Why Are Beveridge-Nelson and Unobserved-Component Decompositions of GDP So Different?

By James C. Morley[*], Charles R. Nelson[**], and Eric Zivot[**]

[*]Department of Economics, Washington University in St. Louis
[**]Department of Economics, University of Washington

First draft: July 23, 1999
This draft February 23, 2001

PRELIMINARY, Not to be quoted without authors' permission.

Copyright © 1999, 2000, 2001 by James C. Morley, Charles R. Nelson, and Eric Zivot

The authors acknowledge with thanks support from the National Science Foundation under grants SES-9818789 and SBR-9711301 and from the Van Voorhis and Ensley endowments at the University of Washington. We received helpful comments from many individuals and particularly Charles Engel, Andrew Harvey, Chang-Jin Kim, Chris Murray, Jeremy Piger, Simon Potter, Glenn Rudebusch, Richard Startz and Mark Wohar, but responsibility for errors is entirely the authors'. Kyongwook Choi provided valuable research assistance.

Address correspondence to: Charles R. Nelson, Box 353330, Univ. of Washington, Seattle, WA 98195, USA; cnelson@u.washington.edu.

Abstract

This paper reconciles two widely-used decompositions into trend and cycle that are both model-based but yield starkly contrasting results. Beveridge-Nelson implies that stochastic trend accounts for most of the variation in output, while Unobserved Components implies that cyclical variation is dominant instead. Which characterization is correct has broad implications for identifying the source of aggregate fluctuations, and the conflicting results remain a puzzle in the literature. We show that these differ because of the restriction usually imposed in the UC framework that trend and cycle innovations are uncorrelated. When this restriction is relaxed the two approaches yield identical decompositions, and the zero restriction is rejected for U.S. quarterly GDP.
JEL classification: C22, C5, E32.

The decomposition of real GDP into trend and cycle remains a problem of considerable practical importance, but two widely used methods yield starkly different results. The unobserved component approach, introduced by Harvey (1985) and Clark (1987), implies a very smooth trend with a cycle that is large in amplitude and highly persistent. In contrast, the approach of Beveridge and Nelson (1981) implies that much of the variation in the series is attributable to variation in the trend while the cycle component is small and noisy. This conflict is apparent in Figures 1 and 2 in this paper where the two cycle components are plotted, and has been widely noted; see Watson (1986), Stock and Watson (1988) among others.

It should surprise us that the unobserved component (UC) and Beveridge-Nelson (BN) methods produce very different trend-cycle decompositions since both are model-based. Each implies an ARIMA representation. Neither imposes smoothness in the trend component *a priori* as does the smoother of Hodrick and Prescott (1997) or as in the polar case of a linear trend that forces all variation, save constant growth, into the cycle. The UC and BN both "let the data speak for itself" in this regard. While it is often stated that BN assumes a perfect negative correlation between trend and cycle innovations, that is a property of the *estimated* trend and cycle, not the unobserved components, and it is a property shared with the UC decomposition. This paper attempts to find out why we do not, after decades of research, have a consistent picture of how variation in a series like real GDP should be allocated between trend and cycle.

Briefly, section 1 demonstrates the theoretical equivalence between the approaches. Section 2 investigates the source of the difference observed in practice. Section 3 concludes.

1. Theoretical Equivalence of the Beveridge-Nelson and Unobserved Component Estimates of Trend and Cycle

The detrending problem is motivated by the idea that the log of aggregate output is usefully thought of as the sum of a nonstationary trend component that accounts for long-term growth and a stationary component that allows for transitory deviations of output from trend. We follow custom in referring to the latter as the “cycle” even if it is not periodic. The UC model takes the form:

(1a) $$y_t = \tau_t + c_t$$

(1b) $$\tau_t = \tau_{t-1} + \mu + \eta_t; \qquad \eta \sim i.i.d.\ N(0, \sigma_\eta^2)$$

(1c) c_t is stationary and ergodic

where $\{y_t\}$ is the observed series, $\{\tau_t\}$ is the unobserved trend assumed to be a random walk with average growth rate μ, and $\{c_t\}$ is the unobserved stationary cycle component.[1] The UC-ARMA adds the condition that $\{c_t\}$ is a stationary and invertible ARMA(p,q) process with innovations that may be contemporaneously cross-correlated with trend innovations,

(1d) $$\phi_p(L)\, c_t = \theta_q(L)\, \varepsilon_t;\ \varepsilon \sim i.i.d.\ N(0, \sigma_\varepsilon^2);\ Cov(\varepsilon_t, \eta_{t+k}) = \sigma_{\varepsilon\eta} \text{ for } k=0;\ 0 \text{ otherwise.}$$

In some implementations the rate of drift μ is allowed to evolve as a random walk, although this has little influence on the estimated cycle component for U.S. GDP, and an additional irregular term is sometimes added. Harvey (1985) and Clark (1987), and Harvey and Jaeger (1993) suggest specifying p=2 which allows the cycle process to be periodic in the sense of having a peak in its spectral density function. They and others further assume that the trend and cycle innovations are uncorrelated, setting

[1] As noted by Blanchard and Quah (1989), the structural trend of output does not necessarily follow a random walk. Therefore, a cycle component that represents the deviation from a random walk trend may reflect both transitory effects of supply shocks and cyclical movements due to demand shocks.

(1e) $s_{eh} = 0$,

thereby casting the UC model in state-space form by treating (1a) as the measurement equation and (1b) as the state transition equation, see the technical appendix for details. We denote this constrained zero-covariance UC-ARMA model as UC-0.

