Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study
Author(s): A. C. Harvey and P. H. J. Todd
Source: *Journal of Business & Economic Statistics*, Vol. 1, No. 4 (Oct., 1983), pp. 299-307
Published by: Taylor & Francis, Ltd. on behalf of American Statistical Association
Stable URL: http://www.jstor.org/stable/1391661
Accessed: 08-12-2016 19:12 UTC

©Journal of Business & Economic Statistics, Vol. 1, No. 4, October 1983

Forecasting Economic Time Series With Structural and Box-Jenkins Models: A Case Study

A. C. Harvey
Department of Statistics, London School of Economics, London, WC2A 2AE

P. H. J. Todd
H. M. Treasury, London SW1P 3AG

The basic structural model is a univariate time series model consisting of a slowly changing trend component, a slowly changing seasonal component, and a random irregular component. It is part of a class of models that have a number of advantages over the seasonal ARIMA models adopted by Box and Jenkins (1976). This article reports the results of an exercise in which the basic structural model was estimated for six U.K. macroeconomic time series and the forecasting performance compared with that of ARIMA models previously fitted by Prothero and Wallis (1976).

KEY WORDS: Forecasting; ARIMA models; Structural models; Unobserved components; Kalman filter; Macroeconomic time series.

1. INTRODUCTION

The autoregressive-integrated-moving average (ARIMA) processes introduced by Box and Jenkins (1976) provide a wide class of models for univariate time series forecasting. In the traditional Box-Jenkins framework, the main tools for specifying a suitable model are the correlogram and, to a lesser extent, the sample partial autocorrelation function. However, the correlogram and sample partial autocorrelation function are not always very informative, particularly in small samples. Furthermore, the difficulties in interpretation are compounded when a series has been differenced, and differencing is the rule rather than the exception in economic time series. The result is that inappropriate models are often fitted. Attempts to select ARIMA models by an automatic procedure, based on, say, the Akaike Information Criterion can lead to even worse results; see the examples cited by Jenkins (1982).

Experienced ARIMA model builders usually take into account the type of forecast function that their models imply; see Box, Hillmer, and Tiao (1978) and Jenkins (1982). They also tend to be aware of the type of time series structure that their models imply. Hillmer and Tiao (1982) attempt to make this last point more explicit by defining what constitutes an acceptable decomposition into trend, seasonal, and irregular components. They examine three of the ARIMA models commonly fitted to economic time series and show that certain restrictions must be placed on the range of parameter values for such a decomposition to exist. However, anyone reading the Hillmer and Tiao article, or the related article by Burman (1980), will realize that the relationship between an ARIMA model and the corresponding decomposition is often complex.

An alternative way of proceeding is to formulate models directly in terms of trend, seasonal, and irregular components. This necessarily limits the choice to those models that have forecast functions satisfying any prior considerations. Such models will be termed *structural* models. The fact that the individual components in a structural model have a direct interpretation opens up the possibility of employing a more formal model selection strategy. The question of developing such a strategy will not, however, be pursued in this article. For many economic time series we believe that one of the simplest structural models, which we call the basic structural model, will be adequate.

In this article we report the results of fitting the basic structural model to a number of economic time series, and then compare the predictions with those obtained using the Box-Jenkins models selected by Prothero and Wallis (1976). The idea of the exercise is not to prove that one method yields better forecasts than the other, but rather to show that for these series at least, the forecasts given by the two methods are comparable.

Having demonstrated the viability of one of the simplest models within the structural class, we feel that the case for using this class as the basis for univariate time series modeling is a strong one.

From the technical point of view, all aspects of structural models can be handled by putting them into state space form. In particular, the likelihood function can be constructed in terms of the prediction error decomposition by using the Kalman filter; see Harvey (1982). Once estimates of the parameters have been computed, optimal predictions of future observations, together with their conditional mean square errors, can be obtained using the Kalman filter. Finally, optimal estimates of the individual components can be computed using a smoothing algorithm.

2. STRUCTURAL MODELS

Let y_t be the observed variable. The basic structural model has the form

$$y_t = \mu_t + \gamma_t + \epsilon_t, \quad t = 1, \ldots, T, \tag{2.1}$$

where μ_t, γ_t, and ϵ_t are trend, seasonal, and irregular components, respectively.

The process generating the trend is of the form

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad t = 1, \ldots, T \tag{2.2a}$$

and

$$\beta_t = \beta_{t-1} + \zeta_t, \qquad t = 1, \ldots, T, \tag{2.2b}$$

where η_t and ζ_t are normally distributed independent white noise processes with zero means and variances σ_η^2 and σ_ζ^2, respectively. The essential feature of this model is that it is a local approximation to a linear trend. The level and slope both change slowly over time according to a random walk mechanism.

The process generating the seasonal component is

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t, \quad t = 1, \ldots, T, \tag{2.3}$$

where $\omega_t \sim \mathrm{NID}(0, \sigma_\omega^2)$, and s is the number of "seasons" in the year. The seasonal pattern is thus slowly changing but by a mechanism that ensures that the sum of the seasonal components over any s consecutive time periods has an expected value of zero and a variance that remains constant over time. This specification could be modified by replacing the white noise disturbance term by a moving average process.[1] The advantage of doing this is that it allows a smoother change in the seasonal pattern than that permitted by (2.3). However, for the small sample sizes considered in this article we felt it best to restrict our attention to one simple model.

