

The Two Interpretations of the Beveridge-Nelson Decomposition

James C. Morley[†]
Washington University

July 2008

ABSTRACT: The Beveridge-Nelson decomposition provides measures of trend and cycle for integrated time series. However, there are two ways to interpret the results from the decomposition. One interpretation is that the long-run forecast (minus any deterministic drift) used to calculate the Beveridge-Nelson trend corresponds to an *estimate* of an unobserved permanent component. The other interpretation is that the long-run forecast *defines* an observable permanent component. This paper examines some issues surrounding these two interpretations and presents empirical support for the Beveridge-Nelson trend and cycle corresponding to estimates when considering with macroeconomic data.

JEL classification: C22; C32; C50

Keywords: Beveridge-Nelson Decomposition; Trend/Cycle Decomposition; State-Space Models; Unobserved Components Models

[†] Department of Economics, Box 1208, Washington University, One Brookings Drive, St. Louis, MO 63130-4899, USA. Email: morley@wustl.edu. Phone: (314) 935-4437. Fax: (314) 935-4156. I thank Eric Zivot for helpful comments. All errors are mine alone.

“The Beveridge-Nelson decomposition can be seen as an ingenious decomposition of an $I(1)$ variable, but it does not properly fit into the unobserved components framework, since the components are, in fact, observable...The assumption ... that the permanent and transitory component share, at every period, the same innovation is a strong assumption, of limited appeal.” – Maravall (1995)

1. Introduction

In the literature on trend/cycle decomposition, the Beveridge-Nelson decomposition has been subject to two very distinct interpretations. One interpretation, emphasized by Watson (1986) and Morley, Nelson, and Zivot (2003), is that the long-horizon forecast (minus any deterministic drift) used to calculate the Beveridge-Nelson trend corresponds to an estimate of the permanent component of an integrated time series. The Beveridge-Nelson decomposition provides a sensible estimate because, under the assumption that the permanent component follows a random walk (with drift) and the transitory component is stationary with an unconditional mean of zero, the long-horizon forecast will be equal to the conditional expectation of the permanent component. A second interpretation, emphasized in the original paper by Beveridge and Nelson (1981), is that the Beveridge-Nelson trend provides a possible definition of the permanent component of an integrated time series. In this case, the permanent component is observable because it is the Beveridge-Nelson trend.

In this paper, I compare the two interpretations of the Beveridge-Nelson decomposition. While many other studies, including Watson (1986), Maravall (1995), Harvey and Koopman (2000), Proietti and Harvey (2000), Proietti (2006), and Oh, Zivot,

and Creal (2006), have alluded to the two interpretations of the Beveridge-Nelson decomposition, I consider exactly how the interpretations are related to each other and whether there is any empirical reason to prefer one interpretation over the other. In particular, I demonstrate that, even though the two interpretations can be observationally equivalent in a univariate setting, they are empirically distinguishable in a multivariate setting. Importantly, this empirical distinction is possible even if the autocovariance structures of the series under consideration can be completely captured by univariate time series models. Then, in an application to U.S. macroeconomic data, I find support for the interpretation that the Beveridge-Nelson decomposition provides estimates of trend and cycle. This finding is important because the view that the Beveridge-Nelson decomposition defines the trend persists in many applied studies (e.g., Clarida and Taylor, 2003, and Anderson, Low, and Snyder, 2006) and forms the basis for skepticism about its general relevance (e.g., the above quote from Maravall). By contrast, the view that the Beveridge-Nelson decomposition provides estimates of trend and cycle implies a highly general and practical method of trend/cycle decomposition.

2. The Beveridge-Nelson Decomposition

The Beveridge-Nelson (BN) trend of an integrated time series y_t is given as follows:

$$BN_t = \lim_{M \rightarrow \infty} E[y_{t+M} - M\mu | \Omega_t], \quad (1)$$

where $\mu = E[\Delta y_t]$ is the deterministic drift and Ω_t is the information set used to calculate the conditional expectation. In words, the BN trend is the long-horizon

conditional point forecast of the time series process $\{y_t\}$, with any future drift removed.

Meanwhile, the BN cycle is simply the difference between the series and the BN trend.

In practice, the BN decomposition is often calculated using an autoregressive moving-average (ARMA) model that is designed to capture the autocovariance structure of $\{\Delta y_t\}$. For example, given an AR(1) model $\Delta y_t = \mu + \phi(\Delta y_t - \mu) + e_t$, where $|\phi| < 1$ and $e_t \sim i.i.d.N(0, \sigma^2)$, the BN trend is $BN_t = y_t + \frac{\phi}{1-\phi}(\Delta y_t - \mu)$ (see Morley, 2002). Note that the BN trend for y_t is observed at time t .

In order to compare the two interpretations of the BN decomposition within a unified framework, I consider a state-space representation for y_t . In particular, assuming a known Gaussian ARMA structure for the first differences $\{\Delta y_t\}$, the level y_t can be thought of as made up of a permanent component, denoted τ_t , and a transitory component, denoted c_t :

$$y_t = \tau_t + c_t, \quad (2a)$$

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

$$\phi(L)c_t = \theta(L)\varepsilon_t, \quad \varepsilon_t \sim i.i.d.N(0, \sigma_\varepsilon^2), \quad (2c)$$

$$\text{corr}(\eta_t, \varepsilon_t) = \rho_{\eta\varepsilon}. \quad (2d)$$

According to (2b), the permanent component follows a random walk with drift and, according to (2c), the transitory component follows a stationary ARMA process with a mean of zero. Note that this is a general state-space representation rather than an unobserved components (UC) representation because the permanent and transitory components could be observable, as discussed in more detail below.

Within the context of the state-space representation in (2), the two interpretations of the BN decomposition can be understood as follows:

Interpretation #1: “BN-as-estimate”

Under this interpretation, the permanent component of a time series is unobservable due to the assumed presence of transitory shocks that have no impact on the permanent component. In particular, the serially-uncorrelated innovation to c_t in (2c) can be rewritten in the following way:

$$\varepsilon_t = \alpha\eta_t + \varepsilon_t^*, \quad \varepsilon_t^* \sim i.i.d.N(0, \sigma_{\varepsilon^*}^2), \quad (3a)$$

Then, as long as the variance of the transitory shocks is positive (i.e., $\sigma_{\varepsilon^*}^2 > 0$), there will be imperfect correlation between the innovations to the permanent and transitory components:

$$|\rho_{\eta\varepsilon}| < 1. \quad (3b)$$

With imperfect correlation, (2) becomes a UC representation and, following the analysis in Watson (1986) and Morley, Nelson, and Zivot (2003), the BN trend provides an optimal (minimum mean squared error) estimate of the permanent component under the assumption that it follows a random walk and the unconditional expectation of $\{c_t\}$ is zero.