In practice, the parameters are unknown and are estimated from the data series $(y_1, .., y_n)$ by putting the model in state-space form and using the maximum likelihood method of Harvey (1981) based on the prediction error decomposition generated by the Kalman Filter. Given the parameters, the Kalman filter is used to compute the expectation of the stochastic trend component conditional on data through time t:

$$\hat{t}_{t|t} = E\left[t_t \middle| Y_t\right], \text{ where } Y_t = (y_1,..,y_t)\,.$$

Alternatively, the BN estimate of stochastic trend for an I(1) time series $\{y_t\}$ is defined to be the limiting forecast as horizon goes to infinity, adjusted for the mean rate of growth; so

$$BN_t = \lim_{M \to \infty} E\left[y_{t+M} - Mm \middle| Y_t\right].$$

BN showed that the time series $\{BN_t\}$ will be a random walk with drift, the deviation from trend is a stationary process, and that the innovations of $\{BN_t\}$ and $\{y_t - BN_t\}$ are perfectly correlated. The series $\{BN_t\}$ is typically calculated from an estimated ARIMA representation of $\{y_t\}$, which in principle is unique after cancellation of any redundant AR and MA factors[2].

It is well known that the UC–ARMA model always implies a univariate ARIMA representation for $\{y_t\}$. This is what Nerlove, Grether, and Carvalho (1979) refer to as the

[2] The theoretical justification for the BN decomposition and its relationship to Martingale decompositions is given in Phillips and Solo (1992). A corresponding decomposition for seasonal time series is given in Box, Pierce, and Newbold (1987).

canonical form of the UC model, and it may be useful to think of it as the reduced form. Substituting (1b) and (1d) into (1a), taking first differences, and rearranging we obtain

$$\text{(2a)} \qquad f_p(L)\,(1\text{-}L)y_t = f_p(1)m + f_p(L)\,h_t + q_q(L)\,(1\text{-}L)e_t\,.$$

Recognizing that the right hand side will have non-zero autocorrelations through lag max(p, q+1), Granger's Lemma (see Granger and Newbold (1986) pg. 29) implies that the univariate representation will be

$$\text{(2b)} \qquad f_p(\mathrm{L})\,(1\text{-}\mathrm{L})y_t = m^* + q^*_{q^*}(L)\;u_t;\; u \sim i.i.d.\; N(0,\; s^2_u);\; q^* = max(p,\, q+1)$$

where the coefficients of $q^*_{q^*}(L)$ and s^2_u are obtained by matching the autocovariances of the right hand side of (2a) and (2b); see Watson (1986). This ARIMA reduced form fully describes the joint distribution of the $\{y_t\}$ and therefore the conditional distribution of future observations given the past and is unique. Note that the BN trend for the reduced-form ARIMA model (2b) may be derived from the Wold representation of (2b) and expressed as

$$\text{(2c)} \qquad BN_t = BN_{t-1} + y(1)u_t = y(1)\sum_{j=1}^{t} u_j\,,$$

where $y(1) = q^*_{q^*}(1)/f_p(1)$ and $BN_0 = 0$. From (2a) it can be seen that the variance of the innovation to the BN trend is $y(1)^2 s^2_u$. The BN cycle is obtained by subtracting from y_t the BN trend and other deterministic components.

Correspondingly, there is always at least one UC representation of any given ARIMA process; as Cochrane (1988) pointed out, the existence of the BN decomposition guarantees this. In general, however, there will not be a *unique* UC representation since the parameters may not all be identified. For example, consider the ARIMA(0,1,1) process so that in the notation of (2b) the orders are p=0 and q^*=1. By inspection of (2a) it is clear that this implies q=0, hence the ARIMA(0,1,1) process has a UC representation as a random walk plus noise, and the first differences are autocorrelated at lag one only. However, the parameters of the UC representation of the ARIMA(0,1,1) process are not all separately identified. This is easy to see from the relation between the two non-zero autocovariances γ_j at lags 0 and 1 (values of which can be inferred from data), and the UC parameters as follows:

$$\gamma_0 = \sigma_\eta^2 + 2\sigma_\varepsilon^2 + 2\sigma_{\varepsilon\eta}$$
$$\gamma_1 = -\sigma_\varepsilon^2 - \sigma_{\varepsilon\eta}$$
$$\gamma_j = 0, j \geq 2$$

While there are three parameters in this case, there are only two pieces of information. Note, however that the variance of the trend innovations may be inferred by adding $2\gamma_1$ to γ_0, while the variance of the cycle innovations and the covariance between them are not separately identified, only their sum is. This reflects a basic theme of this paper: the trend is identified from the univariate properties of the series though the cycle may not be. It is also possible to infer some inequality restrictions; see the discussion in Nelson and Plosser (1981). It is easily shown that there will be at least as many non-zero moment relations as parameters if $p \geq q + 2$, a fact which we use later to identify the covariance in the Harvey-Clark UC model. However, the fact that the covariance is not identified in a case as familiar as the random walk plus noise may well have contributed to the impression evident in the literature that zero covariance is always an identifying restriction for the UC model.

Given that a time series will not in general have a unique UC representation, it seems surprising to us that *the BN trend is the conditional expectation of the random walk component of an I(1) process*. As pointed out by Watson (1986), this is true regardless of the covariance structure of the unobserved components. To see this, consider the unconstrained UC model defined by (1a)-(1c), so cycle and trend innovations may be cross-correlated. The conditional expectation of the trend component at time *t* is

$$E[\tau_t | Y_t] = E[\tau_t + c_{t+M} | Y_t]$$

for large enough M, since the cycle, by its ergodicity, has expectation zero far enough in the future. Further, the expected value of any future innovation in the trend is zero, so we have

$$E[\tau_t | Y_t] = E\left[\tau_t + \sum_{j=1}^{M} \eta_{t+j} + c_{t+M} | Y_t\right].$$

Recognizing that the terms of the right include all the elements of y_{t+M} except the accumulated drift, we have

$$E[\tau_t | Y_t] = E\left[\tau_t + \sum_{j=1}^{M} \eta_{t+j} + c_{t+M} | Y_t\right] = E[y_{t+M} - M\mu | Y_t] = BN_t .$$

Then the conditional expectation of the cycle at time t is simply

$$E[c_t | Y_t] = y_t - E[\tau_t | Y_t] = y_t - BN_t .$$

Thus, we can always compute conditional expectation estimates of trend and cycle at any point in time from the ARIMA representation of the observed series. The two assumptions, (a) the trend is a random walk, and (b) the cycle is ergodic, are sufficient to identify the components, and this does not depend on knowing the covariance between trend and cycle innovations. Intuitively, the forecast at a long enough horizon reflects only the permanence of the random walk trend. Stronger assumptions may be needed to identify the parameters of a UC representation, but they are irrelevant if the only objective is to estimate trend and cycle. In that case, only the conditional expectation of the future given the past is required, and the ARIMA reduced form provides the relevant conditional distribution.

