1 Introduction

Of the diverse problems faced in everyday applied work, those resulting from the reporting frequency of economic time series provided by most statistical agencies are possibly the most severe.[1] The problems arise when the underlying temporal frequency of economic propositions is higher than the reporting frequency of the data available for econometric analysis, a typical case being most macroeconomic variables. The effects on model specification, estimation, identification and inference have been widely reported in the econometrics literature. Some known examples are :

a. Brewer [1973] analysed how temporal aggregation for the ARIMA model altered the orders of the AR and MA components; this problem was also studied by Tiao [1972] and Wei [1981]; Weiss [1984] focused instead on the impact due to systematic sampling; whereas Engle and Liu [1972] and Geweke [1978], among others, analysed temporal aggregation in ARMA models which included lagged exogenous terms.

b. The analysis in a multivariate framework, more akin to economic models, was discussed by Lutkepohl [1987; 1991] for a class of vector ARIMA processes, focusing on information loss and the effect of aggregation on the forecast-error variance for the aggregate model. Nijman and Palm [1990], in contrast, looked at the prediction of the augmented series from aggregate data.

c. Identification and estimation of ARIMA models subject to temporal aggregation was also discussed by Palm and Nijman [1984] and Nijman and Palm [1990].

d. On the problems of inference, Campos *et al.* [1990] reported the impact of temporal aggregation on weak exogeneity; Lippi and Reichlin [1991] looked at aggregation and persistence measures in trade cycles; Granger and Siklos [1995] showed how spurious cointegration may appear under temporal aggregation; and Pierse and Snell [1995] analyzed the incidence of aggregation on unit root tests.

e. Finally, a recent paper by Rossana and Seater [1995] provides a broad discussion, and a good deal of factual evidence, on the links between temporal aggregation, model structure and biased inference.

Regardless of the overwhelming evidence, testing the invariance of economic propositions to temporal aggregation is often neglected in applied work.

To overcome the above situation, two possible (and, to some extent similar) routes are available. Firstly, the analyst may either try to identify those features of the hypothesized original data generation process (DGP) which are invariant to aggregation or, alternatively, derive the theoretical DGP for the aggregate process given a DGP for the high frequency data and the particular temporal aggregation scheme. Secondly, a method that yields data at the appropriate level of temporal aggregation (as implied by the set of economic propositions to be tested) may be devised. The former approach ties economic propositions to the data by moving from the original DGP, belonging to some model class, to its aggregate at the frequency of reported data. For a certain class of models, one example being vector ARIMA models, this seems to be a natural way to deal with aggregated data. The alternative approach, as noted before, is to construct the higher frequency data adopting a procedure that links the original and aggregate data DGPs, and directly use such data for analysis. This paper follows the latter route, but restricts the presentation to the derivation of high frequency data and do not consider the (adequate) use of such data constructs in empirical models.

Another problem associated with temporal aggregation relates to the *ad hoc* use of partial disaggregate information related to the series of interest and available to the researcher. For example, to produce short-term

[1] Griliches [1994, section V] gives a comprehensive view of the present situation on data availability for applied economic research. The discussion is nevertheless quite general, and the consequences of temporal aggregation are considered only implicitly.

forecasts of economic activity, City commentators and (to some extent) academia watch the monthly movements of retail sales or trade or manufacturing output, believed to be the best available indicators for broader measures of demand and output.[2] We argue in Salazar *et al.* [1997] that those procedures, at best, may be not using information efficiently and, at worse, may be positively misleading. A similar argument was provided by Liu and Hwa [1974, pp. 328] to justify the construction of their monthly econometric model.

Having a formal method to approximate the higher frequency components of reported economic time series, therefore, seems relevant. An estimate of the reliability of the data thus generated may be important since any estimation using interpolated data should take account of measurement errors arising from interpolation.

In this paper we present a thorough Monte-Carlo analysis of a model-based method of interpolation. We first discuss existing econometric approaches to interpolation, and show how these extend naturally to a situation in which there a dynamic relationship between the interpoland and indicator variables. We derive theoretically both interpolands and their variance-covariance matrix. We then present a Monte-Carlo study which suggests that, in samples of reasonable length, our parameter estimates and interpolands are close to their true values. Estimates of their variance prove to be a useful guide to the experimental variance of both the parameters of the model and the interpolands. The method of interpolation suggested here appears a viable means of dealing with the problem.

2 Existing approaches to interpolation

2.1 Brief background

Economists and statisticians showed considerable interest in the estimation of unobserved data[3] during the early 1960s. The developments in the methodology may be classified, following Nijman [1985], as either *data-based* or *model-based*.[4] As an example of a data-based approach, we may wish to estimate an unobserved random vector[5]

$$\mathbf{y}_h = (y_1, y_2, ..., y_{nk})'$$

on the basis of knowing some linear aggregates

$$y_t = \sum_{u=1}^{k} c_u y_{k(t-1)+u}$$

or $y_l = \mathbf{C}\mathbf{y}_h$, where $\mathbf{C} = \mathbf{I} \otimes \mathbf{c}'$ and $\mathbf{c}$ a vector of dimension $k \times 1$. If the aggregate is a flow variable, then $\mathbf{c}$ is a unit vector; if it is a rate of flow or an index number, then $\mathbf{c}$ is the unit vector divided by k. If the aggregate is a stock, then the last element of $\mathbf{c}$ is 1 and the other elements are zero.

In the above notation, n indexes low-frequency periods (say, quarters) and k indexes sub-divisions within each low-frequency period (say, months within each quarter). As discussed by Peña and Guerrero [1994], one can assume $\mathbf{y}_h$ admits an ARIMA representation,

$$\mathbf{y}_h - E[\mathbf{y}_h|\mathbf{y}_{0h}] = \mathbf{V}\mathbf{e}_h$$

[2] A typical example in the United Kingdom are the numerous series that analysts (and journalists) employ as a measure of the so-called *feel good* factor, itself a proxy of the level of economic activity.

[3] Our preference is for the term *unobserved* data rather than *missing* data, as mostly adopted in the literature. The latter defines a rather general case of incomplete information either in survey data or due to defective sampling in time series.

[4] Data-based methods in general rely on mathematical interpolation to determine the unobserved series, or optimisation relative to some loss function. The interpoland may be subsequently modulated by information contained in related series. Model-based techniques, on the contrary, make explicit use of conditional expectations.

[5] Where the h subscript indicates, in rather obvious notation, that the vector is of high frequency, and implicitly assume the highest *sampling* frequency equals the highest DGP frequency.

where $\mathbf{e}_h \sim \mathrm{WN}(0,\sigma^2)$, $\mathbf{V}$ is a lower-triangular matrix containing the moving average weights, and in addition $E[\mathbf{e}_h|\mathbf{y}_{0h}] = 0$, $E[\mathbf{e}_h\mathbf{e}_h'|\mathbf{y}_{0h}] = \sigma^2\mathbf{I}$. The vector $\mathbf{y}_{0h} = (...,y_{-1},y_0)'$ denotes the infinite past of $\mathbf{y}_h$, which in practice is unknown. Early methods to augment $\mathbf{y}_l$ assumed that $\mathrm{E}[\mathbf{y}_h|\mathbf{y}_{0h}]$ followed some simple structure, say $\mathrm{E}[\mathbf{y}_h|\mathbf{y}_{0h}] = 0$, and that $\mathbf{V}$ could be derived considering $\mathbf{y}_h$ followed a first or (at the most) second order model; examples of this approach are the papers by Lisman and Sandee [1964], Glesjer [1966], Boot, Feibes and Lisman [1967] and Cohen, Muller and Padberg [1971]. Or the theoretical relationship that links $\mathbf{y}_h$ to $\mathbf{y}_l$ could be explicitly used, as done by Stram and Wei [1986] and Wei and Stram [1990].[6]

Alternatively, the researcher may have data on $\mathbf{y}_h$ available either at the beginning or the end of the series. That assumption is adopted by Harvey and Pierse [1984]; Harvey [1989] and Gomez and Maravall [1994] provide some useful extensions. They do so by recasting the problem in state-space form, using the Kalman filter and the fixed point smoother to estimate the model parameters and the interpoland, respectively. This methodology draws much from the paper by Jones [1980] on estimation of ARMA models with missing observations. A drawback of the Kalman filter implementation is the requirement of observations on $\mathbf{y}_h$ to be available at some point in the interpolation period, when normally it would be an unobserved vector.

In a model-based approach, we substitute the vector $\mathbf{y}_{0h}$ by a linear combination of variables which contain information that may help to explain $\mathbf{y}_h$; if we denote as $\mathbf{x}_h$ the vector (or also the matrix) containing such indicator variables, we now interpret $\tilde{\mathbf{y}}_h = \mathrm{E}[\mathbf{y}_h|\mathbf{x}_h]$ as a *preliminary* series. The basic idea behind this method was presented by Friedman [1962]; however, Chow and Lin [1971, 1976] grounded the theory (and estimation method) used in much of the subsequent applied work.[7] Parallel developments are found in the papers by Denton [1971] and Ginsburgh [1973]. Later methods using model-based techniques were, for the most part, refinements of Denton's and Chow and Lin's contributions.

2.2 A regression-based approach

Since our procedure follows the model-based tradition, a natural starting point is Chow and Lin's approach. Their method can be formalized as follows. Suppose a vector series $\mathbf{y}$ is observed at $t \in Z$ regularly spaced periods, but higher frequency measurements are needed. Information on several, or one related series to $\mathbf{y}$ is available, with periodicity $tu \in Z$. Denote by $\mathbf{x}$ a set of (strictly exogenous) related variables to $\mathbf{y}$, observed in sub-period $u = 1,...,k$ of period $t = 1,...,n$. Therefore, we can define a $nk \times p$ matrix $\mathbf{X}_h$ with typical column $\mathbf{x}_{i,h} = (\mathbf{x}_{i,11}, \mathbf{x}_{i,12},...,\mathbf{x}_{i,1q}, \mathbf{x}_{i,21},...,\mathbf{x}_{i,nk})'$, for $i = 1,..,p$ and the subscript h identifies a series as being of 'high frequency' (with the subscript l indicating 'low frequency' where appropriate). Assume an hypothetical $\mathbf{y}_h$ vector and the related variables in $\mathbf{X}_h$ can be linked by a linear model

$$\mathbf{y}_h = \mathbf{X}_h\beta + \mathbf{v}_h \tag{1}$$

where β is the $p \times 1$ vector of parameters, $\mathbf{v}_h$ is a $nk \times 1$ stationary vector of error terms, such that $\mathrm{E}(\mathbf{v}_h) = 0$ with covariance $\mathrm{E}(\mathbf{v}_h\mathbf{v}_h') = \mathbf{S}_h$. Assume, in addition, that a time-invariant constraint links the vectors $\mathbf{y}_h$ and $\mathbf{y}_l$

$$y_t = \textstyle\sum_{u=1}^{k} c_u y_{t,u} \quad \text{or} \quad \mathbf{y}_l = \mathbf{C}\mathbf{y}_h \tag{2}$$

(requiring monthly estimates to 'add-up' to quarterly figures, for example) where the c_u weights are defined in section 2.1. We can therefore write

[6] Stram and Wei, however, considered a model for the stationary series $(1-\mathbf{L})^d\mathbf{y}_h$ and $(1-\mathbf{B})^d\mathbf{y}_l$, where $\mathbf{L}$ and $\mathbf{B}$ are the lag operators on the high and low frequency data, respectively.