Interpretation #2: “BN-as-definition”

Under this interpretation, the BN trend is the permanent component of a time series because there is assumed to be only one type of shock driving $\{y_t\}$. That is, the serially-uncorrelated innovation to c_t in (2c) can be rewritten in the following way:

$$\varepsilon_t = \alpha \eta_t, \quad (4a)$$

where α is a scalar that allows permanent and transitory innovations to have different signs and variances despite being driven by the same underlying shock. With only one underlying shock, the innovations to the permanent and transitory components will be perfectly correlated:

$$|\rho_{\eta\varepsilon}| = 1. \quad (4b)$$

In this case, the permanent and transitory innovations are observable and can be measured using the forecast error from the reduced-form ARMA representation for $\{\Delta y_t\}$, with the resulting permanent component equal to the BN trend.

It should be noted that these two interpretations still apply even if the structure of $\{\Delta y_t\}$ is better captured by a multivariate time series model, rather than a univariate ARMA model. In such a case, the *BN-as-estimate* interpretation would correspond to the idea that the permanent and transitory innovations are imperfectly correlated across series, as well as with each other. Meanwhile, the *BN-as-definition* interpretation would correspond to the idea there are as many underlying shocks as there are series under examination, with the shocks being observable and proportional to the forecast errors from the multivariate model. In particular, for a given series, the permanent and transitory innovations would be proportional to the same linear combination of the forecast errors and, therefore, would remain perfectly correlated.

3. Univariate Observational Equivalence

One problem in discriminating between the two interpretations of the BN decomposition presented in the previous section is that, despite the apparent restrictiveness in terms of the correlation in (4b) compared to (3b), the two interpretations can be observationally equivalent in terms of the implied univariate autocovariance structure for $\{\Delta y_t\}$.

An empirical example may help clarify the distinction between the two interpretations and illustrate the problem of observational equivalence in the univariate setting. Consider the following estimated equations for a UC model of the natural logarithms of U.S. real GDP over the sample period of 1947:Q1 to 1998:Q2 taken from Morley, Nelson, and Zivot (2003):

$$y_t = \tau_t + c_t, \quad (5a)$$

$$\tau_t = \underset{(0.09)}{0.82} \tau_{t-1} + \eta_t, \quad \eta_t \sim N(0, \underset{(0.15)}{1.24^2}), \quad (5b)$$

$$c_t = \underset{(0.15)}{1.34} c_{t-1} - \underset{(0.08)}{0.71} c_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \underset{(0.16)}{0.75^2}), \quad (5c)$$

$$\text{Corr}(\eta_t, \varepsilon_t) = \underset{(0.07)}{-0.91}, \quad (5d)$$

where standard errors are in parentheses and the log likelihood value is -284.65 . Because of the assumed UC structure, the permanent component must be estimated. Given the parameter estimates, one way to estimate the permanent component is by using the Kalman filter. In particular, the Kalman filter calculates $E[\tau_t | \Omega_t]$, where $\Omega_t = (y_1, \dots, y_t)$.

It turns out that the estimated UC model in (5) places no binding restrictions on the autocovariance structure of $\{\Delta y_t\}$ beyond those implied by an estimated reduced-

form ARMA(2,2) model. As shown in Morley, Nelson, and Zivot (2003), the equivalent autocovariance structure means that $E[\tau_t | \Omega_t]$ from the Kalman filter is identical to the BN trend given the ARMA(2,2) model. Meanwhile, any restriction on the correlation parameter in (5d) would place implicit restrictions on the parameters for the reduced-form ARMA(2,2) model, resulting in a different BN trend than implied by the unrestricted ARMA(2,2) model. The example that Morley, Nelson, and Zivot consider is the widely-used restriction that the correlation between permanent and transitory innovations is zero (i.e., $\rho_{\eta\epsilon} = 0$). The resulting values for $E[\tau_t | \Omega_t]$ are very different from those for the general model in (5). In particular, the estimates of the permanent component are less volatile and implicitly correspond to the BN trend for a restricted version of an ARMA(2,2) that reflects the zero correlation assumption for the UC model. Likewise, it is possible to consider a restricted version of the state-space model in which the correlation between permanent and transitory innovations is restricted to be -1 instead of its estimated value of -0.91 in (5d). While it might seem that this version of the state-space model corresponds to the *BN-as-definition* interpretation, it actually produces a different permanent component than the unrestricted UC model in (5). The reason is that, given the AR(2) structure for the transitory component, the restriction of perfect negative correlation (i.e., $\rho_{\eta\epsilon} = -1$) implies a restricted version of the ARMA(2,2) model for $\{\Delta y_t\}$. Thus, the BN trend for the restricted model would be different from the BN trend for the unrestricted model.

Given that imposing perfect negative correlation would meaningfully restrict the model in (5), the obvious question is why there is an observational equivalence between the two interpretations of the BN decomposition. The answer lies in Anderson, Low, and

Snyder's (2006) insight that reduced-form ARMA models are equivalent to state-space models with only one type of shock, but comparatively complicated dynamics. For example, they show that an unrestricted ARMA(2,2) model for $\{\Delta y_t\}$ is equivalent to a state-space model for $\{y_t\}$ in which innovations between permanent and transitory components are perfectly correlated and the transitory component follows an ARMA(2,1) process, instead of the AR(2) process in (5c). Using the same data as Morley, Nelson, and Zivot (2003), the maximum likelihood estimates for this alternative state-space model are given as follows:

$$y_t = \tau_t + c_t, \quad (6a)$$

$$\tau_t = \underset{(0.09)}{0.82} \tau_{t-1} + \eta_t, \quad \eta_t \sim N(0, \underset{(0.08)}{1.24^2}), \quad (6b)$$

$$c_t = \underset{(0.06)}{1.34} c_{t-1} - \underset{(0.10)}{0.71} c_{t-2} - \underset{(0.01)}{0.22} \eta_t + \underset{(0.04)}{0.30} \eta_{t-1}, \quad (6c)$$

where the log likelihood value is -284.65 . Note that the log likelihood is the same as for the model in (5), meaning that the models in (5) and (6) are observationally equivalent in terms of fitting the sample data.