The disturbances η_t, ζ_t, and ω_t are independent of each other and of the irregular component that is a normally distributed white noise process, that is, $\epsilon_t \sim \mathrm{NID}(0, \sigma^2)$. Although the model as a whole is relatively simple, it contains the main ingredients necessary for a time series forecasting procedure in that it projects a local linear trend and a local seasonal pattern into the future.[2] It will be adequate for many economic time series, and it has the attraction that it involves no model selection procedure whatsoever.

The model can be written in the form

$$y_t = \frac{\xi_t}{\Delta^2} + \frac{\omega_t}{\mathrm{S(L)}} + \epsilon_t, \quad t = 1, \ldots, T, \tag{2.4}$$

where L is the lag operator, Δ is the first difference operator, S(L) is the seasonal operator

$$\mathrm{S(L)} = \sum_{j=0}^{s-1} \mathrm{L}^j, \tag{2.5}$$

and ξ_t is equivalent to an MA(1) process since it is defined by

$$\xi_t = \eta_t - \eta_{t-1} + \zeta_{t-1}. \tag{2.6}$$

The first component on the right side of (2.4) is the trend, while the second is the seasonal component. The fact that the operators Δ^2 and S(L) do not have a root in common is important, because it means that the minimum mean square estimates of both components have finite variance; see Pierce (1979). Put in a more informal way this amounts to saying that changes in the seasonal pattern are not confounded with changes in the trend. Note that the same operators are used in the decompositions proposed by Hillmer and Tiao (1982) and Burman (1980).

Expressing the model in the form (2.4) makes it clear that it belongs to the class of unobserved component ARIMA (UCARIMA) models; compare Engle (1978). Models of this kind are discussed at some length in the book by Nerlove, Grether, and Carvalho (1979), but in their work attention is focused on stationary models fitted to the residuals from a polynomial regression. Thus a deterministic trend is adopted as a matter of course. In (2.2), on the other hand, a deterministic (linear) trend only emerges as a limiting case when $\sigma_\eta^2 = \sigma_\zeta^2 = 0$.

2.1 State Space Form

Suppose for simplicity that $s = 4$. The trend and seasonal components can be written in the form

$$\begin{bmatrix} \mu_t \\ \beta_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_{t-2} \end{bmatrix} = \left[\begin{array}{cc:ccc} 1 & 1 & & 0 & \\ 0 & 1 & & & \\ \hdashline & & -1 & -1 & -1 \\ & 0 & 1 & 0 & 0 \\ & & 0 & 1 & 0 \end{array}\right] \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ \gamma_{t-1} \\ \gamma_{t-2} \\ \gamma_{t-3} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \\ \omega_t \\ 0 \\ 0 \end{bmatrix} \tag{2.7}$$

[1] The coefficients of the moving average model can be specified on a priori grounds or treated as additional parameters to be estimated. In the latter case the order of the process must be restricted to $s - 2$ or the model as a whole ceases to be identifiable. In the acceptable decomposition of Hillmer and Tiao (1982) the order of the MA component can be $s - 1$, but only because of the introduction of an additional restriction requiring that the variances of the trend and seasonal disturbance terms be minimized.

[2] A similar model is employed by Kitagawa (1981) except that he has σ_η^2 constrained to be zero.

or, more compactly, as

$$\boldsymbol{\alpha}_t = \mathbf{C}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\tau}_t, \tag{2.8}$$

where $\boldsymbol{\alpha}_t = (\mu_t, \beta_t, \gamma_t, \gamma_{t-1}, \gamma_{t-2})'$, and so on. Defining

$$\mathbf{z}_t = (1 \quad 0 \quad 1 \quad 0 \quad 0)', \tag{2.9}$$

(2.1) can be written as

$$y_t = \mathbf{z}_t'\boldsymbol{\alpha} + \epsilon_t, \quad t = 1, \ldots, T. \tag{2.10}$$

Equations (2.8) and (2.10) can be regarded as the transition and measurement equations of a state space model; see, for example, Harvey (1981a, Ch. 4).

2.2 Estimation

Maximum likelihood estimators of the parameters in structural models can be computed either in the time domain or in the frequency domain. The most attractive time domain procedure is based on the state space representation of the model. As (2.7) makes clear, the state vector in the basic structural model is nonstationary, but starting values for the Kalman filter can be constructed from the first $s + 1$ observations. The likelihood function for $y_{s+2}, \ldots, y_T$ is then given by the prediction error decomposition, that is,

$$\log \mathrm{L} = -\frac{(T-s-1)}{2} \log 2\pi - \frac{1}{2} \sum_{t=s+2}^{T} \log f_t^* - \frac{1}{2} \sum_{t=s+2}^{T} \frac{\nu_t^2}{f_t^*}, \tag{2.11}$$

where ν_t is the one-step-ahead prediction error at time t, and f_t^* is its variance; compare Harvey (1981a, pp. 204–207). From the practical point of view an easy way of calculating close approximations to the starting values is to initiate the Kalman filter at $t = 0$ with a diagonal covariance matrix in which the diagonal elements are large but finite numbers.

If the variances of η_t, ζ_t, and ω_t are expressed relative to σ^2, the variance of ϵ_t (i.e., σ_η^2/σ^2, σ_ζ^2/σ^2 and σ_ω^2/σ^2), then the likelihood function can be written in the form

$$\log \mathrm{L} = -\frac{(T-s-1)}{2} \log 2\pi - \frac{(T-s-1)}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=s+2}^{T} \log f_t - \frac{1}{2\sigma^2} \sum_{t=s+2}^{T} \frac{\nu_t^2}{f_t}, \tag{2.12}$$

where the variance of ν_t is $\sigma^2 f_t$. It now becomes possible to concentrate σ^2 out of the likelihood function, leaving

$$\log \mathrm{L} = -\frac{(T-s-1)}{2}(\log 2\pi + 1) - \frac{(T-s-1)}{2} \log \tilde{\sigma}^2 - \frac{1}{2} \sum_{t=s+2}^{T} \log f_t, \tag{2.13}$$

where

$$\tilde{\sigma}^2 = (T-s-1)^{-1} \sum_{t=s+2}^{T} \frac{\nu_t^2}{f_t}. \tag{2.14}$$

The advantage of concentrating σ^2 out of the likelihood function is that numerical optimization can be carried out with respect to three parameters rather than four. The disadvantage is that the ML estimator of σ^2 is sometimes equal to zero. When this is the case, the relative variances in the concentrated likelihood tend to infinity.