[7] As Peña and Guerrero [1994] correctly note, Chow and Lin did not condition on $\mathrm{E}[\mathbf{y}_h|\mathbf{x}_h]$ but focused on the simultaneous estimation of $\mathbf{y}_h$ and the regression model linking $\mathbf{x}_h$ to $\mathbf{y}_h$, therefore ignoring the autocorrelation structure in the errors.

$$\mathbf{C}\mathbf{y}_h = \mathbf{C}\mathbf{X}_h\beta + \mathbf{C}\mathbf{v}_h \quad \text{or} \quad \mathbf{y}_l = \mathbf{X}_l\beta + \mathbf{e}_l \tag{3}$$

with $\mathbf{e}_l = \mathbf{C}\mathbf{v}_h$ and covariance matrix $\mathbf{G}_l = \mathbf{C}\mathbf{S}_h\mathbf{C}' = \sigma^2\mathbf{V}(\rho)$. Chow and Lin [1973] assume that $\mathbf{V}(\rho)$ is a known function of ρ, a known parameter. Note in equation [3] that $\mathbf{y}_l$ is the observed series, and $\mathbf{X}_l$ contains the observed values of the related series at the same frequency as $\mathbf{y}_l$. Taking into account the constraints [2], β and $\mathbf{y}_h$ are estimated as

$$\widehat{\beta} = (\mathbf{X}_l'\mathbf{G}^{-1}\mathbf{X}_l)^{-1}\mathbf{X}_l'\mathbf{G}^{-1}\mathbf{y}_l \tag{4}$$

and

$$\widehat{\mathbf{y}}_h = \mathbf{X}_h\widehat{\beta} + (\mathbf{S}_h\mathbf{C}'\mathbf{G}_l^{-1})\widehat{\mathbf{e}}_l \tag{5}$$

with $\widehat{\mathbf{e}}_l = \mathbf{y}_l - \mathbf{X}_l\widehat{\beta}$. The whole procedure, however, depends on the specification of error process $\mathbf{v}_h$ (which defines ρ) in equation [1]. Chow and Lin [1973] assumed $\mathbf{v}_h$ followed an AR(1) process with no initial conditions; Fernandez [1981] discussed the case of $\mathbf{v}_h$ being a random walk with $\mathbf{v}_0 = 0$; Litterman [1983] considered an AR(2) process, $\mathbf{v}_h = \phi_1\mathbf{v}_{h,-1} + \phi_2\mathbf{v}_{h,-2} + \varepsilon_h$ imposing the restriction $\phi_2 = 1 - \phi_1$, $\mathbf{v}_0 = 0$ and $\mathbf{v}_1 = \varepsilon_1$. The researcher has a good margin of choice in terms of the error specification for $\mathbf{v}_h$ and, provided the parameters are admissible, the estimation of $\mathbf{G}_l$ and thus $\widehat{\mathbf{y}}_h$ is straightforward.

Three problems are clearly identifiable in Chow and Lin's procedure. First, the choice of the error process $\mathbf{v}_h$ is heuristic; a search over possible processes and parameters is, clearly, impracticable. Secondly, equation [1] does not accommodate dynamics, except as implied by the specification of $\mathbf{V}$. Current econometric practice suggests, however, that dynamic structures may be very important. In the following section we show a relatively simple way to generalise the model [1], assuming the error process $\mathbf{v}_h$ is white noise and where lagged dependent variables may be used; the inclusion of dynamics, in addition, allows us to take account of (possible) cointegrating relationships linking the low and high-frequency data. Thirdly, their method relies on [1] being a regression in the *levels* of the variables. For example, the practice of using logarithmic transformations to deal with heteroskedasticity is almost universal. We set out below a general method which deals with these problems.

3 A dynamic generalisation of Chow and Lin's method

In this section we present a dynamic generalisation of equation [1]. Working in scalar terms for ease of exposition, we rewrite equation [1] as

$$f(y_{t,u}) = \sum_{j=1}^{m}\alpha_j L^j f(y_{t,u}) + \sum_{i=1}^{p}\beta_i x_{i,tu} + v_{t,u} \tag{6}$$

where the error terms v_{tu} are white noise and L is the lag operator on the high-frequency data, $Ly_{tu} = y_{t,u-1}$. We replace h by tu and l by t where appropriate to indicate specific elements of the relevant vectors. In our notation t enumerates the low-frequency periods and u the high-frequency subdivisions within a particular period. The use of $f(y_{t,u})$ indicates that we may consider non-linear transformations of $y_{t,u}$. We have not indicated any transformations or lags of the x variables since these will normally be specified *a priori* and do not affect the estimation of the model. Note also equation [6] is the replica of [1], appended to include lagged values of the endogenous vector $y_{t,u}$ We can calculate the roots ρ_j of the auxiliary equation $1 - \sum_{j=1}^{m}\alpha_j z^j = 0$, and write [6] as

$$\prod_{j=1}^{m}(1-\rho_j L)f(y_{t,u}) = \sum_{i=1}^{p}\beta_i x_{i,tu} + v_{t,u} \tag{7}$$

As suggested by Oguchi and Fukuchi [1990], we premultiply [7] by $\prod_j(1+\sum_{l=1}^{k}\rho_j^l L^l)$ to obtain

$$\prod_j(1-\rho_j^k L^k)f(y_{t,u}) = \prod_j(1+\sum_{l=1}^{k-1}\rho_j^l L^l)(\sum_i \beta_i x_{i,tu}+v_{t,u}) \tag{8}$$

Denoting $1-\sum_j \theta_j L^{kj}=0$ as the polynomial whose roots are ρ_j^k, we can rewrite equation [8] as

$$f(y_{t,u}) = \sum_j \theta_j L^{kj} f(y_{t,u}) + \prod_j(1+\sum_{l=1}^{k-1}\rho_j^l L^l)(\sum_i \beta_i x_{i,tu}+v_{t,u}) \tag{9}$$

We can now premultiply equation [9] by the aggregator $C=(1+\sum_{l=1}^{k-1}L^l)=\sum_t$ to estimate the model using the available data. This is in fact a regression equation for low-frequency moving averages. Note that only every one in k of the elements $(1+\sum_{l=1}^{k-1}L^l)f(y_{t,u})$ corresponds to an observed low-frequency period. Taking care to avoid overlap of low-frequency periods in the summation, it follows that

$$f(y_{t,u}) = \sum_j \theta_j B^j f(y_{t,u}) + \sum_t\left[\prod_j(1+\sum_{l=1}^{k-1}\rho_j^l L^l)(\sum_i \beta_i x_{i,tu}+v_{t,u})\right] \tag{10}$$

where $B[f(y_{t,u})]=f(y_{t-1,u})$ is the lag operator on the low-frequency data.

The roots of $\prod_j(1+\sum_{l=1}^{k-1}\rho_j^l L^l)$ can be calculated, to transform this expression into the more conventional structure

$$\prod_j(1+\sum_{l=1}^{k-1}\rho_j^l L^l) = (1+\sum_j \gamma_j L^j) \tag{11}$$

and the equation which we estimate is, then,

$$\sum_{u=1}^{k} f(y_{t,u}) = \sum_j \theta_j B^j \sum_{u=1}^{k} f(y_{tu}) + \sum_{u=1}^{k}\left\{\sum_i \beta_i\left[\sum_t(1+\sum_k \gamma_k L^k)x_{i,tu}\right] + \sum_t(1+\sum_j \gamma_j L^j)v_{t,u}\right\} \tag{12}$$

Therefore, we have to replace each $x_{t,u}$ in the high-frequency equation by $\sum_t(1+\sum\gamma_j L^j)x_{t,u}$ in the quarterly equation. The regression has a moving-average error process (MA(2) in the above example), but since the $\gamma's$ are functions of the $\theta's$ (which, in turn, are derived from the $\alpha's$) the error structure depends on the regression coefficients of the lagged values of y_t. To derive the equations if $y_{t,u}$ is expressed in first differences, we simply set $\alpha_1=1$ and $\gamma(L)=-L^k$.

If equation [9] were expressed in terms of y_{tu} rather than $f(y_{tu})$, it would now be defined in the low-frequency values of the endogenous variable. To obtain an operational formulation of [12], we exploit the mean value theorem and express $f(y_{t,u})=f(\bar{y}_t)+f'(y^*_{t,u})(y_{t,u}-\bar{y}_t)$, where $\bar{y}_t=y_t/3$ is the monthly average in quarter t and $y^*_{t,u}$ lies between $y_{t,u}$ and $\bar{y}_t$. If the error of approximation $y_{t,u}-\bar{y}_t$ is relatively small, we have

$$\sum_{u=1}^{3} f(y_{t,u}) \doteq 3f(\bar{y}_t) \tag{13}$$

note that the errors of approximation sum to zero, $\sum_{u=1}^{3}(y_{t,u}-\bar{y}_t)=0$. For a logarithmic transformation, the approximation becomes

$$\sum_{u=1}^{3}\ln y_{t,u} \doteq 3\ln y_t - 3\ln 3 \tag{14}$$

which can be seen to be equivalent to replacing the quarterly value $\sum_{u=1}^{3}\ln y_{t,u}$ by three times the geometric mean of the monthly values $y_{t,u}$, $u=1,2,3$. The geometric mean is never larger than the arithmetic mean, but, if monthly movements are small compared with the monthly average, the approximation error introduced should be of little importance.

4 Estimation by Generalized Least Squares

The model to be estimated is therefore given by equation [12], which is the equation (in terms of the data) we can actually observe. This could be done by iterating over α, and thus θ and γ to find a GLS solution, or by maximisation of the associated likelihood of [12]. We discuss in this section the GLS procedure, and in the next section the maximum likelihood route. In both cases we condition on $\{\mathbf{v}_0, \mathbf{v}_{-1}=0\}$.

