ESTIMATION OF AN EVOLVING SEASONAL PATTERN AS AN APPLICATION OF STOCHASTICALLY VARYING PARAMETER REGRESSION

Adrian Pagan

ECONOMETRIC RESEARCH PROGRAM
Research Memorandum No. 153
October 1973

The research described in this paper was supported by NSF Grant GS 32003X.

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, New Jersey

ESTIMATION OF AN EVOLVING SEASONAL PATTERN AS AN APPLICATION OF STOCHASTICALLY VARYING PARAMETER REGRESSION

Adrian Pagan

The seasonal adjustment of economic time series has exercised the mines of statisticians and economists for a goodly number of years and has resulted in numerous proposals for its accomplishment. In a recent article, Hannan, Terrell and Tuckwell [10] address themselves to the vexed question of estimating an <u>evolving</u> seasonal pattern and present a model that enables the construction of an "optimal" filter for this purpose. More specifically, a series $y(t)$ is assumed composed (additively) of trend, $T(t)$, seasonal, $S(t)$, and irregular, $I(t)$, factors, where each of these evolves according to[1]

$$y(t) = T(t) + S(t) + I(t) \tag{1a}$$

$$T(t) = T(t-1) + e(t) \tag{1b}$$

$$S(t) = \sum_{j=1}^{2} S_j(t) \tag{1c}$$

---

[1]This model differs from [10] in two respects. Firstly, it is relevant to quarterly rather than monthly data and, secondly, the trend is in first rather than second difference form. It will become obvious later that this does not affect generality but facilitates exposition.

$$S_j(t) = \alpha_j \cos t\lambda_j + \beta_j \sin t\lambda_j \tag{1d}$$

$$\alpha_j(t) = \alpha_j(t-1) + e_j(t) \tag{1e}$$

$$\beta_j(t) = \beta_j(t-1) + \eta_j(t) \tag{1f}$$

where $e(t) \sim N(0,\sigma^2)$ $I(t) \sim N(0,\sigma_I^2)$, $e_j(t)$ and $\eta_j(t)$ are $N(0,\sigma_j^2)$, $E(e_j(t)\ \eta_j(t)) = 0$, $E(e(t)\ e_j(t)) = 0$, $E(e(t)\eta_j(t)) = 0$ and $\lambda_j = j\pi/2$.

When $e_j(t) = \eta_j(t) = 0$ the seasonal in (1d) is constituted from constant amplitude sine and cosine waves so that, if the trend was removed separately, extraction of the seasonal follows from an harmonic regression i.e., it is the spectral analogue of [13]. An application of this approach may be found in [3] while investigations into the distributional properties and robustness of the OLS estimator of the parameters α_j and β_j are contained in [27] and [28]. Examination of this polar case makes it clear that the consequence of a non-zero $e_j(t)$ and $\eta_j(t)$ is an harmonic <u>regression with time-varying parameters</u> and, as such, may be studied under the genus "estimation of regression equations with stochastically-varying coefficients,"e.g. see [2], [24] and the survey [23].

To comprehend the objectives of this paper it is sufficient to note three limitations of the study in [10].

(i) "The second simplification which we adopt is a technique which treats each $S_j(t)$ separately ... . In principle it is not necessary to adopt this procedure of taking each λ_j separately ... In practice the problem of computing the optimal coefficients becomes very great unless these things are done, as high order polynomials have to be factored," [10, p. 29].

(ii) For the same motives as expressed in (i) generalization to (say) second order differencing schemes for the evolution of $\alpha_j(t)$ and $\beta_j(t)$ is not easily realized.

(iii) As the estimation of the parameters σ_j^2 is done in a heuristic fashion the statistical properties of the estimates of $T(t)$ and $S(t)$ are unknown.[2]

Central to the paper is the relaxation of (i)-(iii). By expressing the model set out in (1) in linear system form

---

[2]The parameter σ^2 is not estimated cf. "there is little point in considering the estimation of σ^2 since the model is not sufficiently realistic for that. However, in the case of $S(t)$, the model is sufficiently close to the truth to justify considering the estimation of σ_j^2 ." [10, p. 36]. Generally, σ^2 is set to obtain a response for the trend filter that does not have much influence at the seasonal frequencies - see [10] and [26] for more details -- but if desired, estimation of σ^2 is easily done in the framework described later.

[17, p. 39] it is possible to apply the large corpus of filtering and estimation theory that exists for such a representation. At the same time the paper aims to shed some light on the choice of algorithm for the estimation of regression models with stochastically varying parameters.