Approximate ML estimates[3] can be obtained by expressing the likelihood function in terms of the periodogram of the differenced observations, $\Delta\Delta_s y_t$. These differenced observations are stationary and can be expressed as

$$\Delta\Delta_s y_t = \Delta_s \eta_t + (1 + \mathrm{L} + \ldots + \mathrm{L}^{s-1})\zeta_{t-1} + \Delta^2 \omega_t + \Delta\Delta_s \epsilon_t. \tag{2.15}$$

The spectral density of the right side of (2.15) is relatively easy to construct using the autocovariance generating function. Frequency domain methods of this kind have been used quite successfully in the estimation of UCARIMA models; see Nerlove, Grether, and Carvalho (1979).

The autocorrelation structure implied by the basic structural model can be derived directly from (2.15). If $\gamma(\tau)$ denotes the autocovariance of $\Delta\Delta_s y_t$ at lag τ, then for quarterly data

$$\begin{aligned}
\gamma(0) &= 2\sigma_\eta^2 + 4\sigma_\zeta^2 + 4\sigma_\omega^2 + 4\sigma^2, \\
\gamma(1) &= 3\sigma_\zeta^2 - 4\sigma_\omega^2 - 2\sigma^2, \\
\gamma(2) &= 2\sigma_\zeta^2 + \sigma_\omega^2, \\
\gamma(3) &= \sigma_\zeta^2 + \sigma^2, \\
\gamma(4) &= -\sigma_\eta^2 - 2\sigma^2, \\
\gamma(5) &= \sigma^2, \\
\gamma(\tau) &= 0, \quad \tau \geq 6.
\end{aligned} \tag{2.16}$$

These equations can be used as the basis for constructing estimators of the unknown parameters from the sample autocovariance function or from the correlogram. However, since there are six nonzero autocovariances and only four unknown parameters, there is no unique way of forming such estimators. Even in special cases where the number of parameters is equal to the number of nonzero autocovariances, efficient estimators cannot be obtained, just as they cannot be obtained for an MA model. Nevertheless, estimates computed from the correlogram may still be useful as preliminary estimates in a maximum likelihood procedure.

2.3 Prediction and Signal Extraction

Once the parameters in the model have been estimated, predictions of future values, together with their conditional mean square errors, can be made from the

[3] The likelihood in (2.11) is exact if the first $s + 1$ observations are taken to be fixed, although other assumptions can be made; see Harvey (1982). The frequency domain likelihood would be exact if the differenced observations were generated by a circular process.

state space form. The forecast function consists of the local trend with the local seasonal pattern superimposed upon it. If y_t is in logarithms, the estimator of β_t at time T can be regarded as the current estimator of the growth rate. This is of considerable importance to policy makers. The fact that it is immediately available in the structural model, together with its conditional MSE, is a great advantage.

Optimal estimates of the trend and seasonal components throughout the series can be obtained by applying a smoothing algorithm. This is sometimes known as signal extraction. In the present context it can be used to provide a method of model-based seasonal adjustment.

2.4 A Class of Structural Models

Although this article is primarily concerned with the basic structural model, it is worth noting how the model can be generalized. In the first place the trend component can be extended so that it yields a local approximation to any polynomial; see Harrison and Stevens (1976). Second, a more elaborate seasonal model can be fitted, as was observed in the discussion after Equation (2.3). In addition, the seasonal pattern in the eventual forecast function can be made to change over time by adding to (2.3) a component, γ_t^*, which satisfies the condition that $S(L)\gamma_t^*$ is white noise. In the third place the irregular component can be modeled by any stationary ARMA process. Finally, a cyclical component can be brought into the model. This can be done by adding it directly to (2.1) or by incorporating it into the trend.[4]

3. CRITERIA FOR MODEL EVALUATION

Models can be evaluated and compared on the basis of goodness of fit both inside and outside the sample period. The criteria employed in our study are set out in this section.

3.1 Prediction Error Variance

The prediction error variance, that is, the variance of the one-step-ahead prediction errors in the models, is a basic measure of goodness of fit within the sample. For an ARIMA model, the prediction error variance is given directly by the estimator of the variance of the disturbances. For a structural model, the corresponding estimator is given by

$$\tilde{\sigma}_p^2 = \tilde{\sigma}^2 \bar{f}, \tag{3.1}$$

where $\tilde{\sigma}^2$ is given by (2.14) and $\bar{f}$ is defined by

$$\bar{f} = \lim_{t \to \infty} f_t. \tag{3.2}$$

The value of $\bar{f}$ can be found by running the Kalman filter until it reaches a steady state. It can usually be approximated by f_T, although there is an important distinction between $\sigma^2\bar{f}$ and $\sigma^2 f_T$ in that the latter is the *finite* sample prediction error variance.

If the variances in the model are not expressed relative to σ^2, as in (2.11), then $\tilde{\sigma}_p^2 = \bar{f}^*$, where $\bar{f}^*$ is defined analogously to $\bar{f}$.

