Economic Applications: The Beveridge–Nelson Decomposition

Co-authors: Chin Nam Low¹ and Heather M. Anderson²

Two features that characterize most macroeconomic time series are sustained long run growth and fluctuations around the growth path. These features are often called “trend” and “cycle” respectively, and the macroeconomic, econometric and statistical literatures contain a variety of techniques for decomposing economic time series into components that are roughly aligned with these notions. Most popular in the macroeconomic literature is the use of the Hodrick–Prescott (1980) filter for trend-cycle decomposition, followed by the Beveridge–Nelson (1981) decomposition, the decomposition implied by Harvey’s (1985) unobserved component model, and a myriad of other contenders. Some, but not all, of these decomposition methods are based on statistical models, and there is vigorous debate about which of these methods leads to series that best capture the concepts of economic growth and business cycles. Canova (1998) provides an excellent survey of various methods that are used to decompose economic data, and he also outlines the motivational and quantitative differences between them. He is careful to point out that the “cycles” which result from statistical filters need not have a close correspondence with the classical ideas that underlie business cycle dating exercises undertaken by think-tanks such as the National Bureau of Economic Research in the USA. He also emphasizes that such a correspondence is not even desirable. Alternative decomposition techniques extract different types of information from the data, and each can be used to focus on different aspects of economic theory. Which filter is appropriate depends on the question at hand.

The Beveridge–Nelson (1981) (BN) decomposition developed from the observation that economic growth is not predictable, in the sense that it has an intrinsic stochastic component. Deterministic models of trend in economic output, often represented by polynomial functions of time, are

¹ Dr. Chin Nam Low, Aretae Pty Ltd, Singapore.

² Professor Heather Anderson, School of Economics, Australian National University, Australia.

then inappropriate indicators of “trend.” Related work undertaken by Nelson and Plosser (1982) provided compelling empirical evidence that nearly all macroeconomic time series contain a (single) unit root with drift, and this finding had a profound impact on economists, first because stochastic trends are not predictable, and second because innovations to stochastic trends are not dampened but persist into the future. The innovations could be interpreted as technological advances, consistent with standard models of economic growth (as in Solow 1956). Innovations in this context could also be interpreted as fiscal or monetary policy shocks that had permanent long-run effects on the economy. Of particular importance was the implication that innovations to a stochastic trend had a persistent effect, and this, together with Friedman’s (1957) concept of “permanent” and “transitory” components in income, led Beveridge and Nelson to develop their decomposition of macroeconomic time series into “permanent” and “transitory” components.

Economists use BN decompositions to provide measures of persistence in economic output (GDP). They also use the BN decomposition to date and predict various features of the business cycle. An interesting statistical aspect of this decomposition is that it implies perfect correlation between innovations to the permanent and transitory components, and this perfect correlation implies an innovations (single source of error) state-space representation. An interesting aspect of the “trend” and “cycle” interpretation of this decomposition is that the perfect correlation between the innovations to each component will imply that shocks to an economic variable will affect both trend and cycle.

Economists are often interested in the relative contributions of trend and cycle to the total variation in macroeconomic variables, and base their measures of relative contribution on BN components, so as to account for stochastic trends in the data. Empirically, the variation in the BN permanent component is usually very close to the total variation, and the contribution of the transitory component is usually very small. Stock and Watson (1988) discuss this finding, pointing out that economists need to recognize the substantial trend component in output, even if they are primarily interested in short term variation. In practice, the short term variation in the transitory component (cycle) is almost negligible, and its serial dependence properties are typically very weak.

Another interesting property of the BN decomposition arises from the inclusion of (positive) drift in the BN “trend.” This drift ensures that economic growth (i.e., the BN trend) will be positive more often than it is negative, automatically creating an asymmetry in the growth process. This is consistent with the classical economic view of the business cycle that associates recessions with lower growth. It runs counter to a more statistical view that there are asymmetries in business cycles, but one can reconcile the apparent anomaly as a simple artifact of the use of different definitions.

Some economists view recessions as periods of low growth, while others view recessions as low points in cycles.