1. Linear System Formulation

The linear system form relates a $(p \times 1)$ vector of observed variables $y(t)$ to a $(q \times 1)$ unobserved vector $z(t)$ which evolves in a Markov fashion, i.e.,

$$\begin{aligned} y(t) &= H(t)\, z(t) + B\, v(t) \\ z(t) &= M\, z(t-1) + \Gamma\, u(t) \end{aligned} \tag{2}$$

where B and Γ are $(p \times m)$ and $(q \times \ell)$ matrices respectively, $u(t)$ and $v(t)$ are mutually uncorrelated, and normally and independently distributed error terms with covariance matrices Q and R respectively. Equation (3) shows the conversion of (1) into the format of (2).[3]

[3] Notice that, as $\lambda_2 = \pi$ and $\sin(t\pi) = 0$, the term $\beta_2 \sin(t\lambda_2)$ is omitted.

$$y(t) = [1 \ \cos t\lambda_1 \quad \sin t\lambda_1 \quad \cos t\lambda_2] \begin{bmatrix} T(t) \\ \alpha_1(t) \\ \beta_1(t) \\ \alpha_2(t) \end{bmatrix} + I(t)$$

$$\begin{bmatrix} T(t) \\ \alpha_1(t) \\ \beta_1(t) \\ \alpha_2(t) \end{bmatrix} = \begin{bmatrix} T(t-1) \\ \alpha_1(t-1) \\ \beta_1(t-1) \\ \alpha_2(t-1) \end{bmatrix} = \begin{bmatrix} e(t) \\ e_1(t) \\ \eta_1(t) \\ e_2(t) \end{bmatrix} \qquad (3)$$

A comparison of (2) and (3) reveals equivalence when $M = I$, $\Gamma = I$, $B = 1$, $H(t) = [1 \ \cos t\lambda_1, \ \sin t\lambda_1, \ \cos t\lambda_2]$.[4]

2. Filtering and Estimation Theory for Linear Systems

Consider (2). Assuming prior knowledge of $H(t)$ the unknown parameters are R, Q, M and $z(0),\ldots,z(T)$, albeit, in some instances, elements of these matrices may be prescribed, e.g., setting M equal to the identity matrix in the above example. To retain generality it is best to treat all elements as unknown and then employ the following partitioning of the task.

---

[4] As any high order difference equation can be reduced to a first order one by re-definition of variables, there are no conceptual difficulties in expressing (say) second differences in the trend in the form of (2) i.e. $H(t)$, M and $z(t)$ must merely be re-defined and given increased dimensions.

(a) Expression of $z(1),\ldots,z(T)$ in a functional form conditional on R, Q, M and $z(0)$.

(b) Unconditional estimation of R, Q, M and $z(0)$.

The solution to (a) is well known ([14] and [15]) yielding the optimal (minimum mean square error) filtered estimate of $z(t)$, given the realization $y(1),\ldots,y(t)$, as the solution to the recursive equations [17, p. 176]

$$\hat{z}(t|t) = M\,\hat{z}(t-1|t-1) + K(t)\,\phi(t) \tag{4a}$$

$$\phi(t) = y(t) - H(t)\,M\,\hat{z}(t-1|t-1) \tag{4b}$$

$$K(t) = P(t|t-1)\,H'(t)\,V^{-1}(t) \tag{4c}$$

$$V(t) = H(t)\,P(t|t-1)\,H'(t) + B\,R\,B' \tag{4d}$$

$$P(t|t-1) = M\,P(t-1|t-1)\,M' + \Gamma\,Q\,\Gamma' \tag{4e}$$

$$P(t|t) = [I - K(t)\,H(t)]\,P(t|t-1)\ , \tag{4f}$$

while the optimal estimate of $z(t)$, given the complete realization $y(1),\ldots,y(T)$, is [18, p. 6]

$$\hat{z}(t|T) = \hat{z}(t|t) + P(t|t)\,M'\,\lambda(t) \tag{5a}$$

$$\lambda(t-1) = [I - K(t)\,H(t)]\,M'\,\lambda(t) + H'(t)\,V^{-1}(t)\,\phi(t) \tag{5b}$$

$$\lambda(T) = 0\ . \tag{5c}$$

To solve (4) it is necessary that there be a normal estimate, with mean $z(0)$ and covariance matrix $P(0)$, of the initial state vector to begin the recursion. Classical approaches to the problems raised by unknown pre-period values e.g. [20], begin with the presumption that these unknowns are fixed constants (i.e., $P(0) = 0$) and hence a point estimate of $z(0)$ is required. Acceptance of this methodology leaves the unknown parameter set at $\theta = \{Q,R,M, z(0)\}$; given values for the elements of θ enables the calculation of $\hat{z}(t|t)$, $\phi(t)$, $V(t)$ and $K(t)$, which quantities are then inputs into the backward recursion of (5) to eventually yield $\hat{z}(t|T)$.