3.2 Post-Sample Predictions

Once the parameters of a model have been estimated within the sample period, predictions can be made in a post-sample period. The sum of squares of the one-step prediction errors then gives a measure of forecasting accuracy. These quantities can then be compared for rival models.

The prediction errors in the post-sample period can also be compared with the prediction errors within the sample. A test statistic can be employed to test whether the prediction errors in the post-sample period are significantly greater than the prediction errors within the sample period. If they are, we can draw three possible conclusions:

1. The variances of the disturbances are increasing over time;
2. The process generating the observations has changed in some way, possibly due to the impact of certain outside interventions; or
3. The fit achieved in the sample period is to some extent a product of data mining.

If the variances of the disturbances are increasing over time, this should normally be detected within the sample period when the residuals are examined. The heteroscedasticity can often be removed by a suitable transformation such as taking logarithms. Predictive failure due to a changing data generation process simply shows up the weakness of univariate time series models, and there is little that can be done about it apart from extending the models to include explanatory variables. The third reason for predictive failure is the most relevant in the present context, since one of the objections to the Box-Jenkins methodology is that the cycle of identification, estimation, and diagnostic checking can lead to models that, while they give a good fit in the sample period, are inappropriate for making predictions in the future. Such models are usually, though not necessarily, overparameterized.

The mechanics of carrying out a post-sample predic-

[4] In a study of annual U.S. economic time series over a period of 100 years, Nelson and Plosser (1982) found that the correlograms had a pattern consistent with the first differences being stationary about a non-zero mean. In all cases the lag one autocorrelation of first differences was positive. A referee has pointed out that series with this property could not have been generated by an annual structural model in which the trend is (2.2) with $\sigma_\zeta^2 = 0$, and the irregular component is white noise. However, for the U.K. series studied later in this paper, we found that the lag one sample autocorrelation for differenced annual data was negative for all series but one, and in that particular case it was less than .05. This behavior is probably accounted for by the relative stability of the U.K. economy over the relatively short sample period covered (the 1950s and 1960s). Were we to consider modeling longer time series of the kind studied by Nelson and Plosser, we would probably do so by setting up a model in which a stochastic cyclical component was built into the trend.

tive test are as follows. Consider the basic structural model and suppose that the relative variances of η_t, ζ_t, and ω_t are *known*. In this case $\nu_t \sim \text{NID}(0, \sigma^2 f_t)$ for $t = s + 2, \ldots, T$. If the model is correct, the prediction errors in the post-sample period, ν_t, $t = T + 1, \ldots, T + l$, are distributed in a similar way and so

$$\xi(l) = \frac{(\sum_{t=T+1}^{T+l} \nu_t^2/f_t)/l}{\sum_{t=s+2}^{T} (\nu_t^2/f_t)/(T - s - 1)} \sim F_{l,T-s-1}. \quad (3.3)$$

In the special case when $\sigma_\eta^2 = \sigma_\zeta^2 = \sigma_\omega^2 = 0$, the model is a linear regression with time trend and seasonal dummies, and the test based on (3.3) is then identical to the Chow test.

If T is reasonably large, the Kalman filter will be virtually in a steady state with $f_{T+j} \simeq \bar{f}$ for $j = 1, \ldots, l$. Therefore,

$$\xi(l) \simeq \sum_{t=T+1}^{T+l} \frac{\nu_t^2}{l\tilde{\sigma}_p^2}. \quad (3.4)$$

When the relative variances are estimated, the statistic $l \cdot \xi(l)$ has a χ_l^2 distribution under the null hypothesis. However, testing $\xi(l)$ against an F-distribution is still legitimate and may be more satisfactory in small samples.

A post-sample predictive test statistic for an ARIMA model can be derived in a similar way; compare Box and Tiao (1976). The distinction between (3.3) and (3.4) can again be made if a finite sample prediction algorithm is employed; see Harvey (1981b).

3.3 Unconditional Post-Sample Predictions

Another useful measure of forecasting performance is the sum of squares of the prediction errors in the post-sample period for the unconditional predictions. The unconditional predictions are the predictions made for $t = T + 1$ to $T + l$ using the observations up to time $t = T$ only. As pointed out by Box and Tiao (1976), the only formal statistical test of the adequacy of the model is the one based on one-step-ahead predictions. However, looking at predictions several steps ahead is useful as a check that the form of the forecast function is sensible.

4. MODELING MACROECONOMIC TIME SERIES

Prothero and Wallis (1976) fitted Box-Jenkins seasonal ARIMA models to quarterly observations on a number of U.K. economic time series. Their purpose was to compare the performance of these models with that of a small-scale econometric model devised by Hendry. Our purpose is to compare their models with the basic structural model. A subsidiary aim was simply to gain some experience of the problems involved in fitting structural models to relatively short time series.

For each series, Prothero and Wallis presented results for a number of models. In most cases they do not state unequivocally that any one of the models is the preferred specification. However, in light of their comments we have chosen one model in each case. Note that it is the ambiguity surrounding the choice of a suitable ARIMA model that is one of the weaknesses of the whole Box-Jenkins approach. The cycle of identification, estimation, and diagnostic checking is not only time consuming, but it can also on occasion produce poor results through an excess of data mining. The results for Series 5 (Imports) provide a good example.

We must stress again that only one structural model, the basic structural model, was fitted to each series. Hence no model selection was involved at all. It is quite likely that we could have obtained an even better performance by working with the wider class of models sketched out at the end of Section 2. However, in only one case, Series 4, did we feel that restricting ourselves to the basic model was a significant limitation and even in that case the performance of the model was quite reasonable.

