Diagnostic Checking of Unobserved-Components Time Series Models

Andrew C. Harvey and Siem Jan Koopman
Department of Statistics, London School of Economics, London WC2A 2AE, United Kingdom

Diagnostic checking of the specification of time series models is normally carried out using the innovations—that is, the one-step-ahead prediction errors. In an unobserved-components model, other sets of residuals are available. These auxiliary residuals are estimators of the disturbances associated with the unobserved components. They can often yield information that is less apparent from the innovations, but they suffer from the disadvantage that they are serially correlated even in a correctly specified model with known parameters. This article shows how the properties of the auxiliary residuals may be obtained, how they are related to each other and to the innovations, and how they can be used to construct test statistics. Applications are presented showing how residuals can be used to detect and distinguish between outliers and structural change.

KEY WORDS: Misspecification; Outliers; Signal extraction; Smoothing; Structural change; Structural time series model.

Diagnostic checking of the specification of a time series model is normally carried out using the innovations—that is, the one-step-ahead prediction errors. In an unobserved-components model, other residuals are available. These *auxiliary residuals* are estimators of the disturbances associated with the unobserved components. The auxiliary residuals are functions of the innovations, but they present the information in a different way. This can lead to the discovery of features of a fitted model that are not apparent from the innovations themselves. Unfortunately, the auxiliary residuals suffer from the disadvantage that they are serially correlated, even in a correctly specified model with known parameters. The purpose of this article is to show how the properties of auxiliary residuals may be obtained, how they are related to each other and to the innovations, and how they can be used to construct test statistics. The methods extend straightforwardly to models containing observed explanatory variables.

Section 1 derives the properties of the auxiliary residuals using the classical approach based on a doubly infinite sample. This follows Maravall (1987), except that in his article attention is restricted to the irregular component in the decomposition of an autoregressive integrated moving average (ARIMA) model. Although we initially give general results, our main interest lies in structural time series models since, in our view, these models provide the most satisfactory framework for exploring issues concerning outliers and structural change. Structural time series models are now quite widely used, and a full description can be found in the work of Harvey (1989).

Section 2 derives various relationships between the auxiliary residuals in finite samples. We then discuss a general algorithm that can be used to compute the auxiliary residuals in finite samples in any linear state-space model. This algorithm, the full details of which were given by Koopman (in press), is a development of earlier work by De Jong (1989) and Kohn and Ansley (1989). The efficiency and speed of the algorithm makes the computation of diagnostic procedures based on the auxiliary residuals a viable proposition.

The interpretation of the auxiliary residuals means that they are potentially useful, not only for detecting outliers and structural changes in components but for distinguishing between them. Thus we extend the work of Kohn and Ansley (1989), which was concerned only with the residuals that are estimators of the irregular disturbances and the way in which these residuals may be used to detect outliers. Sections 3 and 4 discuss diagnostics. It is shown how the Bowman–Shenton test can be modified to take account of the serial correlation in the auxiliary residuals, and Section 5 applies it to several data sets. Related modifications can also be made to certain tests for heteroscedasticity, but this particular issue is not pursued here.

1. PROPERTIES OF RESIDUALS IN LARGE SAMPLES

Classical results in signal extraction can be used to derive the properties of various auxiliary residuals in a doubly infinite sample. Let the observed series, y_t, be the sum of $m + 1$ mutually uncorrelated ARIMA processes μ_{it}; that is,

$$y_t = \sum_{i=0}^{m} \mu_{it} = \sum_{i=0}^{m} \frac{\theta_i(L)}{\phi_i(L)} \xi_{it}, \tag{1.1}$$

where $\theta_i(L)$ and $\phi_i(L)$ are polynomials in the lag operator and the ξ_{it} are mutually and serially uncorrelated random variables with zero means and constant variances σ_i^2 ($i = 0, \ldots, m$). The autoregressive polynomials may contain unit roots. The minimum mean squared linear estimator (MMSLE) of μ_{it} is

$$\hat{\mu}_{it} = \left\{|\phi_i(L)|^{-2}|\theta_i(L)|^2\sigma_i^2 \Big/ \sum_{j=0}^{m} |\phi_j(L)|^{-2}|\theta_j(L)|^2\sigma_j^2\right\}y_t; \tag{1.2}$$

see Bell (1984). If the reduced form is

$$y_t = \phi^{-1}(L)\theta(L)\xi_t, \tag{1.3}$$

where ξ_t is white noise, with variance σ^2, then

$$\hat{\mu}_{it} = \{|\phi_i(L)|^{-2}|\theta_i(L)|^2\sigma_i^2/|\phi(L)|^{-2}|\theta(L)|^2\sigma^2\}y_t. \tag{1.4}$$

Since the MMSLE of ξ_{it} is given by $\hat{\xi}_{it} = \phi_i(L)\theta_i^{-1}(L)\hat{\mu}_{it}$, we have, from (1.2), (1.3), and (1.4),

$$\begin{aligned}\hat{\xi}_{it} &= \left\{\phi_i^{-1}(L^{-1})\theta_i(L^{-1})\sigma_i^2 \Big/ \sum_{j=0}^{m}|\phi_j(L)|^{-2}|\theta_j(L)|^2\sigma_j^2\right\}y_t \\ &= \{\phi_i^{-1}(L^{-1})\theta_i(L^{-1})\sigma_i^2/|\phi(L)|^{-2}|\theta(L)|^2\sigma^2\}y_t \qquad (1.5) \\ &= \{\phi_i^{-1}(L^{-1})\theta_i(L^{-1})\sigma_i^2/\phi^{-1}(L^{-1})\theta(L^{-1})\sigma^2\}\xi_t.\end{aligned}$$

The last expression may be written as

$$\hat{\xi}_{it} = \frac{\phi(F)\theta_i(F)}{\phi_i(F)\theta(F)}\frac{\sigma_i^2}{\sigma^2}\xi_t, \qquad i = 0, \ldots, m, \tag{1.6}$$

where $F = L^{-1}$ is the lead operator. Unit roots in $\phi_i(F)$ will cancel with unit roots in $\phi(F)$, so, if time is reversed, $\hat{\xi}_{it}$ is seen to be an autoregressive moving average (ARMA) process, driven by the innovations ξ_t. The process is stationary but, due to the possibility of unit roots in $\phi(F)$, not necessarily strictly invertible.

The autocovariance function (ACF) of $\hat{\xi}_{it}$ may be evaluated from a knowledge of the ARMA process implied by (1.6). Alternatively, we may note that the autocovariance generating function of $\hat{\xi}_{it}$ is

$$\hat{g}_i(L) = \frac{|\phi(L)\theta_i(L)|^2}{|\theta_i(L)\theta(L)|^2}\frac{\sigma_i^4}{\sigma^4} = \frac{|\phi_i(L)|^{-2}|\theta_i(L)|^2}{g(L)}\frac{\sigma_i^4}{\sigma^2},$$

where $g(L) = \Sigma|\phi_i(L)|^{-2}|\theta_i(L)|^2\sigma_i^2$. Given a method of computing $\hat{g}_i(L)$, the autocovariances may be obtained.

We now apply these results to some of the principal structural time series models.

1.1 Local Level

The local-level model is

$$y_t = \mu_t + \varepsilon_t \tag{1.7a}$$

and

$$\mu_t = \mu_{t-1} + \eta_t, \tag{1.7b}$$

where ε_t and η_t are mutually uncorrelated white-noise processes with variance σ_ε^2 and σ_η^2. The reduced form is the ARIMA (0, 1, 1) model

$$\Delta y_t = (1 + \theta L)\xi_t, \qquad -1 \le \theta \le 0, \tag{1.8}$$

with

$$\theta = (\sqrt{(q^2 + 4q)} - 2 - q)/2, \tag{1.9}$$

where $q = \sigma_\eta^2/\sigma_\varepsilon^2$.

Writing the model as

$$y_t = \frac{\eta_t}{\Delta} + \varepsilon_t,$$

and applying (1.6) gives

$$\hat{\varepsilon}_t = \frac{(1 - F)}{1 + \theta F}\frac{\sigma_\varepsilon^2}{\sigma^2}\xi_t \tag{1.10}$$

and

$$\hat{\eta}_t = \frac{1}{1 + \theta F}\frac{\sigma_\eta^2}{\sigma^2}\xi_t. \tag{1.11}$$

Thus both $\hat{\varepsilon}_t$ and $\hat{\eta}_t$ depend on future innovations and, if time is reversed, it can be seen that $\hat{\eta}_t$ follows an (autoregressive) AR (1) process with parameter minus θ, whereas $\hat{\varepsilon}_t$ follows a strictly noninvertible ARMA (1, 1) process. Note that the effect of serial correlation is to make the variance of $\hat{\varepsilon}_t$ less than that of ε_t. In fact, it can be shown that $\text{var}(\hat{\varepsilon}_t)/\sigma_\varepsilon^2 = -2\theta/(1 - \theta) \le 1$ for $-1 \le \theta < 0$.