To start the iterations, we set $\alpha=0$ and therefore impose an identity variance matrix $\mathbf{V}$; see equation [4]. The first iteration yields an estimate of θ from which α can be recovered directly. This is used in the next iteration to construct the covariance matrix of the residuals (which, as we have seen, depends on α and, therefore, on θ). We then proceed in a manner similar to Chow and Lin, minimising the sum of squared residuals weighted by the inverse of the covariance matrix, iterating only over α, given the $\beta's$ are linear in α. The remaining parameters in the model are estimated simultaneously estimated with α; see equation [4]. The process continues until convergence in α.

The covariance matrix found at that stage is consistent with the lag structure estimated by generalised least squares. But the GLS estimation of the parameters in the model, however, ignores the uncertainty in the covariance matrix, arising from the estimation of α. This problem is conveniently addressed by estimating [12] using maximum likelihood, which we consider below.[8]

5 Maximum likelihood solution

To estimate equation [12] by maximum likelihood, we should consider[9]

$$\mathcal{L}(\mathbf{f}_l;\mathbf{x}_l,\alpha,\beta) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\log|\mathbf{G}(\alpha)| - \frac{1}{2}\mathbf{e}_l'\mathbf{G}(\alpha)^{-1}\mathbf{e}_l \tag{15}$$

where $\mathbf{e}_l=\mathbf{C}\mathbf{S}_h\mathbf{v}_h+\mathbf{v}_0$ for $\mathbf{v}_0$ initial conditions, $\mathbf{f}_l=f(\mathbf{y}_l)$, $\mathbf{E}(\mathbf{e}_l\mathbf{e}_l')=\mathbf{G}=\sigma^2\mathbf{V}(\alpha)$, $\mathbf{V}(\alpha)=\mathbf{C}\mathbf{S}_h\mathbf{S}_h'\mathbf{C}'$ with σ^2 unknown. The matrix $\mathbf{S}_h$ has a band-diagonal structure, and is completely defined by the expression $1+\sum\gamma_j L^j$ which picks the moving average process induced by aggregating [6]. The γ_j are themselves functions of the lag parameters in the underlying equation, α_j (see section 3). The matrix $\mathbf{V}(\alpha)$ is positive definite by construction and assumed twice differentiable in α.

Simple algebra enables us to rewrite the log-likelihood [15] as

$$\mathcal{L}(\mathbf{f}_l;\mathbf{x}_l,\alpha,\beta) = C + \log|\mathbf{W}(\alpha)| - \frac{T}{2}\log\mathbf{e}_l'\mathbf{V}(\alpha)^{-1}\mathbf{e}_l \tag{16}$$

using the decomposition $\mathbf{G}^{-1}=\sigma^{-2}\mathbf{W}'\mathbf{W}$ (dependence on α will be left implicit, when it leads to simpler formulae); $\mathbf{W}$ can be obtained either from a Cholesky factorisation of $\mathbf{V}^{-1}$ or from the eigenvalues and associated eigenvectors of $\mathbf{V}$.[10] One difficulty with the log-likelihood [16] is the dependence of $\mathbf{V}$, and hence $\mathbf{W}$ on α, which makes the first and second-order conditions somewhat expensive to solve. In the case where the lag is only first order, $m=1$, the first and second-order conditions can be solved analytically (see the Appendix

[8]Even though GLS and ML are asymptotically equivalent, it is desirable to evaluate the ML solution to to find the asymptotic standard errors of the parameter estimates from the information matrix.

[9]The sample likelihood for the model [33] is simply

$$\mathcal{L}(\mathbf{f}_l;\mathbf{x}_l,\lambda,\beta) = (2\pi)^{-\frac{T}{2}}|\mathbf{G}(\lambda)|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\mathbf{e}_l'\mathbf{G}(\lambda)^{-1}\mathbf{e}_l\right]$$

from which the log-likelihood follows.

[10]The choice of method depends, clearly, on the costs of inverting $\mathbf{V}$. Using the eigenvalues $\mathbf{P}$ and eigenvectors $\mathbf{Q}$ of $\mathbf{V}$ implies that $\mathbf{W}=\mathbf{P}\mathbf{Q}^{-0.5}\mathbf{P}'$ and therefore $\mathbf{W}'\mathbf{W}=\mathbf{V}^{-1}$. The Cholesky factorisation applies to $\mathbf{V}^{-1}$ rather than $\mathbf{V}$, but yields an upper triangular $\mathbf{W}$ matrix so $|\mathbf{W}|$ in [16] simplifies to $\Pi_{i=j}h_{ij}$, the product of the elements along the main diagonal of $\mathbf{W}$.

for details). We opted, however, for a direct-search method based on the algorithm of Dennis and Torczon [1991]; see also Higham [1993] for a discussion[11]. The GLS solution provided a starting point for the maximum-likelihood estimation. A somewhat similar procedure was proposed by Magnus [1978, pp. 289], to maximise a regression model by maximum likelihood when the disturbance covariance matrix depends on an unknown set of parameters.[12] Once the algorithm converges, we can, if $m=1$ evaluate the Hessian matrix, and the associated standard errors, using their analytic derivation. This provides a useful check in our subsequent simulations.

Clearly from equation [33], the $\beta's$ are linear in the components of α. The implication is that we can restrict the parameter search to the α only, resulting in a significant gain in computation speed. Estimates of β are obtained by application of generalised least-squares to [33] once the α is concentrated out.

If we only consider a single lag in the interpoland, by differentiating the log-likelihood [16] with respect to the vector β for a fixed α we obtain the familiar result

$$\widehat{\beta}_k = [\mathbf{X}_l'\mathbf{V}(\widehat{\alpha}_r)^{-1}\mathbf{X}_l]^{-1}\mathbf{X}_l'\mathbf{V}(\widehat{\alpha}_r)^{-1}[\mathbf{I}-\widehat{\alpha}_r\mathbf{L}]\mathbf{y}_l \tag{17}$$

for an estimator $\widehat{\alpha}_k$ of α in iteration r. On each step of the optimisation algorithm, given $\widehat{\alpha}_k$ the matrices $\mathbf{S}_h$ and $\mathbf{V}$ are calculated, and by application of [17] we recover the remaining parameters. Once the algorithm converges, a matrix $\mathbf{V}$ has been found which is consistent with the lag structure estimated by MLE, so the $\widehat{\beta}_{\max(\mathcal{L})}$ vector estimated by [17] is consistent with $\widehat{\alpha}_{\max(\mathcal{L})}$.

6 Reconciliation of the interpolands

The estimators of the parameters of the monthly regression equation [6] may then be used to produce fitted values of the interpolands $y_{t,u}$. These fitted values, however, need to be reconciled with the observed quarterly data y_t. Our estimate of $y_{t,u}$ minimises the sums of squares of the residuals in the regression equation [6] subject to the constraint that the interpolated monthly values in each quarter sum to the known quarterly totals, that is, $\sum_{u=1}^{3} y_{t,u} = y_t$.

For simplicity, we confine attention to the first order case, and thus set $m=1$. We assume that observations are available on the quarterly totals y_t for quarters $t=1,...,T$. Therefore, we can rewrite [6] as

$$f(y_{t,u}) = \alpha_1 f(y_{t,u-1}) + \beta_0 + \sum_{i=1}^{p} \beta_i(L)x_{i,tu} + v_{t,u} \tag{18}$$

where, for the first quarter, $t=1$, $u=2,3$ and, for the remainder, $u=1,2,3$, $t=2,...,T$. A set of consistent θ and β estimates (at least asymptotically) is obtained from the quarterly regression [12]. The problem then reduces to optimising the Lagrangian

$$\sum_{u=2}^{3} v_{1,u}^2 + \sum_{t=2}^{T}\sum_{u=1}^{3} v_{t,u}^2 + \sum_{t=1}^{T} \lambda_t \left(\sum_{u=1}^{3} y_{t,u} - y_t\right), \tag{19}$$

where λ_t is the Lagrange multiplier associated with the constraint $\sum_{u=1}^{3} y_{t,u} = y_t$, $t=1,...,T$. The first-order conditions are given by

$$\nabla f(y_{t,u})(v_{t,u} - \alpha_1 v_{t,u+1}) + \lambda_t = 0, \tag{20}$$

[11] The choice of method was driven by its ease of implementation and good performance, in terms of numerical stability and convergence.

[12] Recall that derivative-free methods have, in general, slower convergence rates than algorithms where gradient information is used. Magnus [1978, *op. cit*, footnote 5] suggests that the inversion of the Hessian matrix implied by log-likelihoods like [16] may be computationally expensive, to the point where the speed benefits from gradient-based algorithms are lost. This proved to be also our case. An iterative approach is also suggested in Oberhofer and Kmenta [1974].

where $v_{1,1}=0$, $v_{T+1,1}=0$, $u=1,2,3$, $t=1,...,T$, and ∇ is the derivative operator.

Equation (20) can be solved jointly with the adding-up constraints, $\sum_{u=1}^{3} y_{t,u} = y_t$, $t=1,...,T$, to produce estimates of the interpolands $\hat{y}_{t,u}$, $u=1,2,3$, $t=1,...,T$, and the Lagrange multipliers, $\hat{\lambda}_t$, $t=1,...,T$. The solution is inherently nonlinear because the derivatives $\nabla f(.)$ in (20) are a function of the estimated interpolated data $\hat{y}_{t,u}$, $u=1,2,3$, $t=1,...,T$, which, in principle, necessitates the use of iterative methods. However, when the transformation $f(.)$ is logarithmic, our experience indicates that the derivatives $\nabla f(.)$ in [20] may be satisfactorily evaluated at the monthly average $\bar{y}_t$ of the corresponding quarterly total y_t, hence avoiding further iteration.

