SIMULTANEOUS USE OF ANNUAL AND QUARTERLY DATA IN THE ADAPTIVE EXPECTATIONS MODEL

By DIKAIOS E. TSERKEZOS

1. INTRODUCTION

Economists in building a quarterly model consider the typical case in which one series of interest is on an annual basis. In these circumstances the economists will generally choose between estimating an annual basis model over the entire period or a quarterly model over the subperiod (if such exists) for which full Quarterly data are available. In either case he discards sample information. Alternatively he may incorporate the missing quaterly data on some ad hoc basis.

In this paper we consider estimates for the Adaptive Expectations Model for which data on the dependent variable are available for a subset of the sample period on only an annual basis.

Our aproach is to treat the missing (Quarterly) observations as unknown parameters which estimated simultaneously with the other unknown parameters of the model. We minimise the constrained Residuals Sum of Squares (taking into account the available annual observations) with respect to the unknown parameters and the missing quarterly data. If we assume normality our estimates are Maximum Likelihood Estimates.

In section 2 we set up the model and some important time Aggregation notation. In section 3 we use the Maximum Likelihood method to obtain an

This article is an extension of my M.A. (Econ) dissertation submetted to the Victoria University of Manchester. All errors, of course, remain my own.

Estimator for the Adaptive Expectation model. An illustrative example and conclusions are in sections 4 and 5 respectively.

2. THE MODEL

The adaptive expectations model postulates that changes in y_t are related to changes in the «expected» level of the explanatory variable Xt

$$yt = a + bX^e_t + u_t \quad (2.1)$$

With the ut's are NID (0, o^2) and X^e represents the desired or expected leve of X_t. Since X^e_t is not directly observable, a reasonable hypothesis concerning the manner in which expectations are generated must be formulated. In the adaptive expectations model we put

$$(X^e_t - X_t{}^e{}_{-1} = (1-k) \qquad (Xt - X_t{}^e{}_{-1}) \quad (2.2)$$

with $0 < k < 1$

Rearranging (2.2) and rewriting it using the Lag Operator L gives

$$X^e_t = (1-k)X_t / (1-kL) \quad (2.3a)$$

Substituting for X_t in (2.1) then yields

$$y_t = a + \frac{b(1-k)}{(1-kL)} X_t + u_t \quad (2.3b)$$

multiplying through by(1-kL) and rearanging we obtain :

$$yt = b_0+b_1X_t+ky_{t-1}+v_t \quad (2.4)$$

with $v_t = ut-ku_{t-1}$ (2.5)

and $b_0 = a\ (1 - k)$, $b_1 = b(1 - k)$ (2.6)

which is an ARMAX (1) model with the special characteristic that the coefficient of y_{t-x} is equal to the coefficient of U_{t-1} .This alows us to define $w_t = y_t - u_t$; then

$$w_t = kw_{t-1}+b_0+b_1X_t \quad (2.7)$$

which after successive substitution for $w_{t-1;}$ yields

$$y_t = w_0 k_t + \begin{bmatrix} Z_{1t}(k) & Z_{2t}(k) \end{bmatrix} \begin{vmatrix} b_0 \\ b_1 \end{vmatrix} + u_t \quad (2.8)$$

with $Z_{1t}(k) = (1+k+k2+k3+ \ldots\ldots\ldots\ldots +k^{t-1})$

$Z_{at}(k)= (X_t+kX_{t-1}+k2X_{t-2} +,,. ,\ldots\ldots .. .+k^{t-1}x_1)$

Since now the error term in (2.8) is NID $(0, \sigma^2)$ We minimise

$$\sum_{t=1}^{T} u_t^2 = \sum_{t=1}^{T} (y_t-W_0k^t-b_0Z_{1t}(k)-b_1Z_{2t}(k))^2 \quad (2.9)$$

with respect to k. w_0 and b_1 to obtain ML estimates. This can be conveniently done by regressing y_t on k^t, $Z_{1t}(k)$ and $Z_{2t}(k)$ for various values of k and then choosing the value of k and the associated estimates of w_0, b_0 and b_1 for which the residual sum of squares is a minimum.