The log likelihood associated with (2) is as derived by Schweppe [25]

$$\log L(\theta|y(1),\ldots,y(T)) = \text{const} + \sum_{t=1}^{T} (\log \det V^{-1}(t) - \phi'(t)\, V^{-1}(t)\, \phi(t)) \,. \quad (6)$$

Maximization of (6) subject to the constraints in (4) is conceptually feasible but, in practice, may involve a large number of parameters if the state vector $z(t)$ is large (as it will be if a high order system has to be reduced to a first order one). To overcome this obstacle Rosenberg [22] has derived the maximum likelihood estimate of $z(0)$, conditional on Q, R and M, as

$$\hat{z}(0) = \left(\sum_{t=1}^{T} F(t) \right)^{-1} \sum_{t=1}^{T} f(t) \tag{7}$$

where

$$F(t) = T'(t)\, V^{-1}(t)\, T(t) \tag{8a}$$

$$f(t) = T'(t)\, V^{-1}(t)\, \phi(t) \tag{8b}$$

$$T(t) = H(t)\, E(T|t-1) \tag{8c}$$

$$E(t|t-1) = M\, E(t-1|t-1) \tag{8d}$$

$$E(t|t) = E(t-1|t-1) - K(t)\, T(t) \tag{8e}$$

and $E(0|0) = I$.

Summing up, the estimation strategy has four stages:

(a) For any values of Q, R, and M (say Q^* , R^* and M^*) $z^*(0)$ is found from (7).

(b) Combining $z(0)$ with Q^*, R^* and M^* allows $\phi(t)$, $V(t)$ and $z(t|t)$ to be computed from (4), thereby yielding the likelihood corresponding to Q^*, R^* and M^* (see (6)).

(c) Determine, in some systematic way, the values of Q, R and M that maximize the likelihood by repeating (a) and (b). Methods for doing this are discussed below.

(d) If Q, R and M are the values that maximize the likelihood, $z(1|T),\ldots,z(t|T)$ are evaluated from (5).[5]

---

[5] One further simplification is possible in that R and Q may be divided by a constant. When $y(t)$ is a scalar this constant may be regarded as R so that Q will be variance ratios. Rosenberg [22] shows that the maximum likelihood estimate of R in this case is $T^{-1} \Sigma\, \phi'(t)\, V^{-1}(t)\, \phi(t)$.

Now the only remaining difficulty is to perform (c). A number of iterative techniques suggest themselves e.g. the algorithms of Powell [21], Davidon, Fletcher and Powell (D F P) [4] and Goldfeld and Quandt's Quadratic Hill Climbing (Q H C) [5]. Initial experiments revealed that the choice of algorithm was important owing to:

(a) The slowness of function evaluations -- about four a second for $T = 80$ and $q = 5$ on an IBM 360|91 -- causing selection to be conditioned by the need to minimize the number of function evaluations taken to reach a maximum if computational cost was not to be prohibitive.

(b) The need for a wide domain of convergence i.e., convergence to a maximum from a variety of starting values.

Generally, use of Powell and DFP satisfied (a) but there could be bad failures with respect to (b), while the converse held for Q H C. In particular, it was found that, when incorrect specification led to the maximum being located on one of the boundaries,[6] both Powell and D F P would terminate in the internal parameter space, thereby suggesting that the

---

[6] Rather than estimate the variances σ_j^2, the transformation $\beta_j = \sigma_j^2 | (1-\sigma_j^2)$ was adopted and the constraints were then $0 \leq \beta_j < 1$. This aids in the computation of derivatives as well as restricting the parameter space. Later tables present β_j rather than σ_j^2.

maximum occurred at reasonable parameter values. Q H C always located the boundary maximum.

