

"This is the peer reviewed version of the following article: [GRANT, A. L. and CHAN, J. C.C. (2017), A Bayesian Model Comparison for Trend-Cycle Decompositions of Output. Journal of Money, Credit and Banking, 49: 525–552. doi:10.1111/jmcb.12388.], which has been published in final form at [doi: 0.1002/hec.3276] This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#)."

# A Bayesian Model Comparison for Trend-Cycle Decompositions of Output

Angelia L. Grant

Research School of Economics,  
Australian National University

Joshua C.C. Chan\*

Research School of Economics,  
Australian National University

June 2016

## Abstract

We compare a number of widely used trend-cycle decompositions of output in a formal Bayesian model comparison exercise. This is motivated by the often markedly different results from these decompositions—different decompositions have broad implications for the relative importance of real versus nominal shocks in explaining variations in output. Using US quarterly real GDP, we find that the overall best model is an unobserved components model with two features: 1) a nonzero correlation between trend and cycle innovations; 2) a break in trend output growth in 2007. The annualized trend output growth decreases from about 3.4% to 1.2%–1.5% after the break. The results also indicate that real shocks are more important than nominal shocks. The slowdown in trend output growth is robust when we expand the set of models to include bivariate unobserved components models.

Keywords: Bayesian model comparison, unobserved components, structural break, business cycle

JEL classification: C11, C52, E32

---

\*Joshua Chan would like to acknowledge financial support by the Australian Research Council via a Discovery Early Career Researcher Award (DE150100795).

# 1 Introduction

The decomposition of output into its trend and cyclical components is an important theoretical and empirical problem in the study of macroeconomic fluctuations, business cycles, and monetary and fiscal policy. Important early contributions to the literature include the Beveridge and Nelson (1981) decomposition using an unrestricted ARIMA model and the unobserved components models of Harvey (1985), Watson (1986) and Clark (1987). However, these two widely used trend-cycle decompositions yield markedly different results. The Beveridge-Nelson decomposition attributes most of the variance in output to the variation in trend—the cyclical component is small in amplitude and noisy. In contrast, the cyclical components from the unobserved components models are typically large, highly persistent and account for most of the variation in output.

These apparently conflicting results are reconciled in an important paper by Morley, Nelson, and Zivot (2003). They demonstrate that the difference is entirely due to one restriction imposed in the unobserved components model: the innovations to the trend and cycle are assumed to be uncorrelated. When this restriction is relaxed, they find that the two trend-cycle decompositions are identical (see also Morley, 2011). As a result, both estimation methods imply that real or permanent shocks are important and that cycles are small and noisy, and bear little resemblance to the business cycle chronology dated by the National Bureau of Economic Research (NBER).

However, Perron and Wada (2009) argue that these features of the cycles are artifacts that arise from the neglect of a structural break in trend output growth. When a break in growth is allowed for, the cycle estimates are substantially more persistent and accord well with the NBER chronology. Importantly, their preferred model is one with a deterministic trend, which implies all the variation in output can be attributed to innovations to the cyclical component—i.e., real shocks are unimportant.

This brief overview of the literature underscores the sensitivity of cycle estimates to model choice, with differences in one or two key parameters giving starkly different trend-cycle decompositions. Other model specification choices—such as the dating of a break, which is fixed in 1973Q1 in Perron and Wada (2009)—can also be framed as a model selection problem. The model that is used has broad implications for the conclusions drawn about the relative importance of real versus nominal shocks in explaining variations in output. Hence, it is important to perform a model comparison exercise to select the best model (or average different model estimates across models), but it is seldom done in practice.

We take up this task and use a Bayesian model comparison framework to assess the adequacy of a variety of nested and nonnested models for decomposing US quarterly real GDP. In particular, we compare the unobserved components models of Clark (1987), Morley et al. (2003) and Perron and Wada (2009), and deterministic trend models (with or without a break). By treating each fixed break date as a separate model, we are able to date any change that might have occurred in output growth. There are two closely related papers. The first is Morley and Piger (2012), who consider model comparison

using information criteria and model averaging using an asymptotic approximation to the Bayes factors. In contrast, we provide exact computations of the Bayes factors for either model comparison or model averaging. Consequently, there are notable differences between the results of the two approaches. The second is Luo and Startz (2014), who compare three models using Bayesian methods: the model of Morley et al. (2003) with or without a break and a model with an unknown break date. Their preferred model is the model of Morley et al. (2003) with a break at 2006Q1. We consider a larger class of models in our model comparison exercise. In addition, despite using similar techniques, there are a few noticeable differences in our results, which are detailed below.

