8 Illustrations of the use of the linear model

8.1 Introduction

In this chapter we will give some illustrations which show how the use of the linear model works in practice. State space methods are usually employed for time series problems and most of our examples will be from this area but we also will treat a smoothing problem which is not normally regarded as part of time series analysis and which we solve using cubic splines.

The first example is an analysis of road accident data to estimate the reduction in car drivers killed and seriously injured in the UK due to the introduction of a law requiring the wearing of seat belts. In the second example we consider a bivariate model in which we include data on numbers of front seat passengers killed and seriously injured and on numbers of rear seat passengers killed and seriously injured and we estimate the effect that the inclusion of the second variable has on the accuracy of the estimation of the drop in the first variable. The third example shows how state space methods can be applied to Box–Jenkins ARMA models employed to model series of users logged onto the Internet. In the fourth example we consider the state space solution to the spline smoothing of motorcycle acceleration data. The fifth example provides a dynamic factor analysis based on the linear Gaussian model for the term structure of interest rates paid on US Treasury securities. The software we have used for most of the calculations is *SsfPack* and is described in Section 6.7.

8.2 Structural time series models

The study by Durbin and Harvey (1985) and Harvey and Durbin (1986) on the effect of the seat belt law on road accidents in Great Britain provides an illustration of the use of structural time series models for the treatment of problems in applied time series analysis. They analysed data sets which contained numbers of casualties in various categories of road accidents to provide an independent assessment on behalf of the Department of Transport of the effects of the British seat belt law on road casualties. Most series were analysed by means of linear Gaussian state space models. We concentrate here on monthly numbers of drivers, front seat passengers and rear seat passengers who were killed or seriously injured in road accidents in cars in Great Britain from January 1969 to December 1984. Data were transformed into logarithms since logged values

fitted the model better. Data on the average number of kilometres travelled per car per month and the real price of petrol are included as possible explanatory variables. We start with a univariate analysis of the drivers series. In the next section we perform a bivariate analysis using the front and rear seat passengers.

The log of monthly number of car drivers killed or seriously injured is displayed in Fig. 8.1. The graph shows a seasonal pattern which may be due to weather conditions and festive celebrations. The overall trend of the series is basically constant over the years with breaks in the mid-1970s, probably due to the oil crisis, and in February 1983 after the introduction of the seat belt law. The model that we shall consider initially is the basic structural time series model which is given by

$$y_t = \mu_t + \gamma_t + \varepsilon_t,$$

where μ_t is the local level component modelled as the random walk $\mu_{t+1} = \mu_t + \xi_t$, γ_t is the trigonometric seasonal component (3.7) and (3.8), and ε_t is a disturbance term with mean zero and variance σ_ε^2. Note that for illustrative purposes we do not at this stage include an intervention component to measure the effect of the seat belt law.

The model is estimated by maximum likelihood using the techniques described in Chapter 7. The iterative method of finding the estimates for σ_ε^2,

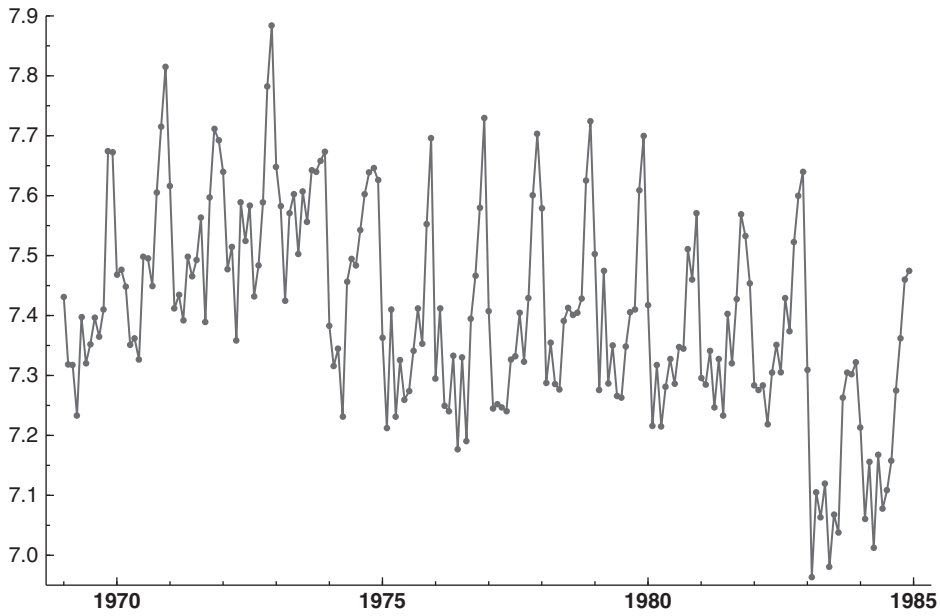

Fig. 8.1 Monthly numbers (logged) of drivers who were killed or seriously injured (KSI) in road accidents in cars in Great Britain.

σ_ξ^2 and σ_ω^2 is implemented in *STAMP* 8.3 based on the concentrated diffuse loglikelihood as discussed in Subsection 2.10.2; the estimation output is given below where the first element of the parameter vector is $\phi_1 = 0.5 \log q_\eta$ and the second element is $\phi_2 = 0.5 \log q_\omega$ where $q_\xi = \sigma_\xi^2/\sigma_\varepsilon^2$ and $q_\omega = \sigma_\omega^2/\sigma_\varepsilon^2$. We present the parameter estimation results obtained from the *STAMP* package of Koopman, Harvey, Doornik and Shephard (2010). The estimates are obtained from a small number of cycles of univariate optimisations with respect to one parameter and for which the other parameters are kept fixed to their current values, starting with arbitrary values for the univariate optimisations. The initial parameter estimates obtained from this procedure are usually good starting values for the simultaneous optimisation procedure that produces the maximum likelihood estimates. In this procedure, one variance parameter is concentrated out. The resulting parameter estimates are given below. The estimate for σ_ω^2 is very small but when it is set equal to zero, evidence of seasonal serial correlation is found in the residuals. Therefore we keep σ_ω^2 equal to its estimated value.

```
Estimated variances of disturbances

Component                    Value    (q-ratio)
Level                  0.000935852    ( 0.2740)
Seasonal              5.01096e-007  (0.0001467)
Irregular               0.00341598     ( 1.000)
```

The estimated components are displayed in Fig. 8.2. The estimated level does pick up the underlying movements in the series and the estimated irregular does not cause much concern to us. The seasonal effect hardly changes over time.