On comparing (1.10) and (1.11), we see that

$$\hat{\eta}_t = \hat{\eta}_{t+1} + q\hat{\varepsilon}_t, \qquad 0 \le q \le \infty. \tag{1.12}$$

The theoretical cross-correlation function, $\rho_{\varepsilon\eta}(\tau)$, can be evaluated from the preceding equation. The cross-covariance is $\gamma_{\varepsilon\eta}(\tau) = E\{(\hat{\eta}_t - \hat{\eta}_{t+1})\hat{\eta}_{t-\tau}/q\}$ ($\tau = 0, \pm 1, \pm 2, \ldots$), so, for $-1 < \theta < 0$,

$$\rho_{\varepsilon\eta}(\tau) = (-\theta)^\tau\sqrt{(1 + \theta)/2}, \qquad \tau \ge 0, \tag{1.13}$$

and

$$\rho_{\varepsilon\eta}(-\tau) = -\rho_{\varepsilon\eta}(\tau - 1), \qquad \tau > 0. \tag{1.14}$$

As σ_ε^2 becomes smaller, θ tends toward 0 and $\rho_{\varepsilon\eta}(0)$ tends towards .707. Thus, although ε_t and η_t are assumed to be uncorrelated, their estimators may be quite highly correlated.

1.2 Local Linear Trend

The local-linear-trend model consists of Equation (1.7a) with the trend having a slope. Thus

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{1.15a}$$

and

$$\beta_t = \beta_{t-1} + \zeta_t, \tag{1.15b}$$

where η_t and ζ_t are mutually uncorrelated white-noise processes with variances σ_η^2 and σ_ζ^2. The reduced form is the ARIMA (0, 2, 2) model

$$\Delta^2 y_t = (1 + \theta_1 L + \theta_2 L^2)\xi_t. \tag{1.16}$$

If the structural form is expressed as

$$y_t = \frac{\eta_t}{\Delta} + \frac{\zeta_{t-1}}{\Delta^2} + \varepsilon_t, \tag{1.17}$$

we find that

$$\hat{\varepsilon}_t = \frac{(1 - F)^2}{1 + \theta_1 F + \theta_2 F^2} \frac{\sigma_\varepsilon^2}{\sigma^2} \xi_t, \tag{1.18}$$

$$\hat{\eta}_t = \frac{1 - F}{1 + \theta_1 F + \theta_2 F^2} \frac{\sigma_\eta^2}{\sigma^2} \xi_t, \tag{1.19}$$

and

$$\hat{\zeta}_t = \frac{F}{1 + \theta_1 F + \theta_2 F^2} \frac{\sigma_\zeta^2}{\sigma^2} \xi_t. \tag{1.20}$$

Thus $\hat{\varepsilon}_t$, $\hat{\eta}_t$, and $\hat{\zeta}_t$ are ARMA(2, 2), ARMA(2, 1) and AR(2), with $\hat{\varepsilon}_t$ and $\hat{\eta}_t$ being strictly noninvertible. The three processes are stationary provided that $\sigma_\zeta^2 > 0$.

As in the local-level model, Expression (1.12) holds, and in addition

$$\hat{\zeta}_t = \hat{\zeta}_{t+1} + q_\zeta^* \hat{\eta}_{t+1}, \qquad 0 < q_\zeta^* < \infty, \tag{1.21}$$

$$= 2\hat{\zeta}_{t+1} - \hat{\zeta}_{t+2} + q_\zeta \hat{\varepsilon}_{t+1}, \qquad 0 < q_\zeta < \infty, \tag{1.22}$$

where $q_\zeta^* = \sigma_\zeta^2/\sigma_\eta^2$ and $q_\zeta = \sigma_\zeta^2/\sigma_\varepsilon^2$.

In typical applications, the variance of σ_ζ^2 is relatively small. As a result, the moving average polynomial in (1.16) will have one, and possibly two, of its roots close to unity. The $\hat{\zeta}_t$'s will therefore tend to exhibit very strong positive serial correlation. This effect is counteracted in the other auxiliary residuals by the presence of unit roots in the moving average.

1.3 Basic Structural Model

The three methods of modeling a seasonal component γ_t were described by Harvey (1989, chap. 2). All can be expressed in the form

$$\sum_{j=0}^{s-1} \gamma_{t-j} = \theta_\omega(L)\omega_t, \tag{1.23}$$

where ω_t denotes a white-noise disturbance with variance σ_ω^2, s is the number of seasons, and $\theta_\omega(L)$ is a polynomial of order at most $s - 2$. The simplest such model has $\theta_\omega(L)$ equal to 1. Combining with a trend component of the form (1.15) and an irregular yields the basic structural model (BSM). This may be written

$$y_t = \frac{\eta_t}{\Delta} + \frac{\zeta_{t-1}}{\Delta^2} + \frac{\theta_\omega(L)\omega_t}{S(L)} + \varepsilon_t, \tag{1.24}$$

where $S(L) = 1 + L + \cdots + L^{s-1}$. The reduced form is such that

$$\Delta\Delta_s y_t = \theta(L)\xi_t, \tag{1.25}$$

where $\theta(L)$ is of order $s + 1$. Then, from (1.6),

$$\hat{\varepsilon}_t = \frac{(1 - F)(1 - F^s)}{\theta(F)} \frac{\sigma_\varepsilon^2}{\sigma^2} \xi_t, \tag{1.26}$$

$$\hat{\eta}_t = \frac{(1 - F^s)}{\theta(F)} \frac{\sigma_\eta^2}{\sigma^2} \xi_t, \tag{1.27}$$

$$\hat{\zeta}_t = \frac{S(F)F}{\theta(F)} \frac{\sigma_\zeta^2}{\sigma^2} \xi_t, \tag{1.28}$$

Table 1. Theoretical Autocorrelations for the Auxiliary Residuals of a Quarterly Basic Structural Model With $q_\eta = 1$, $q_\zeta = .1$, and $q_\omega = .1$

Lag	$\hat{\varepsilon}$	$\hat{\eta}$	$\hat{\zeta}$	$\hat{\omega}$
0	1	1	1	1
1	−.29	.28	.88	−.44
2	−.14	−.02	.70	−.14
3	.02	−.12	.52	−.24
4	−.18	−.24	.37	.65
5	.07	−.09	.28	−.25
6	.03	−.05	.21	−.14
7	.04	−.05	.15	−.14
8	−.11	−.11	.10	.42
9	.05	−.02	.07	−.14
10	.03	.00	.06	−.13

and

$$\hat{\omega}_t = \frac{(1 - F)^2}{\theta(F)} \frac{\theta_\omega(F)\sigma_\omega^2}{\sigma^2} \xi_t. \tag{1.29}$$

The residuals $\hat{\varepsilon}_t$, $\hat{\eta}_t$, and $\hat{\zeta}_t$ bear exactly the same relationship to each other as in the local-linear-trend model. In addition, note that

$$S(F)\hat{\omega}_t = q_\omega \theta_\omega(F)\hat{\varepsilon}_t, \qquad 0 < q_\omega < \infty, \tag{1.30}$$

where $q_\omega = \sigma_\omega^2/\sigma_\varepsilon^2$.

Explicit expressions for the autocorrelation functions of the ARMA processes followed by the auxiliary residuals are not easy to obtain in this case. Numerical values, however, can be computed for specific parameter values using the algorithm described in Subsection 2.2. As an example, for a quarterly BSM where $\theta_\omega(L) = 1$, $q_\eta = 1$, $q_\zeta = .1$, and $q_\omega = .1$, the first 10 autocorrelations are as shown in Table 1. The ACF's of the irregular and level residuals are not too dissimilar to what one might expect in a local-level model with $q = 1$, although, if anything, the serial correlation in the level is somewhat reduced by the presence of the other components. The high positive serial correlation in the slope residual, to which attention was drawn at the end of the previous subsection, is clearly apparent, but the seasonal residual shows a strong pattern of serial correlation, the most prominent feature of which is the high values at the seasonal lags 4 and 8. As regards cross-correlations (see Table 2) the relatively pronounced patterns for $\hat{\varepsilon}\hat{\eta}$ and $\hat{\eta}\hat{\zeta}$ suggested by the analysis for the local-linear-trend model are still apparent, but the relationships involving $\hat{\omega}$ show seasonal effects.

2. FINITE SAMPLES

Relationships between auxiliary residuals, such as (1.12), are valid for doubly infinite samples. However, exact relationships can be derived for finite samples. The following subsection shows how this may be done using a very simple approach based on an idea of Whittle (1991). Unfortunately this approach does not lead to a viable algorithm for computing the auxiliary residuals and associated statistics such as variances and

Table 2. Theoretical Cross-correlations for the Auxiliary Residuals of a Quarterly Basic Structural Model With q_η = 1, q_ζ = .1, and q_ω = .1

Lag	$\hat{\varepsilon}\hat{\eta}$	$\hat{\varepsilon}\hat{\zeta}$	$\hat{\varepsilon}\hat{\omega}$	$\hat{\eta}\hat{\zeta}$	$\hat{\eta}\hat{\omega}$	$\hat{\zeta}\hat{\omega}$
0	.60	.11	.06	.24	−.07	.07
1	.25	−.01	.06	.38	−.00	.03
2	.08	−.05	−.32	.37	.07	.03
3	.10	−.10	.35	.31	−.31	.07
4	−.12	−.04	−.02	.19	.11	−.08
5	−.03	−.02	−.02	.15	.09	−.03
6	−.00	−.02	−.22	.13	.06	.01
7	.05	−.05	.22	.10	−.20	.05
8	−.08	.01	.01	.05	.06	−.05
9	−.02	.00	−.03	.04	.07	−.02
10	.01	−.00	−.14	.04	.04	.01

autocovariances. The ideas underlying a stable algorithm are sketched out in Subsection 2.2.

