ISSN 1440-771X

MONASH University

Australia

Department of Econometrics
and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Beveridge-Nelson Decomposition with Markov Switching

Chin Nam Low, Heather Anderson, Ralph Snyder

August 2006

Working Paper 17/06

Beveridge-Nelson Decomposition with Markov Switching

Chin Nam Low
The University of Melbourne
Victoria 3010
AUSTRALIA

Heather Anderson*
Australian National University
Canberra, ACT, 0200
AUSTRALIA

Ralph Snyder
Monash University
Clayton, Victoria 3800
AUSTRALIA

July 26, 2006

Abstract

This paper considers Beveridge-Nelson decomposition in a context where the permanent and transitory components both follow a Markov switching process. Our approach incorporates Markov switching into a single source of error state-space framework, allowing business cycle asymmetries and regime switches in the long-run multiplier.

Keywords: Beveridge-Nelson decomposition, Markov switching; Single source of error state space models.

JEL classification: C22, C51, E32

*Correspondence: E-mail: heather.anderson@anu.edu.au

1. Introduction

Modeling the behavior of aggregate output has always been an important goal for macroeconomists, who frequently want to study the characteristics of trends and cycles in the economy. Researchers have often used unobserved component (UC) models in this endeavour, specifying a permanent component to represent trend and a transitory component to represent the cycle. These UC models have often been augmented with Markov switching (MS) processes, so as to incorporate asymmetries associated with business cycles or other types of macroeconomic nonlinearities. See Kim and Nelson (1999), Luginbuhl and De Vos (1999), Kim and Murray (2002) and Kim et al. (2005) for examples. This paper considers a new UC class of MS model that is based on a Beveridge-Nelson (BN) decomposition.

UC models are popular because they allow the direct specification of the permanent and transitory components in state-space form, and they can be estimated quite easily, using maximum likelihood and the Kalman filter. The permanent and transitory components are usually assumed to be driven by independent innovations, but recent work has relaxed this assumption, and allowed these innovations to be correlated. The Beveridge-Nelson decomposition is a very special case of UC modelling in which the innovations for permanent and transitory components are perfectly correlated. This property of perfect correlation is supported by the empirical trend and cycle decomposition of US real GDP undertaken by Morley et al (2003), and it is consistent with intuition that shocks to an economy will affect both trend and cycle. BN decomposition has been popular in the applied macroeconomic literature ever since Beveridge and Nelson first suggested it in 1982, but an estimation difficulty associated with approximating an infinite forecasting horizon has sometimes reduced its appeal.

Recent work by Anderson et al (2006) has simplified the computation of the BN components by working with a single source of error (SSOE) state-space approach. Here, we extend BN decomposition in a way that accounts for business cycle asymmetries by introducing a new class of MS model that is built around a SSOE specification. This model (henceforth called an MS-BN model) incorporates an MS process into *both* permanent and transitory components, thus enabling both short run and long run parameters to switch between regimes. The SSOE framework ensures that the embedded permanent and transitory components turn out to be BN components.

MS-BN models have only a few precedents in the literature. Shami and Forbes (2000) use a SSOE state-space approach to estimate a model in which the drift follows a MS process, but they do not interpret their resulting trend and cycle as BN components. More recently, Chen and Tsay (2006) have investigated business cycle asymmetry within a BN decomposition by incorporating a two-state MS process into their permanent component. Like Shami and Forbes (2000), their transitory component is not regime dependent. Further, Chen and Tsay's (2006) estimation technique differs, in that they use the Newbold (1990) procedure in conjunction with the Hamilton (1989) filter.

MS models depend on using hidden Markov chains as latent processes for transiting from one regime to another, and Hamilton's (1989) filter provides a maximum likelihood based algorithm for estimating the probabilities associated with being in each MS regime at each time. Snyder (1985) provides an algorithm that assumes that the innovations of the unobserved state components in a linear setting are perfectly correlated. We estimate our MS-BN models using a maximum likelihood approach, but we replace the standard Kalman filter used in Kim's (1994) approximation procedure for estimating MS state-space models, with Snyder's (1985) perfectly correlated version.

In the next section we introduce a general SSOE state-space model with Markov switching, and discuss some details associated with estimating these models. This section also outlines the special case of a two-state MS-BN $ARIMA(2,1,2)$ specification, that is potentially useful for studying trends and cycles in macroeconomic time series. We report on the application of this model to study quarterly real GNP in the USA in Section 3, and then provide a brief conclusion in Section 4.