6.1 Solution method

To solve [20], define $n=3T$. It is convenient to stack (18) in vector and matrix form as

$$\mathbf{A}_1\mathbf{f}_1(\mathbf{y}_1^h) + \mathbf{A}_2\mathbf{f}_2(\mathbf{y}_2^h) = \mathbf{X}\boldsymbol{\beta} + \mathbf{v}$$

where $\mathbf{A}_1$ is a $(n-1,3)$ matrix with elements $a_{1,ij} = -\alpha_1$, $i=j$, $a_{1,ij}=1$, $j=i+1$ and $a_{1,ij}=0$ otherwise, $\mathbf{A}_2$ a $(n-1,n-3)$ matrix with $a_{2,ij} = -\alpha_1$, $i=j+3$, $a_{2,ij}=1$, $i=j+2$, and $a_{2,ij}=0$ otherwise, $\mathbf{f}_1(\mathbf{y}_1^h) = (f(y_{1,1}),...,f(y_{1,3}))'$, $\mathbf{f}_2(\mathbf{y}_2^h) = (f(y_{2,1}),...,f(y_{T,3}))'$, $\mathbf{y}_1^h = (y_{1,1},y_{1,2},y_{1,3})'$, $\mathbf{y}_2^h = (y_{2,1},...,y_{T,3})'$, $\mathbf{X}$ the $(n-1, 1+\sum_{j=1}^{k}(q_j+1))$ observation matrix on the constant term and the exogenous indicator variables $x_{t,u}^j$ with the $(1+\sum_{j=1}^{k}(q_j+1))$ parameter vector $\boldsymbol{\beta}$ defined appropriately and $\mathbf{v} = (v_{1,2},...,v_{T,3})'$. Note that both $\mathbf{A}_1$ and $\mathbf{A}_2$ are full column rank. Therefore, the first order conditions of [20] become

$$\mathbf{F}_1(\mathbf{y}_1^h)\mathbf{A}_1'\mathbf{v} + \mathbf{C}_1'\lambda_1 = \mathbf{0}, \mathbf{F}_2(\mathbf{y}_2^h)\mathbf{A}_2'\mathbf{v} + \mathbf{C}_2'\boldsymbol{\lambda}_2 = \mathbf{0} \tag{21}$$

where $\mathbf{F}_1(\mathbf{y}_1^h)$ and $\mathbf{F}_2(\mathbf{y}_2^h)$ are $(3,3)$ and $(n-3,n-3)$ diagonal matrices with elements $\nabla f(y_{1,u})$, $u=1,2,3$, and $\nabla f(y_{t,u})$, $u=1,2,3$, $t=2,...,T$, respectively, $\mathbf{C}_1 = (1,1,1)$, $\mathbf{C}_2 = (\mathbf{I}_{T-1}\otimes(1,1,1))$ are the respective aggregator matrices and $\boldsymbol{\lambda}_2 = (\lambda_2,...,\lambda_T)'$.

Solving for λ_1 and $\boldsymbol{\lambda}_2$ in (21) gives

$$\lambda_1 = -(\mathbf{C}_1\boldsymbol{\Omega}_{1,h}^{-1}\mathbf{C}_1')^{-1}\mathbf{C}_1\boldsymbol{\Omega}_{1,h}^{-1}\mathbf{F}_1(\mathbf{y}_1^h)\mathbf{A}_1'\mathbf{v}, \boldsymbol{\lambda}_2 = -(\mathbf{C}_2\boldsymbol{\Omega}_{2,h}^{-1}\mathbf{C}_2')^{-1}\mathbf{C}_2\boldsymbol{\Omega}_{2,h}^{-1}\mathbf{F}_2(\mathbf{y}_2^h)\mathbf{A}_2'\mathbf{v}$$

where $\boldsymbol{\Omega}_{i,h} = \mathbf{F}_i(\mathbf{y}_i^h)\mathbf{A}_i'\mathbf{A}_i\mathbf{F}_i(\mathbf{y}_i^h)$, $i=1,2$. Hence, the first order conditions [21] are rendered as

$$\left(\mathbf{I}_i - \mathbf{C}_i'(\mathbf{C}_i\boldsymbol{\Omega}_{i,h}^{-1}\mathbf{C}_i')^{-1}\mathbf{C}_i\boldsymbol{\Omega}_{i,h}^{-1}\right)\mathbf{F}_i(\mathbf{y}_i^h)\mathbf{A}_i'\epsilon\mathbf{v} = \mathbf{0}, i=1,2 \tag{22}$$

where $\mathbf{I}_1$ and $\mathbf{I}_2$ are $(3,3)$ and $(n-3,n-3)$ identity matrices respectively.

As noted before, [22] is nonlinear and must be solved by iterative methods. An appropriate value to initiate the search for the solution $y_{t,u}$ to [22] is given by $y_{t,u}^0 = \bar{y}_t$ the monthly average for quarter t, $u=1,2,3$, $t=1,...,T$. Let $\mathbf{y}_{1,0}^h$ and $\mathbf{y}_{2,0}^h$ denote the respective vectors of monthly averages of the quarterly data for quarters $t=1$ and $t=2,...,T$. First-order Taylor series expansions of $\mathbf{f}_1(\mathbf{y}_1^h)$ and $\mathbf{f}_2(\mathbf{y}_2^h)$ about $\mathbf{y}_{1,0}^h$ and $\mathbf{y}_{2,0}^h$ are given by

$$\mathbf{f}_i(\mathbf{y}_i^h) \doteq \mathbf{f}(\mathbf{y}_{i,0}^h) + \mathbf{F}_i(\mathbf{y}_{i,0}^h)(\mathbf{y}_i^h - \mathbf{y}_{i,0}^h), i=1,2 \tag{23}$$

Initialising the former of the revised first-order conditions [22] at $\mathbf{y}_{1,0}^h$ and $\mathbf{y}_{2,0}^h$ yields

$$\left(\mathbf{I}_1 - \mathbf{C}_1'(\mathbf{C}_1\boldsymbol{\Omega}_{1,h}^{0-1}\mathbf{C}_1')^{-1}\mathbf{C}_1\boldsymbol{\Omega}_{1,h}^{0-1}\right)\mathbf{F}_1(\mathbf{y}_{1,0}^h)\mathbf{A}_1'$$
$$\times\left(\mathbf{A}_1\left(\mathbf{f}_1(\mathbf{y}_{1,0}^h) + \mathbf{F}_1(\mathbf{y}_{1,0}^h)(\mathbf{y}_1^h - \mathbf{y}_{1,0}^h)\right) + \mathbf{A}_2\mathbf{f}_2(\mathbf{y}_{2,0}^h) - \mathbf{X}\boldsymbol{\beta}\right) \doteq \mathbf{0}, \tag{24}$$

where $\mathbf{\Omega}^0_{1,h} = \mathbf{F}_1(\mathbf{y}^h_{1,0})\mathbf{A}_1'\mathbf{A}_1\mathbf{F}_1(\mathbf{y}^h_{1,0})$. Hence, from [24], defining $\mathbf{v}_0 = \mathbf{A}_1\mathbf{f}_1(\mathbf{y}^h_{1,0}) + \mathbf{A}_2\mathbf{f}_2(\mathbf{y}^h_{2,0}) - \mathbf{X}\boldsymbol{\beta}$, the solution for $\mathbf{y}^h_1$ is

$$\mathbf{y}^h_1 - \mathbf{y}^h_{1,0} \doteq \left(\mathbf{\Omega}^{0-1}_{1,h} - \mathbf{\Omega}^{0-1}_{1,h}\mathbf{C}_1'(\mathbf{C}_1\mathbf{\Omega}^{0-1}_{1,h}\mathbf{C}_1')^{-1}\mathbf{C}_1\mathbf{\Omega}^{0-1}_{1,h}\right)\mathbf{F}_1(\mathbf{y}^h_{1,0})\mathbf{A}_1'\mathbf{v}_0. \tag{25}$$

A similar initialisation of the second of the revised first order conditions [22] yields the solution for $\mathbf{y}^h_2$ as

$$\mathbf{y}^h_2 - \mathbf{y}^h_{2,0} \doteq \left(\mathbf{\Omega}^{0-1}_{2,h} - \mathbf{\Omega}^{0-1}_{2,h}\mathbf{C}_2'(\mathbf{C}_2\mathbf{\Omega}^{0-1}_{2,h}\mathbf{C}_2')^{-1}\mathbf{C}_2\mathbf{\Omega}^{0-1}_{2,h}\right)\mathbf{F}_2(\mathbf{y}^h_{2,0})\mathbf{A}_2'\mathbf{v}_0, \tag{26}$$

where $\mathbf{\Omega}^0_{2,h} = \mathbf{F}_2(\mathbf{y}^h_{2,0})\mathbf{A}_2'\mathbf{A}_2\mathbf{F}_2(\mathbf{y}^h_{2,0})$. Equations [25] and [26] may then be re-initialised using the solutions $\mathbf{y}^h_1$ and $\mathbf{y}^h$ and iterated until convergence; that is, until the right hand sides of [25] and [26] are less than the desired degree of accuracy.

An alternative iterative scheme based on [22] may be obtained by substitution of both approximations for $\mathbf{f}_i(\mathbf{y}^h_i)$, $i = 1, 2$, given in [23] into $\mathbf{v}$. Defining $\mathbf{P}_i = \mathbf{I}_i - \mathbf{C}_i'(\mathbf{C}_i\mathbf{\Omega}^{-1}_{i,h}\mathbf{C}_i')^{-1}\mathbf{C}_i\mathbf{\Omega}^{-1}_{i,h}$, $i = 1, 2$, $\mathbf{P}^0_i$ as $\mathbf{P}_i$ evaluated at $\mathbf{\Omega}^0_{i,h}$, $i = 1, 2$, and using the adding-up conditions $\mathbf{C}_i\mathbf{y}^h_i = \mathbf{y}^l_i$, $i = 1, 2$ yields

$$\begin{pmatrix} \mathbf{\Omega}^0_{1,h} & \mathbf{P}^0_1\mathbf{F}_1(\mathbf{y}^h_{1,0})\mathbf{A}_1'\mathbf{A}_2\mathbf{F}_2(\mathbf{y}^h_{2,0}) \\ \mathbf{P}^0_2\mathbf{F}_2(\mathbf{y}^h_{2,0})\mathbf{A}_2'\mathbf{A}_1\mathbf{F}_1(\mathbf{y}^h_{1,0}) & \mathbf{\Omega}^0_{2,h} \end{pmatrix}\begin{pmatrix} \mathbf{y}^h_1 - \mathbf{y}^h_{1,0} \\ \mathbf{y}^h_2 - \mathbf{y}^h_{2,0} \end{pmatrix} = \begin{pmatrix} \mathbf{P}^0_1\mathbf{F}_1(\mathbf{y}^h_{1,0})\mathbf{A}_1' \\ \mathbf{P}^0_2\mathbf{F}_2(\mathbf{y}^h_{2,0})\mathbf{A}_2' \end{pmatrix}\mathbf{v}_0, \tag{27}$$

which may be iterated in a similar manner.

At the same time as interpolating the data, we are able to produce estimates of approximate expressions for the variances and covariances of the estimated interpolands $\hat{y}_{t,u}$. Including only terms of order $O_P(1)$, the source of error due to the estimation of the regression parameters is asymptotically irrelevant. Hence, only the random component represented by the error terms $v_{t,u}$ is pertinent. Details of the requisite calculations are provided in the following section.