There are, however, a few issues that should be mentioned regarding the observation equivalence. First, given the same ARMA model for $\{\Delta y_t\}$, the BN trend is the same under either interpretation. Thus, inferences about the variability of the permanent component are not sensitive to the interpretation. This robustness to interpretation stands in contrast to the sensitivity of inferences to different assumptions about the correlation between permanent and transitory innovations for a given UC model. Second, despite implying the same variability of the permanent component, the observational equivalence does not mean that deciding between the two interpretations is

merely a matter of normalization. The two interpretations have very different implications in terms of subsequent econometric analysis. Under the *BN-as-definition* interpretation, the BN trend and the implied transitory component are observable and, therefore, can be treated as regular data in any regression analysis, at least assuming a reasonable model of the autocovariance structure and precise parameter estimates. However, under the *BN-as-estimate* interpretation, the estimated components may contain a large degree of measurement error, even with the correct model of the autocovariance structure and as parameter uncertainty goes away asymptotically.

4. Measurement Error in the Multivariate Setting

In principle, given different implications for the presence of measurement error, it should be possible to empirically discriminate between the two interpretations in a multivariate setting. Specifically, if there is a sizable amount of measurement error, subsequent estimates of relationships between BN trends and/or cycles for different variables will be biased and inconsistent due to the endogeneity that arises from measurement error. Thus, given appropriate instruments, the two interpretations can be compared using a Hausman (1978) test, with evidence for endogeneity supporting the *BN-as-estimate* interpretation.¹ On the other hand, the applicability of an endogeneity test is not entirely obvious. In particular, any apparent evidence of measurement error based

¹ In the related setting of trend/cycle decomposition based on UC models, Watson (1986) compares ordinary least squares (OLS) and instrumental variable (IV) estimates for a test of the permanent income hypothesis using filtered and smoothed estimates of transitory components from UC models, with lagged data serving as instruments. He finds that filtered estimates, which condition on data up to and including the period in which the inferences about trend and cycle are being made, appear to be subject to measurement error in the sense that they are different than smoothed estimates, which condition on the full available sample of data. Thus, given the link between the BN decomposition and filtered estimates discussed in Morley, Nelson, and Zivot (2003), there is a direct suggestion from Watson's analysis that the *BN-as-estimate* interpretation is more appropriate in terms of thinking about "permanent income" as a structural quantity that has a relationship with other macroeconomic variables.

on a particular set of instruments may, in fact, be due to a mistaken exclusion of these variables from the forecasting model used to construct the BN trend and cycle. Thus, a primary purpose of this section is to demonstrate that it is possible to test for measurement error even if the forecasting model used for the BN decomposition includes the instruments used in the subsequent Hausman test for endogeneity.

Monte Carlo analysis provides the ideal means of illustrating the key issues related to the BN decomposition and measurement error. To this end, I consider two stylized data generating processes (DGPs) that correspond to the two interpretations of the BN decomposition. Then, for each DGP, I consider two series that are related to each other and examine the ability to empirically detect the true relationship between the series in different circumstances. The DGPs are given as follows:

DGP #1: “*Permanent and transitory components are unobservable*”

For this DGP, the two underlying series have the same general structure as the model in (5):

$$y_{it} = \tau_{it} + c_{it}, \quad (7a)$$

$$\tau_{it} = 1 + \tau_{i,t-1} + \eta_{it}, \quad \eta_{it} \sim N(0,1), \quad (7b)$$

$$c_{it} = 1.25c_{i,t-1} - 0.75c_{i,t-2} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0,0.5^2), \quad (7c)$$

$$\text{Corr}(\eta_{it}, \varepsilon_{it}) = 0, \quad (7d)$$

where $i = 1, 2$. In this case, the two series are related to each other through the following correlation:

$$\text{Corr}(\varepsilon_{1t}, \varepsilon_{2t}) = 0.5. \quad (7e)$$

That is, only the transitory innovations are positively correlated across the two series.

DGP #2: “Permanent and transitory components are BN trend/cycle”

For this DGP, the two underlying series have the same general structure as the model in

(6):

$$y_{it} = \tau_{it} + c_{it}, \quad (8a)$$

$$\tau_{it} = 1 + \tau_{i,t-1} + \eta_{it}, \quad \eta_{it} \sim N(0,1), \quad (8b)$$

$$c_{it} = 1.25c_{i,t-1} - 0.75c_{i,t-2} - 0.2\eta_{it} + 0.3\eta_{i,t-1}, \quad (8c)$$

where $i = 1, 2$. Meanwhile, the two series are related to each other through the following correlation:

$$\text{Corr}(\eta_{1t}, \eta_{2t}) = 0.5. \quad (8d)$$

That is, permanent and transitory innovations are positively correlated across the two series.

For each DGP, I generate 10,000 samples of simulated observations and then calculate the BN trends and cycles for each sample. For simplicity, the BN calculations are done under the assumption of known parameters in order to clarify the source of estimation uncertainty when looking at characteristics of the BN trends and cycles. I consider sample sizes of 200 and 1000 observations.

Table 1 reports results for the Monte Carlo analysis. Starting with inferences about variation in permanent and transitory components and, for the time being, focusing on the results given BN calculations based on univariate forecasting models, the first thing to notice is that, even when the BN trend is only an estimate, the sample standard deviation of the BN trend innovations provides a consistent estimate of the standard

deviation of the true permanent innovations.² Of course, the sample standard deviation is consistent when the BN trend is the true permanent component, although it is interesting to note that the estimator is equally precise for both DGPs. The second thing to notice is that, when the BN cycle is only an estimate, its sample standard deviation provides an inconsistent estimate of the true standard deviation of the transitory component, with the BN cycle understating the variability of the true transitory component. This result presents the first incidence of why it matters which interpretation of the BN decomposition is considered in practice. While the estimated trend behaves like the true permanent component, meaning that inferences about it are robust to interpretation, the estimated cycle does not, meaning that inferences depend on the interpretation, with estimated cycles displaying different variation than the true transitory components due to the presence of measurement error.