It follows that if a particular time series does have a representation as a UC-0 process, then the Kalman filter and BN estimates of trend and cycle will be the same, as long as the parameters of the ARIMA model are those implied by the UC-0 representation. In that case, BN is just another way to compute $\hat{t}_{t|t}$ and $\hat{c}_{t|t}$; the time series of those estimates will be the same and they will have the same properties. In particular, their innovations will both be functions of the innovation in y_t since it is only new information that will cause the long range forecast, the trend, to change. Thus, UC-0 and BN share the often-noted property of the BN decomposition, that the innovations of the *estimated* trend and cycle series are perfectly correlated, as discussed in the paper. Further, none of these results depend on limiting UC representations that might be modeled to the constrained UC-0 case; the corresponding ARIMA representation will always be an equivalent way of obtaining the information relevant to estimating the trend.

To sum up this section, we have shown that whether one uses the UC approach to trend-cycle decomposition based on a state-space representation and the Kalman filter, or the BN approach based on long-range forecasts from a univariate ARIMA model, the specific results should be the same. UC and BN are simply alternative ways of calculating the same conditional expectation of the unobserved trend and cycle at a point in time. The fact that the two have produced such different estimates of trend and cycle in practice implies, then, that they must be based on conflicting representations of the data. Identifying the source of the conflict is the subject of the next section.

2. In What Way Do UC and ARIMA Models of U.S. Real GDP Conflict?

The results of Section 1 imply that the differing results obtained in practice must be traceable to restrictions on the reduced form ARIMA model implied by the UC approach that are in conflict with the unrestricted ARIMA model used in the BN approach. Those restrictions presumably arise from the restrictions that have been placed on the UC–ARMA representation used in implementing the Kalman filter, the particular

ARMA(p, q) form of the cycle process and the zero correlation between trend and cycle innovations, what we called the UC-0 model. Following Clark (1987) we set p=2, to allow for cyclical dynamics, and q=0 in the UC-0 model and obtain for real GDP 1947:1 – 1998:2, in logs and multiplied by 100, the filtered estimate of $c_{t|t}$ shown in Figure 1.[3]

Confirming results in the literature, the estimated cycle produced by the UC-0 model is both large in amplitude and very persistent.[4] It agrees reasonably well with the NBER dating of the business cycle, although it leads the NBER dating at peaks. Reflecting the smoothness of the trend, the cycle shown here is qualitatively similar to that obtained by simple linear detrending of log output by least squares. For example, both imply that the economy has been below trend throughout the 1990s.

Table 1 reports the maximum likelihood estimates of the parameters and their standard errors for the UC-0 model. Detail of the estimation technique used is given in the technical appendix. The roots of the estimated autoregressive polynomial are complex, implying that the business cycle has a period of almost 8 years with a standard deviation of about 3 percentage points around trend, confirming the visual impression of persistence, periodicity, and amplitude in Figure 1. In contrast, the trend process innovation has a standard deviation of only about 0.7 percentage points.

[3] The data series used is gdpq from the DRI databank. Clark (1987) allowed the drift parameter to evolve as a random walk, but estimates of the variance are small. We have assumed that this parameter remains constant, implying that output is I(1).

[4] The smoothed estimate of the cycle is usually presented, but presents qualitatively the same picture.

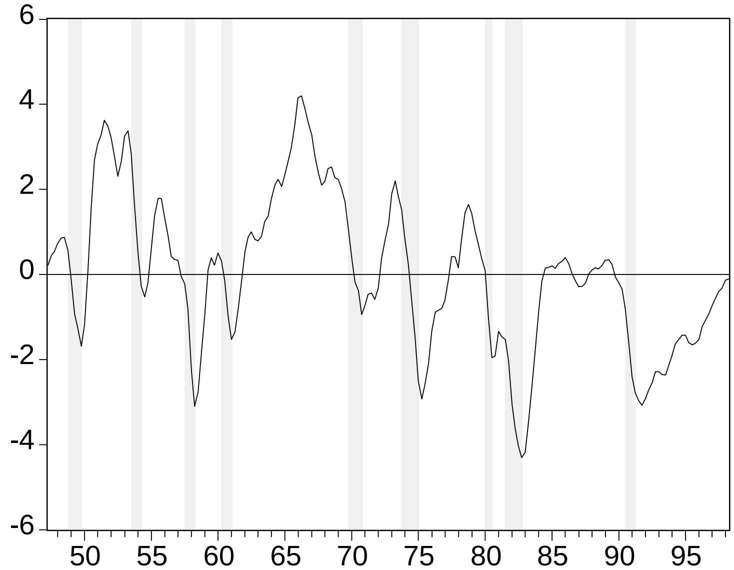

Fig. 1: UC-0 Cycle, U.S. Real GDP.
Percent deviation from trend, NBER recessions shaded

Table 1: Maximum Likelihood Estimates of UC-0 Parameters

	Estimate	Standard Error
Trend process		
Drift: m	0.8119	(0.0500)
Innovation: s_h	0.6893	(0.1038)
Cycle process		
f_1	1.5303	(0.1012)
f_2	-0.6097	(0.1140)
Innovation: s_e	0.6199	(0.1319)

AR Roots (inverted) $0.7652 \pm 0.1558i$
Implied cycle: period 7.7 years, standard deviation .03.