7 The Variance of the interpolands

It is clear from [27] that the estimators $\hat{\mathbf{y}}^h_i$ for the interpolands $\mathbf{y}^h_i$, $i = 1, 2$, satisfy the orthogonality conditions

$$\begin{pmatrix} \hat{\mathbf{\Omega}}_{1,h} & \hat{\mathbf{P}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1)\hat{\mathbf{A}}_1'\hat{\mathbf{A}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2) \\ \hat{\mathbf{P}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2)\hat{\mathbf{A}}_2'\hat{\mathbf{A}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1) & \hat{\mathbf{\Omega}}_{2,h} \end{pmatrix}^{-1}\begin{pmatrix} \hat{\mathbf{P}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1)\hat{\mathbf{A}}_1' \\ \hat{\mathbf{P}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2)\hat{\mathbf{A}}_2' \end{pmatrix}\hat{\mathbf{v}} = \mathbf{0}, \tag{28}$$

where we now explicitly show the estimator $\hat{\mathbf{A}}_i$ for the matrix $\mathbf{A}_i$, $\hat{\mathbf{\Omega}}_{i,h} = \mathbf{F}_i(\hat{\mathbf{y}}^h_i)\hat{\mathbf{A}}_i'\hat{\mathbf{A}}_i\mathbf{F}_i(\hat{\mathbf{y}}^h_i)$, $\hat{\mathbf{P}}_i$ is $\mathbf{P}_i$ evaluated at $\hat{\mathbf{\Omega}}_{i,h}$, $i = 1, 2$, and $\hat{\mathbf{v}} = \hat{\mathbf{A}}\mathbf{f}(\hat{\mathbf{y}}^h) - \mathbf{X}\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{A}} = (\hat{\mathbf{A}}_1, \hat{\mathbf{A}}_2)$, $\mathbf{f}(\hat{\mathbf{y}}^h) = (\mathbf{f}_1(\hat{\mathbf{y}}^h_1)', \mathbf{f}_2(\hat{\mathbf{y}}^h_2)')'$, $\hat{\mathbf{y}}^h = (\hat{\mathbf{y}}^{h\prime}_1, \hat{\mathbf{y}}^{h\prime}_2)'$, with $\hat{\boldsymbol{\beta}}$ the estimator for $\boldsymbol{\beta}$.

In order to derive an expression for the variance of the interpolands, we make use of a number of approximations. Firstly,

$$\hat{\mathbf{v}} \doteq \hat{\mathbf{A}}\left(\mathbf{f}(\mathbf{y}^h) + \mathbf{F}(\hat{\mathbf{y}}^h)(\hat{\mathbf{y}}^h - \mathbf{y}^h)\right) - \mathbf{X}\hat{\boldsymbol{\beta}},$$

where $\mathbf{F}(\hat{\mathbf{y}}^h) = \mathrm{diag}(\mathbf{F}_1(\hat{\mathbf{y}}^h_1), \mathbf{F}_2(\hat{\mathbf{y}}^h_2))$, $\mathbf{f}(\mathbf{y}^h) = (\mathbf{f}_1(\mathbf{y}^h_1)', \mathbf{f}_2(\mathbf{y}^h_2)')'$ and $\mathbf{y}^h = (\mathbf{y}^{h\prime}_1, \mathbf{y}^{h\prime}_2)'$. Secondly, from the orthogonality conditions [28]

$$\hat{\mathbf{y}}^h - \mathbf{y}^h \doteq -\begin{pmatrix} \hat{\mathbf{\Omega}}_{1,h} & \hat{\mathbf{P}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1)\hat{\mathbf{A}}_1'\hat{\mathbf{A}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2) \\ \hat{\mathbf{P}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2)\hat{\mathbf{A}}_2'\hat{\mathbf{A}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1) & \hat{\mathbf{\Omega}}_{2,h} \end{pmatrix}^{-1}\hat{\mathbf{P}}\mathbf{F}(\hat{\mathbf{y}}^h)\hat{\mathbf{A}}'\tilde{\mathbf{v}}, \tag{29}$$

where $\hat{\mathbf{P}} = \mathrm{diag}(\hat{\mathbf{P}}_1, \hat{\mathbf{P}}_2)$ and

$$\tilde{\mathbf{v}} = \hat{\mathbf{A}}\mathbf{f}(\mathbf{y}^h) - \mathbf{X}\hat{\boldsymbol{\beta}}$$

$$= \mathbf{v} + O_P(n^{-1/2})$$

noting that $n^{1/2}(\hat{a}_{k,ij} - a_{k,ij}) = O_P(1)$, $k = 1, 2$, and $n^{1/2}(\hat{\beta} - \beta) = O_P(1)$. Therefore, from [29]

$$\hat{\mathbf{y}}^h - \mathbf{y}^h \doteq -\begin{pmatrix} \hat{\mathbf{\Omega}}_{1,h} & \hat{\mathbf{P}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1)\hat{\mathbf{A}}_1'\hat{\mathbf{A}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2) \\ \hat{\mathbf{P}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2)\hat{\mathbf{A}}_2'\hat{\mathbf{A}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1) & \hat{\mathbf{\Omega}}_{2,h} \end{pmatrix}^{-1}\hat{\mathbf{P}}\mathbf{F}(\hat{\mathbf{y}}^h)\hat{\mathbf{A}}'\mathbf{v}.$$

Consequently, an approximate expression for the variance of $\hat{\mathbf{y}}^h$ is given by

$$\sigma^2_v\mathbf{K}\mathbf{P}\mathbf{F}(\mathbf{y}^h)\mathbf{A}'\mathbf{A}\mathbf{F}(\mathbf{y}^h)\mathbf{P}'\mathbf{K}', \tag{30}$$

where $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$, $\mathbf{F}(\mathbf{y}^h) = \mathrm{diag}(\mathbf{F}_1(\mathbf{y}^h_1), \mathbf{F}_2(\mathbf{y}^h_2))$, $\mathbf{P} = \mathrm{diag}(\mathbf{P}_1, \mathbf{P}_2)$ and

$$\mathbf{K} = \begin{pmatrix} \mathbf{\Omega}_{1,h} & \mathbf{P}_1\mathbf{F}_1(\mathbf{y}^h_1)\mathbf{A}_1'\mathbf{A}_2\mathbf{F}_2(\mathbf{y}^h_2) \\ \mathbf{P}_2\mathbf{F}_2(\mathbf{y}^h_2)\mathbf{A}_2'\mathbf{A}_1\mathbf{F}_1(\mathbf{y}^h_1) & \mathbf{\Omega}_{2,h} \end{pmatrix}^{-1}.$$

Using the approximation $\mathbf{F}(\mathbf{y}^h) \doteq \mathbf{F}(\hat{\mathbf{y}}^h)$, an estimator for [30] is

$$\hat{\sigma}^2_v\hat{\mathbf{K}}\hat{\mathbf{P}}\mathbf{F}(\hat{\mathbf{y}}^h)\hat{\mathbf{A}}'\hat{\mathbf{A}}\mathbf{F}(\hat{\mathbf{y}}^h)\hat{\mathbf{P}}'\hat{\mathbf{K}}', \tag{31}$$

where $\hat{\sigma}^2_v$ denotes a consistent estimator for σ^2_v and

$$\hat{\mathbf{K}} = \begin{pmatrix} \hat{\mathbf{\Omega}}_{1,h} & \hat{\mathbf{P}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1)\hat{\mathbf{A}}_1'\hat{\mathbf{A}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2) \\ \hat{\mathbf{P}}_2\mathbf{F}_2(\hat{\mathbf{y}}^h_2)\hat{\mathbf{A}}_2'\hat{\mathbf{A}}_1\mathbf{F}_1(\hat{\mathbf{y}}^h_1) & \hat{\mathbf{\Omega}}_{2,h} \end{pmatrix}.$$

8 Monte Carlo simulations

In this section, we assess the properties of our interpolation method using an error-correction model (ECM) as the basis for Monte-Carlo experiments. We look at the interpolation of a quarterly series to yield monthly data, and consider a situation where there is a first-order lag. Note the ECM is a simple re-parameterisation of equation [6]. Suppose the interpoland $\mathbf{y}_l = (y_1, ..., y_t)'$ and the single high-frequency indicator series $\mathbf{x}_h = (x_{11}, ..., x_{tu})'$ are co-integrated. We assume the variables are cointegrated in monthly terms, but the identification of the co-integrating vectors can only be made at a quarterly level, between the $\mathbf{y}_l$ series and quarterly aggregates of the monthly indicators, $\mathbf{x}_l = \mathbf{C}\mathbf{x}_h$. The underlying model is

$$Df(y_{tu}) = \beta_0 + Dx_{tu}\beta_1 - \alpha L(f(y_{tu}) - x_{tu}\delta) + v_{tu} \tag{32}$$

where $v_h \sim \mathrm{IN}(0, \sigma^2_v)$, D denotes the difference operator, $Dx_{tu} = x_{tu} - x_{t(u-1)}$, and L the lag operator, $Lx_{tu} = x_{t(u-1)}$, with obvious treatment when $u = 1$.

On the assumption that a single cointegrating parameter δ links the monthly data to the quarterly interpolands, so $f(y_{tu}) = x_{tu}\delta$ defines the cointegrating relationship, we rewrite equation [32] as

$$f(y_{tu}) = Lf(y_{tu})\alpha + \beta_0 + Dx_{tu}\beta_1 + Lx_{tu}\beta_2 + v_{tu} \tag{33}$$

For a given parameterisation of [33] and a series of indicator variables, x_{tu}, we can generate a set of disturbances v_{tu} to simulate y^{mc}_{tu} (the superscript mc indicates the vector is generated by Monte Carlo), and obtain y^{mc}_t. The disturbances v_{tu} are drawn from a $\mathrm{IN}(0, \sigma^2_v)$ distribution to create a sample large enough to be informative about the properties of our interpolation procedure and the behaviour of the relevant statistics, under model [33].

To resemble closely the conditions faced in practical econometric work, we borrowed the structure from a previous estimation of a component of the U.K. National Accounts. In each replication we constructed a synthetic series using a single indicator variable, the UK retail sales index[13] as explanatory variable. We based the exercise on a regression designed to interpolate consumption of services using the retail sales index, and therefore chose as our parameter values those resulting from such a regression. The sample size covers 82 quarters, from July, 1973 to December, 1993. The parameters are shown in Table 1 below.

