No. 343

On Temporal Aggregation of Linear Dynamic Models

by

Noriyoshi Oguchi & Takao Fukuchi

June 1987

(Draft for Discussion Paper)

"On Temporal Aggregation of Linear Dynamic Models"

Noriyoshi Oguchi

Takao Fukuchi

(Institute of Socio-Economic Planning

University of Tsukuba

1-1-1 Tennodai, Sakura

Niihari, Ibaraki 305)

June,1987

1. Introduction

Distributed lags are frequently specified in economic models. However, availables data are often more aggregated with respect to time than suggested by the specification of the true lags. Hence econometricians have to use inappropriate data. The error induced by this practice have been examined by Brewer [1973], Engle and Liu [1972], Moriguchi [1970], Mundlak [1961], Tiao and Wei [1976], and Wei [1978], to mention only a few. These contributions, however, are focused, mainly, on statistical consideratons. One of the difficulties involed, as Griliches [1972] pointed out, is that comparison of the statistical results obtained from the basic model and some analogus aggregate form in parameter estimation, prediction and so on involves errors caused by misspecification as well as tempolal aggregation. In this paper, we investigate the conditon for perfect aggregation for deterministic model. Klein [1974,pp416-422] has discussed the relation between the coefficients of quarterly and annual models in the single equation context. Moriguchi et al. [1966] have extended the discussion to simultaneous equation systems without reaching clear conclusions. We will show that their aggregated model is misspecified.

2. The Model

For simplicity we limit our discussion to the relation between quarterly and annual models. Suppose that the true lag structure is captured by the following quarterly model in k endogenous variables.

$$(1)\quad y_{4t+i} = \alpha_0 + \alpha_1 y_{4t+i} + \sum_{j=1}^{J} \alpha_{2,j}\, y_{4t+i} \qquad t = 0, 1, 2, \ldots \quad i = 0, 1, 2, 3$$

Here y and are k - dimentional vectors, and are k x k matrices. The subscript "i" stands for the relevant quarter, and the year is represented by the subscript "t". The subscript "j" indictes the time lag in quarters and J indicates the longest time lag in the model. Note that usually most elements of are null. Also as we shall see, the length of the lag in each equation is not very important while the length of the lag in the complete model is crucial. Hence we specify only this parameter and assume that the difference equation which can be derived from system (1) is of order q, (q<T). Note finally that in the present model there are no true exogenous variables. The introduction of exogenous variables will be discussed later.

Write the solution of quarterly model (1) as

$$(2)\quad y_{4t+i} = h_0 + h_1 r_1^{4t+i} + \ldots\ldots + h_q r_q^{4t+i}$$

$$= Hr$$

where h_0, h_1, ... h_q are k-dimentional vectors and H is k × (q+1) matrix. Notice that $r = (1\ r_1^{4t+i}\ r_2^{4t+i}\ \ldots\ldots\ldots r_q^{4t+i})'$. r_1, , r_q are the characteristic roots of the solution of the difference equations derived from (1). The constants h_0,....., h_q depend on the initial values of the endogenous variables. If any of the roots is zero, then the solution is equvalent to that of lesser order difference equation. This contradicts the assumption that the order of the difference equation is q. Hence we assume that $r_1, \ldots\ldots\ldots, r_q$ are distinct non-zero roots. We will discuss the repeated root case later. Let the capital letter Y stand for the annual value obtained by summing over the quarters of the true model, i,e.

(3) $Y_t = \sum_{i=0}^{3} y_{4t+i}$.

Substitution of equation (2) into (3) yields:

(4) $$Y_t = 4h_0 + h_1(1+r_1+r_1^2+r_1^3)r_1^{4t} + h_2(1+r_2+r_2^2+r_2^3)r_2^{4t} + h_3(1+r_3+r_3^2+r_3^3)r_3^{4t} + \cdots + h_q(1+r_q+r_q^2+r_q^3)r_q^{4t}$$

$$= H \rho r^{4t}$$

where ρ is a (q+1) x (q+1) diagoanal matrix with non zero elements of the form (1 $1+r_1+r_1^2+r_1^3$ . . . . $1+r_q+r_q^2+r_q^3$) and $r^{4t} = (4\ r_1^{4t}\ .\ .\ .\ .\ .\ .\ .\ r_q^{4t})'$.