Macroeconomists often use “growth cycle” data to study asymmetries in economic time series, where growth cycles are taken to be the cycles discussed in Zarnowitz and Boschan (1977). Growth cycle data are constructed by taking the first differences (of the logarithms) of the raw time series, and the asymmetries are typically modeled using regime switching autoregressive specifications that allow the intercept and/or autoregressive parameters to change over time. Examples include Markov switching (MS) models (Hamilton 1989), threshold autoregressive (TAR) models (Potter 1995), and smooth transition autoregressive (STAR) models used by Teräsvirta and Anderson (1992). Most of these models allow for two regimes, and the regimes are loosely interpreted as recessionary and expansionary phases of the business cycle. Changes in the intercept correspond to changes in growth rates, while changes in the autoregressive parameters indicate changes in the short-run dynamic characteristics of growth. Low et al. (2006) show that it is possible to extend the BN innovations state space representation to allow for asymmetries in growth cycles, and this is briefly discussed at the end of the chapter.

The aim of this chapter is to familiarize readers with the BN decomposition of economic time series, because it provides an interesting application of the linear innovations state space approach to economics. The treatise provided here focusses on the decomposition of a single variable (in practice, this is usually the logarithm of the gross domestic product (GDP) of a country) into just two components. We do not consider more general linear innovations state space approaches that might explicitly account for seasonality, because economic theory has little to say about seasonal effects, and economists typically think in terms of seasonally adjusted data. We also do not consider local trend models, because there is very little empirical evidence that macroeconomic time series follow processes that are integrated of order two, and economic theory does not distinguish between different types of low frequency data. A state space framework that can be used to undertake the BN decomposition is outlined in Sect. 20.1. This framework is based on Anderson et al. (2006) and uses the perfect correlation between permanent and transitory components to recast the BN decomposition as an innovations state space model. It turns out that this state space approach avoids a computational problem associated with other techniques for estimating the BN permanent component (see Miller 1988; Newbold 1990; Morley 2002), and it also facilitates the direct estimation of various measures of “persistence in output.”

In Sect. 20.2 we discuss the application of the state space approach to decomposing data from the USA, the UK and Australia. Some extensions to nonlinear processes are discussed in Sect. 20.3.

20.1 The Beveridge–Nelson Decomposition

The starting point for the Beveridge–Nelson (1981) decomposition is that most economic time series can be approximated by an ARIMA($p, 1, q$) model. The permanent component of the ARIMA($p, 1, q$) series is taken to be the limiting forecast of the series as the forecast horizon goes to infinity (given all history up to time t), and the transitory component is then the difference between the present observed value of the series and the permanent component. The definition of the permanent component embodies a clear focus on forecasting. Further, Beveridge and Nelson (1981) show that the BN transitory component consists of the forecastable momentum of the series at each point in time.

The original derivation of the BN decomposition of a time series y_t assumes that y_t is a linear $I(1)$ variable with a stationary Wold representation given

$$\Delta y_t = b + \gamma(L)\varepsilon_t, \quad (20.1)$$

where b is the long run growth or drift, $\gamma(L)$ is a polynomial in the lag operator L with $\gamma(0) = 1$ and $\sum_{i=0}^{\infty} |\gamma_i| < \infty$, and ε_t is the IID($0, \sigma^2$) one-step-ahead forecast error of y_t . Economists are typically interested in $\gamma(1)$, the multiplier which measures the long-run effect of a shock ε_t on y_t . The 1981 implementation of the decomposition defined the permanent component as

$$\tau_t = \lim_{h \rightarrow \infty} E(y_{t+h} - hb \mid y_t, y_{t-1}, \dots, y_1),$$

and then set the transitory component to be $c_t = y_t - \tau_t$. Beveridge and Nelson then showed that $\Delta \tau_t = b + \gamma(1)\varepsilon_t$, which implied that this permanent component followed a random walk with drift. They also showed that the transitory component was stationary.