As the disadvantages of Q H C are intimately connected with the numerical evaluation of the second derivative matrix that serves to provide a set of weights for the gradients, it seemed that further developments should attempt to approximate this matrix without performing the arduous calculations.[7] If the function to be minimized has the form $S = \epsilon'\epsilon$, the modified Gauss-Newton (GN) method [11] provides a sequence of iterations $\theta^{(0)}, \theta^{(1)}, \ldots.$ governed by

$$\theta^{(n)} - \theta^{(n-1)} = - \left(\frac{\partial \epsilon'}{\partial \theta} \frac{\partial \epsilon}{\partial \theta}\right)^{-1} \frac{\partial \epsilon'}{\partial \theta} \epsilon \qquad (9)$$

to minimize S with respect to θ . (9) would be Q H C if the weighting matrix were $\frac{\partial^2 S}{\partial \theta \partial \theta'}$ and it can be shown that $\operatorname{plim} \frac{\partial^2 S}{\partial \theta \partial \theta'} = \frac{\partial \epsilon'}{\partial \theta} \frac{\partial \epsilon}{\partial \theta}$, i.e., asymptotically the GN and Q H C algorithms yield an identical step length for any iteration. In as much as the time required to effectuate a function evaluation varies directly with T and q^2 , substantial savings could be made in large samples if (9) were adopted. To convert (6) to the form required for the application of GN consider the negative of the log likelihood (with constant term omitted), i.e.

---

[7] The QHC method is essentially a Newton-Raphson technique (see [5, p. 4-9]) with some modifications that make it more robust. In later discussion it will be regarded as a Newton-Raphson algorithm.

$$S = \sum_{t=1}^{T} \phi'(t)\, V^{-1}(t)\, \phi(t) - \sum_{t=1}^{T} \log \det V^{-1}(t) \quad . \qquad (10)$$

Let b and a be $T \times 1$ vectors with $b(j) = \phi(j)\, V^{-1/2}(j)$ and $a(j) = \text{abs}[\log \det V^{-1}(j)]^{1/2}$ respectively, then Appendix 1 shows that, for $y(t)$ a scalar, (10) may be written as

$$S = b'b + a'a$$

or

$$S = \epsilon'\epsilon$$

where $\epsilon = [b \,:\, a]$.

The GN algorithm has a number of advantages:

(i) The number of function evaluations necessary to compute the weighting matrix $(\frac{\partial\epsilon'}{\partial\theta} \frac{\partial\epsilon}{\partial\theta})$ varies directly with the number of parameters in θ (rather than the square as with QHC).

(ii) With modifications e.g. Marquardt [16] it has been successfully applied to a variety of non-linear equations describing economic behavior e.g. [1], [8].

(iii) As the equations to be solved at each iteration have the same form as those for OLS, efficient numerical methods such as the Q-R algorithm [7] or singular value decomposition may be employed (see [12] for more on this).

3. An Example

To illustrate the foregoing methodology the model of (1a) - (1f) was fitted to a quarterly series with (1b) replaced by

$$T(t) = 2T(t-1) - T(t-2) + e(t) \qquad (1b)'$$

i.e. a second difference for the trend. The variance of $e(t)$ was set at 10 and the chosen series was Quarterly Personal Consumption Expenditure on Food ($m) -- one of the thirteen constituents of consumption expenditure in the Australian National Accounts. Table 1 contains estimates of the variance ratios (see f.n. 6) λ_1, λ_2 and the initial state vector.

Table 1

Parameter Estimates

λ_1	λ_2	$\alpha_1(0)$	$\beta_1(0)$	$\alpha_2(0)$	$T(0)$	$T(-1)$	L
0.72	0.70	-8.78	4.00	-0.59	215.6	199.5	-436.9

For the current example all algorithms converged to the above answers but, as Table 2 illustrates, there was wide variation in the requisite number of function evaluations.[8]

Table 3 contains the estimates of $\alpha_1(t|t)$, $\alpha_2(t|t)$, $\beta_1(t|t)$, $S_1(t|t)$ and $S_2(t|t)$ while Table 4 has $T(t|t)$ and $S(t|t)$.

[8] Starting values were zero in every case.

Table 2

Number of Function Evaluations to Reach a Maximum for Different Algorithms

Algorithm	Number
GN	28
Powell	55
DFP	97
QHC	147

Although there is evidence of a strong upward movement in $\alpha_1(t)$ and $\alpha_2(t)$ (and hence $S(t)$) it may be felt that there is too much variation -- particularly for $\beta_1(t)$. A warning that this pattern may be expected if the first difference is selected is to be found in [24] and it is argued there that a second difference may be more appropriate for economic data. Nevertheless, the model illustrates the approach and it is of some interest to note that the correlogram of the estimated irregular factor did not reveal any noticeable autocorrelation.