In order to compare the various trend-cycle decompositions, we develop new Bayesian estimation techniques using Markov chain Monte Carlo (MCMC) methods (see, e.g., Koop and Korobilis, 2010, for a general introduction of Bayesian methods for empirical macroeconomics). A key novel feature of our approach is that it builds upon the band and sparse matrix algorithms for state space models developed in Chan and Jeliaskov (2009), McCausland, Miller, and Pelletier (2011) and Chan (2013), which are shown to be more efficient than the conventional Kalman filter-based algorithms. In addition, due to the modular nature of MCMC algorithms, it is relatively straightforward to extend the estimation methods to regime-switching models or models with non-Gaussian innovations.

Our main results can be summarized as follows. First, allowing for a nonzero correlation between the permanent and transitory shocks substantially improves model fit. This is in line with the finding in Morley et al. (2003), who report an estimate of  $-0.9$  for the correlation parameter, but in contrast to the posterior mean of  $0.18$  reported in Luo and Startz (2014) under their preferred model. Second, the correlated unobserved components model of Morley et al. (2003) dominates any deterministic trend models with or without a break in trend growth.

Among correlated unobserved components models, a break in trend growth is likely to have occurred in 2007. The annual trend output growth is estimated to have decreased from about  $3.4\%$  to  $1.2\%$ – $1.5\%$  after the break. The latter figures are close to the forecast of US potential output growth from 2007 to 2032 given in Gordon (2014).<sup>1</sup> This is in contrast to the even lower trend growth estimate of  $0.89\%$  after 2006 reported in Luo and Startz (2014). Our result of markedly slower trend output growth after 2007 is robust when we expand the set of models to include bivariate correlated unobserved components models of GDP and inflation, and GDP and the unemployment rate.

The remainder of this article is organized as follows. Section 2 first discusses a variety of trend-cycle decompositions of output that are widely used in the literature. It then outlines the Bayesian estimation methods used to fit these models. In Section 3 we give an overview of Bayesian model comparison using the marginal likelihood, as well as an importance sampling approach to compute this quantity. Then, in Section 4 we compare the performance of the various models in fitting US real GDP. Trend-cycle decompositions and other parameter estimates for selected models are also reported. We then investigate

---

<sup>1</sup>See footnote 4 in Gordon (2014), which gives a forecast of potential output growth of  $1.5\%$  to  $1.55\%$ .

the robustness of the decrease in the trend output growth after 2007 by estimating four additional bivariate correlated unobserved components models. These bivariate models give very similar trend output growth estimates, confirming our baseline results. Lastly, Section 5 concludes and briefly discusses some future research directions.

## 2 Trend-Cycle Decomposition using UC Models

In this section we discuss a variety of trend-cycle decompositions of output based on unobserved components models and the Bayesian estimation methods used to fit the models (see, e.g., Koop, Poirier, and Tobias, 2007, for an introduction of Bayesian computations in econometrics).

The estimation in the literature on trend-cycle decompositions using unobserved components models typically uses the maximum likelihood method. However, one issue with this approach is the so-called “pile-up” problem, whereby the maximum likelihood estimates take values at the boundary of the parameter space. This can occur, for example, if a variance parameter is estimated to be zero. The “pile-up” problem makes inference more difficult, as the usual asymptotic properties of the maximum likelihood estimator no longer hold.

Moreover, trend-cycle decompositions in the literature are typically obtained conditional on the maximum likelihood estimates. As such, there is no accounting for parameter uncertainty. However, given that trend-cycle decompositions can be sensitive to the values of a few key parameters, it is crucial to take parameter uncertainty into account. The models in Clark (1987) and Morley et al. (2003) highlight the importance of parameter uncertainty—the models differ in the value of only one parameter, but the trend-cycle decompositions from the two models are drastically different.

We adopt the Bayesian approach in which inference is based on the joint posterior distribution of the parameters. Using the posterior mean as the point estimate avoids the “pile up” problem—as long as a nondogmatic prior is used, by construction the posterior mean is away from the boundary of the parameter space. In addition, the trend-cycle decomposition is constructed by averaging the parameter values with respect to the joint posterior distribution of the parameters—hence, the decomposition does not depend on a particular set of parameter values. Furthermore, the Bayesian approach facilitates comparing nonnested models, which is discussed in more detail in Section 3.