In Fig. 8.3 the estimated level is displayed; the predicted estimator is based on only the past data, that is $\mathrm{E}(\mu_t|Y_{t-1})$, and the smoothed estimator is based on all the data, that is $\mathrm{E}(\mu_t|Y_n)$. It can be seen that the predicted estimator lags the shocks in the series as is to be expected since this estimator does not take account of current and future observations.

The model fit of this series and the basic diagnostics initially appear satisfactory. The standard output provided by *STAMP* is given by

```
Diagnostic summary report

Estimation sample is 69. 1 - 84.12. (T = 192, n = 180).
Log-Likelihood is 435.295 (-2 LogL = -870.59).
Prediction error variance is 0.00586717
Summary statistics
            drivers
Std.Error  0.076597
N            4.6692
H( 60)       1.0600
r( 1)      0.038621
Q(24,22)     33.184
```

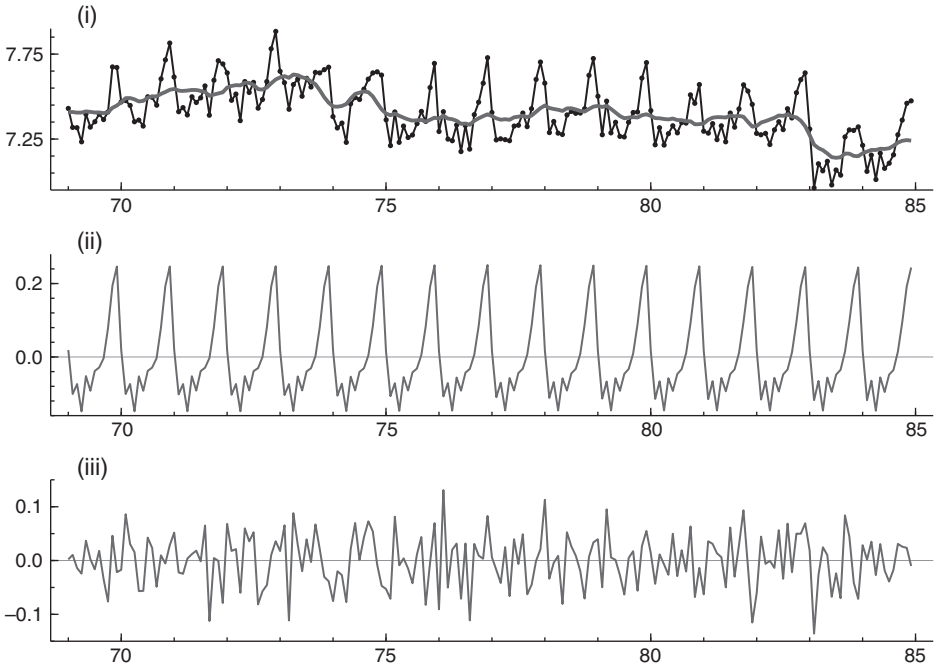

Fig. 8.2 Estimated components: (i) level; (ii) seasonal; (iii) irregular.

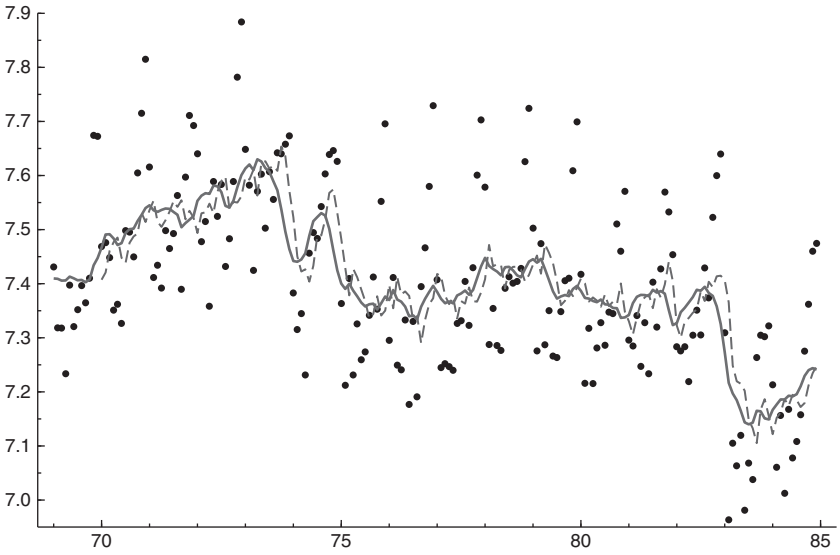

Fig. 8.3 Data (dots) with predicted (dashed line) and smoothed (solid line) estimated level.

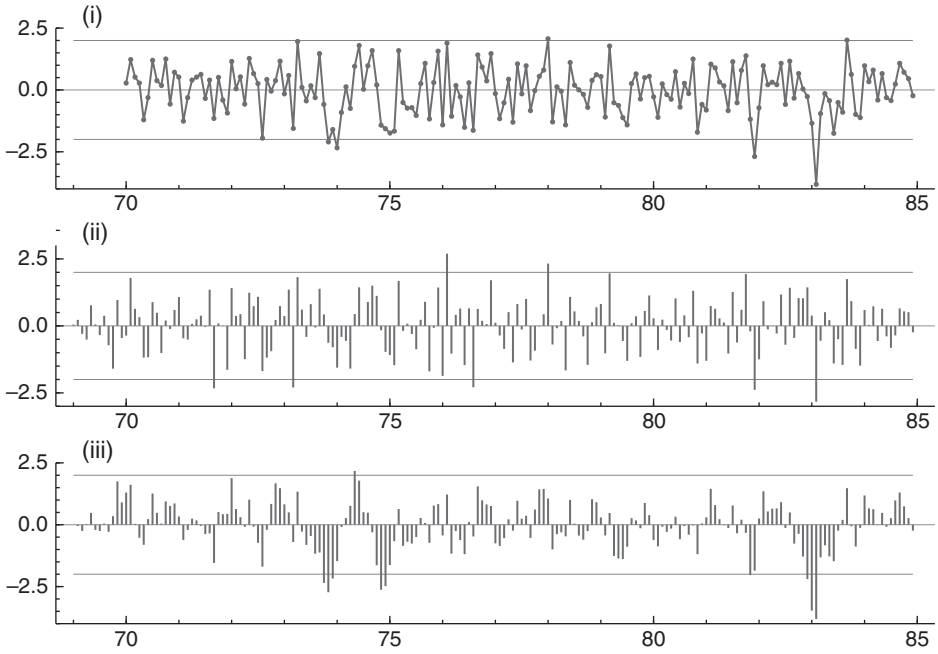

Fig. 8.4 (i) The one-step ahead prediction residuals (time series plot); (ii) auxiliary irregular residuals (index plot); (iii) auxiliary level residuals (index plot).