The next thing to notice is the OLS inferences about the relationship between transitory components for two related series. As with inferences about variation of the transitory components, OLS estimates are biased and inconsistent when the BN cycles are estimates rather than true values. This result is a simple example of the classic errors-in-variables problem. For the BN calculations based on univariate forecasting models, the measurement error in the estimated cycles produces downwardly biased estimates of the true relationship between the transitory components. Meanwhile, not surprisingly, the OLS inferences are consistent when the BN cycles are the true transitory components.

² For the Monte Carlo, the univariate BN decompositions are calculated by directly applying the Kalman filter to separate state-space models for each series in (7) and (8). Given the DGPs, this approach is equivalent to solving for the reduced-form forecasting model for each series, casting it into state-space form, and applying the state-space approach to the BN decomposition presented in Morley (2002) to obtain the corresponding BN trend and cycle. Also, because the two series for each DGP have the same variance parameters, I make no distinction between the two series for a given DGP when discussing inferences about variation in their permanent and transitory components.

The results so far suggest that the reliability of OLS inferences about the variation of the transitory components and their relationship across series depends heavily on which interpretation of the BN decomposition is appropriate. As mentioned above, models corresponding to the two interpretations have the same in-sample fit, so it might appear to be completely a matter of identification regarding the nature of the permanent and transitory components of a time series process. However, the errors-in-variables problem for the estimated trend and cycle directly implies a way to move beyond the problem of observational equivalence in the univariate setting, although it requires the existence of an additional time series that can serve as an instrument for the transitory component in one of the original two series. For the errors-in-variables problem, a good instrument will be correlated with the transitory component, but uncorrelated with the measurement error in the BN cycle. For Monte Carlo analysis, I add a third series to each DGP that is stationary and imperfectly correlated with the transitory component of the second series according to the following structure:

$$y_{3t} = c_{2t} + u_t, \quad u_t \sim N(0, \sigma_u^2), \quad (9)$$

where the variance of the error u_t is calibrated to produce a correlation of 50% between the instrument y_{3t} and the transitory component c_{2t} (the variance is different for the two DGPs).

The results for the IV inferences clearly suggest that, given a good instrument, IV estimation works well in the sense that the estimates are consistent even for the DGP in which the permanent and transitory components are unobservable. Thus, in practice, it should be possible to compare OLS and IV inferences in order to determine which interpretation of the BN decomposition is more appropriate for a given set of time series.

If the OLS and IV inferences are significantly different, it supports the *BN-as-estimate* interpretation. If the OLS and IV inferences are essentially the same, with the IV estimates only being somewhat less precise, it supports the *BN-as-definition* interpretation.

Importantly, these results do not merely reflect implicitly different DGPs when a third variable is introduced into the system via IV analysis. In particular, even given BN decompositions based on multivariate forecasting models that take into account the true joint autocorrelation structure of all three series, the Monte Carlo results display the same overall pattern as for the univariate forecasting models.³ To be sure, there are some quantitative differences in the findings for the multivariate case when the BN trends and cycles are estimates. The bias in the inferences about the standard deviation of the transitory components is smaller than in the univariate case, especially for the transitory component of the second series that is better identified by the information in the third series. There is also less bias in the OLS inferences about the relationship between the transitory components, with the direction of the bias switching from before. However, the crucial result is that there is still a difference between the OLS and IV inferences when the BN trends and cycles are estimates, with only the IV inferences being consistent.

The difference between OLS and IV inferences is crucial because, in practice, whether or not there is a change in certain inferences (e.g., the variability of the cycle) when switching between univariate and multivariate forecasting models is not sufficient to discriminate between the two interpretations of the BN decomposition. It only raises

³ For the Monte Carlo analysis, the multivariate BN decompositions are calculated by applying the Kalman filter to a multivariate state-space model that incorporates all three series for each DGP. Again, this approach is equivalent to solving for the reduced-form multivariate forecasting model, casting it into state-space form, and applying the state-space approach to the BN decomposition.

possible doubts about the relevance of the univariate forecasting model (or, on the contrary, whether the multivariate model is overfitting the data). For example, for the DGP where the BN trends and cycles are only estimates, the multivariate BN decomposition is different from the univariate BN decomposition because the third series Granger-causes the second series and it requires a multivariate model to capture the joint autocorrelation structure of the three series. Meanwhile, for the DGP where the BN trends and cycles are the true permanent and transitory components, the multivariate BN decomposition is the same as the univariate BN decomposition because there is no marginal forecasting information in the third series, meaning that the univariate model is sufficient for summarizing the autocorrelation structures of the first and second series (technically, they are both weakly exogenous with respect to each other and the third variable). However, there is nothing about a change in inferences for different forecasting models that favours a particular interpretation of the BN decomposition. It is true that a multivariate forecasting model suggests the presence of multiple shocks, but it does not imply that these shocks will be unobservable given the information set at hand. Instead, it is a difference between OLS and IV inferences given a forecasting model that fully captures the joint autocorrelation structure of any set of time series under examination that suggests the presence of unobservable shocks and provides clear support for the *BN-as-estimate* interpretation over the *BN-as-definition* interpretation.

5. Application to Macroeconomic Data

Given that it is possible to empirically discriminate between the two interpretations of the BN decomposition, an immediate question is which interpretation is more appropriate in practice when dealing with macroeconomic data. In this section, I

consider an application of the OLS and IV analysis in the previous section to U.S. real GDP and Industrial Production, with Capacity Utilization in Manufacturing serving as an instrument. The data are from the St. Louis Fed database and cover the sample period of 1972:1 to 2005:4.⁴

As a first step in motivating the analysis, I test for the presence of stochastic trends in the data series using the basic Augmented Dickey-Fuller unit root test with lag selection based on the Schwarz Information Criterion. At the 5% significance level, I am unable to reject the null of a unit root against the trend stationary alternative for either log real GDP ($t = -3.09$ for one lag, with a p -value of 0.11) or log Industrial Production ($t = -3.08$ for one lag, with a p -value of 0.12). For log Capacity Utilization in Manufacturing, I am able to reject the null of a unit root in favour of a level stationary alternative ($t = -3.74$ for one lag, with a p -value of <0.01). Thus, I consider applying the BN decomposition to the real GDP and Industrial Production series, but not to the Capacity Utilization series. At the same time, the fact that the Capacity Utilization in Manufacturing appears to be stationary suggests that it could be a good instrument for the transitory component of Industrial Production.

