



The Modeling and Seasonal Adjustment of Weekly Observations

Author(s): Andrew Harvey, Siem Jan Koopman, Marco Riani

Source: *Journal of Business & Economic Statistics*, Vol. 15, No. 3 (Jul., 1997), pp. 354-368

Published by: American Statistical Association

Stable URL: <http://www.jstor.org/stable/1392339>

Accessed: 14/11/2008 05:11

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=astata>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



American Statistical Association is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Business & Economic Statistics*.

<http://www.jstor.org>

The Modeling and Seasonal Adjustment of Weekly Observations

Andrew HARVEY

Faculty of Economics and Politics, University of Cambridge, Cambridge CB3 9DD, United Kingdom

Siem Jan KOOPMAN

Department of Statistics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom

Marco RIANI

Dipartimento Statistico, Universita di Firenze, Viale Morgagni 59, 50139 Firenze, Italy

Several important economic time series are recorded on a particular day every week. Seasonal adjustment of such series is difficult because the number of weeks varies between 52 and 53 and the position of the recording day changes from year to year. In addition certain festivals, most notably Easter, take place at different times according to the year. This article presents a solution to problems of this kind by setting up a structural time series model that allows the seasonal pattern to evolve over time and enables trend extraction and seasonal adjustment to be carried out by means of state-space filtering and smoothing algorithms. The method is illustrated with a Bank of England series on the money supply.

KEY WORDS: Calendar effects; Irregularly spaced observations; Kalman filter; Money supply; Moving festival; Periodic spline; Stochastic seasonality; Structural time series model.

The weekly figures on the U.K. money supply are eagerly anticipated in the City of London because they are believed to be an important economic indicator. One of the key series is the value of the Bank of England notes and coins in circulation, plus cash deposits of commercial banks with the Bank of England. This basically corresponds to the measure known as M0, and we will refer to it in this way hereafter. These figures display considerable seasonal fluctuations and are particularly high just before Christmas. As a result there is a need for the Bank of England to produce a seasonally adjusted series for ease of interpretation.

Figure 1 shows a plot of the logarithms of the observations on M0 starting on May 28, 1969. Taking logarithms yields a series with a more stable seasonal pattern. The figures are recorded every Wednesday, except when the Wednesday falls on a public holiday, in which case the figure is recorded on the previous Tuesday (or Monday if Tuesday is also a holiday). The Christmas peak can be clearly seen and, as with many economic time series, it is apparent that the seasonal pattern has evolved over time due to changing institutional and social factors. An excellent discussion of the changing nature of Christmas and its consequent economic impact is that of Scott (1995).

Modeling a changing seasonal component is relatively easy for quarterly and monthly observations, the seasonal component normally being combined with a stochastic trend and an irregular term. This is either done explicitly, as in the structural time series modeling approach, or implicitly, as in the autoregressive integrated moving average (ARIMA) approach. In the latter case, the seasonal component is specified by means of a "canonical decomposition" as shown by Hillmer and Tiao (1982). The seasonal component can be extracted by a state-space smoothing algorithm;

see, for example, Kitagawa and Gersch (1984) or Harvey (1989). Carrying out such *model-based* seasonal adjustment, using either approach, has considerable attractions because the procedure adapts to the particular characteristics of the series involved. The relationship between the model-based procedures and the widely used Bureau of the Census X-11 program was discussed by Maravall (1985).

Seasonal adjustment of weekly data is not an easy task. The first problem is that, because the observations are normally recorded on a particular day of the week rather than on predetermined dates, the fact that there is not an integral number of weeks in the year means that the number of observations in the year varies between 52 and 53. Thus, even if the seasonal pattern were deterministic, it could not be modeled by a set of dummy variables. Furthermore, the position of the dates of the observation days changes with each year so that even with an integral number of weeks in the year the seasonal pattern would change from year to year. It makes a big difference, for example, if the money-supply figure is recorded on the day before Christmas or six days before Christmas. (The former case arises if Christmas is on a Thursday, the latter if it is on a Tuesday.) To make matters worse, these differing seasonal patterns do not even recur every seven years because of leap years.

The other major problem is that the position of Easter changes from year to year. Furthermore, its effect can be different depending on when it occurs. If it is late, its effects can overlap, and possibly interact with, those associated with the May Day public holiday. Of course, the

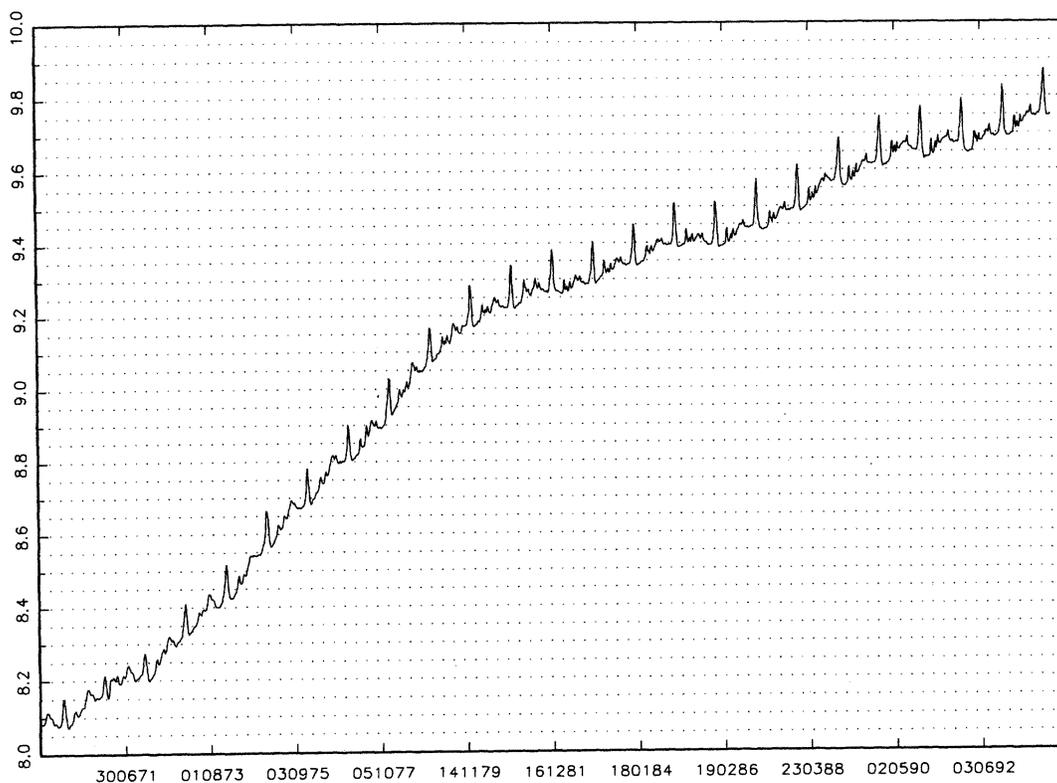


Figure 1. Weekly Observations on the U.K. Money Supply.

position of Easter also affects models for monthly observations, but in this case it is more easily handled, and there is a considerable literature on its treatment; see Bell and Hillmer (1983).

The ARIMA-based procedure does not easily generalize to weekly data. One of the few published articles on weekly model-based seasonal adjustment, that by Pierce, Grupe, and Cleveland (1984), got around some of the problems by using regression to model some of the seasonal effects in a deterministic way and then grafting on stochastic effects using an ARIMA model. Our approach is to attack the problem using structural time series models. Such models can be interpreted as regressions on functions of time in which the parameters are time-varying. This makes them a natural vehicle for handling changing seasonality of a complex form. Once a suitable model has been fitted, the seasonal component can be extracted by a smoothing algorithm.

The plan of the article is as follows. Section 1 briefly reviews the basic structural model as it is typically used for quarterly and monthly data and explains why it cannot be directly applied to weekly data. Sections 2 and 3 describe the two seasonal components that we propose introducing into a structural model to allow it to cope with weekly data, and Section 4 explains how the model is handled statistically. The model is applied to the Bank of England money-supply data in Section 5. Section 6 presents the conclusions.

1. THE BASIC STRUCTURAL TIME SERIES MODEL

The basic structural model (BSM) is formulated in terms of trend, seasonal, and irregular components. All are assumed to be stochastic and driven by serially independent

Gaussian disturbances that are mutually independent. If there are s seasons in the year, the model is

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2), \quad (1.1)$$

where the trend, seasonal, and irregular are denoted by μ_t , γ_t , and ε_t , respectively.

The trend is specified in the following way:

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, & \eta_t &\sim NID(0, \sigma_\eta^2), \\ \beta_t &= \beta_{t-1} + \zeta_t, & \zeta_t &\sim NID(0, \sigma_\zeta^2), \end{aligned} \quad (1.2)$$

where μ_t is the level and β_t is the slope. The disturbances η_t and ζ_t are assumed to be mutually independent. Setting $\sigma_\eta^2 = 0$ gives a trend that is relatively smooth.

The seasonal component is usually set up in terms of stochastic trigonometric functions at the $s/2$ seasonal frequencies, although dummy-variable formulations are also possible. The key point is that, although the seasonal component is nonstationary, it has the property that the expected value of the sum over the previous s time periods is 0. This ensures that seasonal effects are not confounded with the trend. It also means that the forecasts of the seasonal component will sum to 0 over any one-year period.

The statistical treatment of the model is based on the state-space form, with $s+1$ elements in the state vector. Estimation, forecasting, and signal extraction are carried out by means of the Kalman filter and associated algorithms.