Table 5 contains the same variables as Table 4, now estimated from the complete sample, i.e., $\hat{T}(t|T)$ and $\hat{S}(t|T)$.

Table 3

Estimates of Various Parameters for the Food Data

	$\alpha_1(t/t)$	$\beta_1(t/t)$	$\alpha_2(t/t)$	$s_1(t/t)$	$s_2(t/t)$
1950/3	-8.78	4.03	-0.62	4.03	0.62
4	-8.79	4.03	-0.62	8.79	-0.62
1951/1	-8.69	3.85	-0.81	-3.84	0.81
2	-7.60	4.44	0.17	-7.60	0.17
3	-7.28	3.95	0.59	3.95	-0.59
4	-9.77	2.30	2.66	9.77	2.66
1952/1	-10.95	4.03	4.10	-4.03	-4.10
2	-13.48	2.31	2.01	-13.48	2.01
3	-9.81	-3.07	6.45	-3.07	-6.45
4	-9.22	-2.67	5.96	9.22	5.96
1953/1	-10.37	-0.98	7.35	0.98	-7.35
2	-8.86	0.04	8.59	-8.86	8.59
3	-11.15	3.41	5.82	3.41	-5.82
4	-10.66	3.74	5.42	10.66	5.42
1954/1	-12.05	5.79	7.10	-5.78	-7.10
2	-12.66	5.37	6.60	-12.66	6.60
3	-11.05	3.01	8.55	3.01	-8.55
4	-12.81	1.82	9.99	12.81	9.99
1955/1	-11.47	-0.14	8.38	0.14	-8.38
2	-11.13	0.09	8.67	-11.13	8.67
3	-11.71	0.95	7.95	0.95	-7.95
4	-10.13	2.03	6.65	10.13	6.65
1956/1	-8.80	0.08	5.04	-0.08	-5.04
2	-10.83	-1.31	3.36	-10.83	3.36
3	-12.42	1.01	1.45	1.01	-1.45
4	-11.50	1.64	0.69	11.50	0.69
1957/1	-14.28	5.72	4.06	-5.71	-4.06
2	-13.23	6.43	4.92	-13.23	4.92
3	-11.12	3.34	7.48	3.34	-7.48
4	-12.23	2.57	8.40	12.23	8.40
1958/1	-10.15	-0.49	5.88	0.49	-5.88
2	-10.85	-0.96	5.30	-10.85	5.30
3	-10.76	-1.10	5.42	-1.10	-5.42
4	-12.39	-2.22	6.76	12.39	6.76
1959/1	-12.40	-2.20	6.78	2.20	-6.78
2	-13.77	-3.13	5.65	-13.77	5.65
3	-14.59	-1.92	4.66	-1.92	-4.66
4	-15.92	-2.82	5.75	15.92	5.75

Table 3 (cont)

	α_1(t/t)	β_1(t/t)	α_2(t/t)	S_1(t/t)	S_2(t/t)
1960/1	-16.49	-1.99	6.44	1.99	-6.44
2	-15.67	-1.43	7.12	-15.67	7.12
3	-17.06	0.61	5.43	0.61	-5.43
4	-16.36	1.09	4.85	16.36	4.86
1961/1	-18.41	4.09	7.32	-4.09	-7.33
2	-19.71	3.19	6.25	-19.71	6.25
3	-19.54	2.93	6.46	2.93	-6.46
4	-18.15	3.88	5.32	18.15	5.32
1962/1	-18.62	4.56	5.89	-4.56	-5.89
2	-16.95	5.71	7.27	-16.95	7.27
3	-16.71	5.36	7.55	5.36	-7.55
4	-19.99	3.13	10.26	19.99	10.26
1963/1	-18.75	1.30	8.75	-1.30	-8.75
2	-19.09	1.06	8.47	-19.09	8.47
3	-18.75	0.58	8.87	0.57	-8.87
4	-17.30	1.57	7.66	17.30	7.66
1964/1	-18.08	2.73	8.61	-2.73	-8.61
2	-17.26	3.29	9.30	-17.26	9.30
3	-19.27	6.24	6.86	6.24	-6.86
4	-22.89	3.78	9.85	22.89	9.85
1965/1	-26.01	8.36	13.62	-8.36	-13.62
2	-25.32	8.83	14.19	-25.32	14.19
3	-25.92	9.71	13.46	9.71	-13.46
4	-25.83	9.77	13.39	25.83	13.39
1966/1	-27.17	11.74	15.01	-11.74	-15.01
2	-26.89	11.93	15.25	-26.89	15.25
3	-27.48	12.80	14.53	12.80	-14.53
4	-28.17	12.32	15.11	28.17	15.11
1967/1	-26.57	9.97	13.17	-9.97	-13.17
2	-29.29	8.12	10.93	-29.29	10.93
3	-28.73	7.30	11.60	7.30	-11.60
4	-28.40	7.53	11.33	28.40	11.33
1968/1	-27.70	6.50	10.48	-6.50	-10.48
2	-27.47	6.66	10.67	-27.47	10.67
3	-24.83	2.79	13.85	2.79	-13.85
4	-26.44	1.70	15.18	26.44	15.18
1969/1	-26.32	1.52	15.04	-1.52	-15.04
2	-25.98	1.76	15.32	-25.98	15.32
3	-27.66	4.22	15.29	4.22	-13.29
4	-30.13	2.54	15.33	30.13	15.33
1970/1	-30.61	3.24	15.90	-3.24	-15.90
2	-27.66	5.25	18.34	-27.66	18.34