### 2.1 Competing Models

The trend-cycle decomposition of aggregate output is motivated by the idea that it can be usefully viewed as the sum of two separate components: a nonstationary component that represents the long-term trend and a transitory deviation from the trend. More

specifically, let  $y_t$  denote (100 times) the log of real GDP. Then  $y_t$  can be decomposed as

$$y_t = \tau_t + c_t, \quad (1)$$

where  $\tau_t$  is the trend and  $c_t$  is the stationary, cyclical component. The nonstationary trend  $\tau_t$  is modeled as a random walk with drift, whereas the cyclical component  $c_t$  is modeled as a zero mean stationary AR( $p$ ) process:

$$\tau_t = \mu_1 + \tau_{t-1} + u_t^\tau, \quad (2)$$

$$c_t = \phi_1 c_{t-1} + \dots + \phi_p c_{t-p} + u_t^c, \quad (3)$$

where the initial  $\tau_0$  is treated as a parameter to be estimated and for simplicity we assume  $c_{1-p} = \dots = c_0 = 0$ . Note that the drift  $\mu_1$  can be interpreted as the growth rate of trend GDP. Following Morley et al. (2003), we set  $p = 2$  and assume innovations  $u_t^c$  and  $u_t^\tau$  are jointly normal

$$\begin{pmatrix} u_t^c \\ u_t^\tau \end{pmatrix} \sim \mathcal{N} \left( \mathbf{0}, \begin{pmatrix} \sigma_c^2 & \rho \sigma_c \sigma_\tau \\ \rho \sigma_c \sigma_\tau & \sigma_\tau^2 \end{pmatrix} \right).$$

We denote the model in (1)–(3) as UCUR. This model allows a nonzero correlation between the innovations  $u_t^c$  and  $u_t^\tau$ . Hence, it includes the model of Clark (1987) as a special case with  $\rho = 0$ ; this restricted model is denoted as UC0.

Perron and Wada (2009) point out that the trend-cycle decomposition might be sensitive to how the trend is modeled. In particular, they show that when a break is allowed for, the estimates of the cyclical component become larger in magnitude and are more persistent. Hence, we consider specifications with a break in the drift. Specifically, consider replacing (2) by a more general specification

$$\tau_t = \mu_1 1(t < t_0) + \mu_2 1(t \geq t_0) + \tau_{t-1} + u_t^\tau, \quad (4)$$

where  $1(A)$  is the indicator function that takes the value 1 if the condition  $A$  is true and 0 otherwise, and  $t_0$  is a known break point. In other words, the stochastic trend  $\tau_t$  has a growth rate of  $\mu_1$  before the break  $t_0$  and a growth rate of  $\mu_2$  after the break. In the model comparison exercise, we date the break point by comparing models with different  $t_0$ . We denote the UCUR model with a break at time  $t_0$  as UCUR- $t_0$ .

Lastly, we consider a set of models with deterministic trends. This is motivated by the findings in Perron and Wada (2009), where the preferred model is UC0 with a break point in 1973Q1. However, the variance of the innovation to the trend,  $\sigma_\tau^2$ , is estimated to be zero, which is outside of the parameter space—the variance  $\sigma_\tau^2$  should be positive. To circumvent this difficulty, we consider instead the following deterministic trend

$$\tau_t = \mu_1 1(t < t_0) + \mu_2 1(t \geq t_0) + \tau_{t-1}. \quad (5)$$

The cyclical component  $c_t$  is modeled as in (3), with  $u_t^c \sim \mathcal{N}(0, \sigma_c^2)$ . This model is denoted as DT- $t_0$ . We also consider a version without a break, which is denoted as DT.

## 2.2 Bayesian Estimation

In this section we outline the Bayesian estimation methods used to fit the UCUR model given in (1)–(3). More specifically, we develop a Markov sampler to obtain draws from the posterior distribution under the UCUR model. The other unobserved components models can be estimated similarly, and we leave the technical details to Appendix A. A key novel feature of our approach is that it builds upon the band and sparse matrix algorithms developed in Chan and Jeliazkov (2009) and Chan (2013). It is shown in McCausland et al. (2011) that this approach is more efficient compared to the conventional Kalman filter-based algorithms.