The approach used by Anderson et al. (2006) assumes an ARIMA($p, 1, q$) model with drift for y_t , so that

$$\gamma(L) = \frac{\theta_q(L)}{\phi_p(L)} = \frac{1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q}{1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p}.$$

The notation α is used to denote the long-run multiplier given by $\gamma(1) = \theta_q(1)/\phi_p(1)$. Equation (20.1) can then be written as

$$(1 - L)y_t = b + \left(\frac{\theta_q(L)}{\phi_p(L)} - \alpha \right) \varepsilon_t + \alpha \varepsilon_t,$$

so that y_t can be decomposed into two components:

$$y_t = \left(\frac{b}{1 - L} + \frac{\alpha \varepsilon_t}{1 - L} \right) + \left(\frac{\theta_q(L) - \alpha \phi_p(L)}{(1 - L)\phi_p(L)} \right) \varepsilon_t = \tau_t^* + c_t^*. \quad (20.2)$$

It is clear that $\Delta\tau_t^* = b + \alpha\varepsilon_t$, so that $\tau_t^* = \tau_t$, the BN permanent component. It follows that $c_t^* = c_t$, the BN transitory component. Therefore, we simply use the notation τ_t and c_t for these two components from now on. Note that the numerator of the c_t term in the decomposition (20.2) has a unit root by construction.³ It is also useful to note that the expression for c_t in (20.2) measures the short-run effects of ε_t on y_t ; that is, the difference between the total and long-run effects of ε_t on y_t .

Inspection of the two components in (20.2) shows that they are driven by the same innovation, so that innovations to τ_t and c_t are perfectly correlated. This perfect correlation is a by-product of the BN decomposition rather than an assumption, but its presence allows one to model the decomposition within a linear innovations state space framework.

We can rewrite the expression for τ_t as

$$\tau_t = b + \tau_{t-1} + \alpha\varepsilon_t, \quad (20.3)$$

which shows that the permanent component is a random walk with drift b and an uncorrelated innovation given by $\alpha\varepsilon_t$. Thus, what economists call the *permanent component* or a *stochastic trend* corresponds to a *local level with drift* model (see Sect. 3.5.2). We can also rewrite the expression for c_t as

$$c_t = \left(\frac{\theta_q(L) - \alpha\phi_p(L)}{(1-L)\phi_p(L)} \right) \varepsilon_t = \frac{\psi_r(L)}{\phi_p(L)} \varepsilon_t, \quad (20.4)$$

where $\psi_r(0) = 1 - \alpha$, and the order of $\psi_r(L)$ satisfies the condition $r \leq \max(p-1, q-1)$. Letting $\phi_p^*(L) = \phi_1L + \phi_2L^2 + \dots + \phi_pL^p$ and $\psi_r^*(L) = \psi_1L + \psi_2L^2 + \dots + \psi_rL^r$, the expression for the transitory component becomes

$$c_t = \phi_p^*(L)c_t - \psi_r^*(L)\varepsilon_t + (1 - \alpha)\varepsilon_t. \quad (20.5)$$

Substituting (20.3) and (20.5) into $y_t = \tau_t + c_t$ leads to

$$y_t = b + \tau_{t-1} + \phi_p^*(L)c_t - \psi_r^*(L)\varepsilon_t + \varepsilon_t. \quad (20.6)$$

Equations (20.6), (20.3) and (20.5) are a complete representation of a difference-stationary time series. The associated lag polynomials can be of degree greater than one, so it is not a linear innovations state space representation. However, as shown in the next section, equations like this can always be recast into the first-order form associated with linear innovations state space models.

Following the convention of calling the permanent component of the BN decomposition “the trend” and the transitory component “the cycle,” the parameter of interest in empirical studies of (the logarithms of) output is typically α , which measures the long run percentage increase in GDP resulting from a 1% shock in GDP in one quarter. This parameter is often used as a

³ This is confirmed by substituting $L = 1$ into the numerator to give a value of zero.

measure of persistence (see Campbell and Mankiw 1987). It also determines the relative size of (contemporaneous) innovations to each component. In practice, if $\alpha < 1$ then the trend and cycle will have perfect positive correlation and both components will share in the variation of the data. However, if $\alpha > 1$, then the innovations in the trend and cycle will have perfect negative correlation, and the trend τ_t will be more variable than y_t . Some researchers, such as Proietti (2002), have questioned whether one should call τ_t a “trend” when it is more volatile than the output itself, but as was pointed out by Morley et al. (2003) (who observed that $\alpha > 1$ for real US GDP), a shock to output can shift the trend so that the output is behind the trend until it catches up. Thus, it is quite reasonable for “trend innovations” to be negatively correlated with “cycle innovations,” and for the former innovations to be more variable than output innovations.