2. SSOE state-space models with Markov-switching (MS)

2.1. Model specification

The single source of error state-space model for an observable variable y_t is

$$y_t = \beta' x_{t-1} + e_t \tag{2.1a}$$

with

$$x_t = Fx_{t-1} + \alpha e_t, \tag{2.1b}$$

where (2.1a) and (2.1b) respectively specify measurement and state transition equations. The k vector x_t contains the unobserved components at the beginning of period t, α is a fixed k vector of parameters, e_t is an i.i.d. $N\left(0, \sigma^2\right)$ innovation, β is a fixed k vector, and F is a fixed $k \times k$ transition matrix. Often β and F depend on time invariant parameters. The distinguishing feature of this specification is that both equations are driven by the same innovation, and models with this feature are sometimes called "innovations state space models" (see, eg Hannan and Deistler, (1988)). Snyder (1985) adapts the Kalman filter associated with the maximum likelihood estimation of the parameters in (2.1) to explicitly account for this structure in innovations.

Anderson et al (2006) point out that when Δy_t has an ARMA representation, then the perfect correlation between the errors in (2.1) can be exploited to perform a BN decomposition of the variable y_t into its BN trend τ_t and cycle c_t. This is done by including τ_t and c_t in x_t, and appropriately specifying the matrix F. It turns out that the coefficient of α in the trend equation conveniently measures the long run multiplier (i.e. the Campbell-Mankiw (1987) measure of persistence) in this setting.

The addition of an MS process to a SSOE state-space model leads to measurement and state transition equations given by

$$y_t = \beta'_{S_t} x_{t-1} + e_{t,S_t} \tag{2.2a}$$

and

$$x_t = F_{S_t} x_{t-1} + \alpha_{S_t} e_{t,S_t}, \tag{2.2b}$$

in which S_t is an unobserved MS variable that affects both parameters and innovations. For an M-regime first order Markov process, S_t can take just one of M discrete values at time t, and transition between regimes is governed by

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{pmatrix}, \tag{2.2c}$$

where $p_{ij} = \Pr(S_t = j|S_{t-1} = i)$ and $\sum_{j=1}^{M} p_{ij} = 1$ for all i. See Goldfeld and Quandt (1973) and Hamilton (1989) for more details on Markov switching. The k vector x_t in (2.2) contains the unobserved component variables as before, and the single innovation e_{t,S_t} now follows a distribution specified by $e_{t,S_t} \sim N\left(0, \sigma^2_{S_t}\right)$, in which the variance changes with regime. The parameters in $\alpha_{S_t}, \beta_{S_t}$ and F_{S_t} are random variables that depend on the unobserved MS state variable S_t. Like the standard SSOE specification in (2.1), the MS-SSOE specification can be used to perform a BN decomposition, and this potential use leads to our classification of the model specified by (2.2) as an MS-BN model.

2.2. Estimation

The estimation of (2.2) is similar to the estimation of (2.1) in that both involve the calculation of forecasts $x_{t|t-1}$ of the unobserved components x_t, conditional on information available at time $t-1$. However, the estimation of (2.1) just involves the calculation of $x_{t|t-1} = E(x_t|\widetilde{y}_{t-1})$ with $\widetilde{y}_{t-1} = (y_{t-1}, y_{t-2}, ..., y_1)$, whereas the estimation of (2.2) involves the calculation of M^2 forecasts (one for each combination of i and j) of $x^{(i,j)}_{t|t-1} = E(x_t|\widetilde{y}_{t-1}, S_t = j, S_{t-1} = i)$ for each t, which is considerably more complicated.

Kim (1994) outlines an algorithm that is useful for estimating a Markov switching specification that differs from (2.2) in that his error terms are independent (rather than perfectly correlated). His algorithm involves calculating M^2 forecasts $x^{(i,j)}_{t|t-1}$ at each time t, corresponding to every possible combination of i and j, and then using the Kalman filter to update each $x^{(i,j)}_{t|t-1}$ to obtain $x^{(i,j)}_{t|t}$ when y_t becomes available. Kim's algorithm also updates $P^{(i,j)}_{t|t}$, the mean squared error matrix of x_t conditional on $\widetilde{y}_t$. While Kim's algorithm is not directly applicable given that it assumes independent innovations, we adapt it using Snyder's (1985) filtering algorithm for perfectly correlated innovations to obtain