Table 4

Estimates of Trend $T(t|t)$ and Seasonal $S(t|t)$

	Obs.	$S(t\mid t)$	$T(t\mid t)$
1950/3	237	4.66	232.3
/4	257	8.17	248.8
1951/1	263	-3.03	266.0
2	279	-7.43	286.4
3	307	3.36	303.7
4	342	12.43	329.4
1952/1	338	-8.13	346.2
2	346	-11.47	357.6
3	346	-9.52	355.8
4	375	15.18	359.9
1953/1	353	-6.37	359.5
2	366	-0.27	366.2
3	379	-2.42	381.2
4	406	16.08	389.9
1954/1	380	-12.89	393.0
2	391	-6.05	397.1
3	389	-5.54	394.7
4	424	22.80	401.1
1955/1	403	-8.24	411.1
2	417	-2.46	419.4
3	423	-7.00	430.0
4	451	16.78	434.3
1956/1	442	-5.12	447.0
2	443	-7.47	450.6
3	464	-0.43	464.3
4	484	12.20	471.9
1957/1	458	-9.77	468.0
2	465	-8.30	473.2
3	463	-4.14	467.3
4	490	20.64	469.3
1958/1	474	-5.40	479.2
2	477	-5.55	482.6
3	480	-6.52	486.5
4	515	19.15	495.8
1959/1	498	-4.59	502.6
2	497	-8.11	505.2
3	507	-6.58	513.5
4	546	21.68	524.2

Table 4 (cont)

	Obs.	$S(t\mid t)$	$T(t\mid t)$
1960/1	526	-4.45	530.5
2	532	-8.55	540.5
3	551	-4.82	555.7
4	587	21.22	565.8
1961/1	556	-11.42	567.6
2	356	-13.46	569.5
3	569	-3.53	572.5
4	595	23.47	571.6
1962/1	560	-10.45	570.5
2	566	-9.68	575.6
3	575	-2.19	577.2
4	620	30.25	589.6
1963/1	593	-10.05	602.9
2	602	-10.62	612.6
3	613	-8.29	621.3
4	651	24.96	626.1
1964/1	618	-11.34	629.4
2	629	-7.96	636.9
3	652	-0.62	652.4
4	708	32.73	675.1
1965/1	656	-21.98	678.2
2	679	-11.13	690.1
3	700	-3.76	703.7
4	755	39.22	715.8
1966/1	695	-26.75	721.9
2	720	-11.64	731.6
3	742	-1.73	743.7
4	800	43.28	756.7
1967/1	753	-23.14	776.0
2	765	-18.36	783.5
3	788	-4.30	792.3
4	841	39.72	801.3
1968/1	797	-16.98	813.9
2	809	-16.81	825.8
3	814	-11.06	825.3
4	877	41.62	835.3
1969/1	827	-16.56	843.5
2	842	-10.66	852.6
3	860	-9.06	868.9
4	935	45.46	889.4
1970/1	885	-19.14	904.2
2	920	-9.32	929.2