Suppose we specify the annual model as follows.

(5) $Y^*_{m,t} = A_{m0} + A_{m1} Y^{*e}_{m1} + \cdots + A_{mq'} Y^{*e}_{mq'} = A'_m Y^{*e}_m$

where $Y^*_{m,t}$ represents the m-th variable of the annual model and $Y^{*e}_{m1}, \ldots, Y^{*e}_{mq'}$ are explanatory variables. Hence A_m and Y^{*e}_m are (q'+1) dimentional vectors with a unitary first element in Y^{*e}_m. The time subscript for the explanatory variables are not specified since they may refer to lags of different length. It is obvious that $Y^*_{m,t}$ cannot be an explanatory variable while $Y^*_{m,t-j}$ may. We assume finally that the explanatory variables are not exactly collinear.

Theil [1954, pp140] defines aggregation as 'perfect' when "there is no contradiction between the macroequation and the microequations corresponding to it, for whatever values and changes assumed by the microvariables and at whatever point or period of time." We follow this definition by replacing the macro-micro relation with the temporally aggregated-disaggregated one. In our notation this means that the condition $Y_t = Y_t^*$ must hold for all t.

Theorem 1. Necessary and sufficient conditions for perfect temporal aggregation of model (1) are i) the aggregated model must reduce to a system of difference equations of the q-th order with the same characteristic roots as the quarterly model raised to the fourth power; ii) for each equation of the quarterly model containing any lagged explanatory variable in the quarterly model, the annual counterpart must have q explanatory variables and a constant term.

Necessity Suppose that $Y^*_{m,t} = Y_{m,t}$ for $t=0, 1, \ldots, T-1$, and $m = 1, \ldots, k$. where $Y_{m,t}$ is the m-th variable in system (4). Upon substitution of (4) into (5), we have

(6) $Y^{*}_{m,t} = A'_m G_m r^{4t} = Y_{m,t} = H_m \rho r^{4t}$

where G_m is a (q'+1) × (q +1) coefficient matrix on the constant terms and q charateristic roots for the q' explanatory variables used in m-th equation, and where H_m is the m-th row of the matrix H. Since there are only q+1 independent solutions (including the constant term), rank of G_m is at most q+1. Hence by the assumption of no exact collinearity among the explanatory variables, $q' \leq q$. On the other hand, to have q+1 equalities, $A_m G_m = H_m$ in (6), satisfied for all possible $H_m \rho$, there must be q+1 free parameters, A_{m0}, ..., A_q. Hence q' = q.

From equation (6) we can see that the annual variable must be a solution of a difference equation of order q with characteristic roots r^4_1, ..., r^4_q.

Sufficiency If the aggregated model is specified as in equation (5) with q'= q, and the model as a whole reduces to q-th order difference equations with characteristic roots r^4_1,....., r^4_q, then we can write

(7) $Y^{*e}_m = G_m r^{4t}$

as in (6). Substituting (7) into (5), we have

(8) $$Y^*_{m,t} = A'_m G_m r^{4t}$$

For perfect aggregation, $Y^*_{m,t}$ in (8) must be equal to the m-th equation in (4). Hence the following q+1 relations must be satisfied.

(9) $$A'_m G_m = H_m \rho$$

Since there are q+1 free parameters, A_{m0}, ..., A_{mq}, this can be achieved by assigning proper values to Am.