$$x^{(i,j)}_{t|t-1} = F_j x^i_{t-1|t-1},$$

$$P^{(i,j)}_{t|t-1} = F_j P^i_{t-1|t-1} F_j' + \alpha_j \sigma^2_j \alpha_j',$$

$$e_{t|t-1}^{(i,j)} = y_t - \beta_j' x_{t-1|t-1}^i,$$
$$v_{t|t-1}^{(i,j)} = \beta_j P_{t-1|t-1}^i \beta_j' + \sigma_j^2,$$
$$K_{t|t-1}^{(i,j)} = (F_j P_{t-1|t-1}^i \beta_j + \alpha_j \sigma_j^2)(v_{t|t-1}^{(i,j)})^{-1}_{,}$$
$$x_{t|t}^{(i,j)} = x_{t|t-1}^{(i,j)} + K_{t|t-1}^{(i,j)} e_{t|t-1}^{(i,j)},$$

and

$$P_{t|t}^{(i,j)} = P_{t|t-1}^{(i,j)} - K_{t|t-1}^{(i,j)} v_{t|t-1}^{(i,j)} K_{t|t-1}^{(i,j)\,\prime},$$

where $v_{t|t-1}^{(i,j)}$ is the conditional variance of the forecast error $e_{t|t-1}^{(i,j)}$, and $K_{t|t-1}^{(i,j)}$ is the Kalman gain based on information available up to time $t-1$ with $S_{t-1} = i$ and $S_t = j$.

We follow Kim (1994), and simplify the implementation of this algorithm by collapsing the M^2 terms for each of $x_{t|t}^{(i,j)}$ and $P_{t|t}^{(i,j)}$ into M terms for each specified by

$$x_{t|t}^j = \frac{\sum_{i=1}^M \Pr(S_t = j, S_{t-1} = i|\widetilde{y}_t) x_{t|t}^{(i,j)}}{\Pr(S_t = j|\widetilde{y}_t)} \tag{2.3a}$$

and

$$P_{t|t}^j = \frac{\sum_{i=1}^M \Pr(S_t = j, S_{t-1} = i|\widetilde{y}_t)(P_{t|t}^{(i,j)} + (x_{t|t}^j - x_{t|t}^{(i,j)})(x_{t|t}^j - x_{t|t}^{(i,j)}))}{\Pr(S_t = j|\widetilde{y}_t)}, \tag{2.3b}$$

inferring the conditional probabilities in (2.3a) and (2.3b) from a modified version of the Hamilton (1989) filter. As discussed in Kim (1994), the equations in (2.3) are approximations for $E(x_t \mid (\widetilde{y}_t,\ S_t = j))$ and $E[(x_t - x_{t|t}^{;j}) \cdot (x_t - x_{t|t}^{;j})' \mid (\widetilde{y}_t, S_t = j)]$, because $x_{t|t}^{(i,j)}$ and $P_{t|t}^{(i,j)}$ derived from the Kalman filter are only approximations for $E(x_t \mid \widetilde{y}_t, S_t = j,\ S_{t-1} = i)$ and $E[(x_t - x_{t|t}^{;j}) \cdot (x_t - x_{t|t}^{;j})'|(\widetilde{y}_t,\ S_t = j,\ S_{t-1} = i)]$. Nevertheless, these approximations work well in practice, and have little influence on the final estimates.

Estimation of the parameters in (2.1) maximises the log likelihood function given by

$$LL = \sum_{t=1}^T \ln(f(y_t|\widetilde{y}_{t-1})) = \sum_{t=1}^T \ln(\sum_{S_t} \sum_{S_{t-1}} f(y_t|S_t, S_{t-1}, \widetilde{y}_{t-1}) \Pr(S_t, S_{t-1}|\widetilde{y}_{t-1})),$$

in which the conditional density $f(y_t|S_t, S_{t-1}, \widetilde{y}_{t-1})$ given by

$$f(y_t|S_t, S_{t-1}, \widetilde{y}_{t-1}) = (2\pi)^{-\frac{N}{2}} \left| v_{t|t-1}^{(i,j)} \right|^{-\frac{1}{2}} \exp(-\frac{1}{2}(y_t - \beta_j' x_{t-1|t-1}^{(i)})' v_{t|t-1}^{(i,j)} (y_t - \beta_j' x_{t-1|t-1}^{(i)}))$$

is evaluated using the above described filter (along with Kim's (1994) approximations). The conditional joint probabilities $\Pr(S_t, S_{t-1}|\widetilde{y}_{t-1})$ are obtained from recursion of the Hamilton filter.