We will use a tilde to denote finite-sample auxiliary residuals, thereby distinguishing them from the corresponding infinite-sample residuals of Section 1. The properties of the finite-sample residuals in the middle of the sample will be the same as the properties derived for the infinite-sample residuals. Note that both finite- and infinite-sample residuals can be regarded as minimum mean squared estimators of the corresponding disturbances under Gaussianity.

2.1 Relationship Between Auxiliary Residuals

Consider the local-level model (1.7) defined for t = 1 to T, and suppose that the disturbances ε_t and η_t are normally distributed. Suppose also that the initial state is Gaussian with zero mean and a finite variance, p_0; that is, $\mu_0 \sim N(0, p_0)$ and it is independent of the disturbances. The logarithm of the joint density of the observations $y_1, \ldots, y_T$ and the states $\mu_0, \ldots, \mu_T$ is, neglecting constants,

$$J = -\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(\mu_t - \mu_{t-1})^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_{t=1}^{T}(y_t - \mu_t)^2 - \frac{1}{2p_0}\mu_0^2. \quad (2.1)$$

Partially differentiating J with respect to each of the states, $\mu_0, \mu_1, \ldots, \mu_T$ provides a means of evaluating the smoothed estimators, which are, by definition, the expected values (and therefore the modes) of the states conditional on the observations. The result is the backward recursion

$$\tilde{\mu}_{t-1} = 2\tilde{\mu}_t - \tilde{\mu}_{t+1} - q(y_t - \tilde{\mu}_t), \qquad t = T-1, \ldots, 2. \quad (2.2)$$

The initialization, given from $\partial J/\partial\mu_T$, is

$$\tilde{\mu}_{T-1} = \tilde{\mu}_T - q(y_T - \tilde{\mu}_T), \quad (2.3)$$

so (2.2) can be started at $t = T$ by setting $\tilde{\mu}_{T+1}$ equal to $\tilde{\mu}_T$. Letting $p_0 \to \infty$ gives the end condition for a diffuse prior—namely,

$$\tilde{\mu}_2 = \tilde{\mu}_1 - q(y_1 - \tilde{\mu}_1). \quad (2.4)$$

Although (2.2) looks, at first sight, to be an extremely attractive way of computing the smoothed estimators of the μ_t's, it is, unfortunately, numerically unstable, and the $\tilde{\mu}_2$ and $\tilde{\mu}_1$ computed in this way are almost certain to violate (2.4). Nevertheless (2.2) is useful for the theoretical insight it provides. Noting that $\tilde{\eta}_t = \tilde{\mu}_t - \tilde{\mu}_{t-1}$, (2.2) can be rewritten as in (1.12)—namely,

$$\tilde{\eta}_t = \tilde{\eta}_{t+1} + q\tilde{\varepsilon}_t, \qquad t = T, \ldots, 2, \quad (2.5)$$

but with starting value $\tilde{\eta}_{T+1} = 0$. Thus $\tilde{\eta}_t$ is a backward cumulative sum of the $\tilde{\varepsilon}_t$'s; that is,

$$\tilde{\eta}_t = q\sum_{j=t}^{T}\tilde{\varepsilon}_j, \qquad t = 2, \ldots, T. \quad (2.6)$$

Furthermore, from (2.4), $\tilde{\eta}_2 = -q\tilde{\varepsilon}_1$, so, on setting t = 2 in (2.6), it can be seen that

$$\sum_{j=1}^{T}\tilde{\varepsilon}_j = 0. \quad (2.7)$$

It will be recalled that the ordinary least squares regression residuals have this property when a constant term is included.

A similar approach can be used in the local-linear-trend model to show that in a finite sample (2.3) holds and that (1.21) can be initialized with $\tilde{\zeta}_T = 0$. In addition, (2.5) holds, and if $q_\zeta^* > 0$,

$$\tilde{\zeta}_t = q_\zeta^*\sum_{j=t+1}^{T}\tilde{\eta}_j, \qquad t = 3, \ldots, T-1, \quad (2.8)$$

and, provided that $q_\zeta > 0$,

$$\tilde{\zeta}_t = q_\zeta\sum_{j=t+1}^{T}\sum_{i=j}^{T}\tilde{\varepsilon}_i. \quad (2.9)$$

Finally, both (2.6) and (2.7) hold if $\sigma_\varepsilon^2 > 0$.

2.2 Algorithm

Calculation of the auxiliary residuals is carried out by putting the model in state-space form and applying the Kalman filter and smoother. The algorithm described by Koopman (in press) enables the computations to be carried out relatively quickly in a numerically stable manner; see the Appendix. Structural time series models generally contain nonstationary components, and these are handled by means of a diffuse prior on the state. In Koopman's algorithm, the calculations associated with the diffuse prior are carried out exactly.

The theoretical variances of the auxiliary residuals near the middle of the series can be obtained directly from the large-sample theory of Section 1. The variances at the beginning and end of a finite sample are different, however. The exact algorithm is therefore used to standardize all of the residuals before they are plotted.

The theoretical autocorrelations and cross-correlations of the auxiliary residuals can be calculated exactly at any point in time, but for the purposes of the test statistics employed in Section 3 only the autocorrelations appropriate for the middle of the series need be used.

3. DIAGNOSTICS

Within the context of a structural time series model, an outlier arises at time t if the value taken by y_t is not consistent with what might reasonably be expected given the model specification and the way in which this fits the other observations. The best indicator of an outlier should be $\tilde{\varepsilon}_t$; compare Kohn and Ansley (1989). Note that an outlier at time t will not affect the innovations before time t. Therefore, it makes sense that $\tilde{\varepsilon}_t$ depends only on the innovations that are affected by the outlier.

The simplest kind of structural change is a permanent shift in the level of a series that is of a greater magnitude than might reasonably be expected given the model specification and the other observations. Within the context of the local-level model, (1.7), such a shift might be best detected by an outlying value of $\tilde{\eta}_t$. Again, only the innovations at time t and beyond are affected by such a shift, and $\tilde{\eta}_t$ combines these innovations in the most appropriate way.

A sudden change in the slope is likely to be more difficult to detect than a shift in the level. As already noted, the $\tilde{\zeta}_t$'s will typically be very strongly correlated, so a break will spread its effect over several $\tilde{\zeta}_t$'s. Furthermore, the high serial correlation means that the variances of the normality and kurtosis statistics that will be discussed will need to be increased considerably, giving the tests rather low power. The seasonal residuals suffer a similar drawback. Furthermore, it may be difficult to associate a sudden change in the seasonal pattern with a particular disturbance in (1.23). Nevertheless, there may still be some value in using the seasonal auxiliary residuals to detect changes of this kind.

The basic detection procedure is to plot the auxiliary residuals after they have been standardized. (As pointed out in Sec. 2, the residuals at the end and the beginning will tend to have a higher variance.) In a Gaussian model, indications of outliers and/or structural change arise for values greater than 2 in absolute value. The standardized innovations may also indicate outliers and structural change but will not normally give a clear indication as to the source of the problem.

A more formal procedure for detecting the unusually large residuals is to carry out a test for excess kurtosis. If this test is combined with a test for skewness, we have the Bowman–Shenton test for normality. For such tests to be asymptotically valid, it is necessary to make an allowance for serial correlation.

3.1 Tests Based on Skewness and Kurtosis

Let x_t be a stationary, Gaussian time series with autocorrelations ρ_τ ($\tau = 0, 1, 2, \ldots$) and variance σ_x^2. Following Lomnicki (1961), consider the estimated moments about the sample mean

$$m_\alpha = T^{-1} \sum_{t=1}^{T} (x_t - \bar{x})^\alpha, \qquad \alpha = 2, 3, 4, \tag{3.1}$$

and define

$$\kappa(\alpha) = \sum_{\tau=-\infty}^{\infty} \rho_\tau^\alpha, \qquad \alpha = 2, 3, 4. \tag{3.2}$$

Then, if μ_α denotes the theoretical αth moment,

$$\sqrt{T}(m_\alpha - \mu_\alpha) \xrightarrow{L} N[0, \alpha!\, \kappa(\alpha)\sigma_x^{2\alpha}]. \tag{3.3}$$

This result enables asymptotically valid test statistics based on higher order moments to be constructed as follows:

1. *Excess kurtosis test*. The measure of kurtosis is

$$b_2 = m_4/m_2^2. \tag{3.4}$$

Since m_2 is a consistent estimator of σ_x^2, it follows that the excess kurtosis test statistic

$$K = (b_2 - 3)/\sqrt{24\,\kappa(4)/T} \tag{3.5}$$

is asymptotically $N(0, 1)$ under the null hypothesis. An outlier test is carried out as a one-sided test on the upper tail.