We assume proper but relatively noninformative priors for the model parameters  $\boldsymbol{\phi} = (\phi_1, \phi_2)'$ ,  $\sigma_c^2, \sigma_\tau^2, \rho, \mu_1$  and  $\tau_0$ . In particular, we consider a uniform prior on  $(-1, 1)$  for  $\rho$ , and identical uniform priors on  $(0, 3)$  for  $\sigma_c^2$  and  $\sigma_\tau^2$ . The details of the priors are given in Appendix A. Since the marginal likelihood can be sensitive to prior specification, we use exactly the same priors for common parameters across models. For notational convenience, stack  $\mathbf{y} = (y_1, \dots, y_T)'$ , and similarly define  $\boldsymbol{\tau}$ ,  $\mathbf{c}$ ,  $\mathbf{u}^c$  and  $\mathbf{u}^\tau$ . Then, posterior draws can be obtained by sequentially sampling from the following densities: 1.  $p(\boldsymbol{\tau} | \mathbf{y}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0)$ ; 2.  $p(\boldsymbol{\phi} | \mathbf{y}, \boldsymbol{\tau}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0)$ ; 3.  $p(\sigma_c^2 | \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_\tau^2, \rho, \mu_1, \tau_0)$ ; 4.  $p(\sigma_\tau^2 | \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \rho, \mu_1, \tau_0)$ ; 5.  $p(\rho | \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \mu_1, \tau_0)$ ; and 6.  $p(\tau_0, \mu_1 | \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho)$ . Here we discuss how Step 1 can be implemented; the details of other steps are given in Appendix A.

First we write the system (1)–(3) in the following matrix form:

$$\begin{aligned}\mathbf{y} &= \boldsymbol{\tau} + \mathbf{c}, \\ \mathbf{H}_\phi \mathbf{c} &= \mathbf{u}^c, \\ \mathbf{H}\boldsymbol{\tau} &= \tilde{\boldsymbol{\alpha}} + \mathbf{u}^\tau,\end{aligned}$$

where  $\tilde{\boldsymbol{\alpha}} = (\mu_1 + \tau_0, \mu_1, \dots, \mu_1)'$  and

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & -1 & 1 \end{pmatrix}, \quad \mathbf{H}_\phi = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\phi_1 & 1 & 0 & 0 & \cdots & 0 \\ -\phi_2 & -\phi_1 & 1 & 0 & \cdots & 0 \\ 0 & -\phi_2 & -\phi_1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\phi_2 & -\phi_1 & 1 \end{pmatrix}.$$

Note that both  $\mathbf{H}$  and  $\mathbf{H}_\phi$  are band matrices with only a few nonzero elements arranged along the main diagonal. Further, since both are square matrices with unit determinant, they are invertible. Hence, given  $\boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho$  and  $\tau_0$ , we have

$$\begin{pmatrix} \mathbf{c} \\ \boldsymbol{\tau} \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\alpha} \end{pmatrix}, \begin{pmatrix} \sigma_c^2 (\mathbf{H}'_\phi \mathbf{H}_\phi)^{-1} & \rho \sigma_c \sigma_\tau (\mathbf{H}' \mathbf{H}_\phi)^{-1} \\ \rho \sigma_c \sigma_\tau (\mathbf{H}'_\phi \mathbf{H})^{-1} & \sigma_\tau^2 (\mathbf{H}' \mathbf{H})^{-1} \end{pmatrix} \right),$$

where  $\boldsymbol{\alpha} = \mathbf{H}^{-1} \tilde{\boldsymbol{\alpha}}$ . Using the properties of the Gaussian distributions (see, e.g., Kroese and Chan, 2014, Chapter 3.6), the marginal distribution of  $\boldsymbol{\tau}$  (unconditional on  $\mathbf{c}$ ) is

$$(\boldsymbol{\tau} | \sigma_\tau^2, \mu_1, \tau_0) \sim \mathcal{N}(\boldsymbol{\alpha}, \sigma_\tau^2 (\mathbf{H}' \mathbf{H})^{-1}),$$

and the conditional distribution of  $\mathbf{y}$  given  $\boldsymbol{\tau}$  and other parameters is given by

$$(\mathbf{y} \mid \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0) \sim \mathcal{N}(\mathbf{H}_\phi^{-1} \mathbf{a} + \mathbf{H}_\phi^{-1} \mathbf{B} \boldsymbol{\tau}, (1 - \rho^2) \sigma_c^2 (\mathbf{H}_\phi' \mathbf{H}_\phi)^{-1}),$$

where

$$\mathbf{a} = -\frac{\rho \sigma_c}{\sigma_\tau} \mathbf{H} \boldsymbol{\alpha}, \quad \mathbf{B} = \mathbf{H}_\phi + \frac{\rho \sigma_c}{\sigma_\tau} \mathbf{H}.$$