1.1 Trigonometric Seasonality

The trigonometric form of stochastic seasonality used in

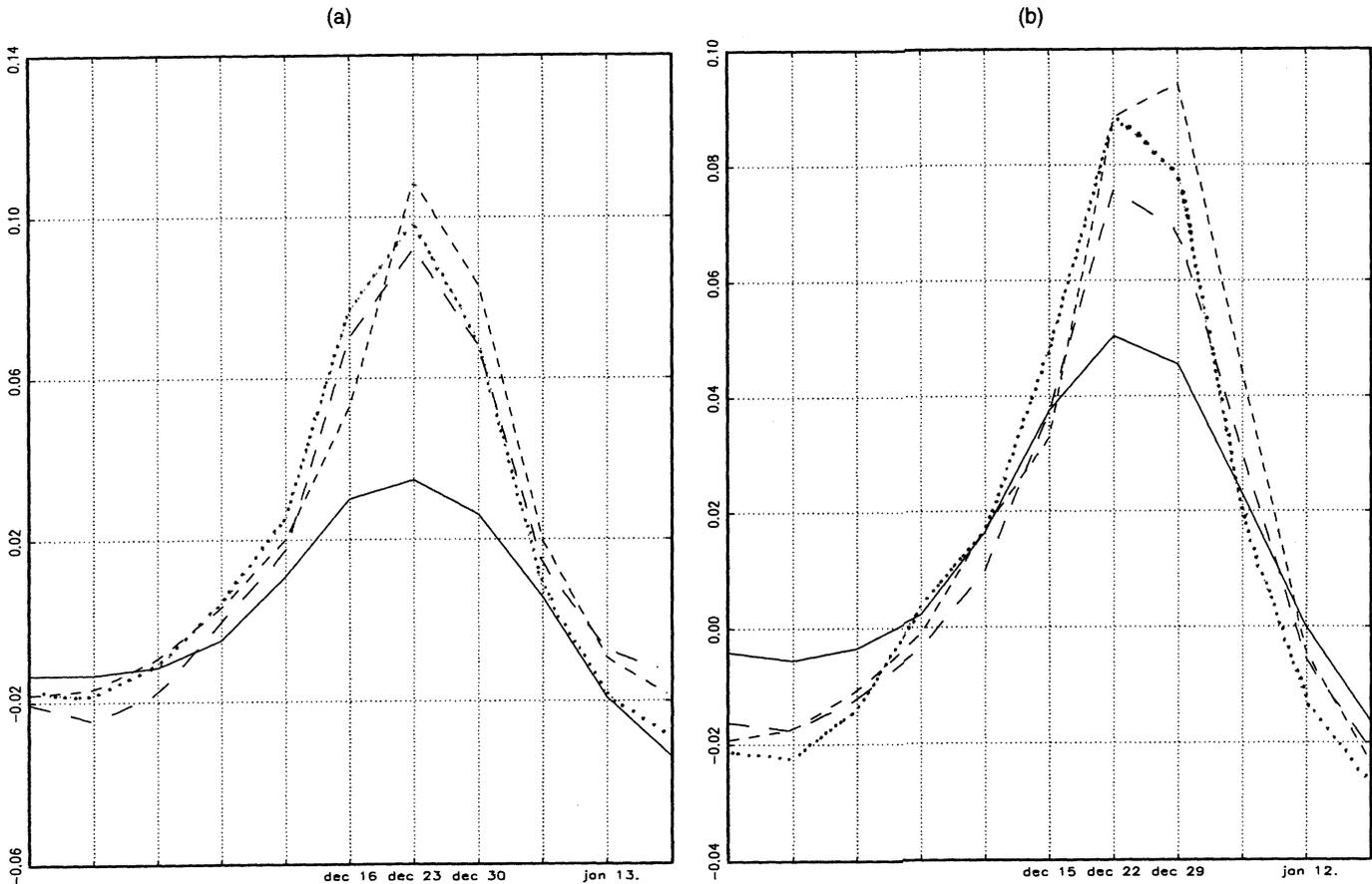


Figure 2. Detrended Series Over the Christmas Period for (a) Observations on December 16, 23, and 30 and January 6 for 1970, --; 1981, - -; 1987, . . .; and 1992, —; (b) Observations on December 15, 22, and 29 and January 5 for 1971, --; 1976, - -; 1982, . . .; and 1993, —.

models of the form (1.1) with s seasons in the year is

$$\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{j,t}, \quad t = 1, \dots, T, \quad (1.3)$$

where each $\gamma_{j,t}$ is generated by

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma_{j,t}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \sin \lambda_j \\ -\sin \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{j,t} \\ \omega_{j,t}^* \end{bmatrix}, \quad (1.4)$$

where $\lambda_j = 2\pi j/s$ is frequency, in radians, for $j = 1, \dots, [s/2]$ and ω_t and ω_t^* are two mutually uncorrelated white-noise disturbances with zero means and common variance σ_ω^2 . For s even, $[s/2] = s/2$, but for s odd, $[s/2] = (s - 1)/2$. Note that $\gamma_{j,t}^*$ is redundant for $j = s/2$.

The BSM consisting of the stochastic trend in (1.2) combined with trigonometric seasonality is easily put in state-space form by defining the $(s + 1) \times 1$ state vector $\alpha_t = (\mu_t, \beta_t, \gamma_{1t}, \gamma_{1t}^*, \gamma_{2t}, \gamma_{2t}^*, \dots)'$. The measurement equation is then

$$y_t = (1, 0, \mathbf{z}_t') \alpha_t + \varepsilon_t, \quad (1.5)$$

where $\mathbf{z}_t' = (1, 0, 1, 0, \dots)$. If the Kalman filter is initiated with a diffuse prior, as shown by De Jong (1991), an estimator of the state with a proper prior is effectively constructed from the first $s + 1$ observations.

1.2 Dummy-Variable Seasonality

The form of dummy-variable seasonality relevant to the development later in the article is one in which each element in an $s \times 1$ vector γ_t represents the effect of a particular month and these effects sum to 0; see Harvey (1989, pp. 40–41) for a discussion of different types of dummy-variable seasonality. The effects evolve over time according to a multivariate random walk

$$\gamma_t = \gamma_{t-1} + \chi_t, \quad (1.6)$$

where χ_t is an $s \times 1$ vector of serially uncorrelated random disturbances with zero mean. The zero sum over the year constraint implies restrictions on the covariance matrix of disturbances. Specifically

$$\text{var}(\chi_t) = \sigma_\chi^2 (\mathbf{I} - (1/s)\mathbf{ii}'), \quad (1.7)$$

where σ_χ^2 is the variance parameter that governs the speed with which the seasonal pattern can change and \mathbf{i} is an $s \times 1$ vector of ones. This covariance matrix enforces the constraint that $\mathbf{i}'\chi_t = 0$ by making its variance 0. Thus, if $\mathbf{i}'\gamma_{t-1} = 0$, then $\mathbf{i}'\gamma_t = 0$.

One of the elements of γ_t can be dropped from Equation (1.6); it can always be recovered as minus the sum of the elements remaining. The state-space form of a BSM with this kind of seasonality is such that $s - 1$ elements of γ_t appear in the state vector and the measurement equation is

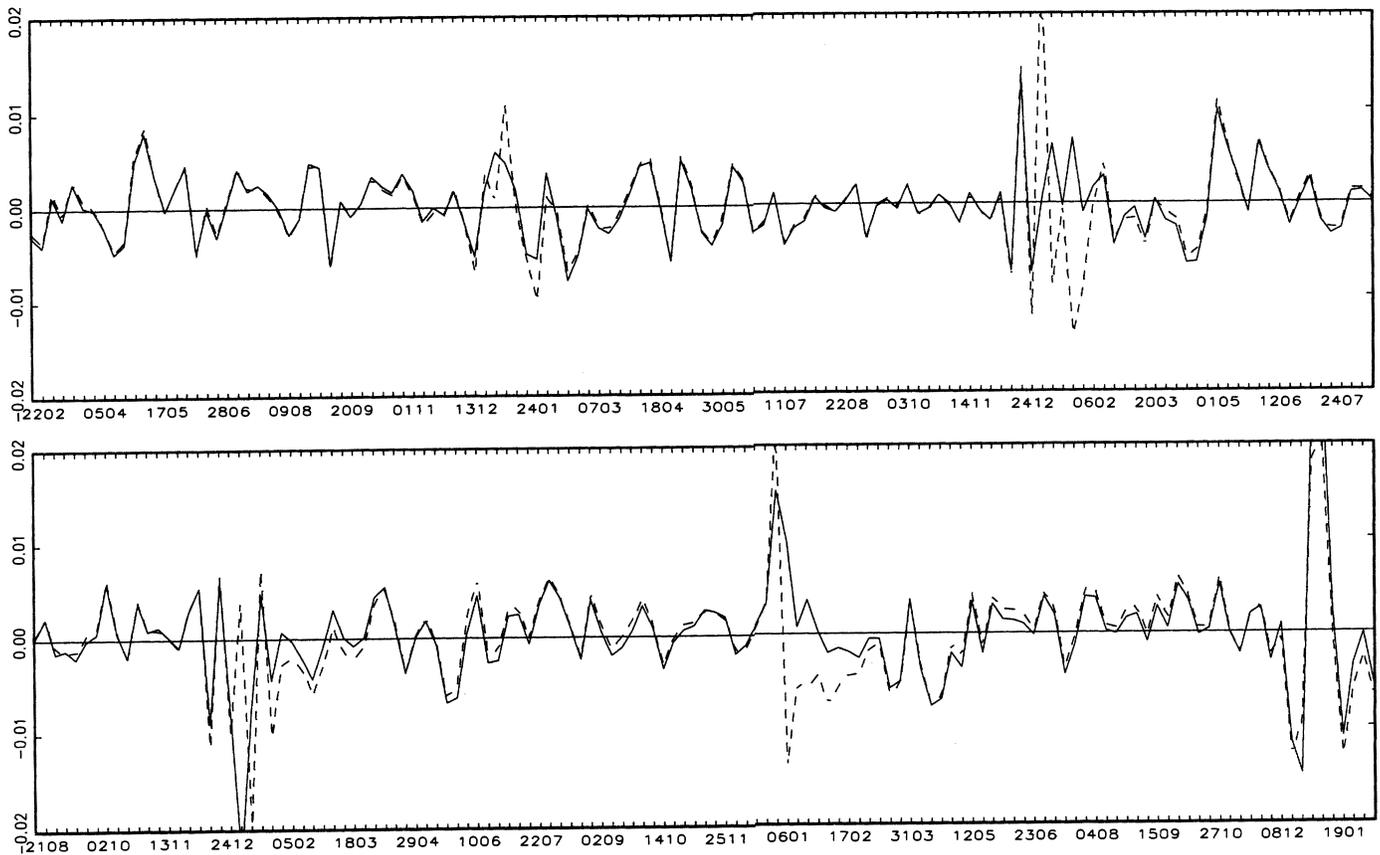


Figure 3. Prediction Errors Over the Last Five Years (broken line is when periodic variance not doubled at Christmas).

as in (1.5) with \mathbf{z}_t being an $(s - 1) \times 1$ vector that yields the effect of the current month. Thus if the s th element of γ_t has been dropped from the state vector, \mathbf{z}_t has a 1 in position j for month $j, j = 1, \dots, s - 1$, and zeros elsewhere, and all elements equal to -1 for month s .

1.3 Weekly Data

The features of weekly data noted in the first section mean that the preceding approach cannot be applied directly. Our solution, like that of Pierce et al. (1984), is to model the seasonal pattern using two components. The first component is a function of the date in the year—that is, the number of days that have passed in the year. Thus, for example, it takes a particular value on day 358 (which happens to be Christmas Eve). The second component is a collection of effects associated with public holidays, such as Easter, that take place on different dates in different years but always fall on the same day of the week. Once these components have been specified as deterministic effects so that they could be handled by regression, it is straightforward to allow them to evolve stochastically over time by casting the whole model in state-space form. This is a con-

siderable advantage over the approach adopted by Pierce et al. (1984) in which a stochastic ARIMA component is added to a regression component with no clear connection between the two.