2. *Normality test*. The measure of skewness is $\sqrt{b_1} = m_3/m_2^{3/2}$. Combining this with b_2 gives the Bowman–Shenton normality test, which when corrected for serial correlation takes the form

$$N = \frac{Tb_1}{6\kappa(3)} + \frac{T(b_2 - 3)^2}{24\kappa(4)}. \tag{3.6}$$

Under the null hypothesis, N is asymptotically χ_2^2; see Lomnicki (1961).

The normality and excess kurtosis tests may be applied to the innovations and auxiliary residuals. In contrast to serial-correlation tests, no amendments are needed to allow for the estimation of unknown parameters; compare Subsection 4.1. The serial-correlation correction terms, the $\kappa(\alpha)$'s, needed for the auxiliary residuals can be computed using the general algorithm of Subsection 2.2. The results in Section 1 are useful in that they enable one to get some idea of the likely size

Table 3. Correlation Factors for Basic Structural Model With $q_\eta = 1$, $q_\zeta = .1$, and $q_\omega = .1$

	$\kappa(3)$	$\kappa(4)$
Irregular	.93	1.02
Level noise	1.01	1.02
Slope noise	3.53	2.90
Seasonal noise	1.49	1.53

of $\kappa(\alpha)$. In the case of the local-level model (1.7) we find that for $\tilde{\eta}_t$

$$\kappa(\alpha) = \frac{1 + (-\theta)^\alpha}{1 - (-\theta)^\alpha} \tag{3.7}$$

This is unity for a random walk—that is, $\theta = 0$—and goes monotonically toward infinity as q tends toward 0; that is, θ tends to minus one. On the other hand, for $\tilde{\varepsilon}_t$,

$$\kappa(\alpha) = 1 + \frac{\{-(1 + \theta)^\alpha}{2^{\alpha-1}\{1 - (-\theta)^\alpha\}}, \tag{3.8}$$

which is greater than or equal to unity for $\alpha = 4$ but less than or equal to unity for $\alpha = 3$. When θ is minus one, it takes the values $-.75$ and 1.125 for $\alpha = 3$ and 4. The kurtosis test statistic, therefore, always becomes smaller after being corrected for serial correlation, and this is also true for the normality statistic when applied to the level residual. The normality test statistic for the irregular may, however, increase. For the irregular, the correction factors are relatively small. The high correction factors for the level residual when θ is close to minus one may appear to make the detection of structural change difficult. If level shifts are introduced into an otherwise well-behaved series, however, the effect is likely to be an increase in the estimate of the relative variance of η_t and hence a corresponding increase in θ.

For more complex models, the correction factors can be computed numerically using the algorithm of Subsection 2.2. Table 3 shows the $\kappa(\alpha)$'s for the four sets of auxiliary residuals from the BSM of Tables 1 and 2, calculated using the first 20 autocorrelations.

3.2 Monte Carlo Experiments

A series of simulation experiments were run to examine the performance of the test statistics discussed in Section 3.1. The experiments were conducted on the local-level model, using a sample size of $T = 150$ and different values of the signal noise ratio, q. The white-noise disturbances, ε_t and η_t, were generated using the Box–Muller algorithm of Knuth (1981). The results presented in Table 4 are based on 1,000 replications and show the estimated probabilities of rejection for tests at a nominal 5% level of significance.

Table 4(a) gives the estimated sizes of the tests. It is known that, for independent observations, the size of the Bowman–Shenton test can be some way from the nominal size for small samples, and Granger and Newbold (1977, pp. 314–315) cited evidence that suggests that serial correlation may make matters even worse. Their remarks, however, are concerned with a test statistic in which the correction factors are based on the correlogram, whereas in our case the correction factor is based on the estimator of a single parameter θ. The figures in Table 4(a) indicate that the estimated type I errors are not too far from the nominal values for both the innovations and the auxiliary residuals.

Table 4(b) shows the estimated powers of the tests when an outlier was inserted three-quarters of the way along the series. The magnitude of the outlier was five times σ^2_ε. As can be seen, the powers of the tests based on the irregular residual are higher than those based on the innovation. As we had hoped, the power of the tests based on the level residual are much lower. The kurtosis test is slightly more powerful than the normality test.

A shift in the level, up by five times σ^2_ε, was introduced three-quarters of the way along the series to generate the results in Table 4(c). The tests based on the level residual are now more powerful.

Overall, the results are very encouraging. They suggest that the tests have acceptable sizes for moderate samples even when serial correlation corrections have to be made. Furthermore, the tests based on auxiliary

Table 4. Estimated Rejection Probability for Tests at a Nominal 5% Level of Significance for a Local-Level Model With T = 150

		q = 2.0		q = .5	
		N	K	N	K
(a) No misspecification*	Innovations	.062	.077	.055	.077
	Irregular	.038 (.036)	.058 (.062)	.039 (.039)	.060 (.062)
	Level	.034 (.038)	.061 (.062)	.037 (.064)	.053 (.065)
(b) Single outlier at $t = 112$	Innovations	.49	.56	.87	.90
	Irregular	.76	.79	.97	.97
	Level	.25	.30	.26	.31
(c) Structural shift on level at $t = 112$	Innovations	.42	.45	.83	.85
	Irregular	.15	.19	.27	.34
	Level	.47	.49	.94	.95

*Uncorrected tests in parentheses.

residuals are reasonably effective in detecting and distinguishing between outliers and structural change.

4. MISCELLANEOUS ISSUES

Several other issues arise in connection with diagnostic checking.

4.1 Tests for Serial Correlation

In a correctly specified model, the standardized innovations are normally and independently distributed (NID) when the parameters are known, and hence a portmanteau test for serial correlation is straightforward to carry out. In the more usual case when parameters have to be estimated, a correction to the degrees of freedom of the relevant χ^2 distribution can be made along the lines suggested by Box and Pierce (1970). As we have seen, the auxiliary residuals are serially correlated even for a correctly specified model with known parameters. We may be alerted to misspecification by the fact that the correlograms of the auxiliary residuals are very different to their implied ACF's; see Maravall (1987). If a formal test of serial correlation is to be based on residuals, however, it would seem that we have no alternative but to *prewhiten* the auxiliary residuals, which, in view of (1.6), means going back to the innovations.

It should perhaps be stressed that if the reduced form of an unobserved-components model is correctly specified, the serial and cross-correlations in the auxiliary residuals tell us nothing whatsoever about the validity of the assumptions underlying the particular unobserved-components model being employed; compare Garcia-Ferrer and del Hoyo (in press). When we talk about misspecification in the previous paragraph, this is to be understood as meaning misspecification of the reduced form. When a particular unobserved-components decomposition is consistent with a correctly specified reduced form, the question of whether the decomposition is a sensible one can only be resolved by an appeal to theoretical arguments concerning the type of properties one wishes components such as trends and seasonals to possess; see Harvey (1989, secs. 6.1 and 6.2).

4.2 Residuals From the Canonical Decomposition

In structural time series modeling, the components are specified explicitly, and the reduced form follows as a result of this specification. This contrasts with the initial specification of an ARIMA model and the subsequent decomposition of this model into unobserved components. The usual way in which this is done is via the canonical decomposition of Pierce (1979) and Hillmer and Tiao (1982), the aim of which is to maximize the variance of the irregular term.

This subsection examines the relationship between the properties of the structural and canonical decomposition auxiliary residuals for observations following an ARIMA(0, 1, 1) process. It is shown that the standardized residuals associated with the irregular term are the same, but the residuals associated with the trend are different. It is then argued that the structural residuals are likely to be more useful for detecting a structural change in the level.

The canonical decomposition of an ARIMA(0, 1, 1) process is such that

$$y_t = \mu_t^* + \varepsilon_t^*, \tag{4.1}$$

where

$$\mu_t^* = \mu_{t-1}^* + \eta_t^* + \eta_{t-1}^*, \tag{4.2}$$

with ε_t^* and η_t^* mutually uncorrelated white-noise processes. The residual estimating ε_t^*, denoted $\tilde{\varepsilon}_t^*$, follows exactly the same process as $\tilde{\varepsilon}_t$, except that its variance is at least as great as that of ε_t, since $\sigma_\varepsilon^2/\sigma^2(= -\theta)$ in (1.10) is replaced by $(1 - \theta)^2/4$; see Maravall (1987, p. 116). The standardized residuals are obviously the same, however. The residuals associated with η_t^*, $\tilde{\eta}_t^*$, on the other hand, follow an ARMA(1, 1) process,

$$\tilde{\eta}_t^* = \frac{1 + F}{1 + \theta F} \frac{\sigma_\eta^{*2}}{\sigma^2} \xi_t, \tag{4.3}$$

where $\sigma_\eta^{*2} = \text{var}(\eta_t^*)$. Comparing this with (1.11), we see that

$$\tilde{\eta}_t^* = (\tilde{\eta}_t + \tilde{\eta}_{t+1})\sigma_\eta^{*2}/\sigma_\eta^2. \tag{4.4}$$

The fact that $\tilde{\eta}_t^*$ is an average of the corresponding structural residuals in the current and next period means that it may provide a less sharply defined tool for detecting structural change.

