A Further Note on the Derivation of Quarterly Figures Consistent with Annual Data

By VICTOR A. GINSBURGH

University of Brussels and CORE

SUMMARY

Several methods have been devised to deal with the problem of constructing quarterly data (a) either when related series are available or (b) when only annual totals exist.

Methods (a) lead to quarterly figures inconsistent with annual totals. Method (b) adds nothing to the economic interpretation of the quarterly phenomenon.

Very little has been done to construct quarterly data using related series such that consistency with annual figures is ensured. This note tries to improve and clarify the interpretation of the only existing method we know of.

Keywords: QUARTERLY INTERPOLATION; ANNUAL TIME SERIES

1. INTRODUCTION

SEVERAL methods have been devised to deal with the problem of constructing quarterly data (a) either when related series are available or (b) when only annual totals exist.

As will be seen, methods (a) lead to quarterly figures inconsistent with (presumably known) annual totals. On the other hand, method (b) adds nothing to the economic interpretation of the quarterly phenomenon.

Only very little has been done to construct quarterly data using related series, in such a way that consistency with known annual figures is ensured. We know of only one such attempt, due to Vangrevelinghe (1966) and utilized by Nasse (1970) and generally in the French quarterly national accounts.

This note tries to improve and clarify the interpretation of Vangrevelinghe's method.

2. THE USE OF RELATED SERIES

The idea is simple and widely applied; let

$$y_i^* = a_0 + a_1 x_i^* + u_i^* \quad (i = 1, 2, \ldots, N)$$

be a relation between annual values y_i^* and x_i^*. The quarterly unknown

$$y_j \; (j = 1, 2, \ldots, 4N)$$

are then simply obtained by

$$y_j = \hat{a}_0 + \hat{a}_1 x_j,$$

given $\hat{a}_0, \hat{a}_1$ and x_j, the related quarterly series.

Clearly, if $u_i^* \neq 0$ for some i,

$$\sum_{j=4i-3}^{4i} y_j \neq y_i^*$$

and the method does not ensure consistency.

3. Purely Mathematical Methods

Lisman and Sandee (1964), followed by Glejser (1966), Boot and Feibes (1967), Boot *et al.* (1967), Doornbos and Lisman (1968), tried to develop methods ensuring consistency.

3.1. *The Method of Lisman and Sandee*

Lisman and Sandee suppose that the quarterly y_j depend linearly on three annual figures y_{i-1}^*, y_i^* and y_{i+1}^*, so that

$$y_j = a_{j,-1} y_{i-1}^* + a_{j,0} y_i^* + a_{j,+1} y_{i+1}^*. \tag{3.1}$$

The twelve constants a_{jk} ($j = 1, 2, 3, 4$; $k = -1, 0, +1$) are found by imposing twelve restrictions on the matrix a_{jk} of which, one at least—the cycle condition—is very artificial; this probably explains why the method does not work very well and sometimes gives strange results—see, for example, Boot and Feibes (1967).

Quarterly data are finally computed as follows:

$$\begin{bmatrix} y_{4i-3} \\ y_{4i-2} \\ y_{4i-1} \\ y_{4i} \end{bmatrix} = \begin{bmatrix} 0{\cdot}0729 & 0{\cdot}1982 & -0{\cdot}0211 \\ -0{\cdot}0103 & 0{\cdot}3018 & -0{\cdot}0415 \\ -0{\cdot}0415 & 0{\cdot}3018 & -0{\cdot}0103 \\ -0{\cdot}0211 & 0{\cdot}1982 & 0{\cdot}0729 \end{bmatrix} \begin{bmatrix} y_{i-1}^* \\ y_i^* \\ y_{i+1}^* \end{bmatrix}. \tag{3.2}$$

The computations are seen to be very easy; however, quarterly interpolations for the first and the last year cannot be obtained.