The definitions of the diagnostics can be found in Section 2.12.

When we inspect the graphs of the residuals in Fig. 8.4, however, in particular the auxiliary level residuals, we see a large negative value for February 1983. This suggests a need to incorporate an intervention variable to measure the level shift in February 1983. We have performed this analysis without inclusion of such a variable purely for illustrative purposes; obviously, in a real analysis the variable would be included since a drop in casualties was expected to result from the introduction of the seat belt law.

By introducing an intervention which equals one from February 1983 and is zero prior to that and the price of petrol as a further explanatory variable, we re-estimate the model and obtain the regression output

```
Estimated coefficients of explanatory variables.

Variable   Coefficient  R.m.s.e.  t-value
petrol     -0.29140     0.09832   -2.96384 [ 0.00345]
Lvl 83. 2  -0.23773     0.04632   -5.13277 [ 0.00000]
```

The estimated components, when the intervention variable and the regression effect due to the price of petrol are included, are displayed in Fig. 8.5.

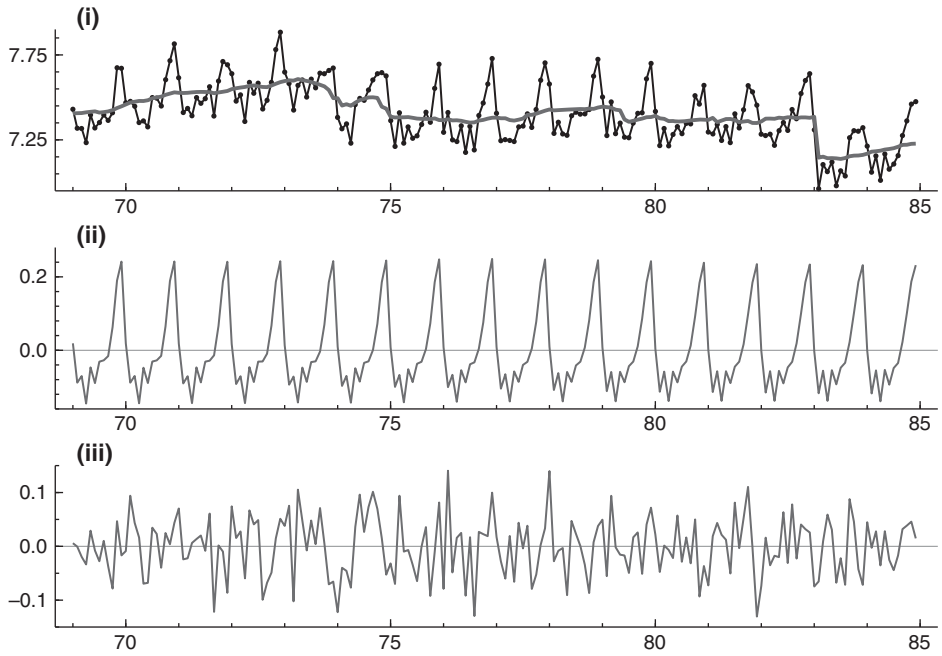

Fig. 8.5 Estimated components for model with intervention and regression effects: (i) level; (ii) seasonal; (iii) irregular.

The estimated coefficient of a break in the level after January 1983 is -0.238, indicating a fall of 21%, that is, $1 - \exp(-0.238) = 0.21$, in car drivers killed and seriously injured after the introduction of the law. The t-value of -5.1 indicates that the break is highly significant. The coefficient of petrol price is also significant.

8.3 Bivariate structural time series analysis

Multivariate structural time series models are introduced in Section 3.3. To illustrate state space methods for a multivariate model we analyse a bivariate monthly time series of front seat passengers and rear seat passengers killed and seriously injured in road accidents in Great Britain which were included in the assessment study by Harvey and Durbin (1986).

The graphs in Fig. 8.6 indicate that the local level specification is appropriate for the trend component and that we need to include a seasonal component. We start by estimating a bivariate model with level, trigonometric seasonal and irregular components together with explanatory variables for the number of kilometres travelled and the real price of petrol. We estimate the model only using

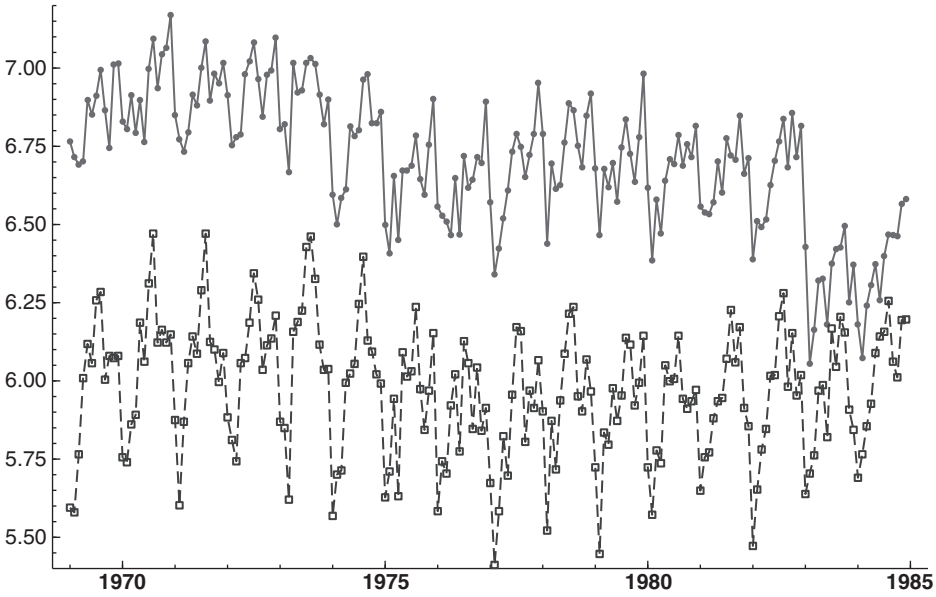

Fig. 8.6 Front seat (grey, dotted line) and rear seat (dashed line with squares) passengers killed and seriously injured in road accidents in Great Britain.