In what follows, we will refer to the first seasonal effect, γ_t , as the *periodic* component and to the second, θ_t , as the *moving festival* component. It is possible to include an additional periodic component if there is a significant intramonthly effect.

Although our model is formulated on a daily basis, we could go further and set up the evolution of the parameters in continuous time. This is quite natural because the periodic component is a continuous function. Although continuous time is an elegant approach, it makes little or no difference to the form of the implied weekly models, however, and its use in the present context should be clear from the general discussion of Harvey (1989, chap. 9).

2. PERIODIC EFFECTS

We wish to model the yearly pattern on a daily basis. For the moment, we will assume that there are no leap years, so each year has 365 days.

The periodic component will be modeled as a linear function of a set of parameters contained in a $g \times 1$ vector γ . If these parameters are fixed, the periodic pattern is fixed, and we may write the periodic effect for the d th day in the year as

$$\gamma_d = \mathbf{z}'_d \gamma, \quad d = 1, \dots, 365, \quad (2.1)$$

Table 1. Estimates of Hyperparameters

Model	q_ζ	q_ν	q_x	σ_ϵ^2
Constant periodic variance	.0149	.0052	.0483	947×10^{-8}
Doubled at Christmas	.0109	.0063	.0401	794×10^{-8}

Table 2. Final State Vector

Description	State	Estimate	t ratio
Level slope	μ_T	9.7814	2,521
	β_T	.0012	1.50
Moving-festival dummies	$\theta_{1,T}$.0122	4.84
	$\theta_{2,T}$.0249	9.80
	$\theta_{3,T}$.0058	2.31
	$\theta_{4,T}$.0120	5.04
	$\theta_{5,T}$.0224	9.22
	$\theta_{6,T}$.0034	1.39
	$\theta_{7,T}$.0188	7.45
	$\theta_{8,T}$.0079	3.09
	$\theta_{9,T}$.0360	13.9
	$\theta_{10,T}$.0368	14.1
	$\theta_{11,T}$.0043	1.66
Coefficients of knots	$\gamma_{1,T}$.0574	13.3
	$\gamma_{2,T}$.0131	3.48
	$\gamma_{3,T}$	-.0200	-7.33
	$\gamma_{4,T}$	-.0336	-9.82
	$\gamma_{5,T}$	-.0281	-6.50
	$\gamma_{6,T}$	-.0189	-4.45
	$\gamma_{7,T}$	-.0025	-.68
	$\gamma_{8,T}$.0054	1.36
	$\gamma_{9,T}$.0177	4.90
	$\gamma_{10,T}$.0034	.92
	$\gamma_{11,T}$	-.0013	-.34
	$\gamma_{12,T}$	-.0068	-1.78
	$\gamma_{13,T}$	-.0051	-1.38
	$\gamma_{14,T}$.0007	.20
	$\gamma_{15,T}$.0298	7.71
	$\gamma_{16,T}$.0540	13.8
	$\gamma_{17,T}$.1040	25.1
	$\gamma_{18,T}$.1058	13.7
Silver Jubilee	ϕ_T	.0122	7.57

where \mathbf{z}_d is a $g \times 1$ vector of known values. The idea is to specify (2.1) so as to have g reasonably small—one hopes much less than 52. There are essentially two options. The first is to let γ_d be a mixture of trigonometric functions. The second is to model it by a periodic spline. In our application the second option seems to offer more scope for a parsimonious parameterization, mainly because of the need to capture the sharp peak at Christmas. The important point, however, is that both approaches can be generalized to allow the seasonal pattern to evolve over time by letting γ be stochastic. Stochastic trigonometric seasonals have long been a part of structural time series modeling methods. Stochastic, or time-varying, periodic splines were first used by Harvey and Koopman (1993) to model intraweekly patterns of hourly electricity demand.

Further scope for cutting down on the number of parameters may be afforded if there is an intramonthly pattern. Again either trigonometric terms or splines may be used.

2.1 Trigonometric Seasonality

A fixed annual pattern may be represented by a trigonometric model as follows:

$$\gamma_d = \sum_{j=1}^k (\gamma_j \cos \lambda_j d + \gamma_j^* \sin \lambda_j d), \quad d = 1, \dots, 365, \quad (2.2)$$

where $\lambda_j = 2\pi j/365$. To include the full set of trigonometric terms, as is normally the case with a monthly or

quarterly model, would mean setting $k = 182$. Pierce et al. (1984) found, however, that setting $k = 8$ is perfectly adequate when combined with intramonthly effects.

Now suppose that the periodic pattern changes over time on a daily basis, irrespective of whether there has been an observation. Each trigonometric component now evolves as in (1.3) and (1.4) with the λ_j 's specified as in (2.2) and the subscript t denoting the seasonal effect on the t th day from the beginning of the sample. Thus the model is modified to

$$\gamma_t = \sum_{j=1}^{[s/z]} \gamma_{j,t}, \quad t = 1, 2, \dots, T_d,$$

where T_d denotes the number of days covered by the sample period. When $\sigma_\omega^2 = 0$, the deterministic model is obtained.

2.2 Periodic Time-Varying Splines

To set up a spline we need to choose h knots in the range $[0, 365]$. Then

$$\gamma_d = \mathbf{w}'_d \boldsymbol{\gamma}^\dagger, \quad d = 1, \dots, 365, \quad (2.3)$$

where \mathbf{w}_d is an $h \times 1$ vector that depends on the position of the knots and is defined in such a way as to ensure continuity of the spline from one year to the next—that is, make it periodic; see the Appendix and Poirier (1976, pp. 43–47). To have the periodic seasonal effects summing to 0 over the year, one of the elements in $\boldsymbol{\gamma}^\dagger$, say the last one, is dropped. Then, in terms of the formulation in (2.1), $\boldsymbol{\gamma}$ consists of the first $g = h - 1$ elements of $\boldsymbol{\gamma}^\dagger$, and the i th element in \mathbf{z}_d is given by

$$z_{di} = w_{di} - w_{dh} w_{*i} / w_{*h}, \quad i = 1, \dots, g, \quad d = 1, \dots, 365, \quad (2.4)$$

where w_{*i} is the i th element of the vector

$$\mathbf{w}_* = \sum_{d=1}^{365} \mathbf{w}_d. \quad (2.5)$$

Note that it is the effects summed over all the days in the year that come to 0 rather than the effects summed over the particular days when there are observations. (If we want to regard d as continuous, then \mathbf{w}_* is an integral; this can be evaluated in practice by summing over many points or by using the formula at the end of Appendix A.)

The splines can be allowed to evolve over time by letting the parameters follow random walks. If we assume that the parameters change every day, irrespective of whether or not there is an observation, we may write

$$\boldsymbol{\gamma}^\dagger_t = \boldsymbol{\gamma}^\dagger_{t-1} + \boldsymbol{\chi}_t, \quad t = 1, 2, \dots, T_d, \quad (2.6)$$

where $\boldsymbol{\chi}_t$ is an $h \times 1$ vector of serially uncorrelated random disturbances with zero mean and covariance matrix

$$\text{var}(\boldsymbol{\chi}_t) = \sigma_\chi^2 (\mathbf{I} - (1/\mathbf{w}'_* \mathbf{w}_*) \mathbf{w}_* \mathbf{w}'_*), \quad (2.7)$$

where σ_χ^2 is the variance parameter that governs the speed with which the spline can change. This covariance matrix

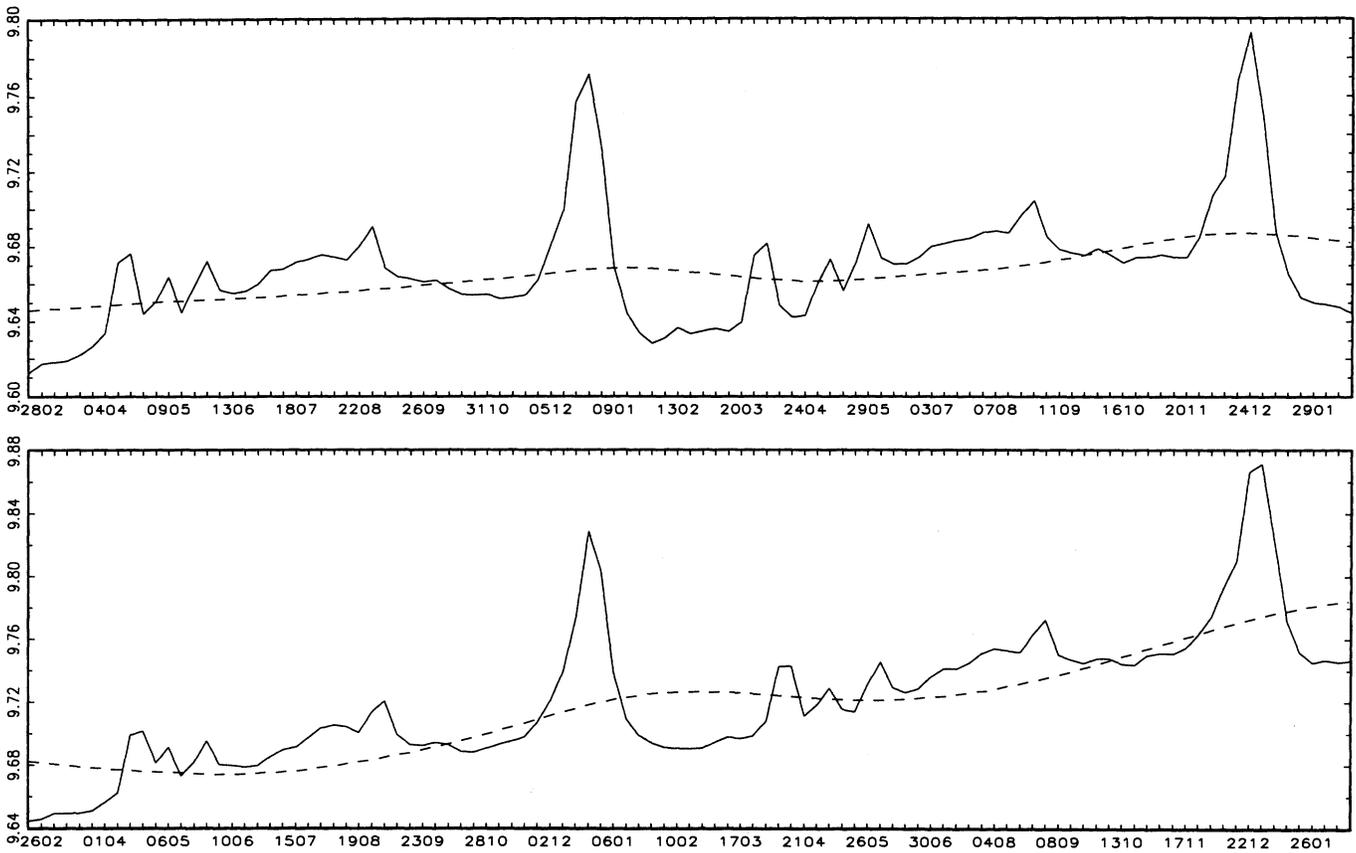


Figure 4. Smoothed Estimate of Trend.

enforces the constraint that $\mathbf{w}'_*\chi_t = 0$. Note that, if there is a knot for each day, the seasonal dummy model of (1.6) is obtained with $h = s$ and $\mathbf{w}_* = \mathbf{i}$.