4.3 Explanatory Variables

Explanatory variables can be added to a structural time series model. Thus we might have

$$y_t = \mu_t + x_t'\delta + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{4.5}$$

where μ_t is a stochastic trend (1.15) and x_t is a $k \times 1$ vector of exogenous explanatory variables with associated $k \times 1$ vector of coefficients, δ. If δ is fixed and known, residuals are constructed exactly as in the corresponding univariate model by treating $y_t - x_t'\delta$ ($t = 1, \ldots, T$) as the observed values. If δ is unknown, the main issue that arises is that two sets of innovations may be calculated, depending on whether or not δ is included in the state vector. If it is, the standardized prediction errors are known as generalized recursive residuals; see Harvey (1989, chap. 7). The distinction between these two sets of residuals is somewhat peripheral to the discussion here since the auxiliary residuals are unaffected. In the example in Section 5, the innovations are calculated by including δ in the state vector.

5. APPLICATIONS

The following examples illustrate the way in which outliers and structural changes may be detected. In all cases, parameter estimation was carried out in the frequency domain, using the method of scoring described by Harvey (1989, chap. 4, sec. 3).

5.1 U.S. Exports to Latin America

The monthly series of U.S. exports to Latin America contains a number of outliers that are easily detected by examining the irregular components $\tilde{\varepsilon}_t$ from a BSM; see the comments by Harvey (1989) on Bruce and Martin (1989). In fact, the principal outliers, which turn out to be due to dock strikes, are easily seen in a plot of the series and also appear quite clearly in the innovations. We therefore aggregated the data to the quarterly level and fitted a BSM of the form described in Subsection 1.3. The outliers are now less apparent in the innovations, though they still emerge clearly in the irregular component; see Figure 1. The kurtosis statistic for the innovations is $K = 2.00$, and the normality statistic is $N = 4.52$. The normality statistic is therefore not statistically significant at the 5% level, whereas the kurtosis is significant on a one-sided test at the 5% level, but not at the 1% level. For the irregular, on the other hand, the raw K and N statistics are 7.92 and 90.73. After correction for serial correlation, these become $K = 7.85$ and $N = 91.55$, both of which are highly significant.

Since σ_η^2 is estimated to be 0, all of the movements in the trend stem from the slope disturbance. The (corrected) K and N statistics for the associated auxiliary residuals are only .18 and .03. The auxiliary residual diagnostics therefore point clearly to the presence of outliers.

5.2 Car Drivers Killed and Seriously Injured in Great Britain

Monthly observations of car drivers killed and seriously injured in Great Britain were used by Harvey and Durbin (1986) in their study of the effects of the seat-belt legislation that took effect at the beginning of February 1983. The seat-belt law led to a drop in the level of the series. We now show how this structural change would be detected by the auxiliary residuals.

To avoid the large fluctuations associated with the oil crisis of 1974, a BSM was estimated using data from 75 M7 to the end of the series in 84 M12. The slope and seasonal variances were both estimated to be 0, so the fitted model is basically a random walk plus noise,

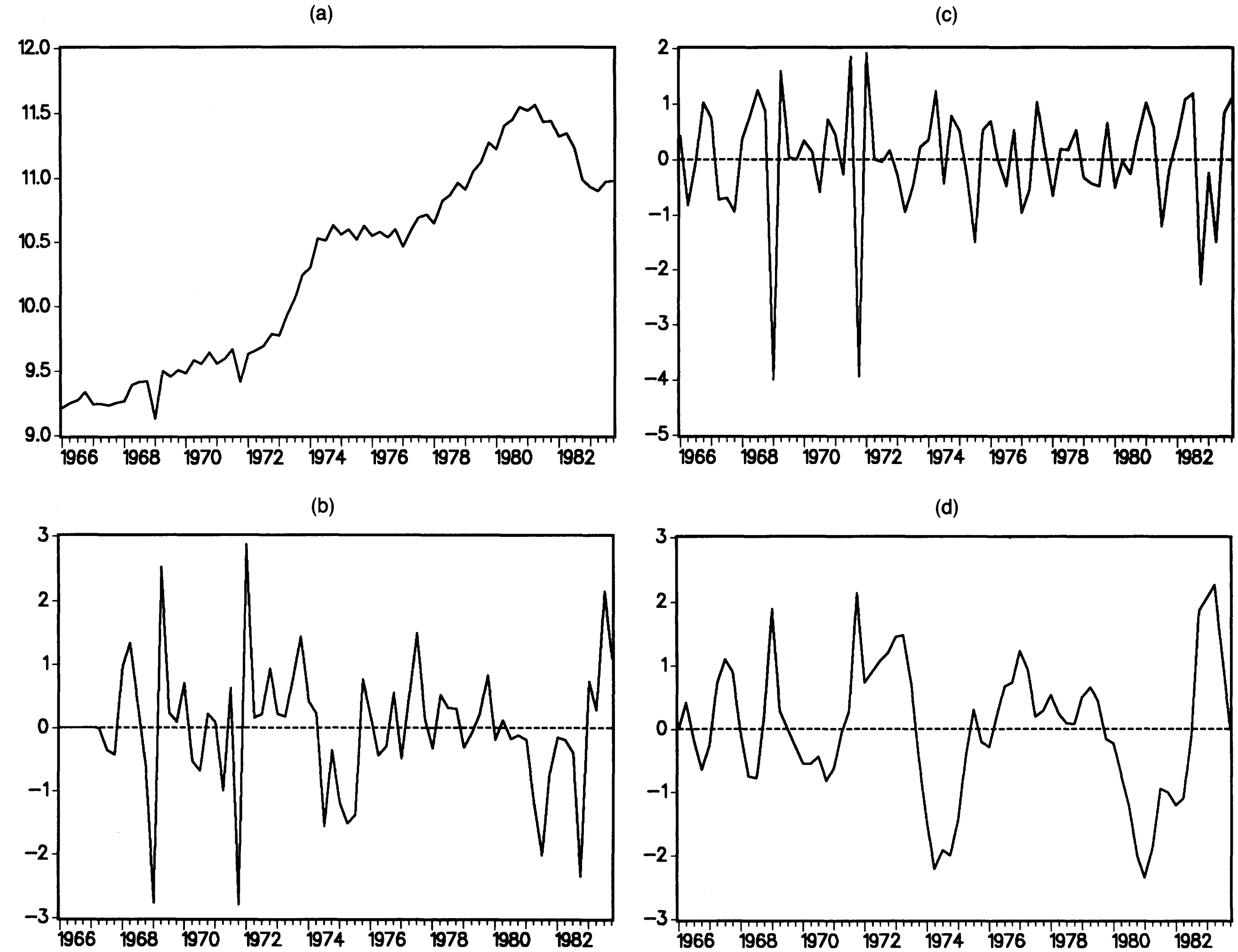

Figure 1. U.S. Exports to Latin America: (a) Observations (in logarithms); (b) Innovations; (c) Irregular; (d) Slope residual.

Table 5. Estimated Hyperparameters ($\times 10^{-5}$) for U.S. Exports to Latin America and Car Drivers Killed and Seriously Injured in Great Britain

Parameters	*Exports*	*Car drivers*
σ_ε^2	314	425
σ_η^2	0	49.5
σ_ζ^2	84	0
σ_ω^2	1	0

with a fixed slope and seasonal; see Table 5. The theory at the end of Subsection 3.1 therefore applies directly, with $q = .118$ and $\theta = -.710$. The correction factors for the irregular are $\kappa(3) = .99$ and $\kappa(4) = 1.00$, but for the level they are $\kappa(3) = 2.12$ and $\kappa(4) = 1.69$.

The kurtosis and normality statistics are shown in Table 6. The innovation statistics clearly indicate excess kurtosis, and the auxiliary residual diagnostics point to this as emanating from a change in level, with the K and N statistics both being statistically significant at the 1% level. The plot of the innovations in Figure 2(b) shows large values in 81 M12 and 83M2 at -3.28 and -3.97. In the irregular residuals, shown in Figure 2(c), both of these months are -2.84, but such a value is not excessively large compared with those for some of the other months. In the level residuals, on the other hand, 83 M2 is -4.46, but 81 M12 is only -1.76.

Table 6. Diagnostic Statistics for Car Drivers

Residual	*K*	*N*
Innovation	2.51*	12.61*
Irregular	.50	.86
Level	4.80*	38.04*

*Significant at 1% level.