The main practical issue in applying the BN decomposition is determining which forecasting model to use. In this paper, I attempt to remain agnostic about which models are most appropriate for the variables under consideration and check the robustness of my findings to different modeling assumptions. The main practical difference between models is in terms of their implied long-horizon predictability of real economic activity.

⁴ The sample period is determined by the availability of the Capacity Utilization series. I convert Industrial Production from monthly to quarterly data by taking average growth rates. I then transform all of the series into natural logarithms and multiply by 100.

For example, low-order ARMA models of output growth imply positive serial correlation at short horizons, but little long-horizon predictability. Meanwhile, higher-order ARMA models and multivariate models can imply a large degree of negative serial correlation at long horizons. In terms of the BN decomposition, these two competing views produce very different looking cycles. Thus, I consider models that accommodate both of these views.

Figure 1 displays the BN cycles for real GDP and Industrial Production based on univariate AR(12) models for the first differences, along with the Capacity Utilization series. These BN cycles reflect negative serial correlation at long horizons and display a strong correspondence to the NBER reference cycle. The two cycles also appear to be related to each other. This relationship is not particularly surprising given the large role manufacturing plays in overall economic activity. However, it is an interesting empirical question as to how closely the manufacturing sector moves with the overall economy. To quantify the relationship, I regress the real GDP cycle on the Industrial Production cycle using OLS:

$$y_t - BN_{y,t} = 0.00 + \frac{0.60}{(0.04)}(ip_t - BN_{ip,t}) + e_t, R^2 = 0.66. \quad (10)$$

Given the standard errors reported in parentheses, the estimates suggest a strong relationship, but it is clearly less than one-for-one.

The main question raised in this paper is whether the estimates in (10) are reliable, even assuming the forecasting models used in the BN decompositions are reasonable. In particular, if the BN trend is an estimate, then the true transitory components will be measured with error and the OLS estimate of their relationship will be biased and inconsistent. To examine this issue, I consider Capacity Utilization in

Manufacturing as an instrument for the transitory component of Industrial Production. Economic considerations suggest that it should be strongly correlated with the true transitory component, but it should presumably be unrelated to any measurement error in the BN cycle. Meanwhile, if the BN cycle is the true transitory component, IV estimates should not be significantly different than the OLS estimates reported in (10).⁵ Figure 1 suggests that there is a strong relationship between the transitory component in Industrial Production and the Capacity Utilization series. Indeed, the sample correlation between the BN cycle for Industrial Production and Capacity Utilization is 70%. Thus, there is little concern about a weak instrument.

The IV regression results for the BN cycles based on AR(12) models are given as follows:

$$y_t - BN_{y,t} = 0.00 + 0.72(ip_t - BN_{ip,t}) - 0.24 \hat{u}_t + e_t, \quad R^2 = 0.69, \quad (11)$$

(0.04) (0.05) (0.07)

where $\hat{u}_t = ip_t - BN_{ip,t} + 69.45 - 0.16cu_t$ is the residual based on a first-stage regression of the BN cycle for Industrial Production on the Capacity Utilization series. The first thing to notice is the coefficient on \hat{u}_t . A t -test for this coefficient is equivalent to the Hausman (1978) test for endogeneity. Thus, the t -statistic of -3.43 corresponds to a strong rejection of the null hypothesis of no endogeneity at the 1% level, which suggests that measurement error is a problem for the BN cycles. That is, this test supports the *BN-as-*

⁵ There are two other possible sources of endogeneity in a regression of one transitory component on the other: omitted variables and simultaneity. However, these sources have to do with a failure to identify a structural relationship in which one transitory component causes the other. By contrast, I am interested in measuring the reduced-form correlation between the transitory components rather than any structural relationships. Capacity utilization is a good instrument for this purpose, not only because it should be uncorrelated with measurement error in the BN cycles, but also because it would not be a good instrument or omitted variable for identifying a structural relationship between the two transitory components. In particular, it should be subject to all of the same structural shocks as the two transitory components. I provide a test of this assumption below.

estimate interpretation. The second thing to notice is that the measurement error has meaningful and predictable effects on estimates of relationship between the transitory components. In particular, the larger coefficient on the BN cycle for Industrial Production in (11) than in (10) suggests a stronger relationship between the true transitory components than between the BN cycles.

Of course, it must be immediately acknowledged that the above results are based on univariate forecasting models, while a multivariate model may be more appropriate. Thus, I also consider a vector autoregressive (VAR) model that includes real GDP, Industrial Production, and Capacity Utilization. Figure 2 displays the BN cycles based on a VAR(12) model, along with the Capacity Utilization series. The two cycles are similar to those in Figure 1, although there seems to be somewhat less of a correspondence between the cycles. The OLS regression results for the relationship between the cycles are as follows:

$$y_t - BN_{y,t} = -0.02 + 0.35(ip_t - BN_{ip,t}) + e_t, R^2 = 0.33. \quad (12)$$

(0.10) (0.04)

As before, the estimated relationship is positive, but weaker than that implied by the cycles based on the AR(12) models.

Given that the VAR(12) models nests the univariate AR(12) models and includes Capacity Utilization, the obvious question is whether there is still any evidence for endogeneity using the BN cycles based on the VAR(12) model and using Capacity Utilization as an instrument. The IV regression results are given as follows:

$$y_t - BN_{y,t} = -0.06 + 0.96(ip_t - BN_{ip,t}) - 0.72 \hat{u}_t + e_t, R^2 = 0.51. \quad (13)$$

(0.09) (0.10) (0.10)

where $\hat{u}_t = ip_t - BN_{ip,t} - 73.30 + 0.17cu_t$. The results are more striking than before.

(15.52) (0.04)

Notably, the t -statistic for the Hausman test of endogeneity is -7.20 , which is significant

at better than the 1% level. Also, the implied relationship between the cycles is much larger than the OLS estimate and close to one-for-one.