4.1 The Data

The data used in the study by Prothero and Wallis (1976) consisted of various U.K. economic time series published in *Economic Trends*. They fitted models to 42 quarterly observations covering the period 1957/3 to 1967/4. In 1969 the Central Statistical Office changed the data base, and this altered the characteristics of the series. This meant that only the first three observations in 1968 could be used for post-sample predictive testing. It was therefore decided to reestimate the preferred ARIMA specifications over the 37 observations from 1957/3 to 1966/3. The same observations were used to fit the basic structural model, while the eight observations 1966/4 to 1968/3 were used for post-sample predictive testing.

Prothero and Wallis fitted their models without taking logarithms. This was done for comparability with Hendry's econometric model. However, there is evidence of heteroscedasticity in some of the series and, other things being equal, one would almost certainly want to consider taking logarithms in these cases. This should be borne in mind when evaluating the results.

4.2 Estimation

The structural model was estimated via the prediction error decomposition using the concentrated form of the likelihood given in (2.13). The likelihood was maximized using the variable metric Gill-Murray-Pitfield algorithm, E04JBP in the NAG library. Analytic derivatives were not used.

The ARIMA models were reestimated using Prothero's own exact ML program, FMLAMS. The randomness of the residuals was assessed by reference to the values of the Box-Pierce Q-statistic.

4.3 Results

The results of fitting basic structural models and the preferred ARIMA specification of Prothero and Wallis—hereafter denoted as P-W—are summarized in Tables 1 and 2. The following points need to be made

Table 1. Estimates of Parameters for the Basic Structural Model and for the Preferred Specification of the ARIMA Model

Series	Structural Model				ARIMA Model[a] and Q-Statistic	
	$\tilde{\sigma}^2_\eta$	$\tilde{\sigma}^2_\zeta$	$\tilde{\sigma}^2_\omega$	$\tilde{\sigma}^2$		
1. Consumer Durables	408.29	.00	181.42	.03	(a)	$\Delta\Delta_4 y_t = (1 - .27L^4)\epsilon_t$, $Q(15) = 10.02$
2. Other Expenditure	305.51	.00	30.01	181.87	(a)	$\Delta\Delta_4 y_t = (1 - .59L^4)\epsilon_t$, $Q(15) = 7.48$
3. Investment	1,392.00	.00	.82	111.90	(d)	$(1 - .27L^4 - .05L^8 - .12L^{12} - .43L^{16})\Delta y_t = \epsilon_t$, $Q(12) = 6.90$
4. Inventory Investment	1,204.08	.00	168.89	371.06	(a)	$\Delta\Delta_4 y_t = (1 - .35L^4 + .36L^2)(1 - .60L^4)\epsilon_t$, $Q(13) = 5.17$
					(e)	$(1 - .25L - .37L^2)(1 - .20L^4 - .06L^8 - .14L^{12} - .51L^{16})y_t = \epsilon_t$, $Q(10) = 4.32$
5. Imports	879.74	.00	.00	268.02	(c)	$(1 + .91L^4 + .94L^8 + .89L^{12} + .21L^{16})\Delta_4 y_t = 219.33 + \epsilon_t$, $Q(11) = 12.13$
6. GDP	3,375.47	.00	599.71	.20	(a)	$\Delta\Delta_4 y_t = (1 - .30L)(1 - .79L^4)\epsilon_t$, $Q(14) = 9.48$

[a] Letters in parentheses denote the specification in Prothero and Wallis (1976).

for the results on individual series.

1. *Consumers' expenditure on durable goods.* The P-W specification (a) was chosen because P-W considered it to be "an obvious choice." The structural model gave a slightly better fit both within and outside the sample period. Both models clearly fail the post-sample predictive test, but the reason for this is almost certainly the change in vehicle registration policy introduced in 1967; see Prothero and Wallis (1976, p. 484).

2. *Consumers' expenditure on all other goods and services.* Again the preferred ARIMA model is of the simple form adopted in the first series. As before the structural model gives a slightly better fit inside the sample period, but the ARIMA model does better outside the sample period. However, the differences are not great, and as with Series 1, the overall conclusion must be that there is little distinction between the two methods in terms of forecasting performance. The relatively high values of the post-sample predictive test statistics are almost certainly explained by the heteroscedasticity in the series. This could probably be rectified by modeling the observations in logarithms.

3. *Investment.* P-W considered four models and chose (d) as being the "most reasonable." Unlike the preferred ARIMA models for the previous two series it contains more parameters than the structural model. Perhaps as a result of this it fits slightly better in the sample period. However, it also fits better in the post-sample period, although as with the other two series, the difference in forecasting performance is not great.

Although the preferred ARIMA model appears to be satisfactory for one-step-ahead forecasting (at least for the post-sample period considered), it is less impressive over a longer time horizon. The sum of squares of the unconditional predictions over the sample period was 15,942, which is approximately twice the figure obtained with the structural model. Futhermore, the eventual

Table 2. Forecasting Performance of the Basic Structural Model and the Preferred Specification of the ARIMA Model

Series	Prediction Error Variance		Post-Sample Prediction Sum of Squares		Predictive F-test[a]		Unconditional Post-Sample Prediction Sum of Squares for Structural Model
	Structural	ARIMA	Structural	ARIMA	Structural	ARIMA	
1. Consumer Durables	1,349	1,509	71,705	78,372	6.67	6.49	46,218
2. Other Expenditure	924	1,084	20,211	18,868	2.75	2.17	22,054
3. Investment	1,823	1,745	13,651	12,065	.95	.86	7,551
4. Inventory Investment	3,274	(a) 3,162 (e) 2,157	62,683	(a) 59,596 (e) 43,381	2.40	(a) 2.36 (e) 2.51	76,929
5. Imports	1,532	1,259	57,428	604,080	4.75	59.98	34,223
6. GDP	7,663	7,733	47,794	33,020	.78	.53	56,091

[a] 5% critical value for $F_{8,32}$ is approximately 2.25.