3.2. *The Methods of Boot, Doornbos, Feibes and Lisman*

The methods work as follows: minimize

$$\sum_{j=2}^{4N} (y_j - y_{j-1})^2 \tag{3.3}$$

under the N constraints

$$\sum_{j=4i-3}^{4i} y_j = y_i^* \quad (i = 1, 2, \ldots, N).\dagger \tag{3.4}$$

The solution of problem (3.3)–(3.4) is straightforward; introducing N Lagrange multipliers 2λ, we have:

$$\begin{bmatrix} \mathbf{B} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} y \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ y^* \end{bmatrix} \tag{3.5}$$

† Or minimize $\Sigma_{j=3}^{4N} (\Delta y_j - \Delta y_{j-1})^2$ under the N constraints (3·4).

where **B** is the following square matrix of order $4N$

$$\mathbf{B} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}$$

and **C** is the following $N \times 4N$ matrix

$$\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \dots & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 1 & 1 & 1 \end{bmatrix}.$$

0 is an N-order zero matrix.

The method gives consistent quarterly data, but fails completely to take into account quarterly information. Further correlation work is hampered for two reasons: first, one can hardly argue that degrees of freedom have been added by this mathematical manipulation; second, the method generates autocorrelated errors.

4. The Combination of Related Series and Mathematical Interpolation

Vangrevelinghe (1966) proposes a method in two steps: first interpolate by using a purely mathematical method; second "modulate" the figures obtained by the first step, by using a related series. More formally, let

$$x_i^* \; (i = 1, 2, \dots, N) \quad \text{and} \quad x_j \; (j = 1, 2, \dots, 4N)$$

be the related annual and quarterly series and y_i^* $(i = 1, 2, \dots, N)$ the series to be interpolated. Thus:

(i) interpolate by Lisman and Sandee's method the series x_i^* and y_i^*; this generates $\hat{x}_j$ and $\hat{y}_j$

(ii) compute an annual least-squares equation $y_i^* = \hat{a}_0 + \hat{a}_1 x_i^*$

(iii) compute the interpolated final y_j as follows:

$$y_j = \hat{y}_j + \hat{a}_1(x_j - \hat{x}_j). \tag{4.1}$$

The method suffers from the same shortcomings as Lisman and Sandee's, especially the loss of eight quarterly figures; moreover, the method seems to be a purely pragmatic device without a clear interpretation.

The change we suggest is to use method (3.3)–(3.4) instead of Lisman and Sandee's to generate $\hat{x}_j$ and $\hat{y}_j$. It will be shown that this apparently minor change permits a clear interpretation of the procedure. Thus:

(i) interpolate x_i^* and y_i^* using (3.3) and (3.4); this generates $\hat{x}_j$ and $\hat{y}_j$

$$\begin{bmatrix} \hat{x} \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ x^* \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \hat{y} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ y^* \end{bmatrix}$$

λ and μ being Lagrange multipliers;

(ii) compute an annual regression equation

$$y_i^* = \hat{a}_0 + \hat{a}_1 x_i^*;$$

(iii) compute the interpolated final y_j as follows

$$\begin{bmatrix} y \\ \xi \end{bmatrix} = \begin{bmatrix} \hat{y} \\ \lambda \end{bmatrix} + \hat{a}_1 \begin{bmatrix} x - \hat{x} \\ -\mu \end{bmatrix} \tag{4.2}$$

with $\xi \equiv \lambda - \hat{a}_1 \mu$.

We now show what this procedure means. Premultiply (4.2) with $\begin{bmatrix} \mathbf{B} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix}$ and rearrange terms; this gives

$$\begin{bmatrix} \mathbf{B} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} y - \hat{a}_1 x \\ \xi \end{bmatrix} \begin{bmatrix} 0 \\ y^* - \hat{a}_1 x^* \end{bmatrix}. \tag{4.3}$$

From (4.3), we have

$$\mathbf{C}y - \hat{a}_1 \mathbf{C}x = y^* - \hat{a}_1 x^*.$$

Since $\hat{a}_1 \mathbf{C}x = \hat{a}_1 x^*$, (4.3) reduces to

$$\begin{bmatrix} \mathbf{B} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} y \\ \xi \end{bmatrix} = \begin{bmatrix} \hat{a}_1 \mathbf{B}x \\ y^* \end{bmatrix}, \tag{4.4}$$

and (4.4) can be seen to be the solution of the following interpolation problem: minimize

$$\sum_{j=2}^{4N} (\Delta y_j - \hat{a}_1 \Delta x_j)^2 \tag{4.5}$$

under the N constraints

$$\sum_{j=4i-3}^{4i} y_j = y_i^* \quad (i = 1, 2, \ldots, N). \tag{4.6}$$

Suppose furthermore the series y_i^* and x_j are expressed in the same units; thus the regression coefficient $\hat{a}_1 = 1$ and the problem can be written: minimize

$$\sum_{j=2}^{4N} (\Delta y_j - \Delta x_j)^2 \tag{4.7}$$

under the N constraints (4.6).