As before, one of the elements of γ_t^\dagger can be dropped to give a $g \times 1$ vector γ_t . The effect in the t th day from the beginning of the sample is then

$$\gamma_t = \mathbf{z}'_{t(d)}\gamma_t, \quad t = 1, \dots, T_d,$$

where the notation $t(d)$ for the subscript of \mathbf{z} stresses dependence on the day of the year.

In some circumstances, a part of the periodic pattern may change more rapidly than the rest of the pattern. For money demand, this seems to be the case with the Christmas effect. This phenomenon can be modeled by letting the parameters at the knots close to the points at which rapid changes take place be subject to relatively larger disturbances. Thus, suppose that the first m elements in γ_t^\dagger have associated with them a variance of σ_m^2 , and the second $n = h - m$ have a larger variance, σ_n^2 . The covariance matrix of χ_t then becomes

$$\begin{aligned} \text{var}(\chi_t) &= \begin{bmatrix} \sigma_m^2 \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \sigma_n^2 \mathbf{I}_n \end{bmatrix} \\ &- \frac{1}{\mathbf{w}'_*\mathbf{w}_*} \begin{bmatrix} \sigma_m^2 \mathbf{w}_m \mathbf{w}'_m & \frac{1}{2}(\sigma_m^2 + \sigma_n^2) \mathbf{w}_m \mathbf{w}'_n \\ \frac{1}{2}(\sigma_m^2 + \sigma_n^2) \mathbf{w}_n \mathbf{w}'_m & \sigma_n^2 \mathbf{w}_n \mathbf{w}'_n \end{bmatrix}, \end{aligned} \tag{2.8}$$

where \mathbf{w}_m consists of the first m elements of \mathbf{w}_* and \mathbf{w}_n contains the last n ; that is, $\mathbf{w}_* = (\mathbf{w}'_m, \mathbf{w}'_n)'$. This additional flexibility is an attractive feature of the spline formulation. Again, $\mathbf{w}'_*\chi_t = 0$.

2.3 Intramonthly Effects

Pierce et al. (1984) observed significant intramonthly effects in U.S. monetary aggregates, primarily due to the higher money supply toward the end of the month when wages are paid. Such effects were captured by the inclusion of trigonometric terms as in (2.2) but with d denoting the day of the month and 365 replaced by the number of days in the month. An intramonthly pattern of this form can be made time-varying exactly as in Subsection 2.1. An additional hyperparameter is needed to fulfil the role of σ_ω^2 .

It is only worth using intramonthly effects if most of the months display a similar pattern. This may well be a reasonable assumption for monetary aggregates, although December may be different if people tend to be paid before the Christmas break.

Intramonthly effects can also be modeled by a time-varying periodic spline. A trigonometric intramonthly component can be used together with an intrayearly spline and vice versa.

2.4 Leap Years

There are two ways to handle leap years. The first is to set the periodic effect for February 29 the same as for

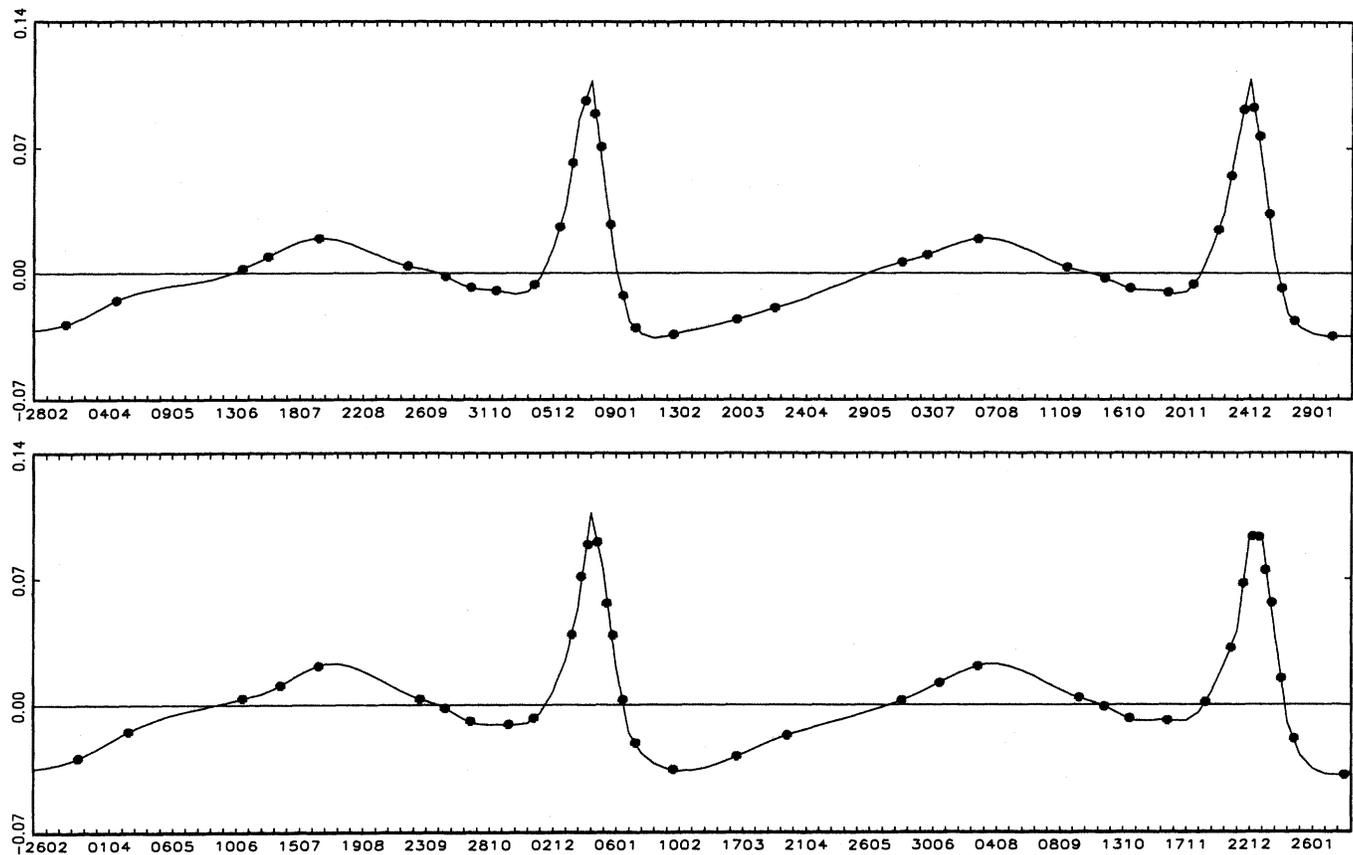


Figure 5. Smoothed Periodic Component Over the Last Four Years, Starting in February 1990, With Position of Knots (•).

February 28—that is, to regard day 59 as occurring twice. By proceeding in this way we ensure that Christmas falls at exactly the same point every year—that is, day 359. Note that day 59 must be counted twice in the summation in (2.5).

A slightly different approach is to let the leap-year effect be spread throughout the whole year. For the trigonometric model, this is easily accomplished by replacing 365 by 366 in the λ_j 's. For the spline, we modify w_d , and hence $z_{t(d)}$, by multiplying the knot positions by 366/365.

3. MOVING FESTIVALS: VARIABLE-DUMMY EFFECTS

The effect of each public holiday may be modeled by a set of dummy variables that are assigned to the surrounding weeks. The day of the year on which the holiday falls, and hence the days on which the surrounding observations fall, depends on the calendar.

Suppose that m dummy variables are used to pick up public-holiday effects. Each effect takes up seven days. Thus the number of days remaining is, averaging over four years,

$$k_{m+1} = 365.25 - 7m. \quad (3.1)$$

These days must be allocated an effect to counterbalance the effect of the public holidays, thereby making the component sum to 0. Averaging over four years avoids a slight end-of-year discontinuity associated with leap years. Thus, if $\theta_1, \dots, \theta_m$ denote the holiday effects, the nonholiday factor

must be

$$\theta_{m+1} = -(\theta_1 + \dots + \theta_m)7/k_{m+1}. \quad (3.2)$$

To allow the dummy-variable effects to evolve over time, we let them follow constrained random walks as in Subsection 2.4. There is no need to include θ_{m+1} because it may be inferred from (3.2). Thus, following the treatment of the daily-effects model as set out by Harvey (1989, pp. 43–44),

$$\theta_{jt} = \theta_{j,t-1} + v_{jt}, \quad j = 1, \dots, m, \quad t = 1, \dots, T_d, \quad (3.3)$$

where v_{jt} is a zero mean, serially uncorrelated disturbance with variance

$$\text{var}(v_j) = \sigma_v^2(1 - 49/K), \quad j = 1, \dots, m, \quad (3.4)$$

where $K = 49m + k_{m+1}^2$. The covariances between disturbances are given by

$$\text{cov}(v_j v_l) = -\sigma_v^2 49/K, \quad j, l = 1, \dots, m. \quad (3.5)$$

The model may be generalized to allow some effects to change more rapidly by giving them a larger variance.