observations available before 1983, the year in which the seat belt law was introduced. The variance matrices of the three disturbance vectors are estimated by maximum likelihood in the way described in Section 7.3. The estimated variance matrix of the seasonal component is small which lead us to model the seasonal component as fixed and to re-estimate the remaining two variances matrices:

$$\widehat{\Sigma}_\varepsilon = 10^{-4}\begin{bmatrix}5.006 & 4.569\\ 4.569 & 9.143\end{bmatrix},\ \widehat{\Sigma}_\eta = 10^{-5}\begin{bmatrix}4.834 & 2.993\\ 2.993 & 2.234\end{bmatrix},\ \hat{\rho}_\varepsilon = 0.675,\ \hat{\rho}_\eta = 0.893,$$

where $\rho_x = \Sigma_x(1,2)\,/\,\sqrt{\Sigma_x(1,1)\,\Sigma_x(2,2)}$ for $x = \varepsilon, \eta$ and where $\Sigma_x(i,j)$ is the (i,j) element of matrix Σ_x for $i,j = 1,2$. The loglikelihood value of the estimated model is 742.088 with AIC equal to -4.4782.

The correlation between the two level disturbances is close to one. It may therefore be interesting to re-estimate the model with the restriction that the rank of the level variance matrix is one:

$$\widehat{\Sigma}_\varepsilon = 10^{-4}\begin{bmatrix}5.062 & 4.791\\ 4.791 & 10.02\end{bmatrix},\ \widehat{\Sigma}_\eta = 10^{-5}\begin{bmatrix}4.802 & 2.792\\ 2.792 & 1.623\end{bmatrix},\ \hat{\rho}_\varepsilon = 0.673,\ \rho_\eta = 1,$$

The loglikelihood value of this estimated model is 739.399 with AIC equal to -4.4628. A comparison of the two AIC's shows only a marginal preference for the unrestricted model.

We now assess the effect of the introduction of the seat belt law as we have done for the drivers series using a univariate model in Section 8.2. We concentrate on the effect of the law on front seat passengers. We also have the rear seat series which is highly correlated with the front seat series. However, the law did not apply to rear seat passengers so the data should therefore not be affected by the introduction of the law. Under such circumstances the rear seat series may be used as a *control group* which may result in a more precise measure of the effect of the seat belt law on front seat passengers; for the reasoning behind this idea see the discussion by Harvey (1996) to whom this approach is owed.

We consider the unrestricted bivariate model but with a level intervention for February 1983 added to both series. This model is estimated using the whole data set giving the parameter estimates:

$$\widehat{\Sigma}_\varepsilon = 10^{-4}\begin{bmatrix}5.135 & 4.493\\ 4.493 & 9.419\end{bmatrix},\ \widehat{\Sigma}_\eta = 10^{-5}\begin{bmatrix}4.896 & 3.025\\ 3.025 & 2.317\end{bmatrix},\ \hat{\rho}_\varepsilon = 0.660,\ \hat{\rho}_\eta = 0.898,$$

The estimates for the level intervention coefficient in both equation are given by

	coeff	rmse	t-value	p-value
front seat	−0.32799	0.05699	−5.75497	0.0000
rear seat	0.03376	0.05025	0.67189	0.50252

From these results and the time series plot of casualties in rear seat passengers in Fig. 8.6, it is clear that they are unaffected by the introduction of the seat belt as we expect. By removing the intervention effect from the rear seat equation of the model we obtain the estimation results:

$$\widehat{\Sigma}_\varepsilon = 10^{-4}\begin{bmatrix}5.147 & 4.588\\ 4.588 & 9.380\end{bmatrix},\ \widehat{\Sigma}_\eta = 10^{-5}\begin{bmatrix}4.754 & 2.926\\ 2.926 & 2.282\end{bmatrix},\ \hat{\rho}_\varepsilon = 0.660,\ \hat{\rho}_\eta = 0.888,$$

with the level intervention estimate given by

	coeff	rmse	t-value	p-value
front seat	−0.35630	0.03655	−9.74877	0.0000

The almost two-fold decrease of the root mean squared error for the estimated intervention coefficient for the front seat series is remarkable. Enforcing the rank of Σ_η to be one, such that the levels are proportional to each other, produces the following estimation results:

$$\widehat{\Sigma}_\varepsilon = 10^{-4}\begin{bmatrix}5.206 & 4.789\\ 4.789 & 10.24\end{bmatrix},\ \widehat{\Sigma}_\eta = 10^{-5}\begin{bmatrix}4.970 & 2.860\\ 2.860 & 1.646\end{bmatrix},\ \hat{\rho}_\varepsilon = 0.659,\ \rho_\eta = 1,$$