Q.E.D

Remark 1. If some of the characteristic roots r_1, - - - - -, r_q are repeated, say $r_1 = r_2 = r_3$, then we proceed by replacing r_2 and r_3 by tr_1 and $t^2 r_1$, respectively, as usually done in solving linear difference equations. In this case equations (2) and (4) will become:

(2)' $$y_{4t+i} = h_0 + h_1 r_1^{4t+i} + h_2(4t+i) r_1^{4t+i} + h_3(4t+i)^2 r_1^{4t+i} + h_4 r_4^{4t+i} + \cdots + h_q r_q^{4t+i}$$

(4)' $$Y_t = 4h_0 + \{h_1(1+r_1+r_1^2+r_1^3) + h_2(r_1+2r_1^2+3r_1^3) + h_3(r_1+4r_1^2+9r_1^3)\} r_1^{4t} + \{h_2(1+r_1+r_1^2+r_1^3) + 2h_3(r_1+2r_1^2+3r_1^3)\} 4t\, r_1^{4t} + 16h_3(1+r_1+r_1^2+r_1^3) t^2 r_1^{4t} + \cdots + h_4(1+r_4+r_4^2+r_4^3) r_4^{4t} + \cdots + h_q(1+r_q+r_q^2+r_q^3) r_q^{4t}$$

since Y_t is still a linear combinations of q independent time paths and G_m is a full rank matrix, the above arguments hold. Similarly, the case of complex roots does not require separate treatment since they are independent solutions.

Remark 2. The theorem is for the estimated equations. Identities need only be adjusted by redefinition. See, for example, equations (14) and (18).

Remark 3. In practice we have no prior knowledge on the characteristic roots. Hence we first estimate aggregated equation (5) with available aggregated data and then derive the difference equations for the aggregated model. Solving this difference equation, we obtain estimates of the characteristic roots.

In conclusion, the main implication of the theorem is that, in general, if the same specification for the annual and for the quarterly model is adopted, then the conditions for perfect temporal aggregation are not satisfied. An example is discussed in section 4.

3. A Model with true exogenous variables

In this section we extend our results to the more realistic case in which there are true exogenous variables, whose time paths are known. We first state the appropriate theorem and then discuss it briefly.

Suppose that the quarterly model is given by

$$(10) \quad y_{4t+i} = \alpha_0 + \alpha_1 y_{4t+i} + \sum_{j=1}^{J} \alpha_{2,j} y_{4t+i-j} + \alpha_3 x_{4t+i} \qquad \begin{matrix} t=0,1,\cdots,T-1 \\ i=0,1,2,3 \end{matrix}$$

where x_{4t+i} is a z x 1 vector of exogenous variables and α_3 is a k x z matrix of coefficients. (the other variables are as in equation (1)). Assume that the system reduces to a set of q-th order difference equations and that all the time paths of exogenous variables are linearly independent of each other and of the solutions for the above difference equations. In place of (2), we have:

$$(11) \quad y_{4t+i} = h_0 + h_1 r_1^{4t+i} + \cdots + h_q r_q^{4t+i} + f_1 x_{1,4t+i} + \cdots + f_z x_{z,4t+i}$$

$$= H r + F x(4t+i)$$

where x(4t+i) is a z dimentional vector function of 4t+i and

$f_0, \ldots, f_q$ are k dimentional vectors of coefficients. The annual aggregates are:

$$(12) \quad Y_t = H \rho r^{4t} + FX(t)$$

where X(t) is the annual aggregates of x(4t+i) and we assume that X(t) is known. Suppose an annual model is specified as follows.

$$(13) \quad Y^*_{t,m} = A'_m Y^{*e}_m + B'_m X_m(t)$$

Where A_m and Y_m are (q'+1) dimentional vectors and B_m and $X_m(t)$ are z' dimentional vectors. Again we assume that there is no exact collinearity among the explanatory variables. Hence $z' \leq z$.

Theorem 2. The conditions for perfect temporal aggregation are now that i) the annual model reduces to a system of difference equations of order q with the same characteristic roots as for the quarterly model raised to the fourth power, ii) the model includes the same set of exogenous variables and, iii) for each estimated equation containing lagged endogenous variables in the quarterly model, the annual counterpart must contain at least q lagged and unlagged endogenous variables for a total of q+z explanatory variables plus a constant term.