Log Likelihood -286.6053

The reduced form ARIMA representation for the UC-0 model corresponding to (2b) is obtained as follows. Taking first differences gives

$$\begin{aligned}\Delta y_t &= (1-L)\boldsymbol{t}_t + (1-L)c_t \\ &= m + \boldsymbol{h}_t + (1-L)(1-f_1 L - f_2 L^2)^{-1}\boldsymbol{e}_t\end{aligned} .$$

Next, multiply both sides by $(1-f_1 L - f_2 L^2)$ to give

$$(3)\quad (1-f_1 L - f_2 L^2)\Delta y_t = m^* + \boldsymbol{h}_t - f_1\boldsymbol{h}_{t-1} - f_2\boldsymbol{h}_{t-2} + \boldsymbol{e}_t - \boldsymbol{e}_{t-1} = m^* + u_t + \boldsymbol{q}_1^* u_{t-1} + \boldsymbol{q}_2^* u_{t-2} .$$

The result in (3) uses the fact that the right-hand side has a representation as an MA(2) by Granger's Lemma with the univariate innovations u_t being i.i.d. N(0, s_u^2), and m* is $m(1-f_1-f_2)$. It is important to recognize that the assumption s_{eh}=0 places complicated nonlinear restrictions on the parameters of the ARIMA(2,1,2) model (3). In particular, Lippi and Reichlin (1992) show that the long-run persistence measure, $y(1) = q(1)^*/f(1)$, will be less than or equal to one. Proietti and Harvey (2000) give further restrictions on the autoregressive parameters. These restrictions are testable implications of the UC-0 model but in empirical work they are almost never tested.

While the reduced form of the UC-0 model is a restricted ARIMA(2,1,2), when we estimate the unrestricted form of that model by exact maximum likelihood and compute the BN cycle component from it we get the very different results seen in Figure 2. As reported in the literature, the estimated BN cycle is small in amplitude compared to the UC-0 cycle and much less persistent. Table 2 reports the maximum likelihood estimates of the parameters for the unrestricted reduced-form ARIMA(2,1,2) model. Confirming the visual impression from Figure 2, the period of the cycle implied by the AR parameters here is much shorter, only 2.4 years instead of 8. The fact that the value of the log likelihood is greater by roughly 2 for the unrestricted ARIMA must reflect restrictions in the UC-0 model not imposed in the reduced form, in particular the zero correlation between trend and cycle innovations. Another indication that the zero correlation restriction may not be supported by the data is that the estimated value of persistence, $y(1)$, is greater than one. To see what correlation is implied by the ARIMA parameters, we next solve for the parameters of the unrestricted UC model of equations (1a)-(1d) that correspond to the estimated unrestricted ARIMA parameters.

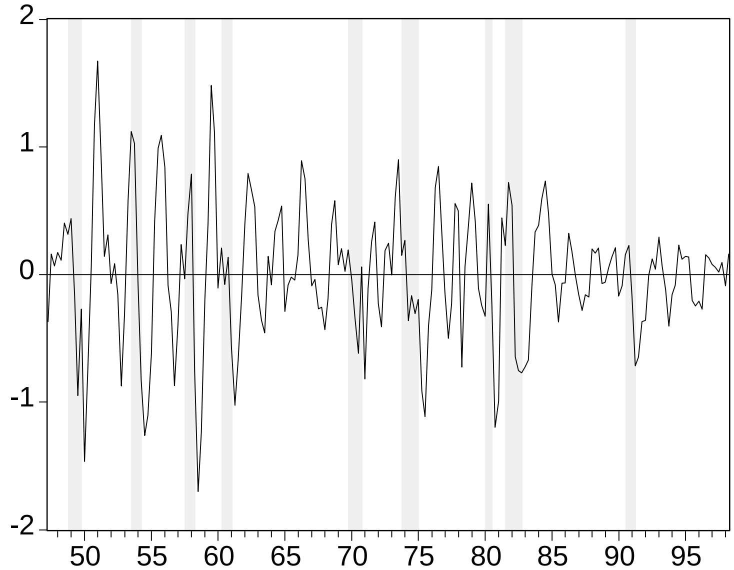

Fig. 2: Beveridge-Nelson Cycle, U.S. Real GDP.
Percent deviation from trend, NBER recessions shaded.

Table 2: Maximum Likelihood Estimates for ARIMA(2,1,2)

	Estimate	Standard Error
Drift m	0.8156	(0.0864)
f_1	1.3418	(0.1519)
f_2	-0.7059	(0.1730)
q_1	-1.0543	(0.1959)
q_2	0.5188	(0.2250)
SE of Regression	0.9694	(0.0478)
$y(1)$	1.2759	(0.1543)
AR roots (inverted)	$0.6709 \pm 0.5057i$	
Implied cycle: period 2.4 years		
MA roots (inverted)	$0.5271 \pm 0.4908i$	
Log Likelihood	-284.6507	

First note that the AR parameters are the same in both the UC and ARIMA reduced form since the AR polynomial on the left side of (2) is the AR polynomial of the UC cycle. Now the observable moments on the MA side of (2) are the mean, which identifies *m*, and the autocovariances:

(4)
$$\begin{aligned} g_0 &= (1+f_1^2+f_2^2)s_h^2 + 2s_e^2 + 2(1+f_1)s_{he} \\ g_1 &= -f_1(1-f_2)s_h^2 - s_e^2 - (1-f_2+f_1)s_{he} \\ g_2 &= -f_2 s_h^2 - f_2 s_{he} \\ g_j &= 0, j \geq 3 \end{aligned}$$

The autocovariances on the left-hand-side of (4) are

$$\begin{aligned} g_0 &= s_u^2(1+q_1^2+q_2^2) \\ g_1 &= s_u^2 q_1(1+q_2) \\ g_2 &= s_u^2 q_2. \end{aligned}$$

The system of equations in (4) can be written in matrix form as

$$\begin{pmatrix} g_0 \\ g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} 1+f_1+f_2 & 2 & 2(1+f_1) \\ -f_1(1-f_2) & -1 & -(1-f_2+f_1) \\ -f_2 & 0 & -f_2 \end{pmatrix} \begin{pmatrix} s_h^2 \\ s_e^2 \\ s_{he} \end{pmatrix}$$

or

$$g = \Phi s .$$

Assuming F is invertible, of which a necessary condition is $f_2 \neq 0$, we can solve for the UC-0 parameters as

(5)
$$s = \Phi^{-1} g .$$

Hence, the three non-zero autocovariance for the MA(2) are just sufficient to identify the three remaining parameters of the UC representation, namely s_h^2, s_e^2, and s_{he}. We note that in a particular case the solution to (4) might not imply a positive definite covariance matrix for the trend and cycle innovations, in which case there would not exist a corresponding UC-ARMA(2,0) representation.