While the multivariate results provide an even clearer indication of measurement error and the *BN-as-estimate* interpretation, it may be that they reflect in-sample overfitting that plagues highly-parameterized forecasting models. Indeed, most model selection procedures would favour lower-order ARMA models for the first-differences of real GDP and Industrial Production. Thus, to continue my check on the robustness of the main results, I also consider AR(2) models. Figure 3 displays the BN cycles for real GDP and Industrial Production based on AR(2) models for the first differences, along with the first differences of the Capacity Utilization series. These BN cycles display much less persistence and amplitude than those in Figures 1 and 2. Also, consistent with the results in the original paper by Beveridge and Nelson (1981), the cycles, defined as the difference between the series and the BN trend, are typically positive during NBER recessions, reflecting the positive momentum structure implicit in the AR(2) models. In this case, the OLS results for the relationship between cycles are given as follows:

$$y_t - BN_{y,t} = \underset{(0.04)}{0.00} + \underset{(0.03)}{0.39}(ip_t - BN_{ip,t}) + e_t, R^2 = 0.56. \quad (14)$$

As before, the results suggest a positive relationship between the transitory components and the relationship is far below one-for-one.

Interestingly, the IV results for the cycles based on AR(2) models and using the Capacity Utilization series as an instrument are considerably weaker than before:

$$y_t - BN_{y,t} = \underset{(0.04)}{0.00} + \underset{(0.14)}{0.52}(ip_t - BN_{ip,t}) - \underset{(0.15)}{0.13}\hat{u}_t + e_t, R^2 = 0.56. \quad (15)$$

where $\hat{u}_t = ip_t - BN_{ip,t} - \underset{(8.26)}{21.01} + \underset{(0.02)}{0.05}cu_t$ is the residual based on a first-stage regression of the BN cycle for Industrial Production on the Capacity Utilization series. In this case, the

t -statistic of -0.87 for the Hausman test is not significant at even the 10% level, although the IV estimates again imply a stronger relationship between the transitory components than the OLS estimates.

However, before declaring the *BN-as-definition* interpretation to be vindicated, it should be noted that the Capacity Utilization series appears to be a poor instrument for the transitory component in Industrial Production. The sample correlation with the BN cycle based on the AR(2) model is -20% . Meanwhile, if the AR(2) model is appropriate for the first differences of the Industrial Production series, it may be inappropriate to think of the Capacity Utilization series as stationary, despite the unit root test results. From Figure 3, it is clear that the first differences of the Capacity Utilization series are stationary and highly (negatively) correlated with the BN cycle for Industrial Production. The sample correlation is -90% . Thus, I also consider the first differences of the Capacity Utilization series as an instrument, with the following results:

$$y_t - BN_{y,t} = 0.00 + 0.45(ip_t - BN_{ip,t}) - 0.33\hat{u}_t + e_t, R^2 = 0.61, \quad (16)$$

(0.03)
(0.03)
(0.08)

where $\hat{u}_t = ip_t - BN_{ip,t} + 0.01 + 0.62\Delta cu_t$. In this case, the t -statistic of -4.13 for the Hausman test is significant at the 1% level.⁶ As in previous cases, accounting for endogeneity produces a larger estimate of the relationship between transitory components than OLS, although the magnitude of the increase is less than before.

As a final robustness check, I consider whether the inclusion of information inherent in the Capacity Utilization series is actually identifying a structural relationship, instead of accounting for measurement error in estimating a reduced-form relationship between transitory components. To examine this possibility, I consider the reduced-form

⁶ This finding of endogeneity is robust for BN cycles based on lower-order VAR models too.

relationship between the first differences of the natural logarithms of real GDP and Industrial Production.⁷ The OLS results for this relationship are given as follows:

$$\Delta y_t = 0.49 + 0.42 \Delta ip_t + e_t, R^2 = 0.59. \quad (17)$$

(0.05) (0.03)

Then, I consider the first differences of the Capacity Utilization series as an instrument and obtain the following results for IV estimation:

$$\Delta y_t = 0.48 + 0.43 \Delta ip_t - 0.10 \hat{u}_t + e_t, R^2 = 0.59, \quad (18)$$

(0.05) (0.03) (0.10)

where $\hat{u}_t = \Delta ip_t - 0.67 - 0.80 \Delta cu_t$. In this case, the t -statistic of -1.00 for the Hausman test means that I cannot reject the null of no endogeneity at even the 10% level. This result is telling because, if the rejection in (16) were driven by structural identification instead of measurement error, this would presumably show up as a rejection in (18) given that the regressions are very similar in every dimension except that there is presumably much less measurement error in the real GDP and Industrial Production data than in their BN cycles.

6. Conclusions

Given a particular autocovariance structure for the growth rate of an integrated time series, there can be multiple state-space representations for the permanent and transitory components of the level of the series.⁸ In one case, the time series is subject to permanent and transitory innovations with imperfect correlation and the permanent and

⁷ The relationship between the first differences mixes the underlying relationships between changes in permanent components and changes in transitory components. However, as long as there is a stable relationship between the relative importance of permanent and transitory components, there should be a stable relationship between the first differences.

⁸ Beyond the two representations considered here, there can be additional representations if more general specifications for the permanent component are considered. See Blanchard and Quah (1989), Quah (1992), and Proietti (1995, 2006).

transitory components are unobservable. In the other case, the time series is subject to observable shocks only, directly implying that the permanent and transitory components are observable. In a univariate setting, it is an identification issue as to which representation should be assumed. However, while both representations correspond to identical inferences about the variability of the permanent component, they have very different implications in terms of the uncertainty about the measure of the permanent component. Meanwhile, the possibility of measurement error suggests that the two interpretations are testable in a multivariate setting. In particular, Monte Carlo analysis shows that instrumental variables analysis can be used to detect the presence or absence of errors-in-variables, even if the instrument is included in the forecasting model used in the Beveridge-Nelson decomposition. An application of the instrumental variables analysis to U.S. real GDP, Industrial Production, and Capacity Utilization in Manufacturing provides support for the practical relevance of the *BN-as-estimate* interpretation for macroeconomic data.