It is not necessary to smooth the unobserved components $x_{t|t}^{(j)}$ when innovations are perfectly correlated, as $x_{t|t}^{(j)}$ converges quickly to $x_{t|T}^{(j)}$, (see Harvey (1989) and Harvey and Koopman (2000)). However, there is still a need to compute the smoothed $\Pr(S_t = j|\widetilde{y}_T)$ to obtain the weighted average unobserved components $x_{t|T}$ at time t. Kim (1994) provides an appropriate smoothing algorithm, with the resulting components being

$$x_{t|T} = \sum_{j=1}^{M} \Pr(S_t = j|\widetilde{y}_T) x_{t|T}^{(j)} .$$

We show below that when Δy_t has an MS-ARMA representation and we define the permanent and transitory components of y_t to be τ_t and c_t respectively, then (2.2a) to (2.2c) can lead to the BN decomposition of y_t. This decomposition simply involves the inclusion of τ_t, and c_t in the component vector x_t, and an appropriate specification of β_{S_t}, F_{S_t} and α_{S_t}.

2.3. SSOE models and the BN decomposition

Anderson et al. (2006) show that if y_t is a I(1) variable with a Wold representation given by $\Delta y_t = \mu + \gamma(L)\varepsilon_t$, where μ is the drift, $\gamma(L) = \frac{\theta(L)}{\phi(L)}$ is an ARMA(p,q) process with $\gamma(0) = 1$ and $\Sigma_{i=0}^{\infty} |\gamma_i| < \infty$, and ε_t is an iid $\left(0, \sigma^2\right)$ innovation, then the BN permanent and transitory components are respectively given by

$$\tau_t = \mu + \tau_{t-1} + \gamma(1)\varepsilon_t \tag{2.4a}$$

and

$$c_t = \phi_p^*(L)c_t + \theta_n^*(L)\varepsilon_t + (1-\gamma(1))\varepsilon_t, \tag{2.4b}$$

where $\phi_p^*(0) = \theta_n^*(0) = 0$, and the orders of $\phi_p^*(L)$ and $\theta_n^*(L)$ are p and n with $n \leq \max(p-1,\ q-1)$. The perfectly correlated innovations in (2.4) fit in with the SSOE framework.

We incorporate an MS process in the permanent and transitory components by specifying

$$\tau_t = \mu_{S_t} + \tau_{t-1} + \alpha_{S_t}\varepsilon_t \tag{2.5a}$$

and

$$c_t = \phi^*_{p,S_t}(L)c_t + \theta^*_{n,S_t}(L)\varepsilon_t + (1-\alpha_{S_t})\varepsilon_t, \tag{2.5b}$$

so that the random parameters $\mu_{S_t}, \phi^*_{p,S_t}(L), \theta^*_{n,S_t}(L)$, and α_{S_t} all depend on S_t. As above, the innovation to y_t is $\varepsilon_t \sim iid(0, \sigma^2)$, and this provides the single source of disturbance. We have restricted σ^2 to be constant in this specification, although in principle σ^2 could depend on S_t without loss of identification. As in (2.4), the perfectly correlated innovations in (2.5) allow us to write the model in SSOE form.

To illustrate the SSOE state space form of an MS-BN model with business cycle asymmetries we note that the incorporation of an MS process into the $ARIMA(2,1,2)$ SSOE model leads to a specification with

$$y_t = \mu_{S_t} + \begin{bmatrix} 1 & -\phi_{1,S_t} & -\phi_{2,S_t} & \theta_{1,S_t} \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ c_{t-2} \\ \varepsilon_{t-1} \end{bmatrix} + \varepsilon_t \tag{2.6a}$$

as the measurement equation, and

$$\begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \\ \varepsilon_t \end{bmatrix} = \begin{bmatrix} \mu_{S_t} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\phi_{1,S_t} & -\phi_{2,S_t} & \theta_{1,S_t} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ c_{t-2} \\ \varepsilon_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{S_t} \\ 1-\alpha_{S_t} \\ 0 \\ 1 \end{bmatrix} \varepsilon_t \tag{2.6b}$$

as the transition equation. The parameters $\mu_{S_t}, \phi_{1,S_t}, \phi_{2,S_t}, \theta_{1,S_t}$ and α_{S_t} are time invariant parameters that depend on the latent MS variable S_t, and one can use the two-dimensional version of (2.2c) and allow this variable to take on two possible values (i.e. $S_t = 1$ or $S_t = 2$), where the two states represent "contractionary" and "expansionary" regimes in the business cycle. Note that the $MA(2)$ parameters in the underlying $ARMA(2,2)$ specification for Δy_t drop out during reparameterisation into SSOE form, being replaced by the α_{S_t} parameters.