The residuals therefore point clearly to a structural break at the beginning of 1983. The role of 81 M12 is less clear. It could be treated as an outlier; in fact, Harvey and Durbin (1986) noted that December 1981 was a very cold month. Even when the model is reestimated with an intervention variable for the seat-belt law, however, it does not give rise to a particularly large

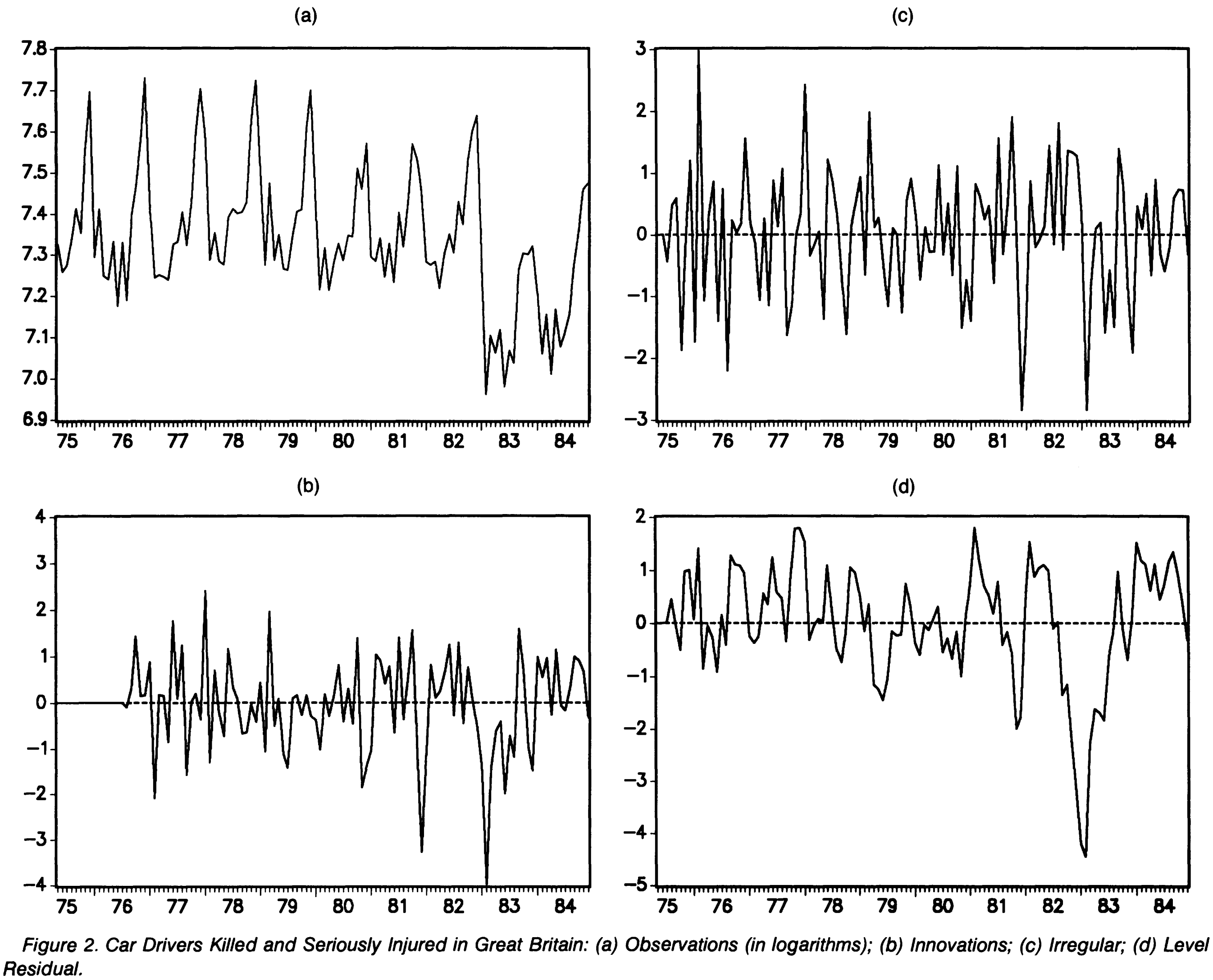

Figure 2. Car Drivers Killed and Seriously Injured in Great Britain: (a) Observations (in logarithms); (b) Innovations; (c) Irregular; (d) Level Residual.

Table 7. Parameters, Diagnostics and Goodness-of-Fit Statistics for Spirits Model before and after Interventions

Parameters/statistics	No interventions	Interventions	
Income	.69 (5.28)	.66 (7.82)	.58 (6.45)
Price	.95 (−13.6)	−.73 (−15.2)	−.53 (−6.31)
1909 level	—	−.09 (−7.90)	−.09 (−8.69)
1915 outlier	—	.05 (5.33)	.06 (4.34)
1916 outlier	—	—	.004 (.30)
1917 outlier	—	—	−.05 (−2.79)
1918 outlier	—	−.06 (−7.47)	−.10 (−6.41)
1919 outlier	—	—	−.01 (−1.25)
σ_ε^2	161×10^{-6}	0	0
σ_η^2	69×10^{-6}	117×10^{-6}	79×10^{-6}
σ_ζ^2	37×10^{-6}	14×10^{-6}	30×10^{-6}
Prediction error SD, σ	229×10^{-4}	166×10^{-4}	144×10^{-4}
R_D^2	.71	.91	.93
Box–Ljung, Q(10)	13.06	5.25	4.93
N	5.87	1.53	.73
K	2.21	1.23	.84
H	2.47	.83	.75

NOTE: Figures in parentheses are t statistics, R_D^2 is the coefficient of determination with respect to the differenced observations as in Harvey (1989, chapter 5), and $Q(P)$ is the Box–Ljung statistic based on the first P residual autocorrelations.

irregular residual, though curiously enough the corresponding innovation is still quite high.

A final point with respect to this example concerns checks for serial correlation. For the innovations, the Box–Ljung statistic based on the first 10 sample autocorrelations is $Q(10) = 8.58$. Thus no serial correlation is indicated. As expected from the argument in Subsection 4.1, the correlograms and theoretical ACF's for the irregular and level residuals are quite similar and hence give no further hint of model misspecification. Nor do the sample and theoretical cross-correlations. Of course, evidence of dynamic misspecification can be masked by outliers and structural breaks, but in this instance there was still no evidence of serial correlation after the inclusion of interventions.

5.3 Consumption of Spirits in the United Kingdom

The per capita consumption of spirits in the United Kingdom for 1870 to 1938 can be explained, at least partly, by income per capita and relative price. A regression formulated in this way, however, shows significant serial correlation even if a time trend is included. Indeed the data set is a classic one and was used as one of the testbeds for the d statistic in the work of Durbin and Watson (1951).

A regression model with a stochastic trend component, as in (4.5), provides a good fit in many respects. It is more parsimonious than the regression model with a quadratic time trend and a first-order autoregressive disturbance reported by Fuller (1976, p. 426), and the stochastic trend can be interpreted as reflecting changes in tastes.

The estimates reported in Table 7 are for the period 1870–1930. As can be seen, the slope is stochastic, so there are three sets of auxiliary residuals. The associated test statistics are in Table 8. Kohn and Ansley (1989) estimated the model without a slope component, so μ_t is just a random walk. Indeed, estimating such a model might not be unreasonable for preliminary data analysis if we wish to focus attention on structural changes that affect the level. In this particular case, however, the kurtosis statistics in Table 8 are high for both the irregular and level residuals, and the presence of the slope makes very little difference.

The plots shown in Figure 3 indicate a shift in the level in 1909, with several candidates for outliers during World War I. We fitted a level intervention first. The 1918 outlier then stood out most clearly in the irregular. On estimating with a 1918 intervention, 1915 stood out most clearly. This led to a model with a 1909 level intervention together with outlier interventions at 1918 and 1915. All of the diagnostics in this model are satisfactory. Table 7 shows the estimated coefficients of the explanatory variables and compares them with the coefficients obtained from the model without interventions. There is a clear improvement in goodness of fit, and this is reflected in the t statistics shown in parentheses. The innovation diagnostics in the intervention model are entirely satisfactory. It is particularly interesting to note the reduction in the value of the Box–Ljung Q statistic based on the first 10 residual autocorrelations, $Q(10)$; in the original model there were

Table 8. Diagnostic Statistics for Spirits

Residual	K	N
Innovation	2.21	5.87
Irregular	7.53	69.76
Level	5.19	31.65
Slope	.32	.45