Now we introduce Some Time - Aggregation notation. The relation between the annual and the quarterly observations for a variable, say y, for a period T_1 is

$$y^a = Cy \tag{2.10}$$

where y^a : annual observations $(T_1/4 \times 1)$ vector

y : quarterly observations $(T_t \times 1)$ vector

and C a $(T_{1/4} \times T_1)$ aggregation matrix of the form

$$C = \begin{vmatrix} 111100000000........0000 \\ 000011110000........0000 \\ 000000001111........0000 \\ 000000000000....11110000 \\ 000000000000....00001111 \end{vmatrix} \tag{2.11}$$

defining $e_4 = (1,1,1,1)$

we may write $C = (I_{T1/4} \otimes e_4')$

with I a $(T_1/_4 \times T_1/_4)$ identity matrix and ® denotes the Kronecker product

We form the following relations

$$CC' = (I_{T^{1/4}} \otimes e_4') \quad (I_{T^{1/4}} \otimes e_4')'$$

$$= (I_{T^{1/4}} \otimes e_4' e_4)$$

$$= 4 \; I_T \tag{2.12}$$

$$CC' \quad (I \quad \otimes e \quad)' \; (I \quad {}_{1/4} \otimes e_4')$$
$$= (I_{T_1/4} \otimes e_4' \;\; e_4) \tag{2.13}$$

$$C' (CC')^{-1} \; C = (I_{T_1/4} \otimes e_4')' \; (4 I_{T_1/4})^{-1} \, (I_{T_1/4} \otimes e_4')$$

$$= 1/4 \; (I_T \otimes e_4' \; e_4)$$

$$= 1/4 \; (C'C)$$

where $C'C = J$ is a $T_1 x \; T_1$ block - diagonal matrix with (4 X 4) blocks of ones down the diagonal and zero elsewere.

$$J = \begin{vmatrix} 11110000 \ldots\ldots\ldots 0000 \\ 11110000 \ldots\ldots\ldots 0000 \\ 11110000 \ldots\ldots\ldots 0000 \\ 11110000 \ldots\ldots\ldots 0000 \\ 00000000 \ldots\ldots\ldots 1111 \\ 00000000 \ldots\ldots\ldots 1111 \\ 00000000 \ldots\ldots 1111 \\ 00000000 \ldots\ldots 1111 \end{vmatrix}$$

using the above notation and relations we may exprees the Time Aggregation relation between the quarterly averages y and the quarterly observations y as :

$$\bar{y} = C' (CC')^{-1} Cy$$

$$= 1/4 \ (I_{T_1/4} \otimes e_4)(I_{T_1/4} \otimes e_4')\ y$$

$$= 1/4 \ (Jy)$$

and similar

$$y^a = C\bar{y}$$

3. MAXIMUM LIKELIHOOD APPROACH

Given k we may write ;

$$X(k) = \quad k^t \quad Z_{1t}(k) \quad Z_{2t} \quad , \quad \delta = (w_0 \quad b_0 \ b_1)'$$

and rewrite the Quarterly model (2.8) as

$$y = X(k)\delta + u \tag{3.1}$$

where $E(u) = 0,$

$$E(X'u) = 0, \tag{3.2}$$

$$E(uu') = \sigma^2 I\tau$$

and where y is a T vector of observations on the dependent variable, and X(k) a T x 3 matrix as defined above, taken to be fixed in repeated samples. We further assume that over the first T_1 quarters only annual observations are available on y. We define $T_2 = T - T_1$, and for algebraic convenience we assume that T_1 is an integer. We therefore split the model

$$\begin{pmatrix} y^1 \\ y^2 \end{pmatrix} = \begin{pmatrix} X^1 \\ X^2 \end{pmatrix} \delta + \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \qquad \textbf{(3.3)}$$

We proceed by obtaining estimates of the parameters of interest by minimising the constrained residual sum of squares of equation (3.3) with respect to all the unknows and y_1. We minimise the constrained RSS equation since we want incorporate all the available informations such as the annual aggregates.

We form the constrained RSS function

$$\varphi = (y - X(k)\delta)' (y - X(k)\delta) - 2\lambda' (y^{1a} - C_y 1) \qquad (3.4)$$

where λ is the $T_{1/4}$ vector of Lagrangian multipliers, and $Cy^1 = y^{1a}$ are the constraints implied from the available by assumption data.