3. Modelling US GNP

This section provides an empirical example of an MS-BN model of the logarithms of real GNP, detailing the characteristics of this model and its implied BN components, and comparing these characteristics with the corresponding linear BN model. An important motivation for this exercise is to determine whether the incorporation of Markov-switching leads to an improved ability to capture asymmetries in business cycles, although we also look at out-of-sample forecast and other aspects of model performance. We focus on the MS-BN ARIMA(2,1,2) model shown in equation (2.6), because researchers often study permanent/transitory decompositions of the linear version of this model.

Our study is based on quarterly seasonally adjusted data that measures (the natural logarithm of) real GNP for the USA from 1947:1 to 2003:1. We use the data for 1947:1 to 2000:1 for estimation, and with-hold the remaining twelve observations for out-of-sample forecast analysis. We estimate the linear BN model first, and retain the estimated coefficients as starting values for corresponding parameter estimates when estimating the MS-BN model. Our estimation of the MS-BN model follows the procedure outlined in Section 2, with the imposition of the condition that $\mu_{S_t=2} = \mu_{S_t=1} + \mu_2$ with $\mu_2 \geqq 0$ so as to identify $S_t = 2$ as the expansionary regime. In light of the well known fact that the likelihood functions of MS models are plagued with numerous local maxima, we experiment with perturbing our starting values and then take parameter estimates corresponding to the highest converged likelihood as our maximum likelihood estimates. Our experiments use starting values of around 0.8 for p_{11} and 0.9 for p_{22}, since these values are close to corresponding estimates in other empirical studies.

3.1. The empirical model

Table 1 presents the maximum likelihood parameter estimates. Since μ_1 is greater than zero, it is appropriate to call $S_t = 1$ a "slow growth" regime rather than a "recessionary" regime. The long-run multipliers measured by α_1 and α_2 are greater than unity, implying that both regimes have strong persistence as measured by Campbell and Mankiw (1987). This persistence measure predicts the long run increase in output resulting from a 1% shock in output in one quarter if the economy was to remain in that state for ever, and our estimates indicate that persistence for the "fast growth" regime is stronger than that for the "slow growth" regime. The persistence measure for the linear model falls between those for the slow and fast regimes. The tendency for the economy to stay in a fast growth regime (p_{22}) is about the same as that found in other empirical studies (i.e. 85%), while the tendency to remain in a slow growth regime is considerably smaller.

The reported R^2 statistics (suggested by Stock and Watson (1988)) measure the proportion of variance in output that can be attributed to variance in the permanent component, and this ratio declines by about 15 percentage points, once the model accounts for Markov-switching. This suggests that the MS process plays an important role in output variation, affecting the transitory component more than the permanent component. However, the latter still plays the dominant role when it comes to explaining changes in output.

The top portions of Figure 1 illustrate the smoothed permanent and transitory components. The transitory component fluctuates considerably, especially when entering and exiting the "slow growth" regime, but the dominant features are two structural changes in variance, with the first occurring in about 1960, and the second occurring in about 1984. This second volatility decline is well documented (see e.g. McConnell and Perez–Quiros (2000)).

The lower portions of Figure 1 presents the smoothed and filtered probabilities of being in the "slow growth" regime, together with peak to trough episodes defined by the NBER. The probabilities of being in the "slow growth" regime for the US peak during all the recession periods dated by NBER. Although the results are less convincing for the recessions in the seventies, they are nevertheless higher than the unconditional probability of 0.28. The probability of being in the "slow growth" regime is only around 0.5 during the 1990-91 recession. This is higher than the unconditional probability of being in the "slow growth" regime, but this recession was not a typical recession, as influence from the political uncertainty caused by the first Gulf War played a role here.

3.2. Model diagnostics

The standard measures of fit reported at the bottom of Table 1 suggest that the MS-BN model fits the data much better than the BN models (see Table 2), but this is hardly surprising, given the inherent flexibility of the MS-BN specification. The question of whether the MS-BN model can "fit" in the sense of capturing features that are actually observed in the data is more important, and we use the parametric encompassing tests suggested by Breunig et al. (2003) to explore this issue. These tests are designed to assess whether an estimated model can capture the mean, variance, and various measures of asymmetry in the data, and they can also provide indirect information on whether the maximum likelihood estimates reflect the true global maximum.