4. STATISTICAL TREATMENT OF THE MODEL

The full daily model is

$$y_t = \mu_t + \gamma_t + \theta_t + \varepsilon_t, \quad t = 1, 2, \dots, T_d, \quad (4.1)$$

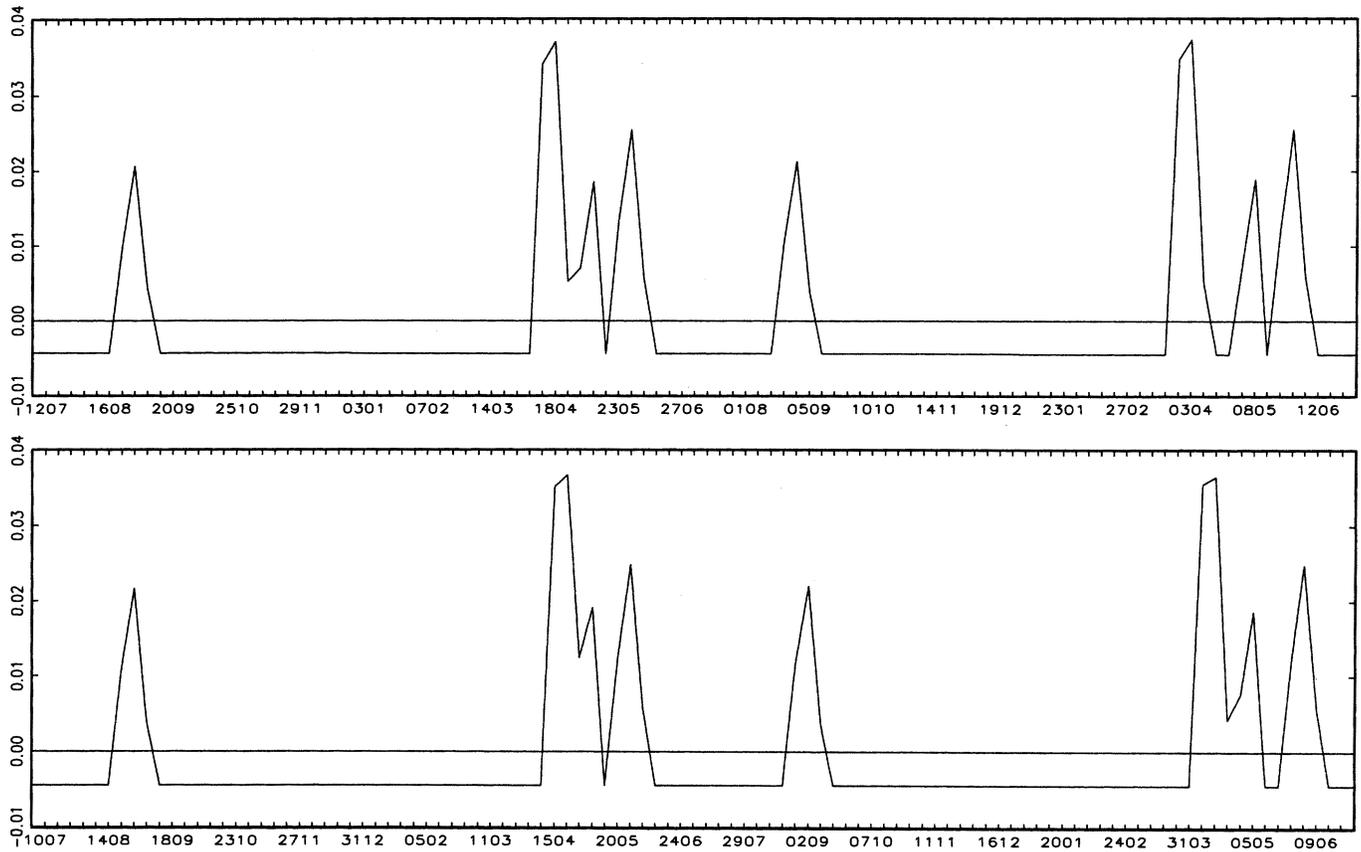


Figure 6. Smoothed Moving-Festival Component Over the Last Four Years, Starting in July 1989.

with the trend defined as in (1.2) and γ_t and θ_t denoting the periodic and moving-festival effects. The irregular term, ε_t , is assumed to be white noise, and the disturbances in the different components are uncorrelated with each other. The model is easily put into state-space form by letting the state vector be $\alpha_t = (\mu_t, \beta_t, \gamma_t, \theta_t)'$. The transition equation is made up of (1.2), (2.6), and (3.3), and the measurement equation is

$$y_t = (1, 0, \mathbf{z}'_{t(d)}, \mathbf{x}'_{t(c)})\alpha_t + \varepsilon_t, \quad t = 1, \dots, T_d, \quad (4.2)$$

where $\mathbf{z}_{t(d)}$ depends on the number of days that have passed in the year and $\mathbf{x}_{t(c)}$ depends on the calendar. The role of $\mathbf{x}_{t(c)}$ is to pick out from the variable dummy vector, θ_t , the appropriate element or elements if there is no direct holiday effect and (3.2) is relevant.

The preceding formulation is independent of the observations. These can be weekly, which is the focus of attention here, or they can arrive on various days with no particular pattern. When there is no observation on a particular day, the Kalman filter simply treats it as a missing observation: There is no difficulty in carrying out prediction, smoothing, and estimation. The hyperparameters—that is, the variances of the disturbances—can be estimated by maximizing the (exact) log-likelihood function computed via the Kalman filter using the prediction error decomposition; see Appendix B. The use of a “square root” filter is recommended because it appears to be much more stable for weekly data. Numer-

ical optimization needs to be carried out with respect to the hyperparameters relative to the variance of the irregular, which can be concentrated out of the likelihood function.

With weekly data, the observations are, for the most part, equally spaced. It is therefore more efficient to convert the model to a weekly basis. If y_τ denotes the observation in week τ of the sample, we can write $y_\tau = \mu_\tau + \gamma_\tau + \theta_\tau + \varepsilon_\tau, \tau = 1, 2, \dots, T$, and the transition equation is modified appropriately. For parameters evolving according to random walks, as in (2.6) and (3.3), all that needs to be done is to observe that the variance for a weekly model will be seven times the variance for a daily model. For the local linear trend, the modification to the covariance matrix of the trend disturbances, η_t and ζ_t in (1.2), was given by Harvey (1989, p. 312). In the case of the trigonometric formulation, the frequencies must be multiplied by 7, and if there is an intramonthly effect, it is necessary to take account of the fact that different months may have different numbers of days. There are occasions in which a figure is not recorded on the usual day of the week due to a holiday. In such cases it is straightforward to modify the state-space formulation to make allowance for the different time intervals involved. This generally involves multiplying disturbance variances (and frequencies, if relevant) by a factor of $p/7$, where p is the number of days since the last observation.

Estimates of the various components in the model using all the observations can be computed by smoothing. The algorithm devised by Koopman (1993) allows smoothing to be carried out with computational efficiency without

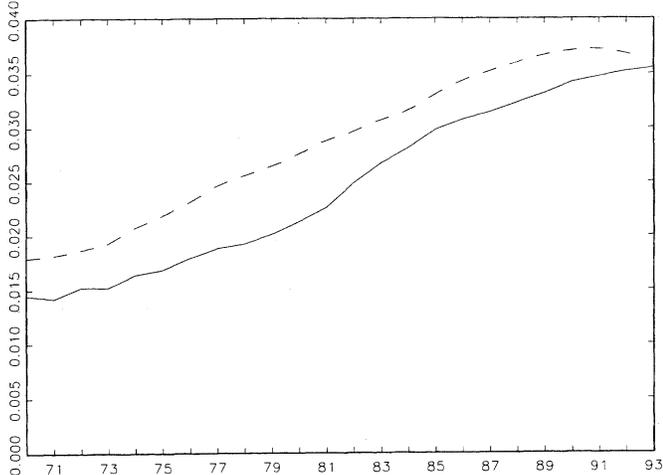


Figure 7. Smoothed Estimates of the Effect of Easter: Wednesday Before (solid line); Wednesday After (dashed line).

excessive storage requirements; see Appendix B. Smoothing forms the basis for seasonal adjustment because all that needs to be done is to remove the periodic and moving-festival smoothed seasonal components. Note that the best estimate of the seasonal effect at the end of the series is given by filtering, so a seasonally adjusted figure can be provided as each new observation becomes available. This figure can subsequently be revised as further observations become available.

5. U.K. MONEY SUPPLY

To get some idea of the seasonal pattern in the M0 series shown in Figure 1, it was subject to simple detrending using the Hodrick–Prescott filter. This filter can be obtained very easily as the smoother for a local linear-trend model—that is, (1.1) and (1.2) without the seasonal—in which $\sigma_{\eta}^2 = 0$ and $\sigma_{\zeta}^2/\sigma_{\varepsilon}^2 = .000625$. A plot of the observations in each year shows clear and permanent changes in the seasonal pattern. This is particularly marked at Christmas. Figure 2, page 356, shows the pattern of detrended observations over the Christmas period when Christmas falls on a Friday and on a Saturday. At the beginning of the sample, the peak is about 5% above the trend, whereas at the end it is about 10%. The same features appear when Christmas falls on other days of the week.

The periodic component was modeled using a time-varying spline. A good deal of experimentation was carried out in positioning the knots and dummy variables. To capture the peaks at times such as Christmas, a relatively large number of knots are needed in a short period. At other times, the seasonal pattern changes quite slowly and only a few knots are needed. Similar considerations applied in modeling the intradaily electricity demand of Harvey and Koopman (1993). Here the situation is more complicated because the interaction between the positioning of the dummy variables needed to capture the moving festivals and the knots used to pick up the rest of the seasonal pattern. The final specification had 19 knots and was decided by factors such as the “*t* ratios” of the knot coordinates and dummies, diagnostics and residual plots, goodness-of-fit statistics, and forecasting performance. Increasing the

Table 3. Diagnostics: Residual Autocorrelations

Model	r_1	r_2	r_3	r_{52}	r_{53}
Constant periodic variance	.06	.08	-.13	.31	.01
Doubled at Christmas	.23	-.16	-.02	.09	-.02

number of knots gives a better fit and reduces the residual serial correlation at lags 1 and 2 and at the annual lag of 52 (and 53). The less smooth the pattern is and the more knots are included, however, the less easy it is to distinguish the periodic pattern from the moving-festival pattern.

All moving public holidays fall on Mondays, except for Good Friday, and the moving-festival dummy variables were specified as follows:

1. Easter—the two weeks before and the week after
2. May Day—the week before and the week after (from 1978)
3. Spring Bank Holiday—two weeks before and the week after
4. August Bank Holiday—two weeks before and the week after

No restrictions were put on these holiday effects, although this is easily done. For example, the same state variable could be used for the Spring and August Bank Holidays. Thus, there are 11 stochastic dummy variables in the state vector. An additional dummy was included in June 1977 to allow for the special holiday for the Queen’s Silver Jubilee.

No evidence was found for a significant intramonthly effect. A smooth trend—that is, σ_{η}^2 set to 0—was preferred because it was not much affected by the seasonal pattern.

The residuals exhibit considerable variability around Christmas. Because of the importance of Christmas and the speed with which the pattern can change, we found that a better model could be obtained by increasing the variance of the disturbances driving the movements in the knots around Christmas; see (2.8). When we doubled the variance of the Christmas knots, we found that the residuals close to Christmas were much more akin to residuals in other parts of the year. Furthermore, the unstandardized prediction errors were also smaller around Christmas; see Figure 3, page 357. (It could be argued that one of the reasons the Christmas effect changes so rapidly is because it is different for Christmas falling on different days of the week. We were unable to capture such an effect by additional dummies; indeed, given that each day occurs only three or four times in our sample, this may be impossible to do. An examination of Fig. 2 and the plots for other days, however, indicates that the evolution over time far outweighs any possible day-of-the-week effect.)

Table 4. Diagnostics: Box–Ljung Statistics

Model	$Q(6)$	$Q(26)$	$Q(53)$	$Q(125)$
Constant periodic variance	16.78	40.88	130.8	271.0
Doubled at Christmas	27.29	55.40	111.2	231.4

NOTE: $Q(P)$ is based on first P residual autocorrelations.