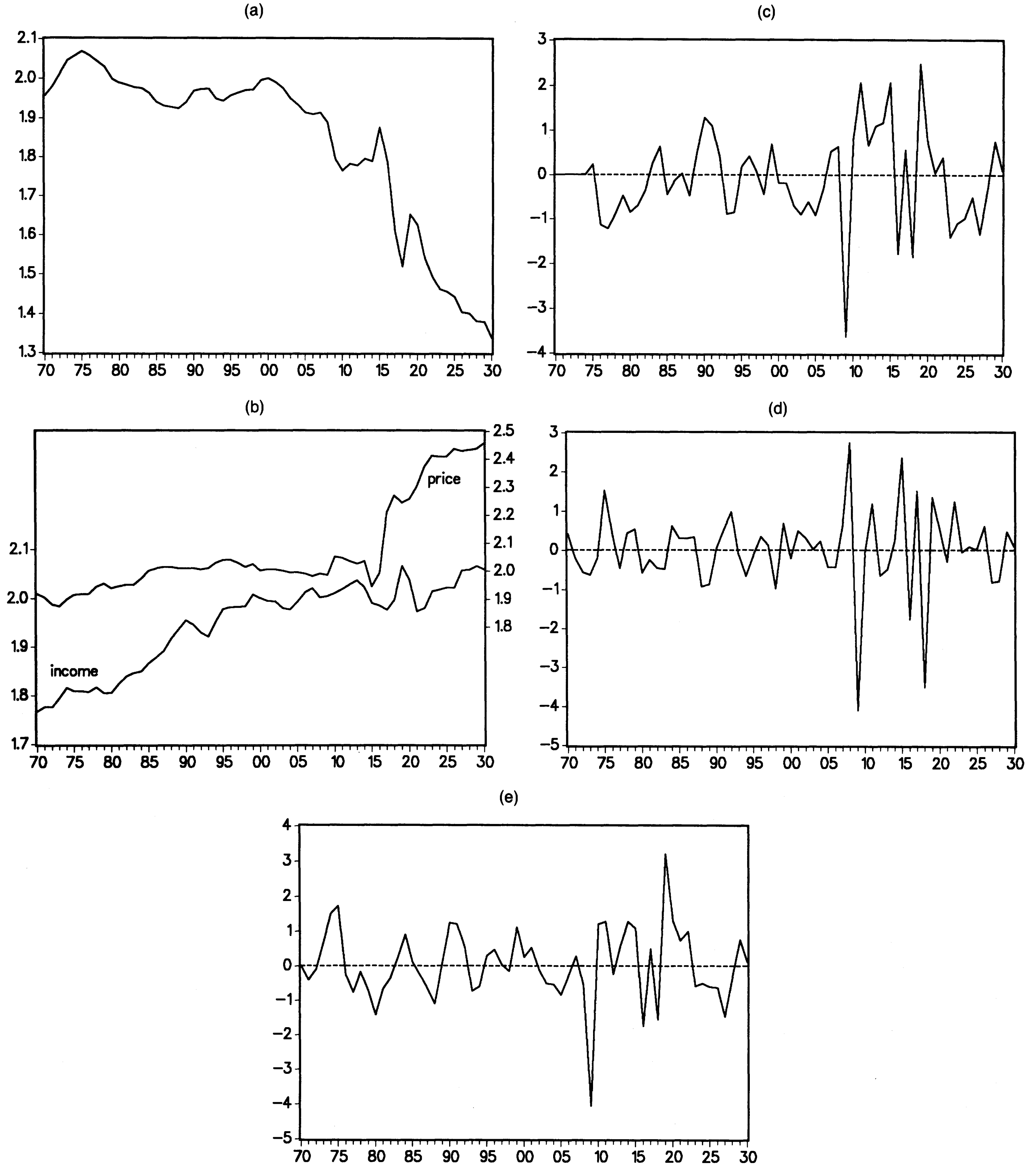

Figure 3. Consumption of Spirits in the United Kingdom, 1870–1930: (a) Observations (in logarithms); (b) Explanatory Variables; (c) Innovations; (d) Irregular; (e) Level Residual.

high autocorrelations at lags 8, 9, and 10, which had no obvious explanation.

Referring back to Prest (1949), who originally assembled the spirits data set, reveals that the figures for 1915–1919 were estimates based on consumption in the British army. Thus they may be considerably less reliable than the other observations, and taking them all out by intervention variables may not be unreasonable. On the basis of Fig. 3(e), there is a case for a structural change in 1919. The general unreliability of the observations in 1915 to 1919, however, makes it difficult to estimate such a change with any degree of confidence.

None of the other results change significantly when the 1919 outlier intervention is replaced by a level shift intervention. The results are shown in the last column of Table 7. The changes in the coefficients of income and price are due to the influence of the observations corresponding to the additional interventions rather than the fact that they may be outliers; see Kohn and Ansley (1989).

The fall in the level in 1909 is highly significant in both of our intervention models and indicates a permanent reduction, other things being equal, of about 9%. It is this feature that is detected by our techniques and that is the prime source of the difference between our model and that of Kohn and Ansley (1989). They identified 1909 as a possible outlier. Their preferred model has outlier interventions for the years 1915–1919 and 1909. Fitting this model, including variations such as the inclusion of a stochastic slope and using time-domain instead of frequency-domain estimation, resulted in a poorer fit than our model and somewhat different coefficients for the explanatory variables. A possible explanation for the shift in 1909 may lie in the program of social reforms begun in that year by Lloyd George; see Tsay (1986, p. 137).

6. CONCLUSIONS

The auxiliary residuals are serially correlated with each other even when the model is correctly specified. Nevertheless, it seems that they are a useful tool for detecting outliers and shifts in the level in structural time series models. Plots of the auxiliary residuals can be very informative, and these can be supplemented with tests for normality and kurtosis corrected to allow for the implied serial correlation. The examples and Monte Carlo experiments illustrate that the techniques work quite well in practice.

ACKNOWLEDGMENTS

We thank Neil Shephard, Peter Robinson, a referee, and an associate editor for helpful comments. All opinions, and any errors, are our responsibility. Support from the Economic and Social Research Council and the Tinbergen Institute is gratefully acknowledged.

APPENDIX: SMOOTHING ALGORITHM FOR COMPUTATION OF THE AUXILIARY RESIDUALS

The unobserved-components time series models discussed in this article can be cast in the state-space form $y_t = z_t'\alpha_t + x_t'\beta + \varepsilon_t$, $\varepsilon_t \sim \text{NID}(0, h_t)$, $\alpha_t = T_t\alpha_{t-1} + W_t\beta + \eta_t$, $\eta_t \sim \text{NID}(0, Q_t)$, $(t = 1, \ldots, T)$, where $\alpha_0 = a + A\delta$, $\beta = b + B\delta$, and δ can be regarded as fixed or diffuse; see the discussion by De Jong (1991). The disturbances ε_t and η_t are assumed to be uncorrelated, but this restriction can be relaxed. A structural time series model is placed in state-space form, with the system vector z_t and system matrix T_t being time-invariant; see Harvey (1989). Regression effects and outlier and structural change interventions are modeled by use of the system vector x_t and system matrix W_t.

Following De Jong (1991), we set up an augmented Kalman filter that gives $a_{t|t-1}$, the estimator, made at time $t-1$, of the state vector at time t when δ is assumed to be 0, together with a matrix $A_{t|t-1}$ that allows for the correction when δ is not 0. Thus the actual estimator of the state vector α_t, made at time $t-1$, is $a_{t|t-1} + A_{t|t-1}\delta = A^{\dagger}_{t|t-1}(1, \delta')'$, where $A^{\dagger}_{t|t-1}$ is the partitioned matrix $(a_{t|t-1}, A_{t|t-1})$. The one-step-ahead prediction errors associated with $A^{\dagger}_{t|t-1}$ are contained in the row vector $v_t^{\dagger} = (v_t, v_t^*)$ so that the scalar v_t and row vector v_t^* correspond to $a_{t|t-1}$ and $A_{t|t-1}$. Thus

$$v_t^{\dagger} = (y_t, \mathbf{0}) - z_t'A^{\dagger}_{t|t-1} - x_t'B^{\dagger}, \qquad t = 1, \ldots, T,$$

where $B^{\dagger}$ is (b, B) and $\mathbf{0}$ is a row vector of zeros. The filter for $A^{\dagger}_{t|t-1}$ is the recursion

$$A^{\dagger}_{t+1|t} = T_{t+1}A^{\dagger}_{t|t-1} + W_{t+1}B^{\dagger} + k_t v_t^{\dagger}, \qquad t = 1, \ldots, T,$$

where $k_t = T_{t+1}P_{t|t-1}z_t'/f_t$ and $f_t = z_tP_{t|t-1}z_t' + h_t$. The mean squared error matrix of the estimated state vector at time $t-1$, $P_{t|t-1}$, is evaluated by the matrix recursion $P_{t+1|t} = T_{t+1}P_{t|t-1}L_t' + Q_t$, where $L_t = T_{t+1} - k_tz_t'$. The starting values of the recursions are $A^{\dagger}_{1|0} = (a, A)$ and $P_{1|0} = 0$.

In addition, we have the following recursion: $M_t = M_{t-1} + v_t^{\dagger}v_t^{\dagger\prime}/f_t$, $(t = 1, \ldots, T)$, where M_t is partitioned as

$$M_t = \begin{bmatrix} q_t & s_t' \\ s_t & S_t \end{bmatrix}$$

and $M_0 = 0$. From this recursion, we obtain the estimator of δ at time t; that is, $m_t = S_t^{-1}s_t$ with mean squared error matrix S_t^{-1}. We also obtain the log-likelihood as $\log L = -\frac{1}{2}T\log 2\pi - \frac{1}{2}\Sigma_{t=1}^{T}\log f_t - \frac{1}{2}(q_T - s_T'S_T^{-1}s_T)$.