20.2 State Space Form and Applications

We now explore the use of the linear innovations state space approach for computing the BN permanent/transitory decompositions for ARIMA(0,1,1), ARIMA(1,1,0) and ARIMA(2,1,2) models of the logarithms of real output for the United States, the United Kingdom and Australia. The US models coincide with those used by Stock and Watson (1988) in their study of the contribution of the trend component to real US GNP, and the scope is broadened to include decompositions for the UK and Australia to demonstrate the relative contribution of trends in other countries.

In this study, we use quarterly GNP data for the USA (from 1947:1 to 2003:1), quarterly GDP data for the UK (from 1960:1 to 2003:1) and quarterly GDP data for Australia (from 1979:3 to 2003:3).⁴ As noted above, the primary parameter of interest is Campbell and Mankiw’s (1987) persistence measure α . Because researchers are often interested in the fraction of the variance of the quarterly change in real output that can be attributed to changes in its stochastic trend, the computed BN trends are used to calculate Stock and Watson’s (1988) R^2 measure of this ratio. The empirical results are presented in Table 20.1, and the details relating to the innovations state space formulation are outlined below.

20.2.1 ARIMA(0,1,1) Model

The BN components of an ARIMA(0,1,1) model are

$$\begin{aligned}\tau_t &= b + \tau_{t-1} + \alpha \varepsilon_t, \\ c_t &= (1 - \alpha) \varepsilon_t,\end{aligned}$$

⁴ The US data were obtained from the Federal Reserve Bank of St Louis, the UK data from the Office of National Statistics, and the Australian data from the Australian Bureau of Statistics. The data were transformed so that y_t is 100 times the natural logarithm of the GNP or GDP in each quarter.

Table 20.1. Measures of the importance of trend in real log GNP/GDP.

Univariate statistical model	Long-run change in GNP predicted from a 1% shock change in GNP in one quarter ($\hat{\alpha}$)	Variance ratios R^2
US GNP (1947:1–2003:1)		
ARIMA(0,1,1)	1.2701 (0.0552)	0.9339
ARIMA(1,1,0)	1.5226 (0.1464)	0.8817
ARIMA(2,1,2)	1.2653 (0.1459)	0.8458
UK GDP (1960:1–2003:1)		
ARIMA(0,1,1)	0.9945 (0.0724)	0.9999
ARIMA(1,1,0)	0.9940 (0.0759)	0.9999
ARIMA(2,1,2)	1.2267 (0.1587)	0.9686
Australia GDP (1979:1–2003:3)		
ARIMA(0,1,1)	1.3000 (0.0878)	0.9175
ARIMA(1,1,0)	1.4942 (0.0110)	0.8882
ARIMA(2,1,2)	1.3733 (0.0460)	0.8822

Standard errors are given in parentheses. The R^2 statistic is obtained by regressing the quarterly change in GNP against the change in the BN trend.

where, in terms of the ARMA coefficients for Δy_t , $\alpha = \gamma(1) = 1 - \theta_1$. These equations can be cast into innovations state space form with

$$y_t = b + [1 \ 0] \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \end{bmatrix} + \varepsilon_t$$

as the measurement equation and

$$\begin{bmatrix} \tau_t \\ c_t \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha \\ 1 - \alpha \end{bmatrix} \varepsilon_t$$

as the transition equation. Forecasts for these state space equations can be computed by using a suitable version of the information or Kalman filters, and the maximum likelihood estimates of the parameters (α and b) are obtained using the prediction error decomposition of the likelihood function. Note that it is α rather than the MA(1) parameter that is directly estimated.