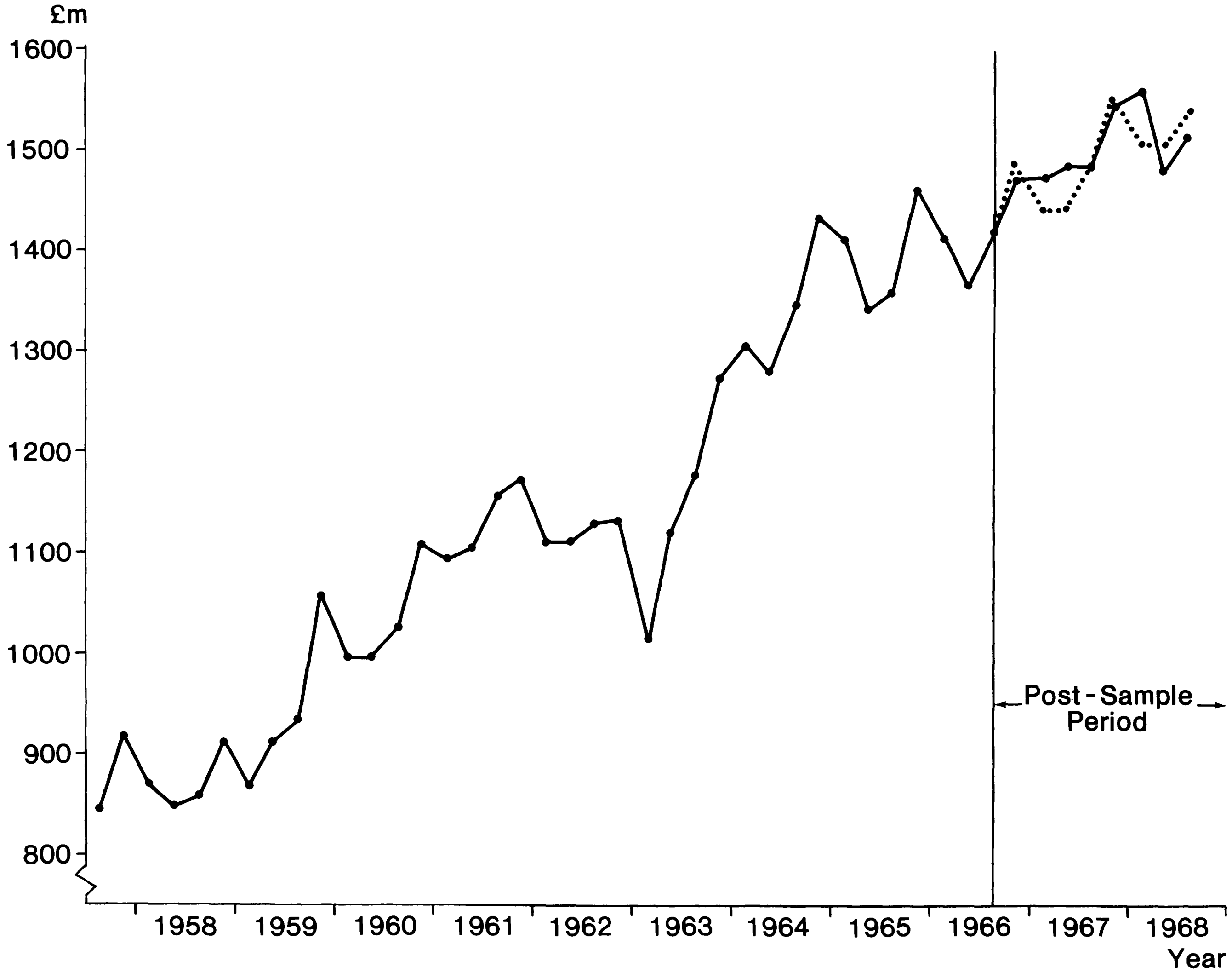

Figure 1. Unconditional Forecasts for Investment Using the Basic Structural Model

forecast function is horizontal, and it seems unlikely that one would want a forecast function of this kind for a series that shows a clear upward movement over time. The forecast function for the structural model over the post-sample period is shown in Figure 1, and one can see that it tracks the series quite well.

4. *Inventory Investment.* P-W observed that an examination of the sample autocorrelation function for various differences of the series suggests the operator $\Delta\Delta_4$. However, the preferred model was one fitted to the undifferenced observations. An examination of the series indicates that this is probably not unreasonable as there are no strong upward or downward movements in the series. Table 1 shows the models P-W fitted to both the differenced ($\Delta\Delta_4$) and undifferenced series. The differenced model has a similar performance to the structural model, but the undifferenced model is clearly superior to both. For this series, therefore, the basic structural model is inadequate and a more general model, in which the irregular component is modeled by an autoregressive process, needs to be employed.