References

- Anderson, H.M., C.N. Low, and R. 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 the 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.
- Blanchard, O. and D. Quah, 1989, The dynamic effects of aggregate demand and supply disturbances, *American Economic Review* 79, 655-73.
- Clarida, R. H. and M.P. Taylor, 2003, Nonlinear permanent-temporary decompositions in macroeconomics and finance, *Economic Journal* 113, C125-C139.
- Harvey, A. C. and S. J. Koopman, 2000, Signal extraction and the formulation of unobserved components models, *Econometrics Journal* 3, 84-107.
- Hausman, J., 1978, Specification tests in econometrics, *Econometrica* 46, 1251-1271.
- Maravall, A., 1995, Unobserved components in economic time series, in: H. Pesaran and M.R. Wickens, eds., *Handbook of Applied Econometrics Vol. 1* (Blackwell Press).
- Morley, J., 2002, A state-space approach to calculating the Beveridge-Nelson decomposition, *Economics Letters* 75, 123-127.
- Morley, J.C., Nelson, C.R. and E. Zivot, 2003, Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different? *The Review of Economics and Statistics* 85, 235-243.
- Oh, K.H., E. Zivot, and D. Creal, 2006, The Relationship between the Beveridge-Nelson decomposition and Unobserved Components Models with Correlated Shocks, Working Paper, University of Washington.
- Proietti, T., 1995, The Beveridge-Nelson decomposition: Properties and extensions, *Journal of the Italian Statistical Society* 4, 101-124.
- Proietti, T., 2006, Trend-cycle decompositions with correlated components, *Econometric Reviews* 25, 61-84.
- Proietti, T., and A. Harvey, 2000, A Beveridge-Nelson smoother, *Economics Letters* 67, 139-146.
- Quah, D., 1992, The relative importance of permanent and transitory components: Identification and some theoretical bounds, *Econometrica* 60, 107-118.
- Watson, M.W., 1986, Univariate detrending methods with stochastic trends, *Journal of Monetary Economics* 18, 49-75.

Table 1
Monte Carlo Analysis

	True Value	Sample Size	Estimator	Mean (Std. Dev.) of Estimator for Univariate BN	Mean (Std. Dev.) of Estimator for Multivariate BN
<i>A. DGP #1 - BN is estimate</i>					
Std. Dev. of Permanent Innovations	1	200	OLS	1.00 (0.05)	1.00 (0.05)
		1000	OLS	1.00 (0.02)	1.00 (0.02)
Std. Dev. of Transitory Components	2.15	200	OLS	1.71 (0.17)	1.74/1.81 (0.18)
		1000	OLS	1.72 (0.08)	1.75/1.82 (0.08)
Slope of Relationship between Transitory Components	0.5	200	OLS	0.39 (0.13)	0.57 (0.11)
			IV	0.50 (0.23)	0.50 (0.12)
		1000	OLS	0.39 (0.06)	0.57 (0.05)
			IV	0.50 (0.10)	0.50 (0.05)
<i>B. DGP #2 - BN is true value</i>					
Std. Dev. of Permanent Innovations	1	200	OLS	1.00 (0.05)	1.00 (0.05)
		1000	OLS	1.00 (0.02)	1.00 (0.02)
Std. Dev. of Transitory Components	0.45	200	OLS	0.45 (0.04)	0.45 (0.04)
		1000	OLS	0.45 (0.02)	0.45 (0.02)
Slope of Relationship between Transitory Components	0.5	200	OLS	0.50 (0.12)	0.50 (0.12)
			IV	0.50 (0.16)	0.05 (0.16)
		1000	OLS	0.50 (0.05)	0.50 (0.05)
			IV	0.50 (0.07)	0.50 (0.07)

Notes: Each Monte Carlo experiment consists of 10,000 simulations. For each experiment, three series were generated for the specified sample sizes, with the first two series generated from either (7) or (8) and the third series generated from (9). Then, the BN decomposition was calculated for the first and second series given known parameters. “Univariate BN” and “Multivariate BN” denote whether the BN decomposition was calculated given a univariate forecasting model or a multivariate forecasting model that includes all three variables. Experiment results separated by a slash (“/”) correspond to different means/standard deviations for the first and the second series.

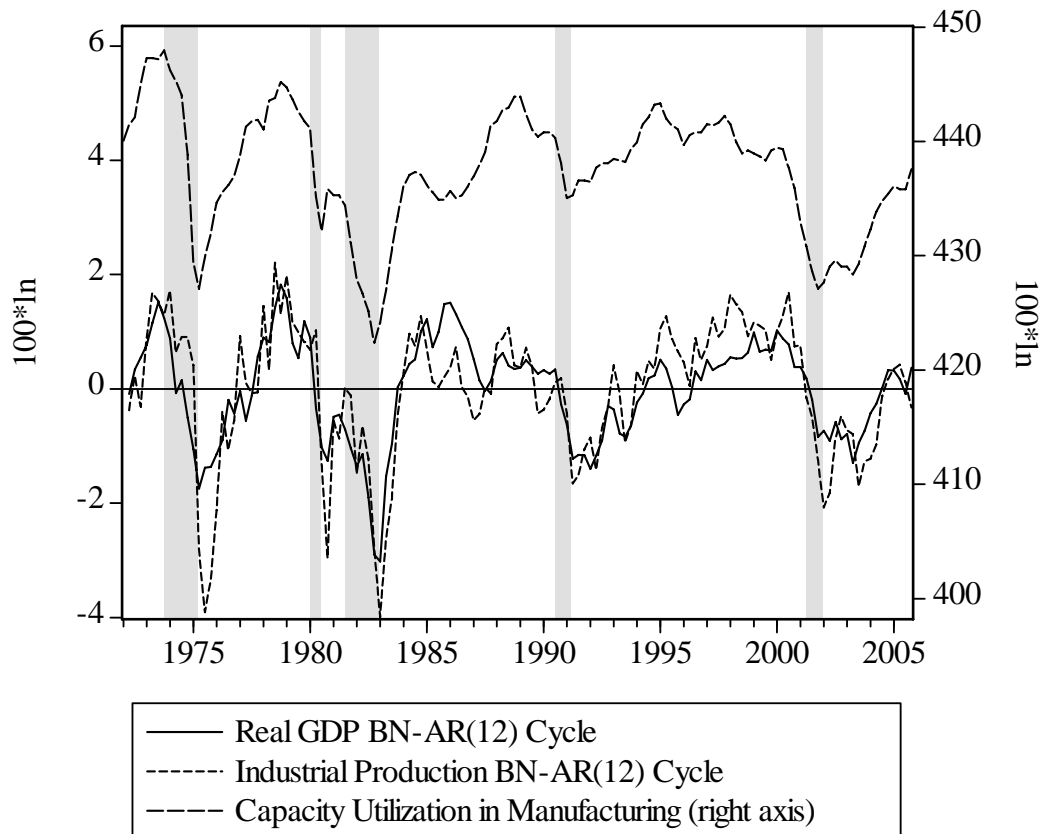


Fig. 1
 Beveridge-Nelson Cycles for U.S. real GDP and Industrial Production Based on AR(2) Models and the Change in Capacity Utilization in Manufacturing (NBER Recessions Shaded)