Table 3 compares the estimates from Table 1 for the UC-0 model with the implied estimates from the unrestricted ARIMA(2,1,2) reduced form obtained from (5). While the parameters for the cycle component are somewhat similar, the unrestricted reduced form implies a standard deviation for the trend innovation that is almost twice as large compared to the UC-0 estimate, and a correlation between trend and cycle innovations that is large and negative instead of zero. It is important to note that the correlation presented in Table 3 is the estimated correlation between unobserved innovations, not the correlation between innovations in the observed estimated series $\hat{\tau}_{t|t}$ and $\hat{c}_{t|t}$. The UC-0 model restricts the correlation between unobserved innovations to be zero, while the unrestricted ARIMA reduced form estimates that correlation implicitly. Thus, the difference in correlation estimates seen in Table 3 reflects a difference between the two models, not a difference in detrending methods.

In contrast, innovations in the estimated components $\hat{\tau}_{t|t}$ and $\hat{c}_{t|t}$ are perfectly negatively correlated regardless of whether they are computed from a UC representation using the Kalman filter or from the corresponding reduced form ARIMA representation using BN. Recall from section 1, that the estimated components obtained by BN and Kalman filter procedures are numerically equivalent if the underlying models have the same reduced form, implying that the well-known perfect negative correlation of estimated BN components is a property shared by UC filtered estimates. Another way to see why UC estimates have the perfect negative correlation property is to inspect the updating equations for the Kalman filter (see Harvey (1981, chapter 4)), noting that the random walk property of the trend component implies that the innovation in the estimated trend component is proportional to the forecast error in predicted y_t, just as for the BN estimate of trend. As Wallis (1995) has pointed out, the distinction between the assumptions of a model (in the UC-0 case that the components are uncorrelated) and the properties of estimates (estimated components are correlated) is perfectly consistent with least squares estimation, but nevertheless has been the source of frequent confusion in the literature.

Table 3: Parameters of UC-0 Model and Those Implied by Unrestricted ARIMA(2,1,2) Reduced Form

	UC-0 Model	UC Model Implied by ARIMA
Trend process		
Drift: m	0.8119	0.8156
Innovation: s_h	0.6893	1.2368
Cycle process		
f_1	1.5303	1.3418
f_2	-0.6098	-0.7059
Innovation: s_e	0.6199	0.7487
Covariance s_{eh}	zero (constrained)	-0.8391
Correlation r_{eh}	zero (constrained)	-0.9062

The fact that s_{eh} is identified in this case implies that we can relax the constraint that it is zero in the UC model and estimate it directly by maximum likelihood. The resulting unconstrained UC-ARMA(2,0) model, denoted UC-UR, is cast in state-space form by including the cycle component along with the trend in the state equations, as noted by Canova (1998) and Nelson and Kim (1999b); see appendix for details. More generally, the order condition for identification of the unconstrained UC-ARMA(*p,q*) model (including the cross-correlation between trend and cycle innovations), in the sense of having at least as many moment equations as parameters, is $p \geq q+2$ and it is just satisfied in this case with $p=2$, $q=0$. Intuitively, when this holds, increasing *p* increases the number of moment equations corresponding to (4) on the MA side of the univariate ARIMA representation without increasing the number of parameters to solve for, since those are identified by the AR side.

Figure 3 displays the filtered estimate $c_{t|t}$ of the transitory component for the unrestricted UC-UR model. The estimated cycle is essentially identical to the estimated cycle from the BN decomposition.[5] This verifies that the filtered estimates from the UC model and the BN estimates are equivalent.

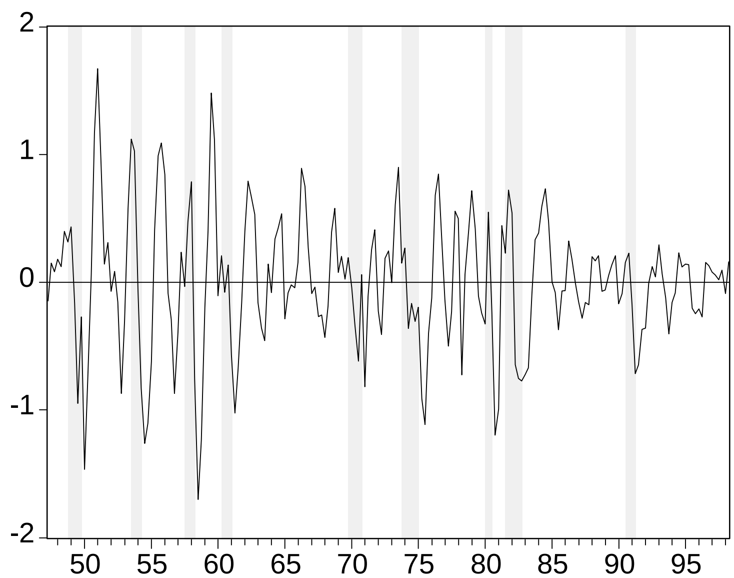

Figure 3 – UC-UR Cycle (NBER dated recessions shaded)

Table 4 reports the maximum likelihood estimates of the parameters for the unconstrained UC model. The unconstrained model was formulated to allow estimation of either the covariance, $s_{\varepsilon\eta}$, or the correlation, $r_{\varepsilon\eta}$, and both estimations produce numerically the same results and both are strongly negative. A striking feature of these estimates is that they are all essentially the same as the implied estimates from the unrestricted ARIMA model reported in Table 3. Note also that the estimated variance of the permanent component, $\hat{s}_h^2 = 1.5296$, is essentially equal to the variance of the innovation to the BN trend from the unrestricted ARIMA(2,1,2) model, $\hat{s}_u^2 \hat{y}(1)^2 = 1.5297$. While the estimated value of f_2 is several times its standard error,

---

[5] The only difference is for the first observation, which is different due to the need to provide an initial guess for the value of the random walk trend in estimation via the Kalman filter.

supporting the $p=2$ specification, we note that this is not a standard testing situation since the model is not identified if the null hypothesis that $f_2=0$ is true.