Letting $\widehat{\theta}$ be the maximum likelihood estimates for the model, the parametric encompassing tests compare a sample moment $\widehat{\gamma}$ for the raw data (eg a sample mean), with the corresponding moment $\gamma(\widehat{\theta})$ for data that has been generated from the estimated model. The test statistic is given by

$$R = (\widehat{\gamma} - \gamma(\widehat{\theta}))'[var(\widehat{\gamma}) - var(\gamma(\widehat{\theta}))]^{-1}(\widehat{\gamma} - \gamma(\widehat{\theta})).var(\gamma(\widehat{\theta})),$$

and it has a $\chi^2_{\dim(\gamma)}$ distribution under the null hypothesis that the model is consistent with the data. Since it is usually difficult to calculate $var(\gamma(\widehat{\theta})$, Breunig et al. (2003) suggest using $var(\widehat{\gamma})$ to approximate $[var(\widehat{\gamma}) - var(\gamma(\widehat{\theta}))]$, thereby making the test more conservative. When testing Markov-switching models, Breunig et al (2003) suggest complementing encompassing tests based on the mean and variance with tests based on

$$q_1 = E[I(\Delta y_{t-2} < 0, \Delta y_t > 0)]$$

and

$$q_2 = E[I(\Delta y_{t-2} > 0, \Delta y_t > 0)],$$

where $I(A)$ is the indicator function, taking the value 1 if event A is true and zero otherwise. These last two moments reflect asymmetries documented in Potter's (1995) study of US real GNP, and encompassing tests based on the corresponding sample moments can indicate whether the model has captured these asymmetries.

We assess our linear and MS-BN models by applying parametric encompassing tests for the mean, variance, q_1 and q_2. Our $\gamma(\widehat{\theta})$ statistics are based on 10,000 replicated samples of the same size as the original data, with starting values fixed at the first observed data point. As in Breunig et al (2003), we obtain robust estimators of $var(\widehat{\gamma})$ by running regressions of the sample γ_t on a constant, using a Newey-West correction that employs 9 lags. The test results are presented in Table 3. These statistics show that although both models can capture the asymmetric characteristics of the data very well, the BN model is unable to capture the variance. The MS-BN model has no trouble in this regard, suggesting that the use of Markov switching improves the modelling of the variance of US GNP. We note, however, that the MS-BN model has a little difficulty in capturing the mean, although this problem is not statistically significant at the 5% level of significance.

3.3. Forecasting performance

We conclude our model analysis with a small out-of-sample forecasting exercise. All forecasts are based on the models estimates derived from the initial samples (i.e. we don't undertake any further estimation), and the forecasts begin with the first observation in the out-of-sample data. We generate a sequence of 1 - 8 step ahead forecasts, roll the forecast origin forward, generate another sequence of 1-8-step ahead forecasts, and repeat this procedure until we have 12 x 1-step ahead forecasts down to 5 x 8-step ahead forecasts for the twelve out-of-sample observations. The forecasts are generated using the standard forecast simulation method with 10,000 replications for each "rolling" forecast. Multi-step ahead forecasts for the MS-BN models are based on

$$E(S_{T+h} = 1|y_T) = S_1 + \lambda^h(\Pr(S_T = 1|y_T) - S_1)$$

where $S_1 = \frac{(1-p_{22})}{(2-p_{11}-p_{22})}$ is the unconditional probability of $S_t = 1$, $\lambda = p_{11} + p_{22} - 1$ and $(\Pr(S_T = 1|y_T)$ is the last filtered probability of $S_T = 1$ conditional on the last in-sample observation y_T. The results of the forecasting exercise are illustrated in Figure 2. The MS-BN model outperforms the BN model for all forecast horizons, although the difference is not statistically significant.

4. Conclusion

This paper has shown that an SSOE specification can provide a useful framework for undertaking BN decompositions when both permanent and transitory components follow a Markov-switching process. The SSOE specification ensures that the permanent and transitory components in the model are BN components, and one can easily adapt the techniques that are typically used to estimate UC and MS models to account for the single source of error. An application to US real GDP shows that an ARIMA(2,1,2) MS-BN model is well specified, and leads to components that reflect recognized "stylized facts".