Differentiating (3.4) with respect to y^1 setting equal to zero and after some algebraic manipulations we obtain the constrained equation

$$y^1 - X^1(k)\delta - C'\lambda = 0 \qquad (3.5)$$

Premultiply by C and recalling that

$$Cy^1 = y^{1a} = C\bar{y}^1$$

we may solve for λ as

$$\lambda = (CC')^{-1}\ C\ (\bar{y}^1 - X^1 \delta) \qquad (3.6)$$

which substituted into (3.5) gives

$$\hat{y}^1 = X^1(k)\,\delta + C'(CC)^{-1} C_{y^{-1}} - C(CC')^{-1} CX^1(k)\,\delta \tag{3.7}$$

using $\quad \bar{y}^1 = c'(cc')^{-1} Cy^1$

equation (3.7) can be written as

$$\hat{y}^1 = \bar{y}^1 + (X^1(k) - \bar{X}^1(k))\,\delta \tag{3.8}$$

Thus conditional upon the estimated vector of coefficients δ given k, we may estimate the missing quaterly observations by adding to the observed annual averages the weighted deviafions of the explanatory variables from their annual averages. For $0< k< 1$ (3.8) suggests the following iterative procedure. Apply OLS to (3.1) for the subperiod T_2 to obtain a starting value for δ, say $\delta\tau$ and tuse his estimate in the iterative process.

For $0 < k < 1$

$$\hat{y}_{r+1} = \bar{y}^{-1} + (X^1(k) - \bar{X}^{-1}(k))\,\hat{\delta}_r \tag{3.9}$$

$$\hat{\delta}_{r+1} = (X(k)'\,X(k))^{-1}\,X(k)'\,\hat{y} \tag{3.10}$$

where $\quad \hat{y} = \begin{pmatrix} \hat{y}_{r+1} \\ y^2 \end{pmatrix}$

and chose the value of k such that the residua) sum of squares is minimum.

We may avoid such expensive computation procedure by noting that

$$\left[\hat{y}^1 - X^1(k)\boldsymbol{\delta} \right] = \left[\bar{y}^{1} - \bar{X}^1(k)\boldsymbol{\delta} \right]$$

and substituting into the unconstrained RSS we obtain

$$\boldsymbol{\varphi} = (y^1 - X^1\ (k)\ \boldsymbol{\delta})'(\bar{y}^1 - \bar{X}^1\ (k)\boldsymbol{\delta}) + (y^2 - X^2(k)\boldsymbol{\delta})'(y^2 - X^2(k)\boldsymbol{\delta}) \tag{3.12}$$

which is immediately recognized as the function that is minimised by OLS estimation of equation (3.1) after replacing of both the missing quaterly observations and the corresponding values of the exogeneous variables by their annual averages.

We may now using the above findings to describe the whole procedure to obtain estimates for δ and consequently for w_0, a and b. For various values of k in the interval (0,1) estimate w_0, b_1 and b_2 by regressing y_t on Xt(k) after replacin g both the missing quarterly data and the corresponding values of the exogeneous variables by the annual averages, and then choose the value of k and the associated estimates of δ for which the residual of squares is a minimum.

To estimate the asymptotic variances of the estimated coefficients $\hat{w}_0$, $\hat{b}_1$, $\hat{b}_2$, $\hat{k}$ and $\hat{\sigma}^2$ we need to construct the information matrix

$$\mathbf{I} = - E \left| \frac{\boldsymbol{\Theta}^2 \log L}{\boldsymbol{\Theta}_{\boldsymbol{\Theta}}\, \boldsymbol{\Theta}_{\boldsymbol{\Theta}'}} \right|$$

where log L is the Log-Likelihood Function

$$\log\ L = -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma^2 - \frac{1}{2\sigma^2} (y - X(k)\boldsymbol{\delta})'\ (y - X(k)\boldsymbol{\delta})$$

and $\boldsymbol{\Theta}' = (\hat{w}_0, \hat{b}_1, \hat{b}_2, \hat{k}, \hat{\sigma}^2)$.

The estimated asymptotic variances of $\hat{w}_0$, $\hat{b}_1$, $\hat{b}_2$, $\hat{k}$ and $\hat{\sigma}^2$ are the diagonal elements of I^{-1}

If instead someone use the iterative scheme given by (3.9) and (3.10) to obtain estimates for 0, then the asymptotic variances must be obtained from the Information marix of the constrained log-likelihood and $\Theta = (\hat{w}_0, \hat{b}_1, \hat{b}_2, \hat{k}, 0^2, \hat{y}^1)$.