Necessity Suppose that $Y^*_{t,m} = Y_{t,m}$ for t=0, 1, ...,T-1, and m=

1, ..., k. where Y_m is the m-th variable in system (12). Upon substitution of (12) into (13), we have

$$(14)\quad Y^*_{t,m} = A'_m (G^1_m r^{4t} + G^2_m X(t)) + B'_m X_m(t)$$
$$= A'_m G^1_m r^{4t} + (A'_m G^2_m + B'_m \lambda) X(t)$$
$$= (A'_m \ B'_m) \begin{pmatrix} G^1_m & G^2_m \\ 0_{z',q+1} & \lambda \end{pmatrix} \begin{pmatrix} r^{4t} \\ X(t) \end{pmatrix} = (A'_m \ B'_m)\, G^*_m \begin{pmatrix} r^{4t} \\ X(t) \end{pmatrix}$$
$$= Y = H \ r^{4t} + F_m X(t)$$

where G^1_m is a (q'+1) × (q +1) and G^2_m is a (q'+1) × z coefficient matrix on the constant term and q charateristic roots for the q' endogenous explanatory variables and where B_m is a z' dimentional vector of coefficient on the z' exogenous variables, X_m, included in the m-th equation. H_m and F_m are the m-th rows of the matrices H and F, respectively. λ is z'× z matrix obtained by eleminating the rows corresponding to exogenous variables which are not included in the m-th equation from z dimentinal identity matrix. $0_{z',q+1}$ is z' × (q+1) null matrix. Since there are only q+z+1 independent solutions (including the constant term), there can be only q+z+1 independent rows in G^*_m. Hence by the assumption of no exact collinearity among the explanatory variables, $q'+z' \leq q+z$. From (14), we have the following q+z+1 equalities,

(15) a) $A'_m G^1_m = H_m \rho$

b) $(A'_m G^2_m + B'_m \lambda) = F_m$

To have (15a) satisfied for all possible $H_m \rho$, there must be at least q+1 free parameters, A_{m0}, ------, A_q. Hence $q' \geq q$. Similarly for (15) to hold for all possible H_m and F_m, there must be q+z+1 free parameters.

From equation (12) we can see that the annual variable must be a solution of a difference equation of order q with characteristic roots $r^4_1, \ldots, r^4_q$.

Sufficiency If the aggregated model is specified as in equation (13) with $q' \geq q$, and $q'+z'=q+z$ and the model as a whole reduces to q-th order difference equations with characteristic roots r_1, --------, r_q, then we can write

(16) $Y_t = H \rho r^{4t} + FX(t)$

as in (12). Substituting (16) into (13), we have

$$(17)\quad Y^{*}_{t,m} = A'_m G^1_m \Gamma^{4t} + (A'_m G^2_m + B'_m \lambda) X(t)$$

For perfect aggregation, $Y^{*}_{t,m}$ in (17) must be equal to the m-th equation in (12). Hence the following q+z+1 relations must be satisfied.

$$(18)\quad A'_m G^1_m = H_m \rho$$

$$(A'_m G^2_m + B'_m \lambda) = F_m$$

Since there are q+z+1 free parameters, A_{m0}, ..., A_{mq}, and B_1,....,B_z, this can be achieved by assigning proper values to Am and Bm.

Q.E.

iii) is satisfied for an equation with q endogenous variables and all the exogenous variables as explanatory variables. This can be easily proved along the same line as sufficiency in theorem 1. However, it is not a necessary condition. If an exogenous variable is included somewhere in the model, then through simultaneity its effect is transmitted to all the endogenous variables. That is, although it is still necessary to have q+z explanatory variables, it is not necessary to include all the exogenous variables in each equation. Hence the necessary and sufficient condition reduces to the selection of a set of q+z endogenous and exogenous variables of which at least q must be endogenous. For large models containing many exogenous variables the following difficulties appear. i) The aggregated model may not directly represent a paticular hypothsis since there must be q+z explanatory variables. ii) The estimation may become more difficult because of collinearity.