It is interesting to observe that even though the perfect correlation between BN permanent and transitory components is normally considered to be just a by-product of BN decomposition, this can be exploited to identify the BN components. The reason for this is that perfect correlation between innovations to the components implies perfect correlation between innovations to trend and *output*, and as noted by Morley et al (2003), the BN trend is always the conditional expectation of the random walk component for any I(1) process. Since the SSOE model explicitly implies perfect correlation between innovations to trend and output, it leads directly to the BN trend.

The SSOE approach is quite easy to work with, and one could easily introduce more sophisticated MS processes into an SSOE model, and then undertake a BN decomposition. Such exercises could be the focus of future research.

References

Anderson, H.M., C.N. Low, and R.D. Snyder (2006), "Single Source of Error State-space Approach to the Beveridge-Nelson Decomposition", *Economics Letters*, 91, 104-109.

Beveridge, S., and C.R. Nelson (1981), "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle", *Journal of Monetary Economics*, 7, 151-174.

Breunig, R.V., S. Najarian and A.R. Pagan (2003), "Specification Testing of Markov-Switching Models", *Oxford Bulletin of Economics and Statistics*, 65, 703-725.

Campbell, J. and N.G. Mankiw (1987), "Permanent and Transitory Components in Macroeconomic Fluctuations", *The American Economic Review*, 77, 111-117.

Chen, C.C. and W.J. Tsay (2006), "The Beveridge-Nelson Decomposition of Markov-Switching Processes", *Economics Letters*, 91, 83-89.

Goldfeld, S.M. and R.E. Quandt (1973), "A Markov Model for Switching Regressions", *Journal of Econometrics*, 6, 3-16.

Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Non Stationary Time Series and the Business Cycle", *Econometrica*, 57, N2, 357-384.

Hannan, E.J. and M. Deistler (1988). The Statistical Theory of Linear Systems, Wiley.

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

Harvey, A.C. and S.J. Koopman (2000). "Signal Extraction and the Formulation of Unobserved Components models", *The Econometrics Journal*, 3, 84-107.

Kim, C.J. (1994), "Dynamic Linear Models with Markov-Switching", *Journal of Econometrics*, 60, 1-22.

Kim, C.J. and C.J. Murray (2002), "Permanent and Transitory Components of Recessions", *Empirical Economics*, 27, 163-183.

Kim, C.J. and C.R. Nelson (1999), **State-Space Models with Regime Switching,** MIT press.

Kim, C.J. and C.R. Nelson (1989), "The Time-Varying-Parameter Model as an Alternative to ARCH for Modelling Changing Conditional Variance: The Case of Lucas Hypothesis", *Journal of Business and Economic Statistics*, 7, 433-440.

Kim, C.J., J.M. Piger and R. Startz (2005), "The Dynamic Relationship Between Permanent and Transitory Components of U.S. Business Cycles", forthcoming in the *Journal of Money Credit and Banking.*

Luginbuhl, R. and A. De Vos (1999), "Bayesian Analysis of an Unobserved Component Time Series Model of GDP with Markov-Switching and Time Varying Growths", *Journal of Business and Economic Statistics*, 17, N4, 456-465.

McConnell, M.M. and G. Perez-Quiros (2000), "Output Fluctuations in the United States: What has Changed since the Early 1980s?", *American Economic Review*, 90, 1464-1476.

Morely, J.C., C.R. Nelson and E. Zivot (2003), "Why are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP so Different?", *The Review of Economics and Statistics*, LXXXV, 235 - 243.

Newbold, P. (1990), "Precise and Efficient Computation of the Beveridge-Nelson Decomposition of Economic Time Series", *Journal of Monetary Economics*, 26, 453 - 457.

Potter, S.M. (1995), "A nonlinear approach to US GNP", *Journal of Applied Econometrics*, 10, N2, 109-125.

Shami, R.G. and C.S. Forbes (2000), " A Structural Time Series Model with Markov Switching", Working Paper 10/2000, Department of Econometrics and Business Statistics, Monash University, Australia.

Sichel, D.E. (1994), "Inventories and the Three Phases of the Business Cycle", *Journal of Business and Economic Statistics, 12*, N3, 269-278.

Snyder, R.D. (1985), "Recursive Estimation of Dynamic Linear Models", *Journal of the Royal Statistical Society, Series B*, 47, 272-276.

Stock, J.H., and M.W. Watson (1988), "Variable Trends in Economic Time Series", *Journal of Economic Perspectives*, 2(3), 147-174.