with level intervention estimate given by

	coeff	rmse	t-value	p-value
front seat	−0.41557	0.02621	−15.85330	0.0000

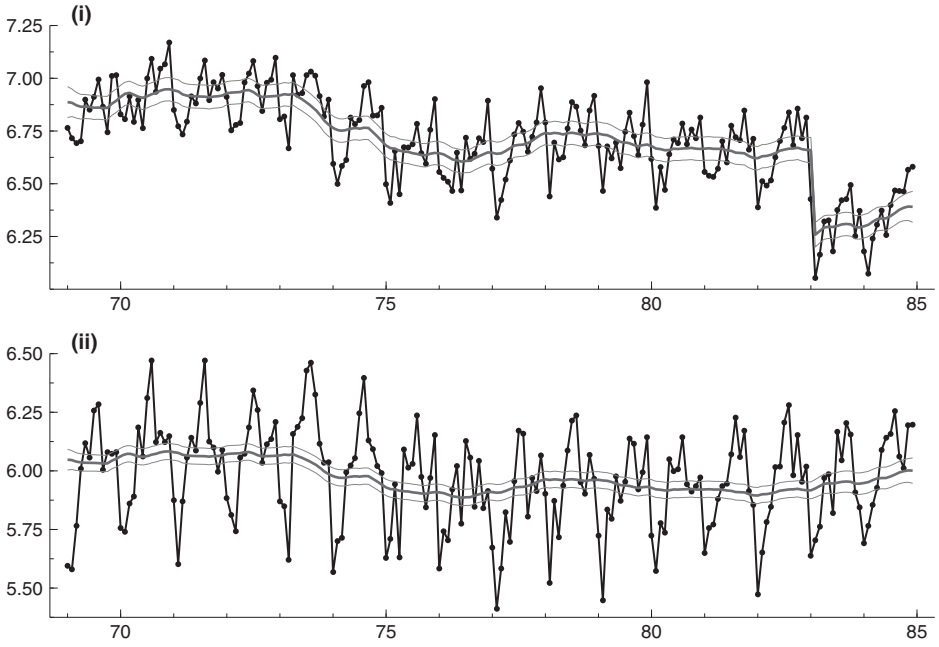

Fig. 8.7 (i) Front seat passengers with estimated signal (without seasonal) and 95% confidence interval; (ii) Rear seat passengers with estimated signal (without seasonal) and 95% confidence interval.

The rank reduction also leads to a large increase (in absolute value) of the t-value. The graphs of the estimated (non-seasonal) signals and the estimated levels for the last model are presented in Fig. 8.7. The substantial drop of the underlying level in front seat passenger casualties at the introduction of the seat belt law is clearly visible.

8.4 Box–Jenkins analysis

In this section we will show that fitting of ARMA models, which is an important part of the Box–Jenkins methodology, can be done using state space methods. Moreover, we will show that missing observations can be handled within the state space framework without problems whereas this is difficult within the Box–Jenkins methodology; see the discussion in Subsection 3.10.1. Finally, since an important objective of the Box–Jenkins methodology is forecasting, we also present forecasts of the series under investigation. In this illustration we use the series which is analysed by Makridakis, Wheelwright and Hyndman (1998): the number of users logged on to an Internet server each minute over 100 minutes.

The data are differenced in order to get them closer to stationarity and these 99 observations are presented in Fig. 8.8(i).

We have estimated a range of ARMA model (3.17) with different choices for p and q. They were estimated in state space form based on (3.20). Table 8.1 presents the Akaike information criteria (AIC), which is defined in Section 7.4, for these different ARMA models. We see that the ARMA models with (p, q) equal to $(1, 1)$ and $(3, 0)$ are optimal according to the AIC values. We prefer the ARMA$(1, 1)$ model because it is more parsimonious. A similar table was produced for the same series by Makridakis, Wheelwright and Hyndman (1998) but the AIC statistic was computed differently. They concluded that the ARMA$(3, 0)$ model was best.

We repeat the calculations for the same differenced series but now with 14 observations treated as missing: $6, 16, 26, 36, 46, 56, 66, 72, 73, 74, 75, 76, 86, 96$. The graph of the amended series is produced in Fig. 8.8(ii). The reported AIC's in Table 8.2 lead to the same conclusion as for the series without missing observations: the preferred model is ARMA$(1, 1)$ although its case is less strong now. We also learn from this illustration that estimation of higher order ARMA models with missing observations lead to more numerical problems.

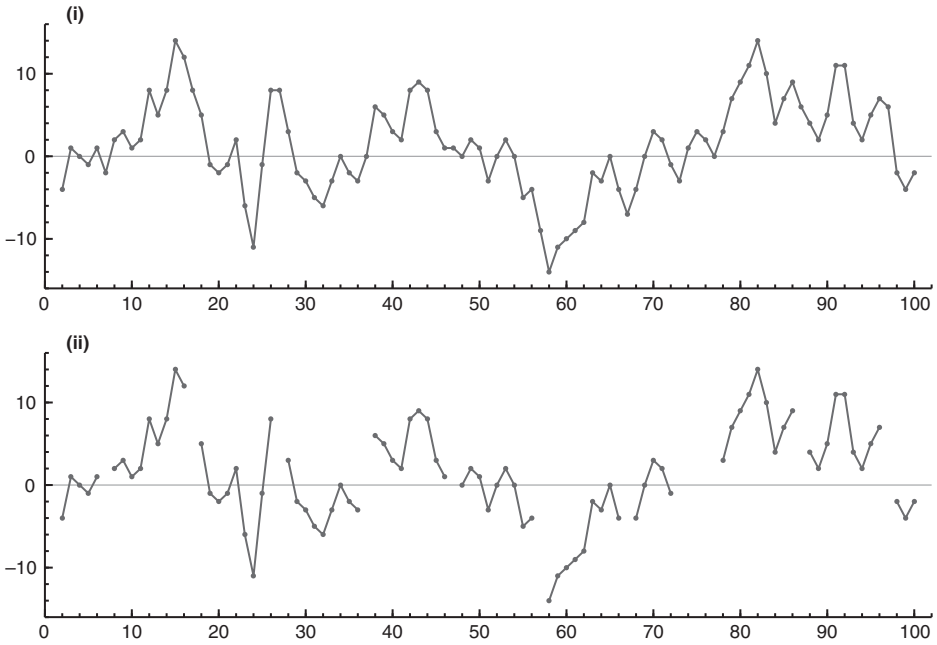

Fig. 8.8 (i) First difference of number of users logged on to Internet server each minute; (ii) The same series with 14 observations omitted.

Table 8.1 AIC for different ARMA models.

q	0	1	2	3	4	5
p						
0		2.777	2.636	2.648	2.653	2.661
1	2.673	2.608	2.628	2.629 (1)	2.642	2.658
2	2.647	2.628	2.642	2.657	2.642 (1)	2.660 (4)
3	2.606	2.626	2.645	2.662	2.660 (2)	2.681 (4)
4	2.626	2.646 (8)	2.657	2.682	2.670 (1)	2.695 (1)
5	2.645	2.665 (2)	2.654 (9)	2.673 (10)	2.662 (12)	2.727 (A)

The value between parentheses indicates the number of times the loglikelihood could not be evaluated during optimisation. The symbol A indicates that the maximisation process was automatically aborted due to numerical problems.