Therefore, the prior density of  $\boldsymbol{\tau}$  and the *conditional* likelihood are given by

$$p(\boldsymbol{\tau} \mid \sigma_\tau^2, \mu_1, \tau_0) = (2\pi \sigma_\tau^2)^{-\frac{T}{2}} e^{-\frac{1}{2\sigma_\tau^2} (\boldsymbol{\tau} - \boldsymbol{\alpha})' \mathbf{H}' \mathbf{H} (\boldsymbol{\tau} - \boldsymbol{\alpha})} \quad (6)$$

$$p(\mathbf{y} \mid \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0) = (2\pi \sigma_c^2 (1 - \rho^2))^{-\frac{T}{2}} e^{-\frac{1}{2(1 - \rho^2)\sigma_c^2} (\mathbf{H}_\phi \mathbf{y} - \mathbf{a} - \mathbf{B} \boldsymbol{\tau})' (\mathbf{H}_\phi \mathbf{y} - \mathbf{a} - \mathbf{B} \boldsymbol{\tau})}. \quad (7)$$

Then, by standard linear regression results (see, e.g., Kroese and Chan, 2014, pp. 237-240), we have

$$(\boldsymbol{\tau} \mid \mathbf{y}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \mu_1, \rho, \tau_0) \sim \mathcal{N}(\hat{\boldsymbol{\tau}}, \mathbf{K}_\tau^{-1}),$$

where

$$\mathbf{K}_\tau = \frac{1}{\sigma_\tau^2} \mathbf{H}' \mathbf{H} + \frac{1}{(1 - \rho^2) \sigma_c^2} \mathbf{B}' \mathbf{B}, \quad \hat{\boldsymbol{\tau}} = \mathbf{K}_\tau^{-1} \left( \frac{1}{\sigma_\tau^2} \mathbf{H}' \mathbf{H} \boldsymbol{\alpha} + \frac{1}{(1 - \rho^2) \sigma_c^2} \mathbf{B}' (\mathbf{H}_\phi \mathbf{y} - \mathbf{a}) \right).$$

Since  $\mathbf{H}$ ,  $\mathbf{H}_\phi$  and  $\mathbf{B}$  are all band matrices, so is the precision matrix  $\mathbf{K}_\tau$ . As such, the precision sampler of Chan and Jeliazkov (2009) can be used to sample  $\boldsymbol{\tau}$  efficiently. We leave the details of Steps 2-6 to Appendix A.

In addition, this approach allows us to derive an analytical expression of the *integrated* or *observed-data* likelihood  $p(\mathbf{y} \mid \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0)$ , which is a crucial quantity for model comparison. We refer the readers to Appendix B for the exact expression and the derivations. Using this analytical expression, the integrated likelihood can then be evaluated quickly using band matrix routines, which is more efficient than using the Kalman filter.

### 3 Model Comparison via the Marginal Likelihood

In this section, we give an overview of Bayesian model comparison using the marginal likelihood. Then, we outline an importance sampling approach based on the improved cross-entropy method to compute the marginal likelihood.

Suppose we wish to compare a set of possibly nonnested models  $\{M_1, \dots, M_K\}$ . Each model  $M_k$  is formally defined by two components: a likelihood function  $p(\mathbf{y} \mid \boldsymbol{\theta}_k, M_k)$  that depends on the model-specific parameter vector  $\boldsymbol{\theta}_k$  and a prior density  $p(\boldsymbol{\theta}_k \mid M_k)$ . The marginal likelihood under model  $M_k$  is defined as

$$p(\mathbf{y} \mid M_k) = \int p(\mathbf{y} \mid \boldsymbol{\theta}_k, M_k) p(\boldsymbol{\theta}_k \mid M_k) d\boldsymbol{\theta}_k. \quad (8)$$



This marginal likelihood can be interpreted as a density forecast of the data under model  $M_k$  evaluated at the actual observed data  $\mathbf{y}$ . Hence, if the observed data are likely under the model, the associated marginal likelihood would be “large”. Since the marginal likelihood is essentially a density forecast evaluation, it has a built-in penalty for model complexity.