Table 5. Diagnostics: Normality Test Statistics

Model	Skewness	Kurtosis	Bowman-Shenton
Constant periodic variance	10.82	93.11	103.9
Doubled at Christmas	1.00	7.94	9.94

Table 1, page 357, shows the estimated hyperparameters for the specification with and without the doubling of the periodic variance at Christmas. The q 's denote hyperparameters relative to the variance of the irregular. Because of the sharp change in the trend in the late 1970s, it turned out to be more satisfactory to drop the first 400 observations in estimating the hyperparameters. They were retained for all other purposes, however. For the reasons given in the previous paragraph and confirmed in the following discussion, the doubled variance model is our preferred specification. Table 2, page 358, shows the estimates of the state for this model at the end of the sample, together with their t ratios—that is, the estimates divided by the corresponding root mean squared errors. In assessing the relative importance of the various estimates from their t ratios, it must be remembered that they are liable to change over time. Thus, although some knots are not significant at the end of the sample, they may have been in the past.

Figure 4, page 359, graphs the smoothed trend. Its relatively slow changes seem to be quite suitable for this series. Figure 5, page 360, shows the smoothed pattern of the periodic component over the last four years, and Figure 6, page 361, shows the moving-festival component. Note that there are very few knots in periods when there are public holidays—for example, in April and May. Small changes in both periodic and moving-festival patterns are apparent even over a short period of time. Figure 7 illustrates the much more dramatic changes that can take place over a longer period. It shows the evolution of the dummies in the weeks before and after Easter over the full sample. As with the Christmas effect, there is a doubling effect, with a movement from around 1.5% of the underlying level to over 3%.

The equation standard error, s , which is the square root of the one-step-ahead prediction-error variance, is normally used as a measure of goodness of fit, but here there is a problem because the nature of the model means that it changes over time and never goes to a steady state. A rough idea of the size of s can be gauged from Figure 3.

Residual serial correlation can be assessed by the autocorrelations at lag τ , denoted $r(\tau)$, and the Box-Ljung statistics, $Q(P)$, based on the first P autocorrelation. Tables 3 and 4 report these statistics for the last five years. It seems to be difficult to eliminate serial correlation completely unless many knots are used, and our preference is to keep the number of knots reasonably small. The model with constant periodic variance shows quite strong serial correlation at lag 52, and the plot of the residual correlogram in Figure 8(a) confirms the impression of some residual seasonal effects. Figure 8(b) indicates that this feature is eliminated in the model with the periodic variance doubled around Christ-

mas. There is now more serial correlation, however, at lags 1 and 2. If this were felt to be of any practical importance, it could be removed by letting the disturbance follow a low-order autoregressive moving average process.

Table 5 reports the skewness and kurtosis moment test statistics and the Bowman-Shenton normality test statistic. When the model is correctly specified with Gaussian disturbances, the skewness and kurtosis statistics are asymptotically distributed as chi squares with 1 df, whereas Bowman-Shenton has a chi-squared distribution with 2 df; see Harvey (1989, chap. 5). The extremely high kurtosis is due to the Christmas effect, but it is reduced to a reasonable level when the periodic variance is doubled around Christmas.

The fact that the model is successful comes out in the predictions over the last few years. Figure 9 shows the one-step-ahead predictions obtained by filtering. This effectively gives the same information as the prediction-error plot in Figure 3, but it brings home more clearly how accurate the predictions are, being less than .5% of the level most of the time. Even more impressive is Figure 10, which shows the multistep predictions made from September 9, 1992. Overall the model is successful in providing a relatively parsimonious representation of the data.

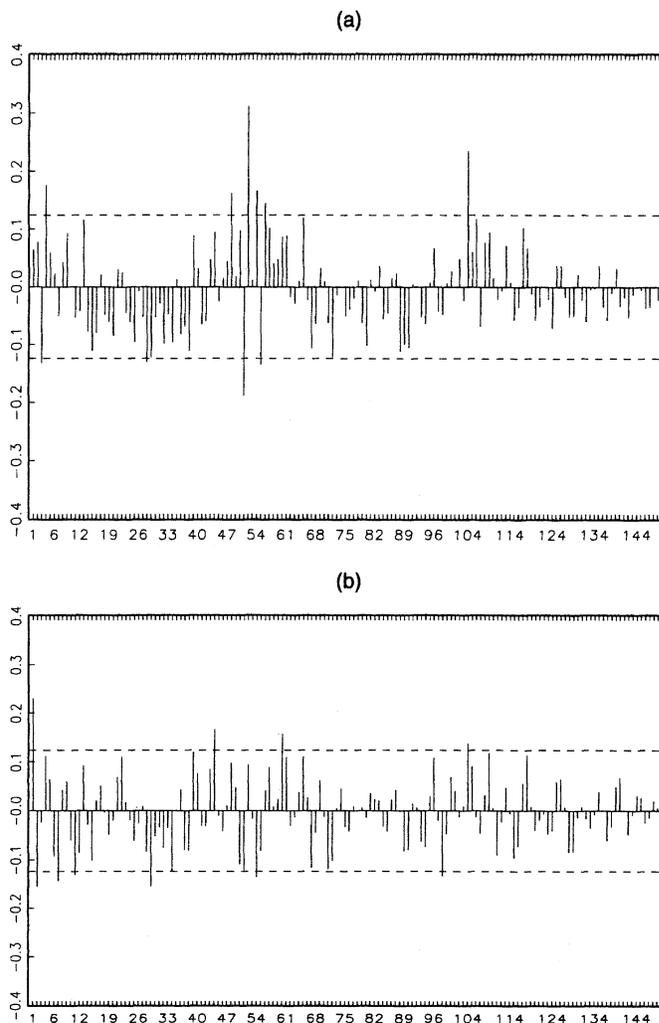


Figure 8. Residual Correlogram: (a) Constant Periodic Variance; (b) Variance Doubled Around Christmas.

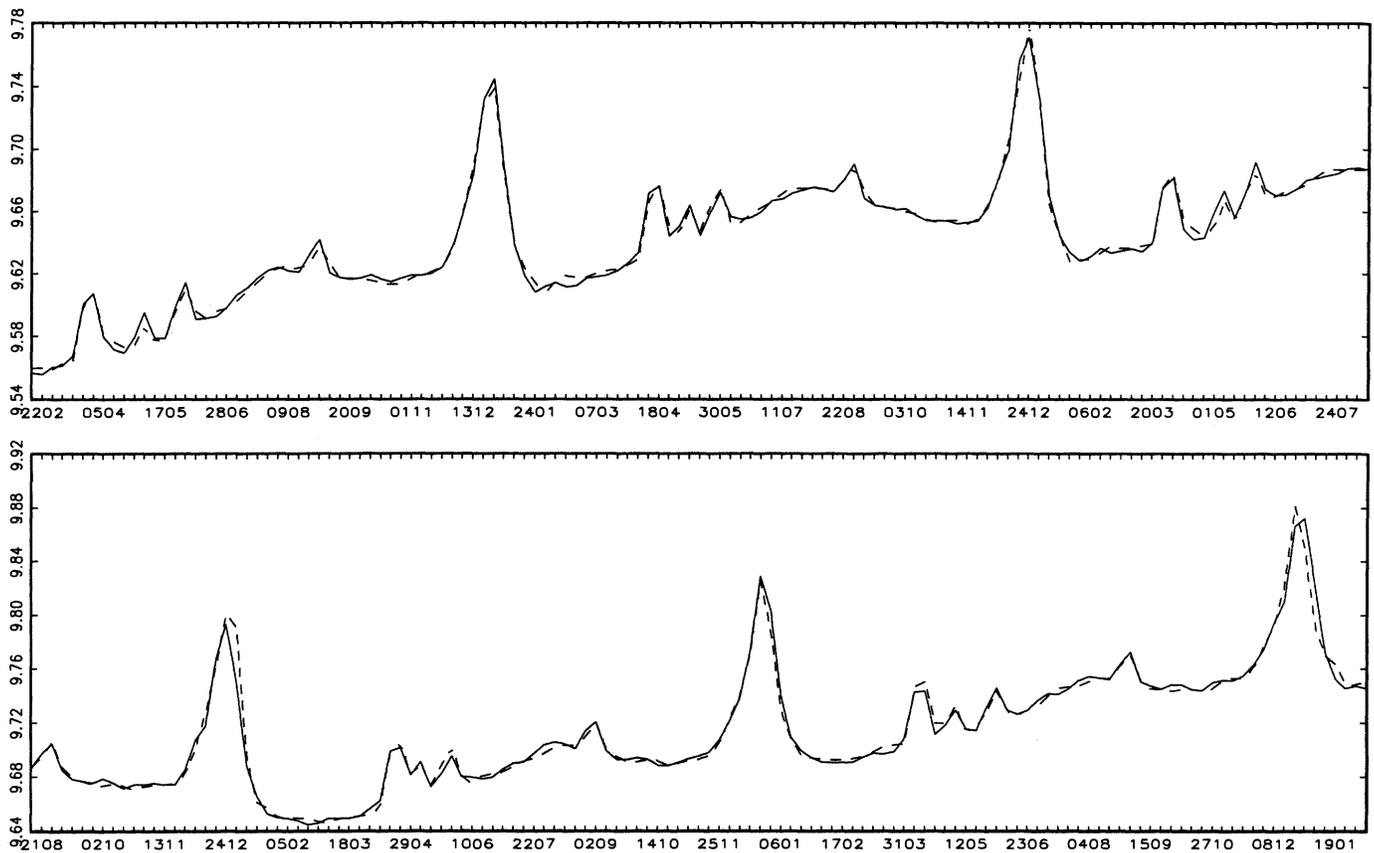


Figure 9. One-Step-Ahead Predictions (---) and Actual Values (—).

6. CONCLUSIONS

This article has set out a method of building a time series model for weekly observations. The key feature of the model is the setting up of the seasonal component in terms of a periodic component and a movable dummy component. The former is parsimoniously modeled with a periodic spline, though a trigonometric formulation could also be adopted and might be preferable in different circumstances. In our application the moving-festival component was primarily needed to deal with observations about Easter. Note, however, that, in countries where carnival is celebrated, the position of Easter may have important effects in February

or early March. In other countries, other moving festivals, such as Chinese New Year, may be more relevant.