5. *Imports of Goods and Services.* For this series the preferred specification of P-W was their model (c) "on account of its small residual variance and small value[5] of the Q-statistic." This model has a smaller prediction error variance than the structural model, but its forecasting performance is disastrous. The devaluation of the pound in 1967 meant that any univariate model would have difficulty in forecasting over the post-sample period with any reasonable degree of accuracy, and this is apparent from the post-sample F-statistic value of 4.75 for the structural model. However, the F-statistic of 59.98 achieved with the ARIMA model is of a completely different order of magnitude. The performance of this model is therefore a particularly dramatic example of the dangers inherent in the Box-Jenkins methodology. We should, however, add that P-W's (e) and (f) specifications—which are somewhat more conventional forms—forecasted in much the same way as the basic structural model.

[5] P-W had $Q(11) = 4.79$. Our estimates are based on fewer observations and Q is rather larger, although still not significant at the 5% level.

6. *Gross Domestic Product.* P-W's model (a) is one of the standard forms of an ARIMA model. The prediction error variance is similar to the one obtained in the structural model, and both perform rather well in the post-sample period.

Overall, the performance of the basic structural model is quite good. For Series 1, 2, 3, and 6, its forecasting performance both inside and outside the sample period is similar to that achieved by the preferred ARIMA model. For Series 4, the ARIMA model is clearly better, but in this case a more general form of the structural model is called for. For Series 5, on the other hand, the forecasting performance of the preferred ARIMA model is disastrous while that of the basic structural model is quite reasonable.

For all of the series, the basic structural model produced a sensible forecast function. This is reflected in the last column in Table 2, which shows the sum of squares of the unconditional predictions over the post-sample period. In half the cases it is actually smaller than the sum of squares of the conditional forecast!

As regards the estimated variances in the structural model, it is interesting to note that in all cases the estimate of σ_ζ^2 is consistent with a steady increase over time, which all of the series display. However, the fact that the estimate goes right to zero may be a reflection of the small sample size and the method of estimation. When exact ML estimation is carried out in the time domain, it is not unusual to find some of the estimates ending up on the boundary of the parameter space; see Sargan and Bhargava (1983) and our Appendix for further details. The properties of approximate ML estimates computed in the frequency domain may be quite different, but this is a matter for future investigation.

5. CONCLUSIONS

The forecast function for a structural model can always be reproduced by an ARIMA model. For example, it is clear from (2.15) and (2.16) that the basic structural model is equivalent to an MA(s + 1) model for $\Delta\Delta_s y_t$ in which the parameters are subject to nonlinear restrictions. The attraction of specifying models in terms of a well-defined structure is that attention is more likely to be confined to models that yield forecast functions of an acceptable form.

The basic model we propose has a similar structure to the Bayesian model of Harrison and Stevens (1976). However, while Harrison and Stevens make assumptions about plausible values for the variances of the disturbances, this article has shown that it is possible to estimate these variances even with a relatively small number of observations. In all the cases examined this led to a sensible forecast function. Furthermore, the forecasting performance of the estimated models compared well with the forecasting performance of ARIMA models selected after the usual process of identification, estimation, and diagnostic checking. Given these results, we feel that the conceptual advantages of structural models make them attractive as a class of univariate time series models.

APPENDIX: BOUNDARY SOLUTIONS OF MAXIMUM LIKELIHOOD ESTIMATES IN STRUCTURAL MODELS

Consider a simple case of the structural model in which there is no seasonal component and no slope; that is,

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim \text{NID}(0, \sigma^2) \qquad \text{(A.1a)}$$

and

$$\mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim \text{NID}(0, \sigma_\eta^2). \qquad \text{(A.1b)}$$

This model is equivalent to an ARIMA (0, 1, 1) model

$$\Delta y_t = \xi_t + \theta\xi_{t-1}, \quad \xi_t \sim \text{NID}(0, \sigma_\xi^2), \qquad \text{(A.2)}$$

in which

$$\theta = [(q^2 + 4q)^{1/2} - 2 - q]/2,$$

where $q = \sigma_\eta^2/\sigma^2$; see Harvey (1981a, p. 170).

For an MA(1) model, Sargan and Bhargava (1983) have shown that there is a relatively high probability that an exact ML estimate of θ will be *exactly* equal to -1, even when the true value is some distance from -1. In Model (A.1), a relatively low value of σ_η^2 corresponds to a value of θ close to -1 in (A.2). In these circumstances, ML estimates of zero will not be uncommon for σ_η^2.

ACKNOWLEDGMENTS

An earlier version of this article was presented at the Second International Symposium on Forecasting, Istanbul, July 1982. We would like to thank the two referees of this journal for their useful comments on the first draft. The views expressed in this article are not necessarily those of H. M. Treasury.

[Received August 1982. Revised March 1983.]

REFERENCES

BOX, G. E. P., HILLMER, S. C. and TIAO, G. C. (1978), "Analysis and Modelling of Seasonal Time Series," in *Seasonal Analysis of Economic Time Series*, ed. A. Zellner, Washington, DC: Bureau of the Census, 309–334.

BOX, G. E. P., and JENKINS, G. M. (1976), *Time Series Analysis: Forecasting and Control* (revised ed.), San Francisco: Holden-Day.

BOX, G. E. P., and TIAO, G. C. (1976), "Comparison of Forecasts With Actuality," *Applied Statistics*, 25, 195–200.

BURMAN, J. P. (1980), "Seasonal Adjustment by Signal Extraction," *Journal of the Royal Statistical Society*, Series A, 143, 321–337.

ENGLE, R. F. (1978), "Estimating Structural Models of Seasonality," in *Seasonal Analysis of Economic Time Series*, A. Zellner (ed.) Washington DC: Bureau of the Census, 281–308.

HARRISON, P. J., and STEVENS, C. F. (1976), "Bayesian Forecasting," *Journal of the Royal Statistical Society*, Series B, 38, 205–247.

