1.4
1.2
1.0
.8
.6
.4
.2
0
-.2
-.4
-.6
-.8
-1.0
2
4
6
8
10
12
14
16
18
λ = 0.3
α = 0.3
ρ = 1.0
(1 - 0.75L) $S_t = \epsilon_t$
(1 - 0.75L)(1 - 0.3L^3)(1 - 0.2L^6)(1 - 0.8L^{12}) $S_t = \epsilon_t$

the sales series was of the moving average, auto-regressive type. The example thus illustrates the general point that is not possible to disentangle the formation of expectations from adjustment without an adequate understanding of the way expectations are formed and what they depend upon. Given a specification of the series S_t, the parameters of the model (α, ρ, λ) could be estimated non-linearly from the coefficients of the lag distribution.

It is not being suggested that choosing and fitting a model for a time series is an easy task. Nevertheless, the structure of the variables used in the forecasting procedure clearly influences the nature of the equation to be estimated and ought to be considered in the formulation of a structural equation.[7]

FOOTNOTES

1. See Whittle ([9] Ch. 8). The two examples covered in Section 1 illustrate this point, as in each case the process has an infinite variance. If one considers $x_t = \alpha x_{t-1} + \epsilon_t - \beta \epsilon_{t-1}$ $|\alpha|, |\beta| < 1$ and derives the least squares forecast of x_{t+1}, then letting α tend towards unity from below, the result will agree with that given above.

2. For a proof see Whittle [9].

3. See Whittle [9]. Note that if $y_t = x_{t+v}$, then $g_{yx}(z) = \sigma^2 z^{-v} B(z) B(z^{-1})$ so the result agrees with that given in equation (2.6).

4. This model and/or more complicated forms of it have been treated a number of times. See, for instance, Holt, Modigliani, Muth, and Simon [2], Childs [1]. The method of solution given here closely follows that used by Whittle [9, Ch. 10].

5. In certain of the calculations an extra parameter δ, $|\delta| < 1$ will be introduced, and then limits will be taken as δ tends up to one. It would be possible to consider the related problem where $P_t = S_t + I_t - \delta I_{t-1}$, obtain its solution and then let δ approach one. For ease of exposition this device will only be used in calculating the transforms needed to obtain the forecasts of the series S_t.

6. The cases λ equal to one or zero are not very interesting. If $\lambda = 1$, all the firm wishes to do is maintain a steady rate of production, and the solution is obviously $P_{t+v} = \bar{S}$. In the case $\lambda = 0$ the firm merely sets production on the basis of the sales expected next period.

7. For an example of choosing the form of an estimating equation using the structure of the variable being forecast see Wallis [8]. For a discussion of the estimation problems involved see Nerlove [6].

REFERENCES

1. Childs, Gerald L., Unfilled Orders and Inventories, Amsterdam: North Holland Publishing Company, 1967.

2. Holt, C.C., F. Modigliani, J.F. Muth, and H. Simon, Planning Production, Inventories, and Work Force, Englewood Cliffs, New Jersey: Prentice Hall, 1960.

3. Jorgenson, D.W., "Rational Distributed Lag Functions," Econometrica, 34 (January 1966).

4. Muth, J.F., "Optimal Properties of Exponentially Weighted Forecasts", The Journal of the American Statistical Association, 55, (June 1960).

5. Nerlove, Marc, The Dynamics of Supply: Estimation of Farmers' Response to Price, Baltimore: John Hopkins, 1958.

6. ____________, "Distributed Lags and Unobserved Components of Economic Time Series", in Fellner et al. Ten Economic Essays in the Tradition of Irving Fisher, New York: John Wiley, 1967.

7. Theil, H., Optimal Decision Rules for Government and Industry, Amsterdam: North Holland Publishing Company, 1964.

8. Wallis, Kenneth F., "Some Econometric Problems in the Analysis of Inventory Cycles", Cowles Foundation Discussion Paper No. 209, (May 1966).

9. Whittle, P., Prediction and Regulation by Linear Least Squares Methods, London: The English Universities Press, 1963.

10. Wold, H., The Analysis of Stationary Time Series, Uppsala: Almquist and Wicksell (1938, 2nd ed., 1954).