The structural time series approach has the advantage that once a regression formulation has been found it can be extended to allow the effects to evolve over time. In other words, a deterministic component can be generalized so that it becomes stochastic. Furthermore, it is possible to build in constraints that ensure that the forecasts of the seasonal component sum to 0 over a year, thereby ensuring that there is no confounding of trend and seasonal effects. Once such a model has been formulated, statistical handling via the state-space form is relatively straightforward. When the parameters have been estimated, seasonal

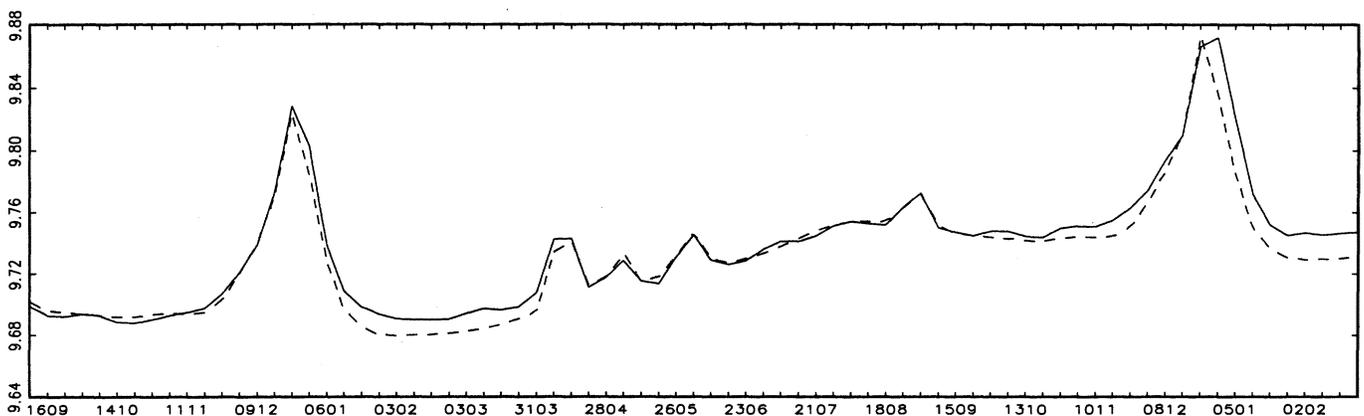


Figure 10. Multistep Predictions From September 9, 1992.

adjustment can be carried out by subtracting the smoothed periodic and moving-festival components. This seasonally adjusted series can be continuously updated as new observations arrive.

The preceding approach is fairly general and should be applicable, with minor modifications, to other weekly series and indeed to other types of series in which data arrive at regular or irregular intervals throughout the year.

ACKNOWLEDGMENTS

We thank John Thorp and Andrew Derry at the Bank of England, for supplying the data and related information, and two referees and Andrew Scott for helpful comments. We are grateful to the Economic and Social Research Council for financial support as part of our project on Interrelations in Economic Time Series, grant R000 23 5330.

APPENDIX A: PERIODIC CUBIC SPLINES

This appendix presents the derivation of a cubic spline function $g(x)$ that approximates an unknown function $f(x)$. Periodic cubic splines are subjected to special end-conditions. The cubic spline is defined as a set of polynomial functions differentiable up to order 3 with the first two derivatives continuous. The set of coordinates $(x_i^\dagger, y_i^\dagger)$, with $i = 0, 1, \dots, k$, will be referred to as the set of knot points associated with the mesh $(x_0^\dagger, \dots, x_k^\dagger)$. In this appendix, the knot coordinates are treated as fixed and known. Section 2.2 describes how the y_i^\dagger 's are estimated when they are allowed to change over time. It will be shown that the cubic spline function is derived as a set of linear equations; see also Poirier (1976).

Let us assume that the mesh is in ascending order,

$$x_0^\dagger < x_1^\dagger < \dots < x_k^\dagger. \quad (\text{A.1})$$

Let $h_j = x_j^\dagger - x_{j-1}^\dagger$, $j = 1, \dots, k$, and let $d_j(x)$ and $d_j^2(x)$ denote the first and second derivatives, respectively, of the spline function $g(x)$ calculated in the interval $[x_{j-1}^\dagger, x_j^\dagger]$. The cubic spline function is a mixture of polynomials of order 3, so its second derivative within $[x_{j-1}^\dagger, x_j^\dagger]$ will be linear; that is,

$$d_j^2(x) = \frac{x_j^\dagger - x}{h_j} a_{j-1} + \frac{x - x_{j-1}^\dagger}{h_j} a_j, \quad j = 1, \dots, k. \quad (\text{A.2})$$

The scalar a_j is the second derivative of the spline function at the knot point x_j^\dagger ; that is, $a_j = d_j^2(x_j^\dagger)$. Starting from (A.2), the cubic spline and the periodic spline are easily derived by the following steps:

1. Expressions for the first derivative function and the cubic spline function $g(x)$ are obtained via standard rules of integration. The spline function is forced to cross the knots; that is, $g_j(x_j^\dagger) = y_j^\dagger$ and $g_j(x_{j-1}^\dagger) = y_{j-1}^\dagger$. This leads

to

$$g_j(x) = (x_j^\dagger - x) \frac{[(x_j^\dagger - x)^2 - h_j^2]}{6h_j} a_{j-1} + \frac{x_j^\dagger - x}{h_j} y_{j-1}^\dagger + (x - x_{j-1}^\dagger) \frac{[(x - x_{j-1}^\dagger)^2 - h_j^2]}{6h_j} a_j + \frac{x - x_{j-1}^\dagger}{h_j} y_j^\dagger. \quad (\text{A.3})$$

and

$$d_j(x) = - \left[\frac{1}{2} \frac{(x_j^\dagger - x)^2}{h_j} - \frac{h_j}{6} \right] a_{j-1} + \left[\frac{1}{2} \frac{(x - x_{j-1}^\dagger)^2}{h_j} - \frac{h_j}{6} \right] a_j \quad (\text{A.4})$$

with $x_{j-1}^\dagger \leq x \leq x_j^\dagger$, $j = 1, 2, \dots, k$.

2. The spline function (A.3) can be expressed in vector notation via

$$g_j(x) = \mathbf{r}'_j \mathbf{y}^\dagger + \mathbf{s}'_j \mathbf{a}, \quad (\text{A.5})$$

where $\mathbf{a} = (a_0, \dots, a_k)'$, $\mathbf{y}^\dagger = (y_0^\dagger, \dots, y_k^\dagger)'$, and the vectors \mathbf{r}_j and \mathbf{s}_j are equal to 0 except for their j th and $(j+1)$ th elements, which correspond to the appropriate weights of (A.3).

3. The continuity restriction also applies to the first derivative (A.4), $d_j(x_j^\dagger) = d_{j+1}(x_j^\dagger)$. After some minor manipulations, we obtain a set of $k-1$ linear restrictions

$$\begin{aligned} & \frac{h_j}{h_j + h_{j+1}} a_{j-1} + 2a_j + \frac{h_{j+1}}{h_j + h_{j+1}} a_{j+1} \\ &= \frac{6y_{j-1}^\dagger}{h_j(h_j + h_{j+1})} - \frac{6y_j^\dagger}{h_j h_{j+1}} + \frac{6y_{j+1}^\dagger}{h_{j+1}(h_j + h_{j+1})}, \\ & j = 1, 2, \dots, k-1. \end{aligned} \quad (\text{A.6})$$

4. A system of $k-1$ linear equations with $k+1$ unknowns a_j cannot be solved unless two linear restrictions are added. Poirier (1976) suggested setting $a_0 = a_k = 0$, which he defined as a *natural* condition for a spline. With these two additional constraints the system of equations (A.6) can be represented in matrix notation as

$$\mathbf{P}\mathbf{a} = \mathbf{Q}\mathbf{y}^\dagger, \quad (\text{A.7})$$

where

$$\mathbf{P} = \begin{pmatrix} 2 & 0 & 0 & \cdots & 0 & 0 \\ \frac{h_1}{h_1+h_2} & 2 & \frac{h_2}{h_1+h_2} & \cdots & 0 & 0 \\ 0 & \frac{h_2}{h_2+h_3} & 2 & \cdots & 0 & 0 \\ 0 & 0 & \frac{h_3}{h_3+h_4} & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & \frac{h_k}{h_{k-1}+h_k} \\ 0 & 0 & 0 & \cdots & 0 & 2 \end{pmatrix} \tag{A.8}$$

and

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ \frac{6}{h_1(h_1+h_2)} & -\frac{6}{h_1 h_2} & \frac{6}{h_2(h_1+h_2)} & \cdots & 0 & 0 \\ 0 & \frac{6}{h_2(h_2+h_3)} & -\frac{6}{h_2 h_3} & \cdots & 0 & 0 \\ 0 & 0 & \frac{6}{h_3(h_3+h_4)} & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{6}{h_{k-1} h_k} & \frac{6}{h_k(h_{k-1}+h_k)} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{A.9}$$

5. The matrices \mathbf{P} and \mathbf{Q} have dimension $(k + 1) \times (k + 1)$. Note that the first and the last rows of \mathbf{P} and \mathbf{Q} represent the *natural* constraints $a_0 = a_k = 0$. The solution for \mathbf{a} is

$$\mathbf{a} = \mathbf{P}^{-1} \mathbf{Q} \mathbf{y}^\dagger \tag{A.10}$$

so that (A.5) becomes

$$g_j(x) = \mathbf{w}'_j \mathbf{y}^\dagger, \tag{A.11}$$

with

$$\mathbf{w}'_j = \mathbf{r}'_j + \mathbf{s}'_j \mathbf{P}^{-1} \mathbf{Q}. \tag{A.12}$$

This shows that, given a set of k knots with a particular vector \mathbf{y}^\dagger , the *natural* spline $g(x)$ can be calculated for any $x_0^\dagger \leq x \leq x_k^\dagger$. Note that if x equals x_j^\dagger , $0 \leq j \leq k$, vector \mathbf{w}'_j is $\mathbf{0}$ except for its j th element, which equals unity and hence $g(x_j^\dagger) = y_j^\dagger$.