The estimated α s and implied variance ratios for the USA, the UK and Australian output are shown in Table 20.1. Here it is interesting to note that

Fig. 20.1. Implied transitory components based on different ARIMA models for the USA, the UK and Australia. The *shaded areas* indicate peak-to-trough episodes (recessions) recorded by the NBER for the USA and by the ECRI for the UK and Australia.

while $\alpha > 1$ for the USA and Australia, implying that innovations to the “trend” and “cycle” are negatively correlated, the same is not true for the UK. In the case of the UK, the standard error associated with the parameter α suggests that the data are consistent with the possibility that $\alpha = 1$. Turning to the R^2 measures of the fraction of the variance in the quarterly change in real output that can be attributed to changes in its stochastic trend, it can be seen that the permanent component makes a relatively lower contribution in the USA and Australia than it does in the UK.

The implied transitory components are illustrated in the left hand side graphs in Fig. 20.1, together with reference recessions published by the NBER⁵ and the ECRI.⁶ While there are often pronounced declines in the transitory components around the NBER/ECRI peak-to-trough episodes, there are also clear differences between BN-cycles based on ARIMA(0,1,1) models of output and conventional business cycles. This is hardly surprising, given

⁵ <http://www.nber.org/cycles.html>.

⁶ <http://www.businesscycle.com>.

that the two types of cycle have been constructed to serve different purposes, and have been based on quite different information sets.

20.2.2 ARIMA(1,1,0) Model

For an ARIMA(1,1,0) model, the permanent trend component is the same as above, although in this case $\alpha = 1/(1 - \phi_1)$ in terms of the ARMA coefficients for Δy_t . The cycle component is given by

$$c_t = \phi_1 c_{t-1} + (1 - \alpha) \varepsilon_t.$$

Arranging the model into innovations state space form, the measurement equation is

$$y_t = b + [1 \ \phi_1] \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \end{bmatrix} + \varepsilon_t$$

and the transition equation is

$$\begin{bmatrix} \tau_t \\ c_t \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & \phi_1 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha \\ 1 - \alpha \end{bmatrix} \varepsilon_t.$$

Estimation of the innovations state space model imposes the identity that $\phi_1 = (\alpha - 1)/\alpha$ (which arises from the observation that $\alpha = 1/(1 - \phi_1)$) and provides a direct estimate of α . Results are provided in Table 20.1, and the implied transitory components are illustrated in the center graphs of Fig. 20.1. As for the ARIMA(0,1,1) model, $\alpha > 1$ for the USA and Australia, while $\alpha < 1$ for the UK. Again, because of the standard error of the estimate of α , it is conceivable that $\alpha = 1$, which suggests that UK GDP is largely governed by a random walk with drift. The implied R^2 values for the USA and Australia are much smaller than that for the UK, reflecting a comparatively more noisy transitory component in the former countries.

20.2.3 ARIMA(2,1,2) Model

The ARIMA(2,1,2) model of output has been used by Morley et al. (2003) for US GDP. If the focus is restricted to just ARIMA(0,1,1), ARIMA(1,1,0) and ARIMA(2,1,2) models, it is the model chosen with the AIC for both the USA and the UK. For Australia, however, the ARIMA(1,1,0) model is selected. As usual, the permanent component is given by (20.3), while the transitory component is given by

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} - \psi_1 \varepsilon_{t-1} + (1 - \alpha) \varepsilon_t.$$

In this case $\alpha = (1 - \theta_1 - \theta_2)/(1 - \phi_1 - \phi_2)$ in terms of the ARMA coefficients for Δy_t , although this relationship does not affect the following estimation.

The model can be cast into an innovations state space form with

$$y_t = b + [1 \ \phi_1 \ 1] \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ d_{t-1} \end{bmatrix} + \varepsilon_t$$

being the measurement equation, and

$$\begin{bmatrix} \tau_t \\ c_t \\ d_t \end{bmatrix} = \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & \phi_1 & 1 \\ 0 & \phi_2 & 0 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ d_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha \\ 1 - \alpha \\ -\psi_1 \end{bmatrix} \varepsilon_t$$

being the transition equation.