Since by assumption the residuals of (3.1) are NID $(0, \sigma^2 I_t)$, the estimates obtained by the above suggested method, if the model correctly specified, are consistent and asymptotically efficient. The distribution of the estimated $\hat{y}^1$ may be deduced from (3.12) by substitution fc $\hat{y}^1$ and subtracting y^1. For more details see my M.A. Thesis.

4. AN ILLUSTRATIVE NUMERICAL EXAMPLE

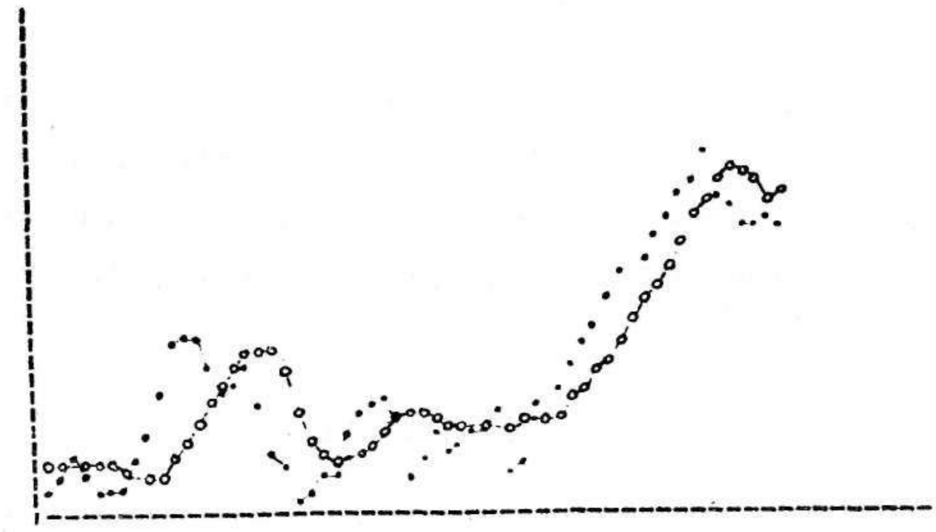
To illustrate our suggestion, we use the capital expenditures (yt) and the net new appropriations (Xt) quarterly basis series of Almon, which span the years 1953 - 1967. From a comparison of the plots of the these time series (fig. I), we see that, expect for a location shift, the behavior of the two is quite similar. As a first approximation, a model of the form (2.1)-(2.2) seems to be a resanoble specification.

(Don't forget that from (2.3b) we may obtain

$$y_t = b(1-k) \sum_{j=0}^{\infty} k^j X_{t-j} + u_t \tag{4.1}$$

with $b_0 = 0$, which is a geometric distributed lag representation).

FIGURE 1.



Capital expenditures (o) and net new appropriations (·).

First we estimated the Adaptive Expectations model using all the available quarterly observations (T=60), and we obtained $\hat{b}$=0.98394369, $\hat{k}$=0.736845
(0.15493....) (0.00217...)
and $\hat{\sigma}^2$=(0.00349. . . . .). Then assuming that quarterly data for the variable yt are not available for the first T_1 = 20 observations (instead Annual observations are available) we reestimate the Adaptive Expectations model to obtain

b=0.9318565, $\hat{k}$= 0.4743 and $\hat{\sigma}^2$= (0.008431...).
0.5430...) (0.0474...)

Finally using our mixed approach, using simultaneously Annual and Quaterly observations we obtain $b = 0.9793$, $k = 0.7204$ and $\hat{\sigma}^2 = (0.00644 \ldots)$
(0.0721...) (0.0885...)

The estimated «weights» $b_j = b(1-k)k^j$ for the three different assumptions about the data availability are given in table I.

TABLE I

Estimated lag weights

Length of Lag	$T = T_1 + T_2 = 60$	$T_2 = 40$	Mixed Approach	
0	0.2589	0.4898	0.273	0.242
1	0.1907	0.2323	0.197	0.195
2	0.1405	0.1102	0.142	0.148
3	0.1035	0.052	0.102	0.110
4	0.0763	0.024	0.073	0.080
5	0.0562	0.011	0.053	0.058
5	0.0414	0.005	0.038	0.042
7	0.0305	0.002	0.027	0.030
8	0.0225	0.001	0.019	0.022
9	0.0165	0.0005	0.014	0.016
10	0.0122	0.0002	0.010	0.011
11	0.0090	0.0001	0.007	0.008
12	0.0066		0.005	0.006
13	0.0048		0.003	0.004
14	0.0036		0.002	0.003
SUM	0.984	0.932	0.979	0.979

In Figure II. we give the graph of the estimated «weights» b_{wj} for the three different assumptions about the data availability.