HARVEY, A. C. (1981a), *Time Series Models*, Deddington: Philip Allan, and New York: John Wiley.

——— (1981b), "Finite Sample Prediction and Overdifferencing,"

Journal of Time Series Analysis, 2, 221–232.
——— (1982), "Estimation Procedures for a Class of Univariate Time Series Models," LSE Econometrics Program Discussion Paper No. 32.
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.
JENKINS, G. M. (1982), "Some Practical Aspects of Forecasting in Organizations," *Journal of Forecasting*, 1, 3–21.
KITAGAWA, G. (1981), "A Nonstationary Time Series Model and Its Fitting by a Recursive Filter," *Journal of Time Series Analysis*, 2, 103–116.
NELSON, C. R., and PLOSSER, C. I. (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139–162.
NERLOVE, M., GRETHER, D. M., and CARVALHO, J. L. (1979), *Analysis of Economic Time Series*, New York: Academic Press.
PIERCE, D. A. (1979), "Signal Extraction Error in Nonstationary Time Series," *Annals of Statistics*, 7, 1303–1320.
PROTHERO, D. L., and WALLIS, K. F. (1976), "Modelling Macroeconomic Time Series," *Journal of the Royal Statistical Society*, Series A, 139, 468–500.
SARGAN, J. D., and BHARGAVA, A. (1983), "Maximum Likelihood Estimation of Regression Models with First-Order Moving Average Errors when the Root Lies on the Unit Circle," *Econometrica*, 40, 617–636.

Comment

Craig F. Ansley
Graduate School of Business, University of Chicago, Chicago, IL 60637

Harvey and Todd (HT) have introduced a simple structural model, a variance components model, which they claim will be adequate for many economic time series. Variants of the same model have been described by Kitagawa (1983), who justifies the models HT(2.2) and (2.3) as smoothness priors, as in Wahba (1978). We will compare the HT model to another that has found wide application for seasonal series, the ARIMA-$(0, 1, 1)(0, 1, 1)_s$ model discussed in Box and Jenkins (1976, Sec. 9.2) and to a related structural model.

HT point out that the variance of ζ_t in their Equation (2.2b) is usually very small, and because of the pile-up effect in maximum likelihood estimation of moving average models, it is more often than not estimated to be zero. The pile-up effect was first described analytically by Cryer and Ledolter (1981); a more complete version appears in their 1980 Technical Report. Monte Carlo evidence of the pile-up effect for other moving average models is given in Ansley and Newbold (1980).

For this reason we consider the simplified version of the model:

$$y_t = \alpha + \beta t + \mu_t + \gamma_t + \epsilon_t, \qquad (1a)$$

$$\mu_t = \mu_{t-1} + \eta_t, \qquad (1b)$$

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t, \qquad (1c)$$

where

$$\epsilon_t \sim N(o, \sigma^2),$$

$$\omega_t \sim N(0, \lambda_1\sigma^2),$$

$$\eta_t \sim N(0, \lambda_2\sigma^2).$$

The linear trend term $\alpha + \beta t$ is included because the transformation operations required for stationarity in this model do not remove the drift β, which must therefore be estimated explicitly. This point is missed by HT.

Applying the difference operators $\Delta \equiv 1 - \mathrm{L}$ and $\mathrm{S(L)}$ defined in HT(2.5), we have

$$\Delta \mathrm{S(L)} y_t = s\beta + \sum_{0}^{s-1} \eta_{t-j} + \omega_t - \omega_{t-1} + \epsilon_t - \epsilon_{t-s}. \qquad (2)$$

Model (2) is stationary, and has four parameters: λ_1, λ_2, σ^2, and β. HT are effectively advocating the transformation $\Delta\mathrm{S(L)}$ to reduce an arbitrary series to stationarity.

Consider now the transformation $\Delta\Delta_s$ used by Box and Jenkins (1976), where $\Delta\Delta_s \equiv 1 - \mathrm{L}^s$. Noting that $\Delta\mathrm{S(L)} = \Delta_s$, we have from (2)

$$\Delta\Delta_s y_t = \eta_t - \eta_{t-s} + \omega_t - 2\omega_{t-1} + \omega_{t-2} + \epsilon_t - \epsilon_{t-1} - \epsilon_{t-s} + \epsilon_{t-s-1}. \qquad (3)$$

We now suggest an alternative structural model,

$$y_t = \alpha + \beta t + \mu_t + \gamma'_t + \epsilon_t, \qquad (4a)$$

$$\mu_t = \mu_{t-1} + \eta_t, \qquad (4b)$$

$$\gamma'_t = \gamma'_{t-s} + \omega_t. \qquad (4c)$$

In this case we have

$$\Delta\Delta_s y_t = \omega_t - \omega_{t-1} + \eta_t - \eta_{t-s} + \epsilon_t - \epsilon_{t-1} - \epsilon_{t-s} + \epsilon_{t-s-1}. \qquad (5)$$

The difference between models (1) and (4) is essentially the terms $\omega_t - 2\omega_{t-1} + \omega_{t-2}$ and $\omega_t - \omega_{t-1}$ in their stationary derivatives (3) and (5) respectively. This arises from the seasonal components (1c) and (4c). In (1c), $\mathrm{S(L)}\gamma_t$ is white noise; in (4c), $\mathrm{S(L)}\gamma'_t$ is a random walk. Thus the models differ in the smoothness of the evolution of $\mathrm{S(L)}\gamma_t$: HT impose no smoothness prior at all on $\mathrm{S(L)}\gamma_t$, while the $\Delta\Delta_s$ transformation implies a model