Component	$\log CServ_{t-1}$	$\log RSales_{t-1}$	Constant	$\Delta \log RSales_{t-1}$
Label from [33]	$\widehat{\alpha}$	$\widehat{\beta}_2$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
parameter value	0.9347	0.0921	0.2783	0.3478

Table 1: Parameters for Monte Carlo simulation

Using the above parameter values and indicator series, the Monte Carlo experiments were set by simulating $\mathrm{IN}(0, s^2)$ random vectors $\mathbf{v}_h$ to provide the disturbances v_{tu}. The Box-Muller transform was used; see Devroye [1986] or Press *et al.* [1992] for further details.[14] The simulations are run assuming $v_{00} = 0$. Each Monte Carlo experiment consisted of 500 replications.

In each experiment, the fact that our lag was only first order made it easy to calculate the first and second-order conditions independently of the optimisation procedure. This gave a check against a numerical evaluation of the Hessian. It was done once every 50 replications in two ways. First, by perturbing the analytic first-order conditions on convergence; secondly, by re-starting the optimisation but using a BFGS algorithm[15], and recovering the numerical Hessian. To avoid distortions, the same termination criterion was used for the direct-search (analytic and perturbation alternatives) and BFGS algorithms.

Table 2 below shows the ML Monte Carlo estimates, their mean bias, the Monte Carlo sampling standard deviations (MCSD), the estimated standard errors (ESE) and the Monte Carlo standard errors (MCSE) for each coefficient.[16]

The results from our MLE Monte Carlo experiment are satisfactory. Except for the constant, the regression estimators are close to the ones from Table1. The ESE and MCSD estimates are close, indicating the conventional variance formula[17] provides a reasonable measure of the true parameter uncertainty, given the model and sample size. In fact, ESE $<$ MCSD in 3 out of the 4 coefficients (and particularly for α), so our estimation procedure if anything suggests a slight tendency to underestimate the true parameter uncertainty, as the MCSD is the correct value of the parameter standard deviation.

[13]Code FAAM in the U.K. Central Statistical Office database.

[14]An alternative is to apply the Box-Jenkins [1973] methodology to the pseudo-random vectors generated by standard computer algorithms.

[15]BFGS stands for the Broyden-Fletcher-Goldfarb-Shanno optimisation procedure. See Gill, Murray and Wright [1982] or Hendry [1995, A5], among others, for a description of optimisation methods.

[16]The terminology follows Hendry [1995].

[17]That is, $\mathrm{V}(\widehat{\alpha, \beta_i}) = s^2(\mathbf{X}_i\mathbf{V}^{-1}\mathbf{X}_i)^{-1}$, with s^2 an estimator of σ^2.

Component	$\log CServ_{t-1}$	$\log RSales_{t-1}$	Constant	$\Delta \log RSales_{t-1}$
Label from [33]	$\widehat{\alpha}$	$\widehat{\beta}_2$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
parameter value	0.9271	0.1014	0.3052	0.3676
MCSD	0.0221	0.0271	0.0804	0.0779
ESE	0.0184	0.0242	0.0790	0.0803
R^2	0.9977			
s^2	0.0082			
Chow test A, F(5,74)	0.9922	[0.43]		
Chow test B, F(4,75)	1.2348	[0.30]		
Bera-Jarque χ^2 (2)	1.7192	[0.42]		
SC, F(1,78)	0.9188	[0.34]		
SC, F(4,75)	1.0321	[0.40]		
ARCH, χ^2 (1)	0.9947	[0.32]		
ARCH, χ^2 (4)	3.6421	[0.46]		

Table 2: Simulation estimates and diagnostics [p-values in brakets] for true $\alpha = 0.934$ (other parameters as in Table 1) using MLE.

Table 3 below presents the mean bias of the parameter estimators and Bera-Jarque $\chi^2(2)$ statistics for normality (based on D'Agostino *et al.*, 1990).

Parameter	$\widehat{\alpha}$	$\widehat{\beta}_2$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
Mean bias	−0.0076	0.0093	0.0269	0.0198
Bera-Jarque χ^2 (2)	31.202 [0.00]	34.093 [0.00]	28.331 [0.00]	3.334 [0.19]

Table 3: Mean bias and normality tests of estimators in Table 2

The mean biases, in turn, are small as a proportion of the coefficient estimates[18] but somewhat large relative to the MCSD. Finally, figure [1] presents the histograms (normalised to have zero mean and unit variance) for the four estimates in Table 2.

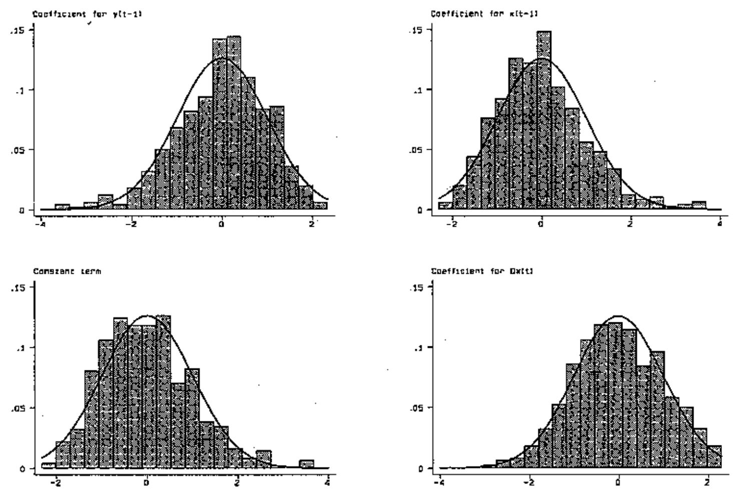

Figure 1: Histograms and density for the parameter estimators of Table 3.

[18]To put the bias $(\overline{\alpha} - \alpha) = -0.0108$ inside a 95% confidence interval, we would need $M \sim 6200$ replications; for $M = 500$, the length of a 95% confidence interval is roughly 0.0382. This means that, even the number of replications is not big, the estimation of α is precise enough.

As expected, the bias in $\widehat{\alpha}$ and $\widehat{\beta}$ offset each other, so to maintain the long-run outcome. The negative skewness in the distribution of $\widehat{\alpha}$ (our parameter of interest) may account in part for the small discrepancy between the MCSD and ESE. Both values, however, are close enough to indicate that model uncertainty is indeed correctly captured, as mentioned above.

Returning to the results in Table 2, we provide several additional diagnostics. The Chow tests A and B are, respectively, a N↑-step and parameter stability tests based on Chow [1960]. The SC and ARCH rows tabulate standard LM-type serial correlation and ARCH tests, respectively. The tests are satisfactory at a 5% level, and overall the equation is well behaved.

We conducted two additional experiments, for different α values but the same β's as in the benchmark regression. The selected values for α in [33] were 0.850 and 0.950 (corresponding to quarterly coefficients of 0.614 and 0.8574 respectively), roughly the lower and upper bounds found for α in the regressions for the components of the U.K. National Accounts. Table 4 presents the results for $\alpha = 0.850$.

Component	$\log CServ_{t-1}$	$\log RSales_{t-1}$	Constant	$\triangle \log RSales_{t-1}$
Label from [33]	$\widehat{\alpha}$	$\widehat{\beta}_2$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
parameter value	0.8317	0.1000	0.2986	0.3520
MCSD	0.0420	0.0246	0.0773	0.0791
ESE	0.0459	0.0281	0.0808	0.0800
R^2	0.9899			
s^2	0.0082			
Chow test A, F(5,74)	0.9433	[0.46]		
Chow test B, F(4,75)	1.0972	[0.36]		
Bera-Jarque χ^2 (2)	1.9172	[0.38]		
SC, F(1,78)	1.0962	[0.30]		
SC, F(4,75)	1.0156	[0.40]		
ARCH, χ^2 (1)	0.8824	[0.35]		
ARCH, χ^2 (4)	3.6373	[0.46]		

Table 4: Simulation estimates and diagnostics [p-values in brakets] for true $\alpha = 0.850$ (other parameters as in Table 1) using MLE.

As in the first experiment, the results are satisfactory. There is no serious bias in the coefficient estimates, and the regression diagnostics show no problems. In Table 5 we report the same statistics as Table 3. Histograms of the four parameter estimators are plotted in Figure [2].

Parameter	$\widehat{\alpha}$	$\widehat{\beta}_2$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
Mean bias	−0.0194	0.0074	0.0211	0.0039
Bera-Jarque χ^2 (2)	52.282 [0.00]	46.463 [0.00]	62.113 [0.00]	0.831 [0.66]

Table 5: Mean bias and normality tests of estimators in Table 5

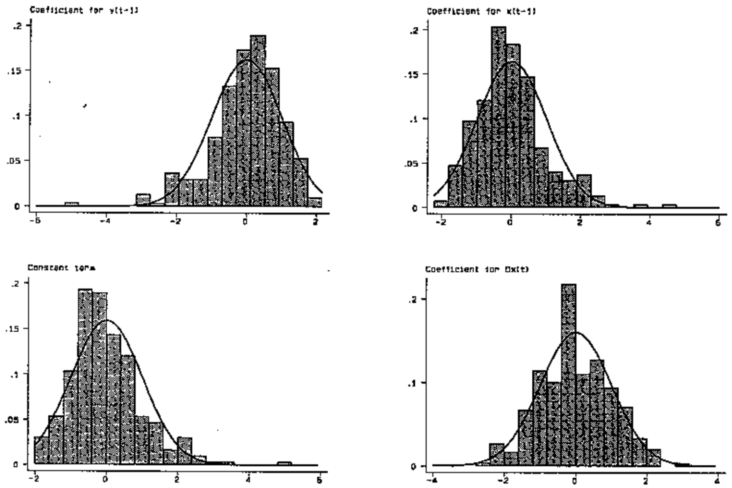

Figure 2: Histograms and density for the parameter estimators of Table 5.

Component	$\log CServ_{t-1}$	$\log RSales_{t-1}$	Constant	$\triangle \log RSales_{t-1}$
Label from [33]	$\widehat{\alpha}$	$\widehat{\beta}_2$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
parameter value	0.9436	0.0998	0.3036	0.3507
MCSD	0.0168	0.0223	0.0812	0.0789
ESE	0.0151	0.0210	0.0765	0.0792
R^2	0.9985			
s^2	0.0083			
Chow test A, F(5,74)	1.0423	[0.40]		
Chow test B, F(4,75)	1.1874	[0.32]		
Bera-Jarque χ^2 (2)	1.7933	[0.41]		
SC, F(1,78)	1.9025	[0.35]		
SC, F(4,75)	1.0702	[0.38]		
ARCH, χ^2 (1)	0.9591	[0.33]		
ARCH, χ^2 (4)	3.8243	[0.43]		

Table 6: Simulation estimates and diagnostics [p-values in brakets] for true $\theta = 0.950$ (other parameters as in Table 1) using MLE.