Table 8.2 AIC for different ARMA models with missing observations.

q	0	1	2	3	4	5
p						
0		3.027	2.893	2.904	2.908	2.926
1	2.891	2.855	2.877	2.892	2.899	2.922
2	2.883	2.878	2.895 (6)	2.915	2.912	2.931
3	2.856 (1)	2.880	2.909	2.924	2.918 (12)	2.940 (1)
4	2.880	2.901	2.923	2.946	2.943	2.957 (2)
5	2.901	2.923	2.877 (A)	2.897 (A)	2.956 (26)	2.979

The value between parentheses indicates the number of times the loglikelihood could not be evaluated during optimisation. The symbol A indicates that the maximisation process was automatically aborted due to numerical problems.

Finally, we present in Fig. 8.9 the in-sample one-step ahead forecasts for the series with missing observations and the out-of-sample forecasts with 50% confidence intervals. It is one of the many advantages of state space modelling that it allows for missing observations without difficulty.

8.5 Spline smoothing

The connection between smoothing splines and the local linear trend model has been known for many years; see, for example, Wecker and Ansley (1983). In Section 3.9 we showed that the equivalence is with a local linear trend formulated in continuous time with the variance of the level disturbance equal to zero.

Consider a set of observations $y_1, \ldots, y_n$ which are irregularly spaced and associated with an ordered series $\tau_1, \ldots, \tau_n$. The variable τ_t can also be a measure for age, length or income, for example, as well as time. The discrete time model implied by the underlying continuous time model is the local linear trend model with

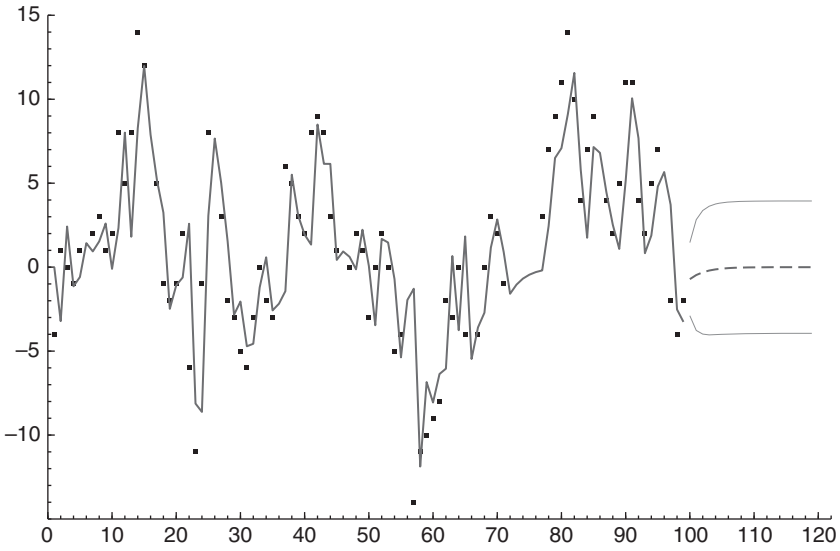

Fig. 8.9 Internet series (solid blocks) with in-sample one-step ahead predictions and out-of-sample forecasts with 50% confidence interval.

$$\text{Var}(\eta_t) = \sigma_\zeta^2 \delta_t^3/3, \qquad \text{Var}(\zeta_t) = \sigma_\zeta^2 \delta_t, \qquad E(\eta_t \zeta_t) = \sigma_\zeta^2 \delta_t^2/2, \tag{8.1}$$

as shown in Section 3.8, where the distance variable $\delta_t = \tau_{t+1} - \tau_t$ is the time between observation t and observation $t+1$. We shall show how irregularly spaced data can be analysed using state space methods. With evenly spaced observations the δ_t's are set to one.

We consider 133 observations of acceleration against time (measured in milliseconds) for a simulated motorcycle accident. This data set was originally analysed by Silverman (1985) and is often used as an example of curve fitting techniques; see, for example, Hardle (1990) and Harvey and Koopman (2000). The observations are not equally spaced and at some time points there are multiple observations; see Fig. 8.10. Cubic spline and kernel smoothing techniques depend on a choice of a smoothness parameter. This is usually determined by a technique called *cross-validation*. However, setting up a cubic spline as a state space model enables the smoothness parameter to be estimated by maximum likelihood and the spline to be computed by the Kalman filter and smoother. The model can easily be extended to include other unobserved components and explanatory variables, and it can be compared with alternative models using standard statistical criteria.

We follow here the analysis given by Harvey and Koopman (2000). The smoothing parameter $\lambda = \sigma_\zeta^2/\sigma_\varepsilon^2$ is estimated by maximum likelihood (assuming normally distributed disturbances) using the transformation $\lambda = \exp(\psi)$. The

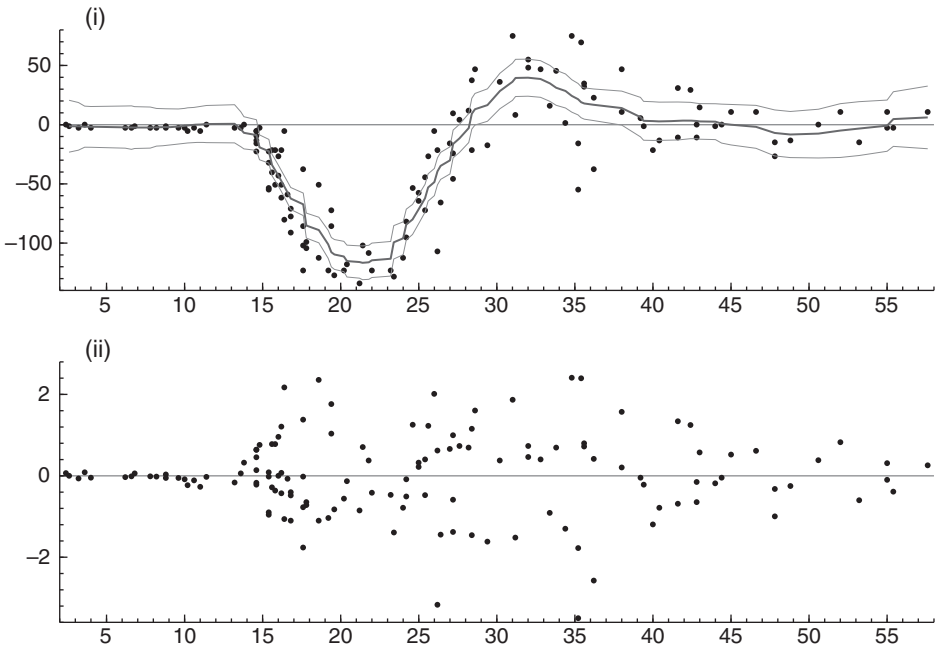

Fig. 8.10 Motorcycle acceleration data analysed by a cubic spline. (i) observations against time with spline and 95% confidence intervals, (ii) standardised irregular.