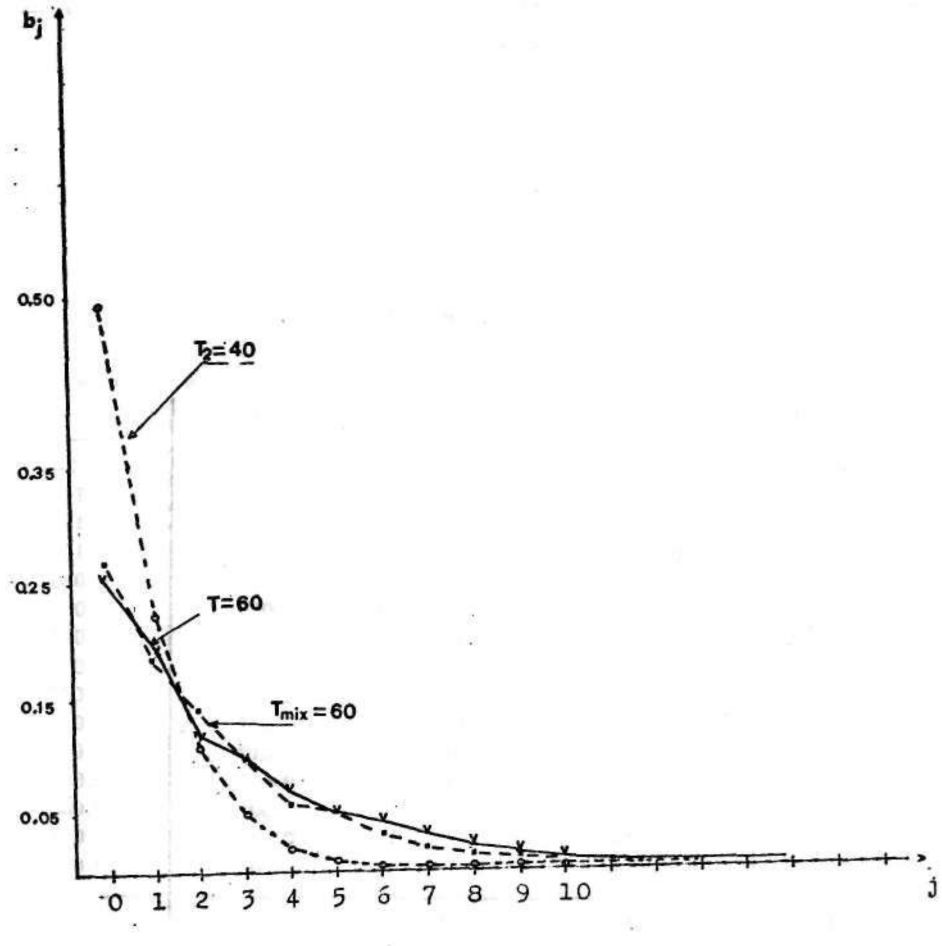


FIGURE II.

In the fifth column of table I we present Schmith's estimates obtained from his paper [21].

Table II presents the mean and variance of the estimated lag distributions.

TABLE II

Average+ lag and variance++ of the Lag Distributions

	Mean lag	Variance
T = 60	2.80	9.20
$T_2 = 40$	0.902	1.72
Mixed Approach	2.58	9.22

+Mean lag = $k/(1-k)$
++Variance = $k/(1-k)^2$

In table III we give the Impact (b_0), Interim (b^+) and Standardized Interim (b_j) multipliers respectively.