The log likelihood value for the UC-UR model is also the same as for the ARIMA model, and significantly larger than for the restricted UC-0 model, thus confirming identification of the covariance between trend and cycle innovations. The likelihood ratio statistic for testing the restriction $\rho_{\varepsilon\eta} = 0$, which may be interpreted as an overidentification test, is 3.909, with a corresponding p-value of 0.048. As a check on the small sample properties of this test, particularly whether this test rejects zero correlation too often, we generated data from the UC model calibrated to the UC-0 estimates, tested the null hypothesis $\rho_{\varepsilon\eta} = 0$, and found the size to be approximately correct. Thus, we can strongly reject the restriction of a zero correlation between permanent and transitory shocks by comparing the results for the UC-0 model with either the results for the reduced-form ARIMA model or the unrestricted UC model.

Note that the estimate of $\rho_{\varepsilon\eta}$ is –0.906 with an estimated standard error of 0.073, so a Wald-type t-test implies a far smaller p-value than the likelihood ratio test reported above. Since the estimated value of $r_{\varepsilon\eta}$ is near the boundary of admissible values the small estimated standard error might give a misleading impression of precision. To reconcile the two tests, we estimated the UC-ARIMA model for fixed values of $r_{\varepsilon\eta}$ = -0.95, -0.9, .., 0.9, 0.95 and, for each model computed the likelihood ratio statistic for the hypothesis that $r_{\varepsilon\eta}$ is equal to the imposed value. Figure 4 gives a plot of these likelihood ratio statistics as a function of the hypothetical $r_{\varepsilon\eta}$. The horizontal line indicates the 95 percent quantile from the chi-square distribution with 1 degree of freedom. The shape of the plot clearly indicates a global maximum of the likelihood at the estimated value of $r_{\varepsilon\eta}$ and the sharpness of the likelihood around that point is reflected in the small standard error. The implied 95 percent confidence interval for $r_{\varepsilon\eta}$ obtained by "inverting" the likelihood ratio statistic is fairly wide but excludes $r_{\varepsilon\eta} = 0$. Thus, the difference between the Wald and likelihood ratio test results is traced to local verses global behavior of the likelihood function.

Table 4: Maximum Likelihood Estimates for the UC-UR Model

	Estimate	Standard Error
Trend process		
Drift: m	0.8156	(0.0865)
Innovation: s_h	1.2368	(0.1518)
Cycle process		
f_1	1.3419	(0.1456)
f_2	-0.7060	(0.0822)
Innovation: s_e	0.7485	(0.1614)
Roots of AR process	0.6710 + 0.5058i	
	0.6710 - 0.5058i	
Covariance: $s_{\varepsilon\eta}$	-0.8389	(0.1096)
Correlation: $\rho_{\varepsilon\eta}$	-0.9063	(0.0728)
Log Likelihood Value	-284.6507	

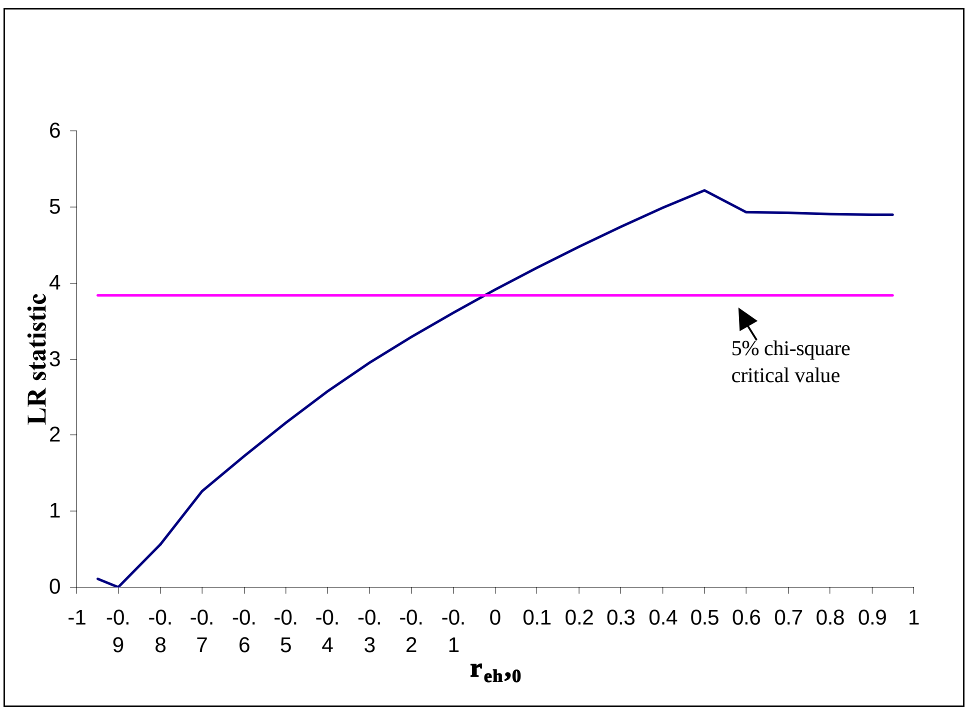

Figure 4 – Likelihood Ratio Statistics for testing $r_{\varepsilon\eta}=r_{\varepsilon\eta,0}$

3. Summary and Conclusions

We have shown that trend-cycle decompositions based on unobserved component models cast in state-space form and on the long run forecast implied by an ARIMA model differ not because they differ in principle but because the underlying empirical models differ. In particular, the restriction that innovations in the unobserved trend and cycle are uncorrelated has been imposed in the former but in the latter. We note that when this restriction is relaxed in the state-space model, the two approaches lead to identical trend-cycle decompositions and identical univariate representations. Further, the restriction of zero correlation is strongly rejected by the data for U.S. real GDP, quarterly 1947-1998.