The outcome of experiment for $\alpha = 0.95$ was satisfactory as well, as seen by the statistics in Table 6. In line with the previous two experiments, the equation parameters have non-normal distributions, but now with the exception of the trend term (and, therefore, a MCSD almost identical to the ESE statistic). The histograms and densities of the model parameters for $\alpha = 0.95$ are shown in figure [3].

Parameter	$\widehat{\alpha}$	$\widehat{\beta}_2$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
Mean bias	−0.0069	0.0077	0.0253	0.0030
Bera-Jarque χ^2 (2)	12.013 [0.00]	15.010 [0.00]	10.572 [0.01]	0.001 [0.99]

Table 7: Mean bias and normality tests of estimators in Table 7

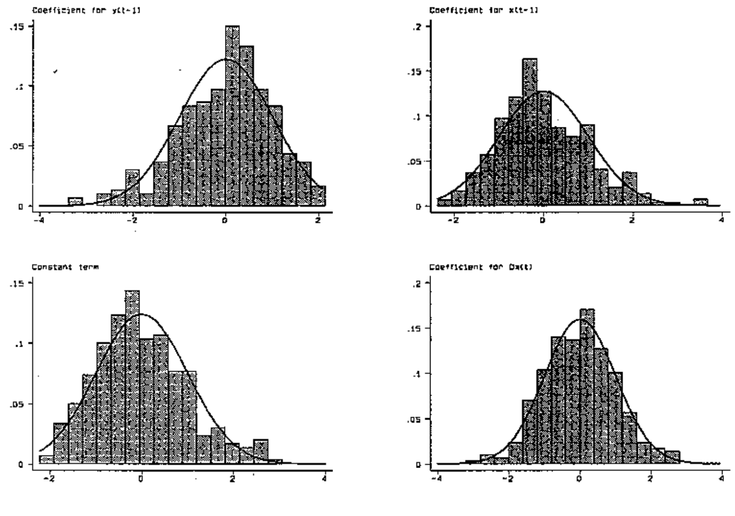

Figure 3: Histograms and density for the parameter estimators of Table 7.

8.1 Testing the validity of restrictions due to aggregation

Another relevant issue of this whole exercise concerns the impact of temporal aggregation on the estimation of the parameters in [33]. We can rewrite [33] as

$$\begin{aligned}(1-\alpha^3L^3)\,f(y_{tu}) &= (1+\alpha L+\alpha^2L^2)\,DX_{tu}\beta-\alpha\,(1+\alpha L+\alpha^2L^2)\,LX_{tu}\delta+\\ &+\ (1+\alpha L+\alpha^2L^2)\,v_{tu}\end{aligned}$$

obtaining, after some simplification,

$$f(y_{tu}) = \varphi_1 L^3 f(y_{tu}) + DX_{tu}\varphi_2 + LDX_{tu}\varphi_3 + L^2DX_{tu}\varphi_4 + LX_{tu}\varphi_5 + (1+\alpha L+\alpha^2L^2)\,v_{tu} \quad (34)$$

Note equation [34] is the unrestricted version of [33]. There are two restrictions implicit in [33]. A Lagrange multiplier test of the form

$$\mathbf{e}_l'\mathbf{V}^{-1}\mathbf{X}_l\left(\mathbf{X}_l\mathbf{V}^{-1}\mathbf{X}_l\right)^{-1}\mathbf{X}_l'\mathbf{V}^{-1}\mathbf{e}_l' \quad (35)$$

can be carried out, distributed in our case as $\chi^2(2)$;[19] see, for example, Davidson and MacKinnon [1993, ch. 3] or Godfrey [1988] for a thorough discussion. The LM test statistic [35] should provide a clear indication if the restrictions imposed to estimate [33] are valid, and therefore it may be regarded as a misspecification test for [33]. Table 8 summarizes the results for the three experiments conducted. In every case, the empirical size of the LM test comfortably falls into the $\chi^2(2)$ critical region, at the selected fractiles.[20]

$\chi^2(2)$ fractile Outcomes	0.90	0.95	0.99
$\alpha = 0.935$ (base case)	0.098	0.044	0.001
$\alpha = 0.850$	0.084	0.044	0.000
$\alpha = 0.950$	0.100	0.046	0.004

Table 8: Empirical size of LM test, for the 0.90, 0.95 and 0.99 fractiles of the $\chi^2(2)$ distribution, 500 replications

8.2 Standard errors of the interpolands

In section 7 we showed how to estimate standard errors of the interpolands. The Monte Carlo standard errors can be compared with the actual standard errors across the different replications, in order to assess how good the theoretical calculations are as an indicator of the true accuracy of the interpolated data.

The following tables presents the interpoland standard errors for the middle and end-sample quarters derived from the replications and then calculated theoretically, with the rows indicating the standard error for the first, second and third month of each quarter. These are shown as proportions of the interpolated data.

[19] We have five unrestricted regressors, φ_1 to φ_5 and three restricted parameters, α, β_1 and δ. Note in [17] we are always working under H_0, so there is no need to estimate [16]; the covariance and error vectors are those obtained from the MLE of [33].

[20] And marginally so for $\alpha = 0.95$ at a 10% significance.

	Mid-sample			End-sample		
	$\alpha = 0.850$	$\alpha = 0.935$	$\alpha = 0.950$	$\alpha = 0.850$	$\alpha = 0.935$	$\alpha = 0.950$
1	0.5367	0.5068	0.5044	0.5605	0.5414	0.5658
2	0.4003	0.3896	0.4000	0.4116	0.3984	0.4107
3	0.5511	0.5056	0.4850	0.5890	0.5682	0.5481

Table 9: Monte Carlo interpoland standard errors for middle and end-sample quarters

	Mid-sample			End-sample		
	$\alpha = 0.850$	$\alpha = 0.935$	$\alpha = 0.950$	$\alpha = 0.850$	$\alpha = 0.935$	$\alpha = 0.950$
1	0.5270	0.5049	0.4871	0.5510	0.5224	0.5375
2	0.4002	0.3618	0.3628	0.3918	0.3920	0.3911
3	0.5216	0.5058	0.4575	0.5700	0.5510	0.5009

Table 10: Asymptotic interpoland standard errors for middle and end-sample quarters

Overall, the second moments both for the middle and end-sample quarters are low, there seems to be no significant end-of-sample problems, and the Monte Carlo values are quite close to the asymptotic ones. The off-diagonal elements in the error covariance matrix decay rapidly to zero.

We have also computed the Bera-Jarque statistic (using the same procedure as in tables 3, 5 and 7) for the interpolated values, finding in all cases that normality cannot be rejected.

	Mid-sample			End-sample		
	$\alpha = 0.850$	$\alpha = 0.935$	$\alpha = 0.950$	$\alpha = 0.850$	$\alpha = 0.935$	$\alpha = 0.950$
1	2.31 [0.32]	1.96 [0.37]	2.12 [0.35]	0.11 [0.95]	0.30 [0.86]	0.36 [0.84]
2	1.97 [0.37]	2.56 [0.28]	2.71 [0.26]	1.38 [0.50]	1.27 [0.53]	1.32 [0.52]
3	1.47 [0.48]	1.31 [0.52]	1.69 [0.43]	2.61 [0.27]	0.36 [0.84]	0.58 [0.75]

Table 11: Normality tests for interpolated data [P-values in brackets]

In addition, to compare the distributions of the outcomes against the true values we present a graphical approach: figures 4 to 6 present 2-way quantile plots for the quarters of interest. Each plot shows the quantiles of $\widehat{\mathbf{y}}_{tu}$ (circles) against those of $\mathbf{y}_{tu}$ (solid line) for month t of quarter u. Given the above statistics, both distributions should be quite close to the 45° line. Clearly from the graphs, in each of the experiments and months compared, the distributions of $\widehat{\mathbf{y}}_j$ and $\mathbf{y}_j$ are, roughly, the same.

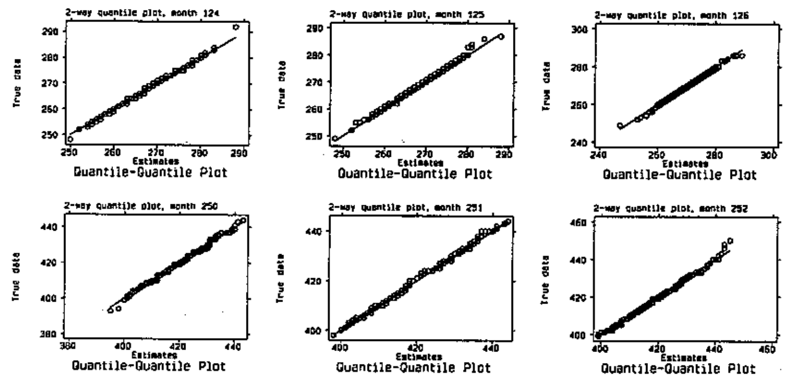

Figure 4: 2-way quantile plot, quarters 41 [top row] and 82 [bottom row], for $\alpha^3 = 0.935$ simulation and 500 replications.

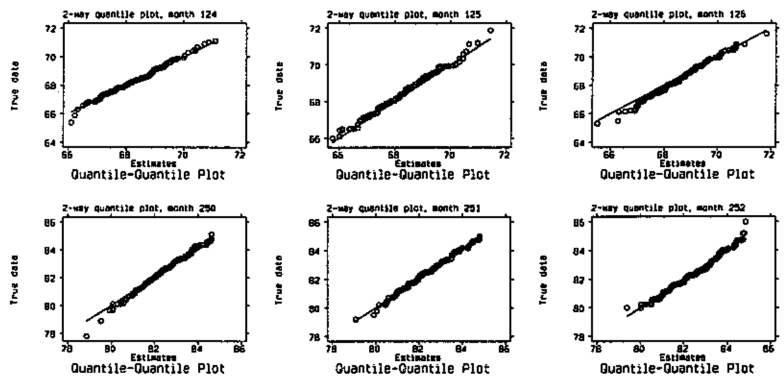

Figure 5: 2-way quantile plot, quarters 41 [top row] and 82 [bottom row], for $\alpha^3 = 0.850$ simulation and 500 replications.