TABLE III (a)

Interim and Standardized Interim multipliers
when $T = T_1 + T_2 = 60$

j	0	1	2	3	4	5	6	7	8
	26 %	45 %	59 %	70 %	78 %	83 %	87 %	91 %	92%

TABLE III (b)

Interim and Standardized Interim multipliers
when $T_2 = 40$

j 0	1	2	3	4	5	6	7	8	9
.489	0.72	0.83	0.88	0.90	0.91	0.92	0.93	0.94	0.92
52 %	77 %	89 %	94 %	97 %	98 %	99 %	99 %	99 %	99 %

TABLE III (c)

Interim and Standardized Interim multipliers
Mixed approach.

j 0	1	2	3	4	5	6	7	8	9
0.273	0.47	0.612	0.714	0.787	0.84	0.878	0.905	0.924	0.938
27 %	48 %	62 %	72 %	80 %	85 %	89 %	92 %	94 %	95 %

The computation formula for the above multipliers is $b_j^+ = \sum_{j=0}^{J}$ bj, J = 0,1,2 ... for the Interim multiplier

and $b_j^+ = \frac{b_j^+}{b_\infty^+}$, J = 0,1,2 ... for the Standardized Interim multiplier.

were $b_\infty^+ = \sum_{v=0}^{\infty} b_v$

5. CONCLUSIONS

Economists frequently encounter data which are subject in different temporal aggregation for different time periods. In this paper give a maximum likelihood estimator using Annual and Quaterly data in the Adaptive Expectations model, although we extended our results to the Geometric declining lags model, using the form (4.1).

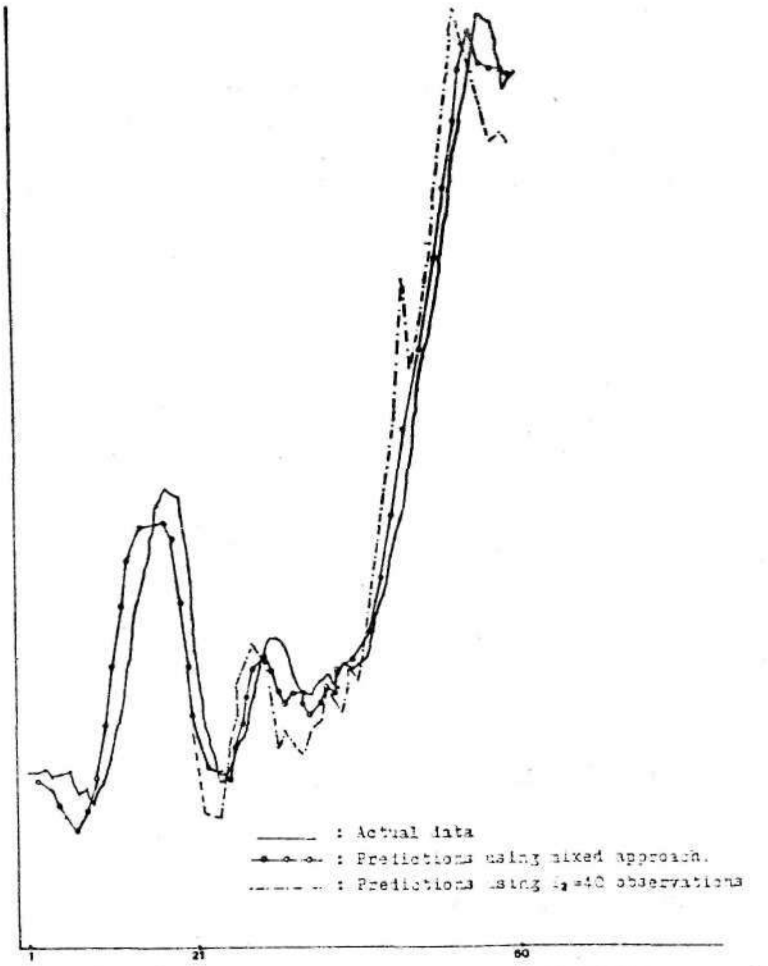


FIGURE III

By writing an iterative non linear regression program in simple BASIC and using Almon's data we illustrate the importance of the suggested technique from the dynamic analysis point of view. This can be seen in table III where we study the pattern of Interim and Standardized multipliers for different values of k, i.e k = 0. 736845, k = 0.4743 and k = 0.7204.

The results are quite obvious. Using only the $T_2 = 40$ available quarterly observations the results about the dynamic responce of the suggested dynamic process may be missleading.

The predictive power of the suggested approach, as illustrated in figure 3 may be helpfull to «Predict» missing quarterly observations. Using (3.8) is not that difficult to obtain «very raliable predictions», for the capital expenditures (yt) in the period Tl where by assumption we have missing quarterly data. Specially for the case of Greece where quarterly data for some macroeconomic variables are available only after 1975, although there are quite raliable annual observations from 1958.