Table 5

Estimates of Trend $T(t|T)$ and Seasonal $S(t|T)$

	$S(t\|T)$	$T(t\|T)$
1950/3	4.59	232.2
4	8.69	248.5
1951/1	-3.86	267.0
2	-8.23	287.4
3	-1.50	308.5
4	14.11	327.7
1952/1	-4.31	341.9
2	-3.85	349.6
3	-6.65	353.0
4	17.80	357.2
1953/1	-9.84	363.1
2	-5.44	371.6
3	-2.35	380.9
4	19.03	386.9
1954/1	-10.46	390.5
2	-2.73	393.7
3	-8.17	397.5
4	19.71	404.2
1955/1	-8.88	411.9
2	-3.44	420.2
3	-4.64	427.8
4	14.98	436.3
1956/1	-4.16	445.9
2	-10.93	454.4
3	0.80	462.6
4	17.29	466.5
1957/1	-9.13	467.5
2	-4.22	468.9
3	-6.69	469.9
4	17.30	472.9
1958/1	-4.23	478.0
2	-6.38	483.3
3	-8.49	488.6
4	20.09	494.9
1959/1	-3.43	501.2
2	-10.00	507.1
3	-6.97	514.1
4	22.94	523.1

Table 5 (cont)

	S(t|T)	T(t|T)
1960/1	-7.20	533.3
2	-11.98	544.0
3	-3.25	554.1
4	25.16	561.7
1961/1	-10.05	566.1
2	-12.66	568.8
3	-1.56	570.3
4	24.86	570.3
1962/1	-11.40	571.4
2	-9.63	575.6
3	-7.09	582.5
4	27.21	592.8
1963/1	-10.68	603.4
2	-9.87	611.6
3	-4.40	617.4
4	28.13	623.3
1964/1	-14.34	632.2
2	-13.67	642.9
3	-3.61	655.6
4	39.06	668.6
1965/1	-23.91	680.2
2	-12.86	691.8
3	-3.02	702.8
4	42.49	712.4
1966/1	-26.21	721.6
2	-13.48	733.4
3	-4.36	746.4
4	40.60	759.3
1967/1	-18.54	771.4
2	-16.89	782.2
3	-5.36	793.2
4	38.01	803.2
1968/1	-16.15	813.0
2	-12.06	820.7
3	-12.48	826.8
4	42.28	834.7
1969/1	-17.59	844.4
2	-13.88	856.2
3	-11.69	871.6
4	45.58	889.3
1970/1	-23.18	908.4
2	-9.32	929.2

Conclusion

There are a number of advantages to treating seasonal adjustment in the above framework. Firstly, it is possible to test alternative hypothesis concerning the nature of the seasonal by standard statistical methods. Secondly, an evolving seasonal pattern in the variables of a regression equation may be handled. Thirdly, as the filtering and estimation theory holds for $y(t)$ a vector i.e., $p > 1$, there is the potential for the simultaneous deseasonalization of closely related series. Finally, in the light of the demonstration elsewhere [19] that the Grether-Nerlove unobserved components model of trend, seasonal and irregular [9] can be given the linear system formulation (2), it is possible to compare the relative effectiveness of the two approaches in a common framework.

APPENDIX 1

To Prove: If $y(t)$ is a scalar

$$S = \sum_{t=1}^{T} \phi'(t)\, V^{-1}(t)\, \phi(t) - \sum_{t=1}^{T} \log \det V^{-1}(t)$$

can be written as

$$S = b'b + a'a$$

where $b(j) = \phi'(j)\, V^{-\frac{1}{2}}(j)$ and $a(j) = \text{abs}[\log \det V^{-1}(t)]^{\frac{1}{2}}$.

Method of Proof: A sufficient condition for the above result to hold is that $\log \det V^{-1}(j) < 0$.

PROOF: From the text $V(j) = H(j)\, P(j|j-1)\, H'(j) + R$ and, as noted in f.n. 5, for scalar $y(t)$ there is no loss of generality in setting $R = 1$. Therefore, $V(j) = H(j)\, P(j|j-1)\, H'(j) + 1$. Consider $P(j|j-1)$. It is a covariance matrix [17, p. 172] and thus positive definite i.e., for any vector α, $\alpha \neq 0$, $\alpha' P(j|j-1)\, \alpha > 0$ [6, p. 34]. Choosing $\alpha = H'(j)$ gives $H(j)\, P(j|j-1)\, H'(j) > 0$ so that $V(j) > 1$, $V^{-1}(j) < 1$ and $\log V^{-1}(j) < 0$.