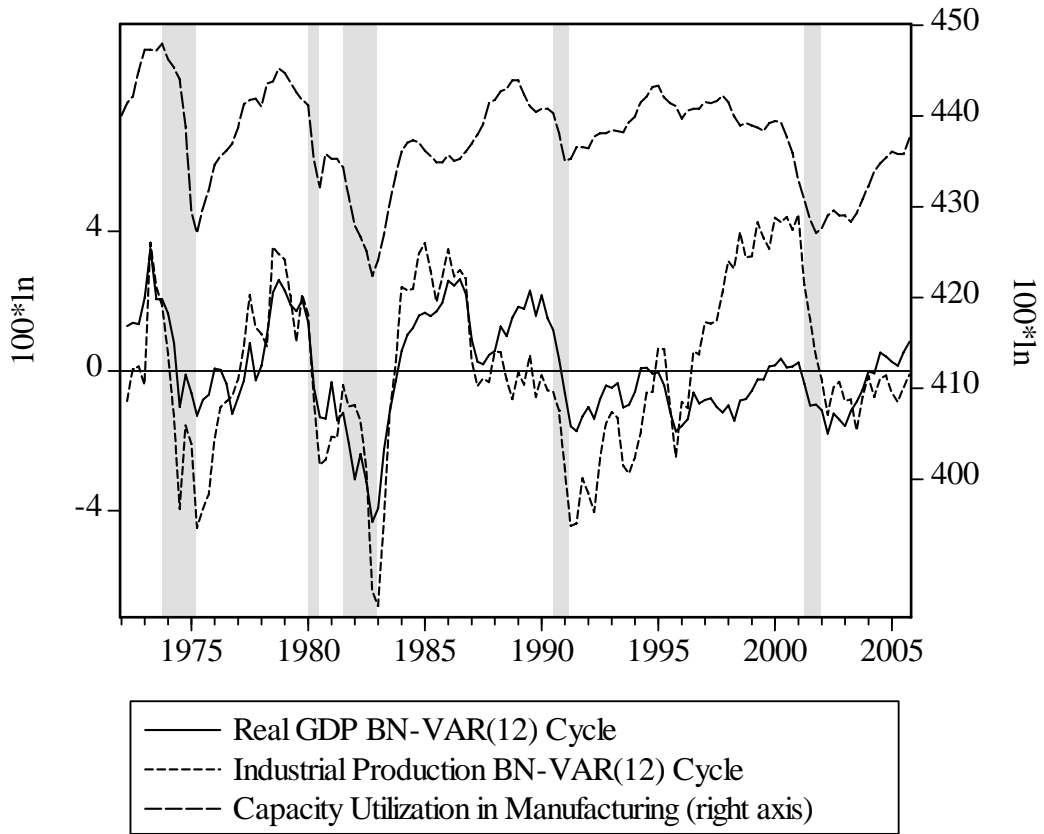


Fig. 2
 Beveridge-Nelson Cycles for U.S. Real GDP and Industrial Production Based on
 VAR(12) Model and Capacity Utilization in Manufacturing (NBER Recessions Shaded)

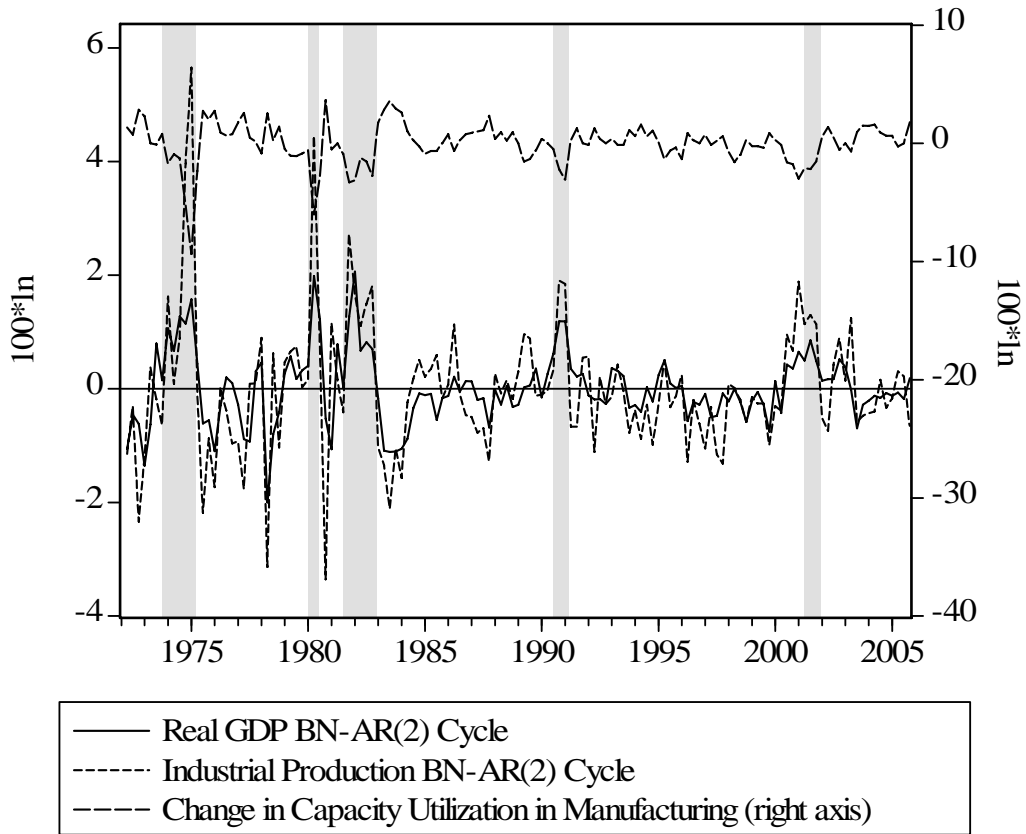


Fig. 3
 Beveridge-Nelson Cycles for U.S. real GDP and Industrial Production Based on AR(2) Models and the Change in Capacity Utilization in Manufacturing (NBER Recessions Shaded)