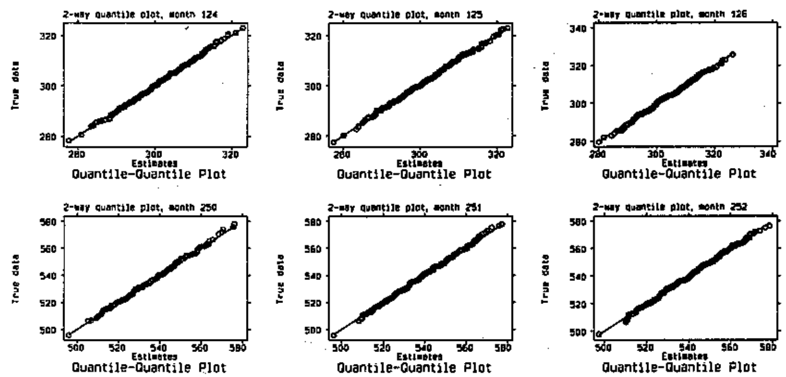

Figure 6: 2-way quantile plot, quarters 41 [top row] and 82 [bottom row], for $\alpha^3 = 0.950$ simulation and 500 replications.

9 Conclusions

The model-based method of interpolation suggested by Chow and Lin [1971] proves to be readily extendible to deal with the case where there is a dynamic relationship between the indicator variables and the interpoland. We have described both Generalized Least Squares and Maximum-Likelihood solution methods.

A Monte-Carlo analysis of a partial adjustment error-correction model suggests that there are no serious problems arising from biases either in the parameters or in the interpolated data. Very importantly, the estimated standard errors of both the parameters and the interpolated data seem to be good indicators of the experimental standard errors. This means that the method described here can be used both to interpolate data and to provide an indicator of the reliability of the resulting interpolands.

A Appendix: Standard errors of Maximum Likelihood Estimators

The estimation of the standard errors involves, as usual, recovering the Hessian matrix of the log-likelihood equation; c.f. [15] or [16]. In our case, the derivation of the Hessian is not straightforward given the dependence of $\mathbf{V}$ on θ, and the inclusion of dynamics.

For ease of exposition. we skip some of the intermediate steps in the derivation of the Hessian matrix. We avoided vectorizing to keep notation understandable, however at the expense of some loss of simplicity in several elements of the Hessian. Vectors are presented in [bold] lower case and matrices in [bold] upper case; non-bold typefaces refer to scalars.

We define the following auxiliary vectors and matrices, where the notation should be evident from the main text.:

1. $\mathbf{X}_1 = \mathbf{CS}_h\mathbf{X}_h$
2. $\mathbf{X}_2 = \mathbf{C}(\partial\mathbf{S}_h/\partial\alpha)\mathbf{X}_h$
3. $\mathbf{F}_1 = [\mathbf{Ly}\ \mathbf{X}_2][3\alpha^2\ \beta]'$
4. $\mathbf{F}_2 = [\mathbf{Ly}\ 2\mathbf{CX}_h][6\alpha\ \beta]'$
5. $\mathbf{H} = \mathbf{V}^{-1}(\partial\mathbf{V}/\partial\alpha)\mathbf{V}^{-1}$
6. $\mathbf{K}_1 = \mathbf{V}^{-1}\mathbf{e}$
7. $\mathbf{K}_2 = \mathbf{V}^{-1}\mathbf{F}_1 + \mathbf{He}$
8. $\mathbf{P} = \mathbf{K}_1'\mathbf{F}_1 + 0.5\mathbf{K}_1'(\partial\mathbf{V}/\partial\alpha)\mathbf{K}_1$

The second-order conditions of the log-likelihood [16], with respect to α and β, are

$$\mathbf{D}_{11} = \frac{2}{Ts^4}\mathbf{P}'\mathbf{P} + \frac{1}{s^2}\left[\frac{1}{2}\mathbf{K}_1'\frac{\partial^2\mathbf{V}}{\partial\alpha\partial\alpha'}\mathbf{K}_1 - \mathbf{K}_1'\frac{\partial\mathbf{V}}{\partial\alpha}\mathbf{K}_2 - \mathbf{K}_2'\mathbf{F}_1 - \mathbf{F}_2'\mathbf{K}_1\right] + \frac{1}{2}\mathrm{tr}\left(\mathbf{H} - \mathbf{V}^{-1}\right)\frac{\partial\mathbf{V}}{\partial\alpha}$$

$$\mathbf{D}_{21} = \frac{2}{Ts^4}\mathbf{X}_1'\mathbf{K}_1\mathbf{P} + \frac{1}{s^2}\left[\mathbf{X}_2'\mathbf{K}_1 - \mathbf{X}_1'\mathbf{V}^{-1}\mathbf{F}_1 - \mathbf{X}_1'\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\alpha}\mathbf{K}_1\right]$$

$$\mathbf{D}_{22} = \frac{2}{Ts^4}\mathbf{X}_1'\mathbf{K}_1\mathbf{K}_1'\mathbf{X}_1 - \frac{1}{s^2}\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1$$

and the Hessian matrix is simply

$$\mathbf{H}(\alpha,\beta) = \frac{1}{s^2}\left[\begin{array}{c|c} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \hline \mathbf{D}_{21} & \mathbf{D}_{22} \end{array}\right]$$

where s^2 is the MLE estimator of σ^2 and $\mathbf{D}_{12} = \mathbf{D}'_{21}$; $\mathbf{H}(\alpha, \beta)$ should be negative definite at the maximum value of the likelihood. Therefore, $\text{Var}(\widehat{\alpha, \beta}) = -\mathbf{H}(\widehat{\alpha, \beta})^{-1}$ is the asymptotic variance of the maximum likelihood estimates.

References

[1] C. G. Boot, W. Feibes, and J. H. C. Lisman. Further comments on the derivation of quarterly figures from annual data. *Applied Statistics*, 16:65–75, 1967.

[2] G. E. P. Box and G. M. Jenkins. *Time Series Analysis (Revised Edition)*. Holden Day, San Francisco, 1973.

[3] K. R. W. Brewer. Some consequences of temporal aggregation and systematic sampling for arma and armax models. *Journal of Econometrics*, 1(2):133–54, 1973.

[4] J. Campos, N. R. Ericsson, and D. F. Hendry. An analogue model of phase averaging procedures. *Journal of Econometrics*, 43(3):275–92, 1990.

[5] A. L. Chow, G C; Lin. Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *Review of Economics and Statistics*, 53(4):372–75, 1971.

[6] A. L. Chow, G C; Lin. Best linear unbiased estimation of missing observations in an economic time series. *Journal of the American Statistical Association*, 71:719–21, 1976.

[7] K. J. Cohen, W. Muller, and M. W. Padberg. Autoregressive approaches to disaggregation of time series data. *Journal of the Royal Statistical Society*, Series C (20):119–129, 1971.

[8] E. de Alba. Temporal disaggregation of time series: A unified approach. *Proceedings of the American Statistical Association*, pages 359–365, 1979.

[9] J. E. Dennis and V. J. Torczon. Direct search methods on parallel machines. *SIAM Journal of Optimization*, 1(4):448–474, 1991.

[10] F. T. Denton. Adjustment of monthly or quartely series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association*, 66(333):99–102, 1971.

[11] L. Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, 1986.

[12] R. B. Fernandez. A methodological note on the estimation of time series. *Review of Economics and Statistics*, 63:471–476, 1981.

[13] M. Friedman. The interpolation of time series by related series. *Journal of the American Statistical Association*, 57:729–757, 1962.

[14] J. F. Geweke. Temporal aggregation in the multiple regression model. *Econometrica*, 46(3):643–61, 1978.

[15] P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, London and New York, 1981.

[16] V. A. Ginsburgh. A further note on the derivation of quarterly figures consistent with annual data. *Journal of the Royal Statistical Society*, Series C (22):368–374, 1973.

[17] H. Glesjer. Une methode d'evaluation de donnees mensuelles a partir d'indices trimestrels ou annuels. *Cahiers Economiques de Bruxelles*, 29:45–54, 1966.

[18] V. Gomez and A. Maravall. Estimation, prediction, and interpolation for nonstationary series with the kalman filter. *Journal of the American Statistical Association*, 89(426):611–24, 1994.

[19] C. W. J. Granger and P. L. S. Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration: Theory and evidence. *Journal of Econometrics*, 66(1-2):357–69, 1995.

[20] V. M. Guerrero and D. Pena. Linear combination of information in time series analysis. Universidad Carlos III Working Paper 95-52; Statistics and Econometrics Series 18, 1995.

[21] A. C. Harvey. *Forecasting, Structural Time Series and the Kalman Filter*. Cambridge University Press, Cambridge, 1989.

[22] A. C. Harvey and R. G. Pierse. Estimating missing observations in economic time series. *Journal of the American Statistical Association*, 79(385):125–31, 1984.

[23] D. F. Hendry. *Dynamic Econometrics*. Oxford University Press; Oxford, pages 869, 1995.

[24] N. J. Highman. Optimization by direct search in matrix computations. *SIAM Journal of Matrix Analysis (Applied)*, 14(2):317–333, 1993.

[25] R. Jones. Maximum likelihood fitting of arima models to time series with missing observations. *Technometrics*, 22:389–395, 1980.

[26] M. Lippi and L. Reichlin. Trend-cycle decompositions and measures of persistence: Does time aggregation matter? *Economic Journal*, 101(405):314–323, 1991.

[27] J. H. C. Lisman and J. Sandee. Derivation of quarterly figures from annual data. *Applied Statistics*, 13:87–90, 1964.

[28] R. B. Litterman. A random walk, markov model for the distribution of time series. *Journal of Business and Economic Statistics*, 1:169–173, 1983.

[29] T. C. Liu and E. C. Hwa. A monthly econometric model of the us economy. *International Economic Review*, 15(2):328–65, 1974.

[30] H. Lutkepohl. *Forecasting aggregated vector ARMA processes*. Lecture Notes in Economics and Mathematical Systems series, no 284 New York; Berlin; London and Tokyo: Springer, 1987, pages 323, 1987.

[31] J. R. Magnus. Maximum likelihood estimation of the gls model with unknown parameters in the disturbance covariance matrix. *Journal of Econometrics*, 7(3):281–312, 1978.

[32] T. E. Nijman. *Missing Observations in Dynamic Macroeconomic Modeling*. Free University Press; Amsterdam, pages 240, 1985.

[33] T. E. Nijman and F. C. Palm. The construction and use of approximations for missing quarterly observations: A model based approach. *Journal of Business and Economic Statistics*, 4(1):47–58, 1986.

[34] T. E. Nijman and F. C. Palm. Predictive accuracy gain from disaggregate sampling in arima models. *Journal of Business and Economic Statistics*, 8(4):405–15, 1990.