If we accept the implication that innovations to trend are strongly negatively correlated with innovations to the cycle, then the case for the importance of real shocks in the macro economy is strengthened. As pointed out by Stock and Watson (1988) in their influential paper, real shocks tend to shift the long run path of output, so short term fluctuations will largely reflect adjustments toward a shifting trend if real shocks play a dominant role. For example, a positive productivity shock, such as the invention of the Internet, will immediately shift the long run path of output upward, leaving actual output below trend until it catches up. This implies a negative contemporaneous correlation since this positive trend shock is associated with a negative shock to the cycle. In contrast, a positive nominal shock, say a shift in Fed policy towards stimulus will be an innovation to the cycle without any impact on trend.

Closing with a few caveats, we note that the decompositions considered here share a common restriction, that the cycle process is symmetric. Recent business cycle research suggests that asymmetry has been an important feature of postwar U.S. experience - recessions being characterized as an occasional sharp drop followed by more gradual recovery; see Neftci (1984), Sichel (1993, 1994), Beaudry and Koop (1993), and Kim and Nelson (1999a). The inference that variation in GDP is dominated by variation in trend may reflect primarily the long periods of expansion when actual output is relatively close to potential and any cycle is short-lived and small in amplitude. Finally, the decompositions considered here are univariate with only two sources of shocks, and additional information introduced in a multivariate setting may affect estimates of trend and cycle.

References

Ariño, M. A. and P. Newbold, 1998, "Computation of the Beveridge-Nelson Decomposition for Multivariate Economic Time Series," *Economics Letters*, 61, 37-42.

Beaudry, P. and G. Koop, 1993, "Do Recessions Permanently Change Output?" *Journal of Monetary Economics*, 31, 149-163.

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," *J. of Monetary Economics*, 7, 151-174.

Blanchard, O., and D. Quah, 1989, "The Dynamic Effects of Aggregate Demand and Suppply Disturbances," *American Economics Review*, 79, 655-73.

Box, G. E. P., Pierce, David A., and P. Newbold, 1987, "Estimated Trend and Growth Rates in Seasonal Time Series, " *Journal of the American Statistical Association*, 82, 276-282.

Canova, F., 1998, "Detrending and Business Cycle Facts," *Journal of Monetary Economics*, 41, 475-512.

Clark, P. K., 1987, "The cyclical component of U.S. economic activity," *The Quarterly Journal of Economics*, 102, 797-814.

Cochrane, J. H., 1988, "How Big Is the Random Walk in GNP?," *Journal. of Political Economy*, 893-920.

Granger, C.W.J., and P. Newbold, 1986, *Forecasting Economic Time Series, Second Edition*, Orlando: Academic Press.

Hamilton, J.D., 1994, *Time Series Analysis*, Princeton: Princeton University Press.

Hamilton, J.D., 1993, "State-Space Models," in R. Engle and D. McFadden, eds., *Handbook of Econometrics*, Vol. 4, New York: North Holland.

Harvey, A. C., 1981, *Time Series Models,* Oxford: Philip Allen.

Harvey, A. C., 1985, "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 3, 216-27.

Harvey, A. C., and A. Jaeger, 1993, "Detrending, Stylized Facts and the Business Cycle," *Journal of Applied Econometrics*, vol. 8 no. 3, July-September, 231-47.

Hodrick, R. J., and E. C. Prescott, 1997 (1981), "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, vol. 29 no. 1, February, 1-16.

Kim, C-J. and C. R. Nelson, 1999a, "Friedman's Plucking Model of Business Fluctuations: Tests and Estimates of Permanent and Transitory Components," *Journal. of Money Credit and Banking*, 31, 317-34.

Kim, C.-J., and C.R. Nelson, 1999b, *State-Space Models with Markov Switching: Classical and Gibbs Sampling Approaches*, Cambridge: MIT Press.

Lippi, M., and L. Reichlin, 1992, "On Persistence of Shocks to Economic Variables. A Common Misconception," *Journal of Monetary Economics*, 29, 87-93.

Morley, J.C., 2001, "A State-Space Approach to Calculating the Beveridge-Nelson Decomposition," Manuscript, Washington University.

Neftci, S. N., 1984, "Are Economic Time Series Asymmetric Over the Business Cycle?" *Journal of Political Economy*, vol. 92, no. 2, 307-328.

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

Nerlove, M., D.D. Grether, and J.L. Carvalho, 1979, *Analysis of Economic Time Series – A Synthesis*, New York: Academic Press.

Proietti, T., and A. Harvey, 2000, "A Beveridge-Nelson Smoother," *Economics Letters*, 67, 139-146.

Phillips, P.C.B., and V. Solo, 1992, "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971-1001.

Sichel, D. E., 1993, "Business Cycle Asymmetry: A Deeper Look," *Economic Inquiry*, April, 224-236.

Sichel, D. E., 1994, "Inventories and the Three Phases of the Business Cycle," *Journal of Business and Economic Statistics*, vol. 12, no. 3, 269-77.

Stock, J H., and M. W. Watson, 1988, "Variable Trends in Economic Time Series," *Journal. of Economic Perspectives*, 2, 147-74.

Wallis, K. F., 1995, “Discussion of 'Trend Extraction: A Practitioners Guide' by M. R. Wickens,” HM Treasury Working Paper No. 67.