Table 1: Estimates of MS-BN model

MS-BN Model		
Parameter	Coeff	Std Error
α_1	1.1446	0.0000
ϕ_{11}	1.2778	0.0001
ϕ_{21}	-0.9912	0.0001
μ_1	0.4994	0.0002
θ_{11}	0.4226	0.0001
α_2	1.3476	0.0001
ϕ_{12}	1.4352	0.0002
ϕ_{22}	-0.8183	0.0002
μ_2	0.9655	0.0000
θ_{12}	0.2526	0.0001
p_{11}	0.6268	0.0000
p_{22}	0.8524	0.0001
R^{2*}	0.7127	
SSE	153.30	
AIC	-0.2162	

* The R^2 statistic is obtained by regressing the quarterly change in GDP against the change in the BN trend component.

Table 2: Estimates of BN model

BN Model		
Parameter	Coeff	Std Error
α	1.2379	0.1419
ϕ_1	1.3724	0.1334
ϕ_2	-0.7760	0.1644
β	0.8520	0.0834
θ	0.3477	0.1154
R^{2*}	0.8493	
SSE	188.90	
AIC	-0.0731	

* The R^2 statistic is obtained by regressing the quarterly change in GDP against the change in the BN trend component.

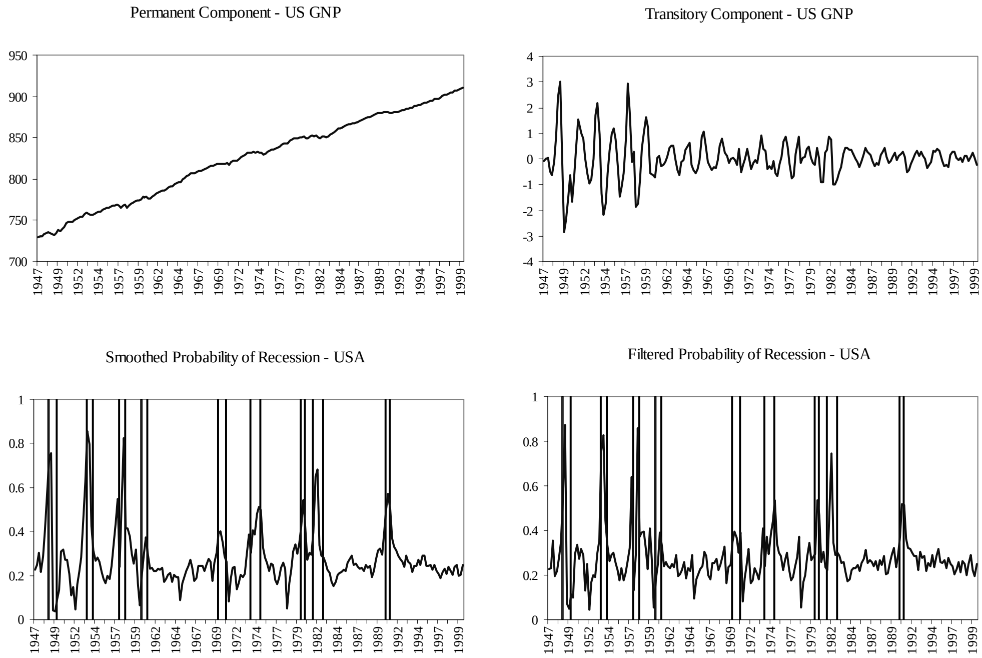
Table 3: Parametric Encompassing Test Results for MS–BN and BN Models

MS-BN Model				
	Model	Data	R-stat	p-value*
Mean	824.18	819.07	3.3222	0.0684
Variance	2665.01	2668.41	0.0316	0.8588
q_1	0.1048	0.1464	0.0761	0.7827
q_2	0.7381	0.6523	0.1539	0.6949

BN Model				
	Model	Data	R-stat	p-value*
Mean	821.18	819.07	0.5685	0.4509
Variance	2791.72	2668.41	41.5654	0.0000
q_1	0.1463	0.1464	0.0000	0.9996
q_2	0.6511	0.6523	0.0000	0.9957

$*$The test statistic is distributed as a χ^2_1

Figure 1: Permanent and Transitory Components, Filtered and Smoothed Probability of Recession



Note: The pair of lines on the graphs indicate peak to trough episodes (recessions) recorded by NBER.

Figure 2: Forecast Performance of MS-BN and BN Models