We may extend the suggested technique when we assume that the residuals are autocorrelated. In that case the suggested estimation technique can be applied but the computation is more difficult. A computational expensive Gauss - Seidel iterative process must be applied between the «forecasted» quarterly observations and the under estimation parameters.

Finally we may extend the suggested approach for different models with different time aggregation and different time periods. For example in the case where we use simultaneously quarterly and monthly data we follow the same approach but the C matrix now is

$$C = (I_{T_1/4} \otimes e'_3) \quad \text{with } e'_3 = (1, 1, 1)$$

For more details see my M.A. Thesis (9), and (23), (24), (25).

REFERENCES

B o o k s

1. Dutta, M., 1974, Econometric Methods (South Western Publish Co)
2. Harvey, ., 1981, The Econometric Analysis of time Series
3. Gamaletsos, T., 1973. Econometrics (Athens)
4. Goldberger, A., 1964, Econometric Theory (New York Wiley)
5. Coult, Hoskins, Milner, Pratt, 1974, Computational Methods in Linear Algsbra
6. Schneeweib, H., Okonometrie (Physica -Verlag, 1978)
7. Schonfeld, P., Methoden der Okonometrie
   Bd. I : Lineare Regressionsmodelle (Berlin)
   Bd. II : Stochastische Regressoren und Simultané Gleichungen (Munchen 1971)
8. Theil, H., 1971, Principles of Econometrics (John Wiley).
9. Tserkesos E. Dikaios., 1982, «The Simultaneous Use of Annual and Quaterly Data in Econometric Models» unpublished M.A (Econ) dissertation submitted to the VICTORIA UNIVERSITY OF MANCHESTER.
10. Tserkesos Ef Dikaios., 1978, «The Consumption function of the Greek Economy 1958 - 1975» An Empirical Analysis. Research Paper No 1. Department of Econometrics, Graduate school of Industrial studies. Piraus.

A r t i c l e s

11. Almon, S. 1965, «The distributed lag between capital appropriations and expenditures» *Econometrica* pp. 178-196.
12. Beach M.C. and Mackinnon., 1978, «A ML procedure for regression with autocorrelated errors» *Econometrica*
13. Beach M.C. and Mackinnon., 1978, «Full Maximum Likelihood estimation of second-order autoregressive error models» *Journal of Econometrics*
14. Dagenais, M.G., 1973, «The use of incomplete observations in multiple regression analysis" A Generalized Least Squares approach, Journal of Econometrics 317-328.
15. Gilbert, C.L., 1975, «Estimation of regression equations using mixed annual and quarterly data» (Uni. of Bristol) and Journal of Econometrics 1977, 221-239.
16. Gilbert, C.L., 1976, «Missing data in regression analysis. The exogeneous case. (Uni. of Bristol).
17. Hsiao, C, 1979, «Linear regression using both temporally aggregated and temporally disaggregated data» Journal of Econometrics, 243 - 252.

18. Pagan Adrian, 1978, «Rational and Polynomial lags» Journal of Econometrics

19. Pagano M and Hartley M., 1981, «On fitting distributed lag models subject to Polynomia restrictions» Journal of Econometrics

20. Pesaran M. Hamsen., 1973, «The small sample problem of truncation remainders in thel estimation of Distributed lag models with autocorrelated errors» International Eco nomic Review.

21. Schmidt Peter., 1974, «An Argument for the usetulness of the gamma distributed lag model» International Economic Review.

22. Sargan, J. D. and E.S. Drettakis, 1974, «Missing data in an autoregressive model» International Economic Review 39 - 58.

23. Tserkesos Ef Dikaios., 1982, «Simultaneous Use of Annual and Quarterly data in Simultaneous equations models» (Unpub ished paper).

24. Tserkesos Ef Dikaios., 1983, «Simultaneous Use of Annual and Quarterly data in the Adaptive Expectations model» (Unpublished paper)

25. Tserkesos Ef. Dikaios., 1984, «Simultaneous Use of Annual and Quarterly data in the Partial Adjustment model» (Unpublished paper).

26. Zellner A and Geisel M., 1970, «Analysis of distributed lag models with Applications to consumption function estimation» EconometricaVol 38 No 6.

27. Zellner, Huang, Chau 1965, «Further analysis of the short- run consumption function with emphasis on role of liquid assets» Econometrica.