Watson, M W., 1986, “Univariate detrending methods with stochastic trends,” *Journal. of Monetary Economics*, 18, 49-75.

Appendix

This appendix contains notes on how to calculate the BN decomposition for an ARIMA model and how to set up the state-space model in order to estimate the UC1 model via the Kalman filter.

BN decomposition

Numerous practical methods for calculation of the exact BN decomposition have been proposed in the literature, including Newbold (1990) and Ariño and Newbold (1998). As discussed in Morley (2001), a particularly convenient way to calculate the exact BN decomposition for an ARIMA model is to first convert the model into its state-space companion form. For the ARIMA(2,1,2) model, this is given as follows:

$$\begin{bmatrix} \Delta y_t - m \\ \Delta y_{t-1} - m \\ e_t \\ e_{t-1} \end{bmatrix} = \begin{bmatrix} f_1 & f_2 & q_1 & q_2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} - m \\ \Delta y_{t-2} - m \\ e_{t-1} \\ e_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} e_t , \tag{A.1}$$

or, more compactly,

$$b_t = F b_{t-1} + v_t \tag{A.1'}$$

when the eigenvalues of F have modulus less than 1 and there are no canceling roots so that F is invertible. Then, the BN trend can be calculated at any given point of time as

$$BN_t = y_t + \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} F (I - F)^{-1} b_{t|t} , \tag{A.2}$$

where $b_{t|t} = E_t[b_t]$ can be obtained from the Kalman filter used in exact maximum likelihood estimation of the ARIMA(2,1,2) model. We define the BN cycle as the gap between real GDP and the BN trend, ($y_t - BN_t$).

The State-Space Model

One possible explanation for prevalence of the unnecessary zero correlation assumption in empirical work is the way the state-space model for a UC model is traditionally set up and estimated. In particular, the random walk trend is often treated as the state variable, while the AR(2) cycle is treated as a residual in the observation equation; for example, see Clark (1987). Given this setup, estimation via the Kalman filter requires an assumption of independence between trend and cycle innovations.

However, as demonstrated in Kim and Nelson (1999), there are other ways of setting up the state-space representation for a UC model. One possibility is to make the observation equation an identity, with both the trend and cycle treated as state variables. This is the approach we take:

<u>Observation Equation</u>:

$$y_t \equiv \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} t_t \\ c_t \\ c_{t-1} \end{bmatrix}, \qquad \text{(A.3)}$$

or, more compactly,

$$y_t = H\boldsymbol{b}_t, \qquad \text{(A.3')}$$

<u>State Equation</u>:

$$\begin{bmatrix} t_t \\ c_t \\ c_{t-1} \end{bmatrix} = \begin{bmatrix} m \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & f_1 & f_2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} t_{t-1} \\ c_{t-1} \\ c_{t-2} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} h_t \\ e_t \end{bmatrix}, \qquad \text{(A.4)}$$

or, more compactly,

$$b_t = \tilde{m} + F b_{t-1} + v_t \,. \qquad \text{(A.4')}$$

Then, we can use the Kalman filter to estimate the model, even if we allow the trend and cycle innovations to be correlated. That is,

$$Q \equiv E[v_t v_t'] = \begin{bmatrix} s_h^2 & s_{he} & 0 \\ s_{he} & s_e^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad \text{(A.5)}$$

The Kalman filter is given by the following six equations:

$$b_{t|t-1} = \tilde{m} + F b_{t-1|t-1} \,, \qquad \text{(A.6)}$$

$$P_{t|t-1} = F P_{t-1|t-1} F' + Q \,, \qquad \text{(A.7)}$$

$$y_t - y_{t|t-1} = y_t - H b_{t|t-1} \,, \qquad \text{(A.8)}$$

$$f_{t|t-1} = H P_{t|t-1} H' \,, \qquad \text{(A.9)}$$

$$b_{t|t} = b_{t|t-1} + K_t (y_t - y_{t|t-1}) \,, \qquad \text{(A.10)}$$

$$P_{t|t} = P_{t|t-1} - K_t H P_{t|t-1} \,, \qquad \text{(A.11)}$$

where $b_{t|t-1} \equiv E[b_t | y_{t-1}]$, for example, is the expectation of b_t conditional on information up to time $t-1$; $P_{t|t-1}$ is the variance of $b_{t|t-1}$; $f_{t|t-1}$ is the variance of $(y_t - y_{t|t-1})$; and $K_t \equiv P_{t|t-1} H' f_{t|t-1}^{-1}$ is the Kalman gain.[6]

[6] For a more general discussion of the Kalman filter and state-space models, as well as details on the derivation of the Kalman gain, refer to Hamilton (1993, 1994) and Kim and Nelson (1999).

Given some initial values $b_{0|0}$ and $P_{0|0}$ (as discussed below), we can iterate through (A.6)-(A.11) for $t = 1, \ldots, T$ to obtain filtered inferences about b_t conditional on information up to time *t*. Also, as a by-product of this procedure, we obtain $(y_t - y_{t|t-1})$ and $f_{t|t-1}$, which we can use to find maximum likelihood estimates of the hyper-parameters based on the prediction error decomposition:

$$\max_{q} l(q) = -\frac{1}{2}\sum_{t=t+1}^{T} \ln(2p f_{t|t-1}) - \frac{1}{2}\sum_{t=t+1}^{T} (y_t - y_{t|t-1})^2 f_{t|t-1}^{-1}, \qquad \text{(A.12)}$$

where $q = (m, f_1, f_2, s_h, s_e, s_{he})'$ and t gives the number of start-up values skipped in the evaluation of the log-likelihood. We set $t = 2$.

In terms of the initial values $b_{0|0}$ and $P_{0|0}$, we specify the initial value of the data for the random walk component, but assign it an extremely large variance. For the transitory component, we use the unconditional mean and variance of the AR(2) process. In the process of the maximizing the log-likelihood function we impose stationarity constraints on the autoregressive parameters and a positive definiteness constraint on the innovation covariance matrix.