The expressions (4.6) and (4.7) can now be interpreted very logically, as the construction of a quarterly series y_j whose variations reflect closely those of the related series x_j, ensuring a smooth adaptation between successive quarters.

5. Empirical Results

The methods described in Sections 3 and 4 have been utilized to generate the quarterly American G.N.P. (Gross National Product) for the period 1955–64, using the index of industrial production as related series. The results of the various interpolations are shown in the Appendix, as well as the actual G.N.P. figures.

In order to summarize the results, two statistics have been computed; the first is Theil's (1958) well-known inequality coefficient

$$U = \frac{\{n^{-1}\sum(p_i - a_i)^2\}^{\frac{1}{2}}}{\{n^{-1}\sum p_i^2\}^{\frac{1}{2}} + \{n^{-1}\sum a_i^2\}^{\frac{1}{2}}} \quad (i = 1, 2, \ldots, n),$$

where p_i represents the prediction (the interpolation in our case) and a_i the actual figure, both expressed in quarter to quarter changes. It is easy to see that $0 \leqslant U \leqslant 1$, small U's being indications of good predictions.

The second statistic also due to Theil (1958) analyses turning points. Let

m_1 = number of turning points, correctly predicted;
m_2 = number of cases where turning points are incorrectly predicted;
m_3 = number of cases where turning points are, incorrectly, not predicted.

Theil then defines two coefficients

$$\phi_1 = m_2/(m_1 + m_2) \quad \text{and} \quad \phi_2 = m_3/(m_1 + m_3)$$

where $0 \leqslant \phi_1, \phi_2 \leqslant 1$; small ϕ's indicate successful turning point prediction.

The various statistics are shown in Table 1.

Table 1

Comparison of interpolations

	n	U	ϕ_1	ϕ_2
Lisman–Sandee	31	0·42	0·50	0·90
Boot–Doornbos–Feibes–Lisman	39	0·36	0·75	0·79
Vangrevelinghe	31	0·34	0·17	0·00
Author's	39	0·36	0·21	0·07

The comparison tends to judge Vangrevelinghe's method the best, although the difference with the author's method is very small, especially if one bears in mind that eight "observations" are lost when Vangrevelinghe's method is utilized.

Acknowledgements

I am indebted to H. Glejser and J. Waelbroeck, University of Brussels, for helpful comments on earlier drafts of this paper.

Note added in proof. Chow and Lin's (1971) paper appeared after this paper was completed and a first draft submitted. The ideas of their paper are close to the ones presented here. It would be interesting to test which method gives a better approximation to reality.

References

BOOT, C. G. and FEIBES, W. (1967). On Glejser's derivation of monthly figures from yearly data. *Cahiers Economiques de Bruxelles*, **36**, 539–546.

BOOT, C. G., FEIBES, W. and LISMAN, J. H. C. (1967). Further comments on the derivations of quarterly figures from annual data. *Appl. Statist.*, **16**, 65–75.

CHOW, G. C. and LIN, AN-LOH (1971). Best linear unbiased interpolation, distribution and extrapolation of time series by related series. *Rev. Econ. Statist.*, **53**, 372–375.

DOORNBOS, R. and LISMAN, J. H. C. (1968). Afleiding van kwartaal cijfers uit jaartotalen. *Statist. Neerland.*, **22**, 199–205.

FRIEDMAN, M. (1962). The interpolation of time series by related series. *J. Amer. Statist. Ass.*, **57**, 729–757.