Table 20.1 reports the estimation results and Fig. 20.1 illustrates the implied transitory components. In the case of the UK, $\hat{\theta}_1$ is statistically insignificant and is set to zero. The reported results are similar to those above. The estimated value of α for the UK is now greater than one, but its standard error suggests the possibility that its value could still be one. Once again, the results suggest that the permanent component in the US and Australian decompositions are relatively less volatile than the corresponding component in the UK decomposition.

20.3 Extensions of the Beveridge–Nelson Decomposition to Nonlinear Processes

The analysis in the previous two sections is based on the assumption that y_t is a linear process, so that all parameters are constant over time and the transitory component is symmetric around zero. This is inconsistent with observed asymmetries in business cycles, and it has led to a small literature on the decomposition of nonlinear time series that has attempted to address this problem.

Clarida and Taylor (2003) have proposed the simulation of the long-run forecast of a variable that follows a nonlinear data generating process. Given an estimated model of the data, the empirical conditional means of simulated long-run forecast densities based on the estimated parameters of this model will deliver the permanent component of the data, provided the data are integrated short-memory in mean. This implies that the long horizon forecast of the first differenced series is a constant (which rules out limit cycles, chaotic behavior and other time variations in the drift). The empirical conditional means of simulated long-horizon forecasts (adjusted for drift) are, of course, analogous to the original definition of the permanent component given by Beveridge and Nelson (1981), in which one had to evaluate the expected conditional mean of the series (adjusted for drift). As before, the transitory component is defined as the difference between the present value of the series and its permanent component.

A few researchers have worked with explicit models, and found direct ways of estimating the permanent-transitory decomposition. For instance, Chen and Tsay (2006) added an exogenous Markov switching (MS) mechanism to the drift in the permanent BN decomposition of GDP, so as to allow for lower growth rates of GDP during recessions. Rather than simulate the implied permanent component, they extended Newbold's (1990) procedure for BN decomposition to allow for regime switches in drift, and found that the transitory component became much less variable as a result.

Low et al. (2006) extended Chen and Tsay's (2006) model by allowing the parameters in both the permanent and transitory components of the decomposition to switch between regimes. As in Chen and Tsay (2006), the regime switches are generated by an exogenously determined MS process, and the regimes are interpreted as periods of recession and expansion. The extended model is given by $y_t = \tau_t + c_t$, with

$$\tau_t = b_{S_t} + \tau_{t-1} + \alpha_{S_t} \varepsilon_t \quad (20.7)$$

and

$$c_t = \phi_{p,S_t}^*(L)c_t - \psi_{r,S_t}^*(L)\varepsilon_t + (1 - \alpha_{S_t})\varepsilon_t, \quad (20.8)$$

in which the subscripted transition variable S_t can take just one of two discrete values at time t , so that the random parameters μ_{S_t} , $\phi_{p,S_t}^*(L)$, $\psi_{r,S_t}^*(L)$, and α_{S_t} all depend on S_t . The transition of S_t between its two regimes is driven by a probability transition matrix P , where

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad (20.9)$$

$p_{ij} = \Pr(S_t = j | S_{t-1} = i)$ and $p_{i1} + p_{i2} = 1$ for all i .

The innovation in y_t is again $\varepsilon_t \sim \text{IID}(0, \sigma^2)$, and this provides the single source of innovation. The variance σ^2 of ε_t is restricted to be constant in this model, although in principle it could depend on S_t without loss of identification. An example, expressed in innovations form, is the Markov switching version of an ARIMA(2,1,2) specification given by

$$y_t = b_{S_t} + \begin{bmatrix} 1 & \phi_{1,S_t} & 1 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ d_{t-1} \end{bmatrix} + \varepsilon_t$$