REFERENCES

[1] Arndt, S.W., "International Short Term Capital Movements: Distributed Lag Model of Speculation in Foreign Exchange," Econometrica, 36(1968), 59-70.

[2] Cooley, T.F. and E.C. Prescott, "Varying Parameter Regression: A Theory and Some Applications," Annals of Economic and Social Measurement (forthcoming).

[3] Doran, H.E. and J.J. Quilkey, "Harmonic Analysis of Seasonal Data: Some Important Properties," American Journal of Agricultural Economics, 54 (1972), 646-651.

[4] Fletcher, R., and M.J.D. Powell, "A Rapidly Convergent Descent Method for Minimization," Computer Journal, 6 (1963), 163-168.

[5] Goldfeld, S.M. and R.E. Quandt, Nonlinear Methods in Econometrics (Amsterdam, North-Holland, 1972).

[6] Goldberger, A.S., Econometric Theory (John Wiley, New York, 1964).

[7] Golub, G.H., "Numerical Methods for Solving Linear Least Squares Problems," Numerische Mathematik, 7 (1965), 206-216.

[8] Gordon, R.J., "The Incidence of the Corporation Income Tax in U.S. Manufacturing, 1925-62," American Economic Review, 57 (1967), 731-758.

[9] Grether, D.M. and M. Nerlove, "Some Properties of 'Optimal' Seasonal Adjustment," Econometrica, 38 (1970), 682-704.

[10] Hannan, E.J., R.D. Terrell and N. Tuckwell, "The Seasonal Adjustment of Economic Time Series," International Economic Review, 11 (1970), 24-52.

[11] Hartley, W., "The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares," Technometrics, 3 (1961), 269-280.

[12] Jennings, L.S. and M.R. Osborne, "Applications of Orthogonal Matrix Transformations to the Solution of Systems of Linear and Nonlinear Equations," Technical Report No. 37, Computer Centre, Australian National University, 1970.

[13] Jorgenson, D.W., "Minimum Variance, Linear, Unbiased Seasonal Adjustment of Economic Time Series," Journal of the American Statistical Association, 59 (1964), 681-724.

[14] Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," Journal of Basic Engineering, 82 (1960), pp. 35-45.

[15] ______________, and R.S. Bucy, "New Results in Linear Filtering and Prediction Theory," Journal of Basic Engineering, 83 (1961), 95-108.

[16] Marquardt, D.W., "An Algorithm for Least Squares Estimation of Non-Linear Parameters," S.I.A.M. Journal, 11 (1963), 431-441.

[17] Meditch, J.S., Stochastic Optimal Linear Estimation and Control (McGraw-Hill, New York, 1969).

[18] Mehra, R.K., "On a Limiting Procedure in Linear Recursive Estimation Theory," Journal of Mathematical Analysis and Applications, 26 (1969), 1-8.

[19] Pagan, A.R., "A Note on the Extraction of Components from Time Series," Econometric Research Program Research Memorandum No. 148 (May 1973), Princeton University.

[20] Phillips, A.W., "The Estimation of Systems of Difference Equations with Moving Average Disturbances," (Paper read to the Econometric Society Meeting, San Francisco, 1966).

[21] Powell, M.J.D., "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives," Computer Journal, 7 (1964), 155-162.

[22] Rosenberg, B., "The Analysis of a Cross Section of Time Series by Stochastically Convergent Parameter Regression," Annals of Economic and Social Measurement (forthcoming).

[23] ____________, "A Survey of Stochastic Parameter Regression," Annals of Economic and Social Measurement (forthcoming).

[24] Sarris, A.H., "A Bayesian Approach to Estimation of Time-Varying Regression Coefficients," Annals of Economic and Social Measurement (forthcoming).

[25] Schweppe, F.C., "Evaluation of Likelihood Functions for Gaussian Signals," I.E.E.E. Transactions on Information Theory, 11 (1965), 61-70.

[26] Terrell, R.D. and N.E. Tuckwell, "Application of an Evolving Seasonal Pattern to an Australian Time Series," Australian Economic Papers, 9 (1970), 186-218.

[27] ____________, "The Efficiency of Least Squares in Estimating a Stable Seasonal Pattern," Journal of the American Statistical Association, 66 (1971), 354-362.

[28] Walker, A.M., "On the Estimation of a Harmonic Component in a Time Series with Stationary Independent Residuals," Biometrika, 58 (1971), 21-36.