4. A numerical example

Let us consider the example discussed in Moriguchi et al. [1966]. The quarterly model is given by the following equations:

$$(19) \quad j_{4t+i} = 0.3*y_{4t+i} - 0.3*h_{4t+i}$$

$$(20) \quad c_{4t+i} = 30 + 0.8*y_{4t+i-1}$$

$$(21) \quad y_{4t+i} = c_{4t+i} + j_{4t+i}$$

$$(22) \quad h_{4t+i} = j_{4t+i} + h_{4t+i-1}$$

For this quarterly model they specified the following annual model and compared the time pathes generated by the quarterly and annual models.

$$(23) \quad J_t = a_0 + a_1 Y_t - a_2 H_{t-1}$$

$$(24)\quad C_t = b_0 + b_1 Y_{t-1}$$

$$(25)\quad Y_t = C_t + J_t$$

$$(26)\quad H_t = J_t + H_{t-1}$$

The quarterly model (19) - (22) reduces to a set of second order difference equations. Equations (23) - (26) of the annual model are specified in the same form as the quarterly model, hence they also reduces to a set of second order difference equations. For perfect aggregation, however, equation (24) in the annual model should have another explanatory variable which does not change the order of difference equation. For example, Y_t satisfies this condition. Since the annual model (19) - (22) does not satsfy the aggregation condition given above, it is natural that the authors detected specification error in the annual model. As an additional variable is added to equation (24), and intial conditions of $h_0 = 50$ and $y_0 = 100$, the following annual model generates the same yearly values as the four quarter aggregates of the quarterly model.

$$J_t = 15.4521 + 0.18356*Y_t - 0.83724*H_{t-1}$$

$$C_t = 45.6048 + 0.74553*Y_t + 0.17847*Y_{t-1}$$

$$Y_t = C_t + J_t$$

$$H_t = J_t + H_{t-1}$$

(University of Tsukuba)
(University of Kyoto)

References

Brewer, K. R. W., "Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models", Journal of Econometrics, 1 (May, 1973), 133-154.

Engle, R. F. and T. C. Liu, "Effects of Aggregation Over Time on Dynamic Characteristics of an Econometric Model", in Econometric Models of Cyclical Behaviors, Vol. 2, Ed. B. G. Hickman, (Columbia University Press, 1972), 673-737.

Griliches, Z., Discussion on the paper by R. F. Engle and T. C. Liu, in Econometric models of Cyclical Behavior, Vol. 2, Ed. B. G. Hickman, (Columbia University Press, 1972), 733-737.

Klein, k., A Textbook of Econometrics, 2nd Edition (Englewood Cliffs, Prentice-Hall, 1974)

Moriguchi, C., T. Nakazawa, and S. Toshida, "Temporal Aggregation and Macroeconometric Models", Keizai Bunseki (Economic Analysis), No.18 (June, 1966), 1-7.

Moriguchi, C.,"Aggregation Over Time in Macroeconomic Relations" International Economic Review, 11 (October, 1970), 427-440.

Mundlak, Y., "Aggregation Over Time in Distributed Lag Models", International Economic Review, 2 (May, 1961), 154-163

Oguchi, N., and T. Fukuchi, "On Temporal Aggregation of Linear Dynamic Systems", Institute of Socio-Economic Planning Discussion Paper No.186, University of Tsukuba, (June, 1983)

Theil, H., Linear Aggregation of Economic Relations, (Amsterdam, North-Holland Publishing Co., 1954).

Tiao, G. C., and W. S. Wei, "Effect of Temporal Aggregation on the Dynamic Relationship of Two Time Series Variables", Biometrika, 63 (December, 1976), 513-523.

Wei, W. S., "The Effect of Temporal Aggregation on Parameter Estimation in Distributed Lag Model", Journal of Econometrics, 8 (October, 1978), 237-246.