6. The cubic spline becomes periodic when the first and the last knots are restricted to be the same. The continuity is enforced by letting the corresponding first and second derivatives to be the same as well; that is,

$$y_0^\dagger = y_k^\dagger, \quad d_1(x_0^\dagger) = d_k(x_k^\dagger), \quad d_1^2(x_0^\dagger) = d_k^2(x_k^\dagger). \tag{A.13} \quad \text{and}$$

The second restriction of (A.13) implies

$$\begin{aligned} & \frac{h_k}{h_1 + h_k} a_{k-1} + 2a_k + \frac{h_1}{h_1 + h_k} a_1 \\ &= \frac{6y_{k-1}^\dagger}{h_k(h_1 + h_k)} - \frac{6y_k^\dagger}{h_1 h_k} + \frac{6y_1^\dagger}{h_1(h_1 + h_k)}, \end{aligned} \tag{A.14}$$

which can be added to the set of linear restrictions (A.6). The last restriction of (A.13) sets $a_0 = a_k$. Therefore, the *natural* constraints are no longer required because we now have a system of k linear restrictions and k unknowns. The $(k \times k)$ matrices \mathbf{P}_p and \mathbf{Q}_p for a periodic spline become

$$\mathbf{P}_p = \begin{pmatrix} 2 & \frac{h_2}{h_1+h_2} & 0 & \cdots & 0 & \frac{h_1}{h_1+h_2} \\ \frac{h_2}{h_2+h_3} & 2 & \frac{h_3}{h_2+h_3} & \cdots & 0 & 0 \\ 0 & \frac{h_3}{h_3+h_4} & 2 & \cdots & 0 & 0 \\ 0 & 0 & \frac{h_4}{h_4+h_5} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & \frac{h_k}{h_{k-1}+h_k} \\ \frac{h_1}{h_1+h_k} & 0 & 0 & \cdots & \frac{h_k}{h_1+h_k} & 2 \end{pmatrix} \tag{A.15}$$

$$\mathbf{Q}_p = \begin{pmatrix} -\frac{6}{h_1 h_2} & \frac{6}{h_2(h_1+h_2)} & 0 & \dots & 0 & \frac{6}{h_1(h_1+h_2)} \\ \frac{6}{h_2(h_2+h_3)} & -\frac{6}{h_2 h_3} & \frac{6}{h_3(h_2+h_3)} & \dots & 0 & 0 \\ 0 & \frac{6}{h_3(h_3+h_4)} & -\frac{6}{h_3 h_4} & \dots & 0 & 0 \\ 0 & 0 & \frac{6}{h_4(h_4+h_5)} & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\frac{6}{h_{k-1} h_k} & \frac{6}{h_k(h_{k-1}+h_k)} \\ \frac{6}{h_1(h_1+h_k)} & 0 & 0 & \dots & \frac{6}{h_k(h_1+h_k)} & -\frac{6}{h_1 h_k} \end{pmatrix}. \tag{A.16}$$

Furthermore, the $(k \times 1)$ vectors $\mathbf{a}, \mathbf{y}, \mathbf{r}_j$, and $\mathbf{s}_j, j = 1, \dots, k$, are adjusted appropriately, also with respect to the first restriction of (A.13).

In the main body of the article vector \mathbf{w}_* is defined in (2.5) as

$$\mathbf{w}_* = \sum_{j=1}^k \mathbf{w}_j, \tag{A.17}$$

where $\mathbf{w}_j, j = 1, \dots, k$, must be computed even if the evaluation of the spline at position j is not required. Therefore, when k is large, the calculation of \mathbf{w}_* may become a computational burden. An alternative expression for \mathbf{w}_* is

$$\mathbf{w}'_* = \mathbf{r}'_* + \mathbf{s}'_* \mathbf{P}_p^{-1} \mathbf{Q}_p, \tag{A.18}$$

where $\mathbf{r}_* = \sum_{j=1}^k \mathbf{r}_j$ and $\mathbf{s}_* = \sum_{j=1}^k \mathbf{s}_j$ can be analytically derived as

$$\mathbf{r}_* = \left(\frac{x_2^\dagger - x_0^\dagger}{2}, \dots, \frac{x_k^\dagger - x_{k-2}^\dagger}{2}, \frac{x_1^\dagger - x_0^\dagger + x_k^\dagger - x_{k-1}^\dagger}{2} \right)', \tag{A.19}$$

and

$$\mathbf{s}_* = \left(\frac{x_2^\dagger - x_0^\dagger - h_2^3 - h_1^3}{24}, \dots, \frac{x_k^\dagger - x_{k-2}^\dagger - h_k^3 - h_{k-1}^3}{24}, \frac{h_1(1-h_1^2) + h_k(1-h_k^2)}{24} \right)'. \tag{A.20}$$

APPENDIX B: STATE-SPACE METHODS AND SEASONAL ADJUSTMENT

The full model (4.1) can be put in state-space form; that is,

$$\begin{aligned} y_t &= \mathbf{Z}_t \boldsymbol{\alpha}_t + \mathbf{G}_t \mathbf{u}_t, & \mathbf{u}_t &\sim NID(\mathbf{0}, \mathbf{I}), \\ \boldsymbol{\alpha}_{t+1} &= \mathbf{T}_t \boldsymbol{\alpha}_t + \mathbf{H}_t \mathbf{u}_t, & t &= 1, \dots, T, \end{aligned} \tag{B.1}$$

where \mathbf{Z}_t is given in (4.2) and the matrices $\mathbf{T}_t, \mathbf{G}_t$, and \mathbf{H}_t are constructed straightforwardly. Note that the matrices \mathbf{G}_t and \mathbf{T}_t are time-invariant, but matrix \mathbf{H}_t is time-varying

because the variances of the periodic effect are forced to increase around the Christmas period. The disturbance vector \mathbf{u}_t is the stack of all disturbances associated with the unobserved components of the model including the irregular ε_t . The initial state $\boldsymbol{\alpha}_1$ is treated as a random vector generated from a diffuse distribution.

The Kalman filter, or its more stable counterpart the ‘‘square root’’ Kalman filter, is used for the evaluation of the likelihood and for the calculation of one-step-ahead prediction errors. The Kalman filter is given by

$$\begin{aligned} v_t &= y_t - \mathbf{Z}_t \mathbf{a}_{t|t-1}, \\ f_t &= \mathbf{Z}_t \mathbf{P}_{t|t-1} \mathbf{Z}_t' + \mathbf{G}_t \mathbf{G}_t', \\ \mathbf{K}_t &= \mathbf{T}_t \mathbf{P}_{t|t-1} \mathbf{Z}_t' f_t^{-1}, \\ \mathbf{a}_{t+1|t} &= \mathbf{T}_t \mathbf{a}_{t|t-1} + \mathbf{K}_t v_t, \\ \mathbf{P}_{t+1|t} &= \mathbf{T}_t \mathbf{P}_{t|t-1} \mathbf{T}_t' - f_t \mathbf{K}_t \mathbf{K}_t' + \mathbf{H}_t \mathbf{H}_t', \end{aligned} \tag{B.2}$$

with the initializations $\mathbf{a}_{1|0} = \mathbf{0}$ and $\mathbf{P}_{1|0} = \kappa \mathbf{I}$, where κ is a suitably chosen large number. The vector $\mathbf{a}_{t|t-1}$ is the one-step-ahead prediction of the state $\boldsymbol{\alpha}_t$ with its mean squared error matrix $\mathbf{P}_{t|t-1}$. The one-step-ahead prediction error and its variance are given by v_t and f_t , respectively. The vector \mathbf{K}_t is referred to as the Kalman gain. Usually, the Kalman filter is computationally not very demanding but Model (4.1) requires a large state vector that leads to a computational effort with respect to $\mathbf{P}_{t|t-1}$. Of course, the computations take longer as the number of observations increases.

A state smoothing algorithm is designed to compute full-sample estimates of the state vector. The estimated trend and periodic components can be extracted from the smoothed state vector, and they can be graphically reproduced as part of a validation procedure of the estimated model. Seasonal adjustment procedures remove seasonal and periodic variation from the observed series and, therefore, in the context of state-space models, they require a

Table B1. Computational Performance of Smoothing

Algorithm	Multiplications	Storage	Seconds
De Jong (1989)	1,024	594	14.7
Koopman (1993)	62	34	1.8

state smoother. Smoothing algorithms are computationally expensive, especially when the state vector is large, and they require much storage space because a selection of Kalman filter quantities needs to be stored for $t = 1, \dots, T$.

We consider two different state smoothers, proposed by De Jong (1989) and Koopman (1993). Table B.1 reports the number of multiplications for each time index t , the number of values to be stored for each time index t , and total computer time required for smoothing a series of 1,500 observations on a Pentium processor with clock speed 90 MHz. The model considered is (4.1) with 20 knots for the periodic spline and 10 stochastic dummy variables for the moving-festival effects. The results show dramatically that the smoothing algorithm of Koopman (1993) outperforms the algorithm of De Jong (1989) with respect to all indicators. Koopman's algorithm does not give the mean squared errors of the smoothed estimators of the state vector, but for many applications this is not necessary. Koopman's smoothing algorithm is a two-step approach. First, a backward disturbance smoother is applied; that is, $e_t = v_t/f_t - \mathbf{K}'_t \mathbf{r}_t$, $\mathbf{r}_{t-1} = \mathbf{Z}'_t e_t + \mathbf{T}'_t \mathbf{r}_t$, $t = T, \dots, 1$, with initialization $\mathbf{r}_T = \mathbf{0}$. The Kalman filter only needs to store the scalar v_t/f_t and the vector \mathbf{K}_t , $t = 1, \dots, T$. The storage space can be overwritten to store the vector $\mathbf{n}_t = \mathbf{H}_t \mathbf{H}'_t \mathbf{r}_t$. Second, the forward recursion $\mathbf{a}_{t+1|T} = \mathbf{T}_t \mathbf{a}_{t|T} + \mathbf{n}_t$, $t = 1, \dots, T-1$, must be used with initialization $\mathbf{a}_{1|T} = \mathbf{a}_{1|0} + \mathbf{P}_{1|0} \mathbf{r}_0$. The storage space can be overwritten by the smoothed state vector $\mathbf{a}_{t|T}$.

[Received May 1995. Revised March 1996.]

REFERENCES

- Bell, W. R., and Hillmer, S. C. (1983), "Modeling Time Series With Calendar Variation," *Journal of the American Statistical Association*, 78, 526-534.
- De Jong, P. (1989), "Smoothing and Interpolation With the State Space Model," *Journal of the American Statistical Association*, 84, 1085-1088.
- (1991), "The Diffuse Kalman Filter," *The Annals of Statistics*, 19, 1073-1083.
- Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, U.K.: Cambridge University Press.
- Harvey, A. C., and Koopman, S. J. (1993), "Forecasting Hourly Electricity Demand Using Time-Varying Splines," *Journal of the American Statistical Association*, 88, 1228-1236.
- Hillmer, S. C., and Tiao, G. C. (1982), "An ARIMA-Model-Based Approach to Seasonal Adjustment," *Journal of the American Statistical Association*, 77, 63-70.
- Kitagawa, G., and Gersch, W. (1984), "A Smoothness Priors-State Space Modeling of Time Series With Trend and Seasonality," *Journal of the American Statistical Association*, 79, 378-389.
- Koopman, S. J. (1993), "Disturbance Smoother for State Space Models," *Biometrika*, 80, 117-126.
- Maravall, A. (1985), "On Structural Time Series Models and the Characterization of Components," *Journal of Business & Economic Statistics*, 3, 350-355.
- Pierce, D. A., Grupe, M. R., and Cleveland, W. P. (1984), "Seasonal Adjustment of the Weekly Monetary Aggregates: A Model-Based Approach," *Journal of Business & Economic Statistics*, 2, 260-270.
- Poirier, D. (1976), *The Econometrics of Structural Change: With Special Emphasis on Spline Functions*, Amsterdam: North-Holland.
- Scott, A. (1995), "Why Is Consumption so Seasonal?" CEP discussion paper, London School of Economics.