estimate of ψ is -3.59 with asymptotic standard error 0.22. This implies that the estimate of λ is 0.0275 with an asymmetric 95% confidence interval of 0.018 to 0.043. Silverman (1985) estimates λ by cross-validation, but does not report its value. In any case, it is not clear how one would compute a standard error for an estimate obtained by cross-validation. The Akaike information criterion (AIC) is 9.43. Fig. 8.10 (i) presents the cubic spline. One of the advantages of representing the cubic spline by means of a statistical model is that, with little additional computing, we can obtain variances of our estimates and, therefore, standardised residuals defined as the residuals divided by the overall standard deviation. The 95% confidence intervals for the fitted spline are also given in Fig. 8.10 (i). These are based on the root mean square errors of the smoothed estimates of μ_t, obtained from V_t as computed by (4.43), but without an allowance for the uncertainty arising from the estimation of λ as discussed in Subsection 7.3.7.

8.6 Dynamic factor analysis

The yield curve, or the term structure of interest rates, describes the relation between the interest rate (cost of borrowing) and the time to maturity of the debt for a given borrower in a given currency. The US dollar interest rates

paid on US Treasury securities for various maturities are closely watched by economists and traders. The shape of the yield curve is particularly scrutinised because it provides an indication of future interest rate change and economic activity. The typical yield curve is one in which longer maturity loans have a higher yield compared to shorter-term loans due to the risks associated with time. A well-known monthly data set of yields consists of US interest rates for 17 differents maturities over the period from January 1985 up to December 2000. The maturities are 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months. The data set is presented in Fig. 8.11. We refer to Diebold and Li (2006) for more details on this data set.

The yield curve is typically a smooth function of maturity. Nelson and Siegel (1987) propose to describe the yield curve by three factors that represent level, slope and curvature of the yield curve. Let y_{it} be the interest rate at time t for maturity τ_i for $t = 1, \ldots, n$ and $i = 1, \ldots, N$. The Nelson–Siegel regression model for the yield curve at time t can be expressed by

$$y_{it} = \beta_{1t} + x_{i2}\beta_{2t} + x_{i3}\beta_{3t} + \varepsilon_{it}, \tag{8.2}$$

with

$$x_{i2} = \frac{1 - z_i}{\lambda\tau_i}, \qquad x_{i3} = \frac{1 - z_i}{\lambda\tau_i} - z_i, \qquad z_i = \exp(-\lambda\tau_i),$$

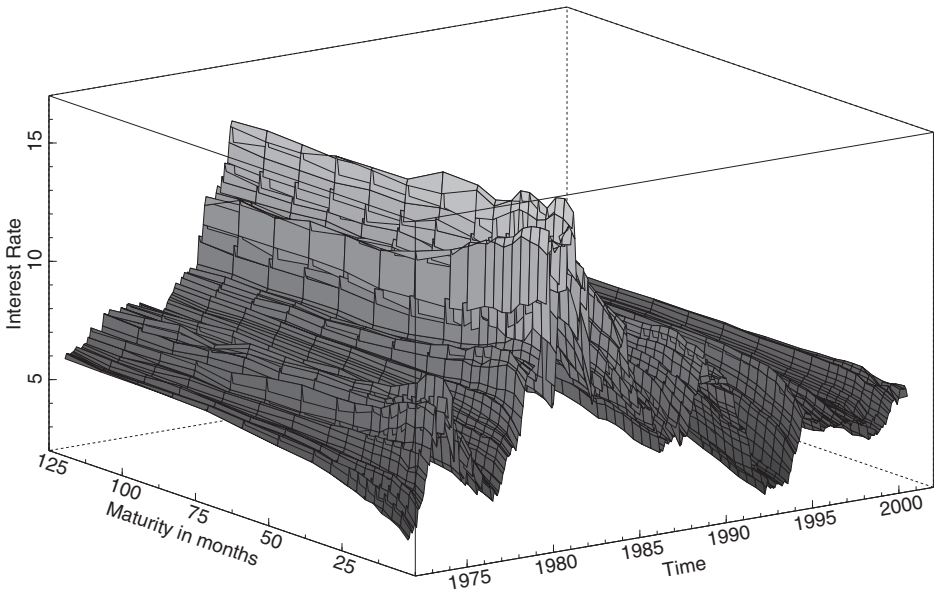

Fig. 8.11 Monthly US interest rates for maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months, for the period from January 1985 up to December 2000.

where β_{jt} can be treated as regression coefficients, $j = 1, 2, 3$, λ is a nonlinear coefficient and ε_{it} is a normally independently distributed disturbance with mean zero and unknown variance σ^2 for $i = 1, \ldots, N$. The construction of the regressors x_{ij} allow β_{jt} to have an interpretation for $j = 1, 2, 3$ and $i = 1, \ldots, N$. The first coefficient f_{1t} represents the level or mean for all interest rates. Since regressor x_{i2} converges to one as τ_i gets smaller and converges to zero as τ_i gets larger, coefficient β_{2t} identifies the slope of the yield curve. The regressor x_{i3} converges to zero as τ_i gets smaller or larger, is concave in τ_i, and therefore coefficient β_{3t} reflects the shape of the yield curve. At time t and for yield observations $y_{1t}, \ldots, y_{Nt}$, with $N > 5$, we can estimate the unknown coefficients via nonlinear least squares. When coefficient λ is set equal to some fixed value, the remaining coefficients can be estimated by ordinary least squares.