GLEJSER, H. (1966). Une méthode d'évaluation de données mensuelles à partir d'indices trimestriels ou annuels. *Cahiers Economiques de Bruxelles*, **29**, 45–54.

KÜHLO, K. C. (1965). Quartalsschätzung für jährliche Zeitreihen. *Z. IFO-Inst. Wirtschaftsfors.*, **14**, 81–122.

LISMAN, J. H. C. and SANDEE, J. (1964). Derivation of quarterly figures from annual data. *Appl. Statist.*, **13**, 87–90.

NASSE, Ph. (1970). Peut-on suivre l'évolution trimestrielle de la consommation? *Econ. Statist.*, **8**, 20–22.

THEIL, H. (1958). *Economic Forecasts and Policy*. Amsterdam: North-Holland.

VANGREVELINGHE, G. (1966). L'évolution à court terme de la consommation des ménages. *Etudes et Conjoncture* (*INSEE*), **9**, 54–102.

APPENDIX 1 (*see overleaf*)

Appendix 1

Actual and Interpolated Figures

		Actual G.N.P. (1958 *prices*)	Industrial *production* (1958 = 100)	*Interpolated G.N.P. obtained with*			
				Method (3.1)—(3.2) (1)	*Method* (3.3)—(3.4) (2)	*Method* (4.1) (3)	*Method* (4.6)—(4.7) (4)
1955	1	428·0	98·3	—	436·5	—	419·0
	2	435·4	102·4	—	437·4	—	435·9
	3	442·1	104·9	—	438·4	—	445·5
	4	446·4	106·7	—	439·6	—	451·6
1956	1	443·6	105·9	443·1	441·4	444·7	446·1
	2	445·6	106·4	445·3	444·1	444·8	446·2
	3	444·5	105·3	447·1	447·5	440·7	439·7
	4	450·3	108·6	448·6	451·1	453·8	452·0
1957	1	453·4	109·2	451·0	453·9	455·8	452·9
	2	453·2	108·7	453·6	454·6	453·2	452·8
	3	455·2	108·6	453·8	452·6	455·9	458·1
	4	448·2	103·6	451·5	448·9	445·1	446·2
1958	1	437·5	96·5	446·4	444·9	430·0	429·1
	2	439·5	95·9	442·4	443·8	436·5	433·2
	3	450·7	101·7	445·3	446·7	457·9	458·1
	4	461·6	105·7	455·2	453·8	464·8	468·9
1959	1	468·6	110·0	466·6	463·7	471·3	474·6
	2	479·9	116·1	475·2	473·4	487·5	491·5
	3	475·0	112·3	480·2	481·0	471·8	469·9
	4	480·4	112·3	481·8	485·8	473·2	467·9
1960	1	490·2	118·6	483·6	487·8	496·7	496·3
	2	489·8	117·1	486·7	488·1	490·9	491·6
	3	487·4	115·9	489·4	487·7	487·2	488·2
	4	483·8	112·4	491·6	487·6	476·5	475·1
1961	1	482·7	110·7	491·8	488·9	472·6	469·6
	2	492·9	116·0	492·2	492·7	493·3	492·9
	3	501·6	119·6	497·5	499·3	506·9	508·0
	4	511·9	122·4	507·6	508·1	516·5	518·6
1962	1	519·7	124·1	518·7	518·1	522·4	523·6
	2	527·9	126·1	527·9	527·3	529·8	530·1
	3	533·6	127·3	534·5	534·6	533·8	533·4
	4	538·5	127·5	538·5	539·7	533·6	532·6
1963	1	541·2	128·9	541·8	543·4	537·8	537·2
	2	544·9	132·6	546·2	547·1	550·3	551·2
	3	553·7	134·0	552·1	551·7	554·6	555·0
	4	560·0	134·9	560·0	557·6	557·1	556·4
1964	1	567·1	137·2	—	565·0	—	563·3
	2	575·9	140·4	—	573·2	—	574·7
	3	582·6	142·8	—	581·7	—	583·5
	4	584·7	144·2	—	590·4	—	588·8

(1) Lisman–Sandee.
(2) Boot–Doornbos–Feibes–Lisman.
(3) Vangrevelinghe.
(4) Author's.