The smoothed estimator of the disturbances and the corresponding mean squared error matrices are obtained by the following algorithm derived by Koopman (1992): $\tilde{\varepsilon}_t = h_t(e_t + e_t^*m_T)$, $\text{MSE}(\tilde{\varepsilon}_t) = h_t - h_t^2(d_t - e_t^*S_t^{-1}e_t^{*\prime})$, and $\tilde{\eta}_t = Q_t(r_{t-1} + R_{t-1}m_T)$, $\text{MSE}(\tilde{\eta}_t) = Q_t - Q_t(N_{t-1} - R_{t-1}S_T^{-1}R_{t-1}')Q_t$ $(t = T, \ldots, 1)$, where the row vector $e_t^{\dagger} = (e_t, e_t^*)$, the scalar d_t, the matrices $R_t^{\dagger} = (r_t, R_t)$, and N_t are calculated by the backward recursions $e_t^{\dagger} = v_t^{\dagger}/f_t - k_t'R_t^{\dagger}$, $d_t = 1/f_t + k_t'N_tk_t$, and $R_{t-1}^{\dagger} = z_t'v_t^{\dagger}/f_t + L_t'R_t^{\dagger}$, $N_{t-1} = z_t'z_t/f_t + L_t'N_tL_t$ $(t = T, \ldots, 1)$, started off with $R_T^{\dagger} = 0$ and $N_T = 0$. The auxiliary residuals are obtained by standardizing $\tilde{\varepsilon}_t$ and $\tilde{\eta}_t$ using their mean squared errors.

The disturbance smoother requires, for a typical structural time series model, about the same number of computations as for the Kalman filter. The storage requirement is limited to $v_t^{\dagger}$, f_t, and k_t. Koopman (in press) also discussed a more general disturbance smoother, efficient methods of calculation, and a quick-state smoother.

[*Received July 1991. Revised March 1992.*]

REFERENCES

Bell, W. R. (1984), "Signal Extraction for Nonstationary Time Series," *The Annals of Statistics*, 13, 646–664.

Box, G. E. P., and Pierce, D. A. (1970), "Distribution of Residual Autocorrelations in Autoregressive-Integrated-Moving Average Time Series Models," *Journal of the American Statistical Association*, 65, 1509–1526.

Bruce, A. G., and Martin, R. D. (1989), "Leave-k-out Diagnostics for Time Series," *Journal of the Royal Statistical Society*, Ser. B, 51, 363–424.

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.

Durbin, J., and Watson, G. S. (1951), "Testing for Serial Correlation in Least Squares Regression II," *Biometrika*, 56, 1–15.

Fuller, W. A. (1976), *Introduction to Statistical Time Series*, New York: John Wiley.

Garcia-Ferrer, A., and del Hoyo, J. (in press), "On Trend Extraction Models: Interpretation, Empirical Evidence and Forecasting Performance" (with discussion), *Journal of Forecasting*, 11.

Granger, C. W. J., and Newbold, P. (1977), *Forecasting Economic Time Series*, New York: Academic Press.

Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, U.K.: Cambridge University Press.

Harvey, A. C., and Durbin, J. (1986), "The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling," *Journal of the Royal Statistical Society*, Ser. A, 149, 187–227.

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.

Knuth, D. E. (1981), *The Art of Computer Programming* (Vol. II), Reading, MA: Addison-Wesley.

Kohn, R., and Ansley, C. F. (1989), "A Fast Algorithm for Signal Extraction, Influence and Cross-validation in State Space Models," *Biometrika*, 76, 65–79.

Koopman, S. J. (in press), "Disturbance Smoother for State Space Models," *Biometrika*, 79.

Lomnicki, Z. A. (1961), "Tests for Departure From Normality in the Base of Linear Stochastic Processes," *Metrika*, 4, 37–62.

Maravall, A. (1987), "Minimum Mean Squared Errors Estimation of the Noise in Unobserved Components Models," *Journal of Business & Economic Statistics*, 5, 115–120.

Pierce, D. A. (1979), "Signal Extraction Error in Nonstationary Time Series," *The Annals of Statistics*, 7, 1303–1320.

Prest, A. R. (1949), "Some Experiments in Demand Analysis," *Review of Economics and Statistics*, 31, 33–49.

Tsay, R. S. (1986), "Time Series Model Specification in the Presence of Outliers," *Journal of the American Statistical Association*, 81, 132–141.

Whittle, P. (1991), "Likelihood and Cost as Path Integrals" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 53, 505–538.

LINKED CITATIONS

Diagnostic Checking of Unobserved-Components Time Series Models

Andrew C. Harvey; Siem Jan Koopman

Journal of Business & Economic Statistics, Vol. 10, No. 4. (Oct., 1992), pp. 377-389.

Stable URL:

http://links.jstor.org/sici?sici=0735-0015%28199210%2910%3A4%3C377%3ADCOUTS%3E2.0.CO%3B2-V

References

Signal Extraction for Nonstationary Time Series

William Bell

The Annals of Statistics, Vol. 12, No. 2. (Jun., 1984), pp. 646-664.

Stable URL:

http://links.jstor.org/sici?sici=0090-5364%28198406%2912%3A2%3C646%3ASEFNTS%3E2.0.CO%3B2-N

Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models

G. E. P. Box; David A. Pierce

Journal of the American Statistical Association, Vol. 65, No. 332. (Dec., 1970), pp. 1509-1526.

Stable URL:

http://links.jstor.org/sici?sici=0162-1459%28197012%2965%3A332%3C1509%3ADORAIA%3E2.0.CO%3B2-A

Leave-k-Out Diagnostics for Time Series

Andrew G. Bruce; R. Douglas Martin

Journal of the Royal Statistical Society. Series B (Methodological), Vol. 51, No. 3. (1989), pp. 363-424.

Stable URL:

http://links.jstor.org/sici?sici=0035-9246%281989%2951%3A3%3C363%3ALDFTS%3E2.0.CO%3B2-O

LINKED CITATIONS

Smoothing and Interpolation with the State-Space Model
Piet de Jong
Journal of the American Statistical Association, Vol. 84, No. 408. (Dec., 1989), pp. 1085-1088.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28198912%2984%3A408%3C1085%3ASAIWTS%3E2.0.CO%3B2-J

The Diffuse Kalman Filter
Piet De Jong
The Annals of Statistics, Vol. 19, No. 2. (Jun., 1991), pp. 1073-1083.
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28199106%2919%3A2%3C1073%3ATDKF%3E2.0.CO%3B2-O

Testing for Serial Correlation in Least Squares Regression. III
J. Durbin; G. S. Watson
Biometrika, Vol. 58, No. 1. (Apr., 1971), pp. 1-19.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28197104%2958%3A1%3C1%3ATFSCIL%3E2.0.CO%3B2-L

The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling
A. C. Harvey; J. Durbin
Journal of the Royal Statistical Society. Series A (General), Vol. 149, No. 3. (1986), pp. 187-227.
Stable URL:
http://links.jstor.org/sici?sici=0035-9238%281986%29149%3A3%3C187%3ATEOSBL%3E2.0.CO%3B2-U

An ARIMA-Model-Based Approach to Seasonal Adjustment
S. C. Hillmer; G. C. Tiao
Journal of the American Statistical Association, Vol. 77, No. 377. (Mar., 1982), pp. 63-70.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28198203%2977%3A377%3C63%3AAAATSA%3E2.0.CO%3B2-9

A Fast Algorithm for Signal Extraction, Influence and Cross-Validation in State Space Models
Robert Kohn; Craig F. Ansley
Biometrika, Vol. 76, No. 1. (Mar., 1989), pp. 65-79.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198903%2976%3A1%3C65%3AAFAFSE%3E2.0.CO%3B2-7

LINKED CITATIONS

Disturbance Smoother for State Space Models
Siem Jan Koopman
Biometrika, Vol. 80, No. 1. (Mar., 1993), pp. 117-126.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28199303%2980%3A1%3C117%3ADSFSSM%3E2.0.CO%3B2-8

Minimum Mean Squared Error Estimation of the Noise in Unobserved Component Models
Agustín Maravall
Journal of Business & Economic Statistics, Vol. 5, No. 1. (Jan., 1987), pp. 115-120.
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28198701%295%3A1%3C115%3AMMSEEO%3E2.0.CO%3B2-O

Signal Extraction Error in Nonstationary Time Series
David A. Pierce
The Annals of Statistics, Vol. 7, No. 6. (Nov., 1979), pp. 1303-1320.
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28197911%297%3A6%3C1303%3ASEEINT%3E2.0.CO%3B2-L

Some Experiments in Demand Analysis
A. R. Prest
The Review of Economics and Statistics, Vol. 31, No. 1. (Feb., 1949), pp. 33-49.
Stable URL:
http://links.jstor.org/sici?sici=0034-6535%28194902%2931%3A1%3C33%3ASEIDA%3E2.0.CO%3B2-P

Time Series Model Specification in the Presence of Outliers
Ruey S. Tsay
Journal of the American Statistical Association, Vol. 81, No. 393. (Mar., 1986), pp. 132-141.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28198603%2981%3A393%3C132%3ATSMSIT%3E2.0.CO%3B2-7

Likelihood and Cost as Path Integrals
P. Whittle
Journal of the Royal Statistical Society. Series B (Methodological), Vol. 53, No. 3. (1991), pp. 505-538.
Stable URL:
http://links.jstor.org/sici?sici=0035-9246%281991%2953%3A3%3C505%3ALACAPI%3E2.0.CO%3B2-H