Diebold, Rudebusch and Aruoba (2006) adopt the Nelson–Siegel framework but specify a dynamic process for the three factors and carry out a state space analysis for the resulting dynamic factor model (3.32) as given by

$$y_t = \Lambda f_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{N}(0, \sigma^2 I_N),$$

where $y_t = (y_{1t}, \ldots, y_{Nt})'$, $f_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$, $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$ and the (ij) element of loading matrix Λ equals x_{ij}, with $x_{i1} = 1$, for $i = 1, \ldots, N$, $j = 1, 2, 3$ and $t = 1, \ldots, n$. The dynamic specification for f_t is generally given by $f_t = U_t \alpha_t$ in (3.32) where U_t is typically a known selection matrix. It allows the vector autoregressive process for the 3×1 vector f_t as proposed by Diebold, Rudebusch and Aruoba (2006) and others, that is

$$f_{t+1} = \Phi f_t + \eta_t, \qquad \eta_t \sim \mathrm{N}(0, \Sigma_\eta),$$

with $U_t = I_3$ and where Φ is the vector autoregressive coeffient matrix and Σ_η is the disturbance variance matrix. It is often assumed that vector f_t follows a stationary process. The disturbance vectors ε_t and η_t are mutually and serially uncorrelated at all times and lags. It follows that the dynamic Nelson–Siegel model is the state space model (3.1) with $Z_t = \Lambda$, $H_t = \sigma^2 I_N$, $T_t = \Phi$, $R_t = I_3$ and $Q_t = \Sigma_\eta$ for $t = 1, \ldots, n$. This dynamic factor model describes a linear relationship between the interest rates and the dynamic level, slope and curvature factors as advocated by Litterman and Scheinkman (1991).

A state space analysis includes the estimation of λ, σ^2, Φ and Σ_η by the method of maximum likelihood for which the Kalman filter is used for likelihood evaluation. Before the Kalman filter is applied, the $N \times 1$ observation vector y_t is collapsed to a 3×1 vector y_t^* as described in Section 6.5. It leads to a computationally efficient method for parameter estimation and analysing the yield curve based on the dynamic Nelson–Siegel model. The maximum likelihood estimation results are

$$\begin{matrix} \hat{\lambda} = 0.078, \\ \hat{\sigma}^2 = 0.014, \end{matrix} \quad \hat{\Phi} = \begin{bmatrix} 0.994 & 0.029 & -0.022 \\ -0.029 & 0.939 & 0.040 \\ 0.025 & 0.023 & 0.841 \end{bmatrix}, \; \hat{\Sigma}_\eta = \begin{bmatrix} 0.095 & -0.014 & 0.044 \\ -0.014 & 0.383 & 0.009 \\ 0.044 & 0.009 & 0.799 \end{bmatrix}.$$

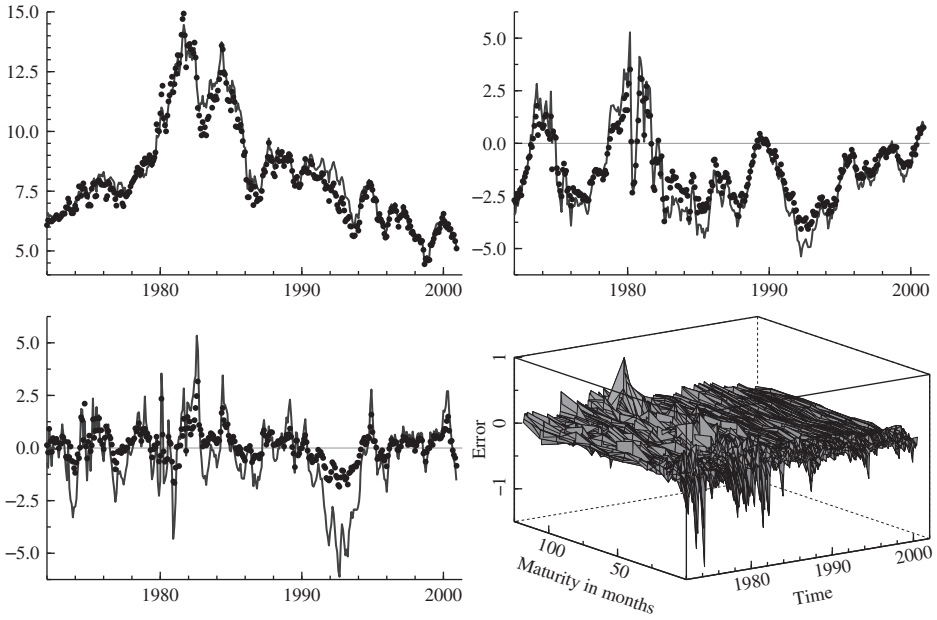

Fig. 8.12 Smoothed estimates of factors f_t and disturbances ε_t for the period from January 1985 up to December 2000: (i) data-proxy of level (dots) and $\hat{f}_{1t}$ (solid line); (ii) data-proxy of slope and $\hat{f}_{2t}$; (iii) data-proxy of slope and $\hat{f}_{3t}$; (iv) smoothed disturbances $\hat{\varepsilon}_{it}$, $i = 1, \ldots, N$ with $N = 17$.

The smoothed estimates of the time-varying factors f_t are displayed in Fig. 8.12. The three smoothed factors are displayed together with their data-proxies. An approximation of the level is the interest rate for a maturity of 10 years. The slope factor can be approximated by the interest rate spread that is defined as the difference between the rates for maturities 3 months and 10 years. The curvature proxy is the difference between the spread associated with 3 months and 2 years, and the spread associated with 10 years and 2 years. The estimated factors follow the same patterns over time in comparison to their data approximations and therefore the interpretations of the factors as level, slope and curvature of the yield curve are justified. The smoothed disturbance vectors $\hat{\varepsilon}_t$ are also displayed in a similar way as the data is presented in Fig. 8.11. We can learn that the model is more successful in fitting the long-term interest rates although most of the larger errors (in absolute values) occur in the months between the early 1970s and the early 1980s, a period of economic turmoil. An analysis with a more comprehensive model for this panel of time series is conducted by Koopman, Mallee and van der Wel (2010).