as the measurement equation, and

$$\begin{bmatrix} \tau_t \\ c_t \\ d_t \end{bmatrix} = \begin{bmatrix} b_{S_t} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & \phi_{1,S_t} & 1 \\ 0 & \phi_{2,S_t} & 0 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ d_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{S_t} \\ 1 - \alpha_{S_t} \\ -\psi_{1,S_t} \end{bmatrix} \varepsilon_t$$

as the transition equation. The parameters in this model can be estimated using a maximum likelihood approach, replacing the standard Kalman filter

used in Kim's (1994) approximation procedure for estimating the MS innovations state space models with a method similar to the Kalman filter described in Sect. 12.7. Low et al. (2006) estimate the above model, using the same US data as that in Sect. 20.2, and they find evidence of switches in both the permanent and transitory regimes. Growth in the expansionary regime is about 3.5%, compared with 2% in the recessionary regime, and the probability of staying in an expansionary regime is about 85%, compared with a probability of about 63% of staying in the recessionary regime. Perhaps one of the most interesting features of this model is that the long-run multiplier changes if there is switching in the transitory components, and the results presented in Low et al. (2006) suggest that this multiplier is 1.35 during expansions, but only 1.14 during recessions.

20.4 Conclusion

An advantage of the innovations state space approach is that it offers a simple and straightforward formulation of the permanent and transitory components of linear economic time series, and it allows direct study of the long-run multiplier α , a key parameter in macroeconomic analysis. Moreover, because it allows α to be estimated directly, it is then possible to obtain auxiliary statistics such as its standard error and t -statistic. Another benefit of the approach, as shown in Sect. 20.3, is that it can easily be adapted to deal with asymmetries in the data generating process. Furthermore, although this has not been studied here, it could be expanded to undertake multivariate decompositions. The interesting aspect of the latter suggestion is that standard real business cycle theories, as in King et al. (1988), assert that output, consumption and investment are all driven by a common unit root process trend, so that economic theory predicts that the permanent and transitory components of these three variables could all be estimated using a single source of error approach.

20.5 Exercises

The following exercises apply to the monthly copper price series in the data set `mcopper`.

Exercise 20.1. Fit a state space form of the Beveridge–Nelson model with an AR(1) transitory component to the logarithm of all but the last 2 years of the series. To simplify matters fix $\hat{c}_0 = 0$.

You should find the following two local minima:

Solution 1: $\hat{\tau}_0 = 5.529$; $\hat{\alpha} = 0$; $\hat{b} = 0.0038$; $\hat{\phi} = 0.9819$; $\hat{\sigma} = 0.0633$;

Solution 2: $\hat{\tau}_0 = 5.524$; $\hat{\alpha} = 1.37$; $\hat{b} = 0.0034$; $\hat{\phi} = -0.01$; $\hat{\sigma} = 0.0595$.

(Try using different starting values in order to find the two minima.) The first local minimum implies that the series is trend stationary; the second that it is difference stationary. The latter, where the series is dominated by a stochastic trend, has the lowest standard error.

Exercise 20.2. Use both fitted models to predict the series (in the original data space; not the log space) from the *fixed* forecast origin 1 January 2005. Show that the MAPEs are 31.17% (solution 1) and 34.46% (solution 2). This confirms that a good fit does not guarantee good forecasts.

Exercise 20.3. Repeat Exercise 20.2 using a rolling origin from 1 January 2005. Show that the MAPEs based on the one-step-ahead prediction errors are 5.24 (solution 1) and 5.06 (solution 2). The second solution adapts slightly better to unexpected changes in the market.

Exercise 20.4. Assume that the time series consists of a stochastic trend and an AR(1) cycle according to the multiplicative model:

$$\begin{aligned}y_t &= \tau_{t-1}(1+b)(1+\phi c_{t-1})(1+\varepsilon_t), \\ \tau_t &= \tau_{t-1}(1+b)(1+\alpha\varepsilon_t), \\ c_t &= \phi c_{t-1} + (1-\alpha)\varepsilon_t.\end{aligned}$$

Fit this model to the original series (withholding the final 2 years of data as before) so as to minimize the augmented sum of squared errors (see 5.4). Compare the results with those from the previous exercises.