Given two models  $M_i$  and  $M_j$ , if the marginal likelihood under model  $M_i$  is larger than that under  $M_j$ —i.e., the observed data are more likely under model  $M_i$  compared to model  $M_j$ —then it is viewed as evidence in favor of model  $M_i$ . The weight of evidence can be gauged by the *posterior odds ratio* between the two models, which can be written as follows:

$$\frac{\mathbb{P}(M_i | \mathbf{y})}{\mathbb{P}(M_j | \mathbf{y})} = \frac{\mathbb{P}(M_i)}{\mathbb{P}(M_j)} \times \frac{p(\mathbf{y} | M_i)}{p(\mathbf{y} | M_j)},$$

where  $\mathbb{P}(M_i)/\mathbb{P}(M_j)$  is the prior odds ratio and the ratio of the marginal likelihoods  $p(\mathbf{y} | M_i)/p(\mathbf{y} | M_j)$  is called the *Bayes factor* in favor of model  $M_i$  against  $M_j$ . If both models are equally probable *a priori*, i.e., the prior odds ratio is one, the posterior odds ratio between the two models is then equal to the Bayes factor. Then, if, for example,  $\text{BF}_{ij} = 50$ , it implies model  $M_i$  is 50 times more likely than model  $M_j$  given the data. For a more detailed discussion of the Bayes factor, we refer the readers to Koop (2003). Next, we outline an importance sampling method for calculating the marginal likelihoods under the various unobserved components models discussed in the previous section.

The computation of the marginal likelihood is in general nontrivial—the integral in (8) is often high-dimensional and cannot be obtained analytically. Here we adopt an improved version of the classic cross-entropy method to estimate the marginal likelihood. The classic cross-entropy method was originally developed for rare-event simulation by Rubinstein (1997, 1999) using a multi-level procedure to construct the optimal importance sampling density (see also Rubinstein and Kroese, 2004, for a book-length treatment). Chan and Kroese (2012) later show that the optimal importance sampling density can be obtained more accurately in one step using MCMC methods. This new variant is applied in Chan and Eisenstat (2015) for marginal likelihood estimation, which is outlined as follows.

Suppose we wish to estimate  $p(\mathbf{y} | M_k)$ , the marginal likelihood under model  $M_k$ . For notational convenience we drop the model index  $M_k$ , and write the marginal likelihood, likelihood and prior as  $p(\mathbf{y})$ ,  $p(\mathbf{y} | \boldsymbol{\theta})$  and  $p(\boldsymbol{\theta})$ , respectively. The ideal zero-variance importance sampling density for this estimation problem is the posterior density  $p(\boldsymbol{\theta} | \mathbf{y})$ . Unfortunately, this density is only known up to a constant and therefore cannot be used directly in practice. Nevertheless, it provides a good benchmark to obtain a suitable importance sampling density.

The idea is to locate a density that is “close” to the ideal importance sampling density. Operationally, we find the density within a convenient family of distributions such that its Kullback-Leibler divergence—or the cross-entropy distance—to the ideal density is minimized. Once the optimal density, say,  $g(\cdot)$ , is obtained, it is used to construct the

importance sampling estimator:

$$\widehat{p(\mathbf{y})} = \frac{1}{R} \sum_{r=1}^R \frac{p(\mathbf{y} | \boldsymbol{\theta}^{(r)})p(\boldsymbol{\theta}^{(r)})}{g(\boldsymbol{\theta}^{(r)})}, \quad (9)$$

where  $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(R)}$  are draws from the importance sampling density  $g(\boldsymbol{\theta})$ . The main advantage of this importance sampling approach is that it is easy to implement and the numerical standard error of the estimator is readily available. We refer the readers to Chan and Eisenstat (2015) for technical details.

For the unobserved components models discussed in Section 2.1, the likelihood—or more accurately, the *integrated* likelihood or *observed-data* likelihood (the density of the data marginal of the latent states)—can in principle be evaluated using the Kalman filter. Here we adopt a more efficient approach, which is substantially faster than the Kalman filter. Specifically, we first derive analytical expressions for the integrated likelihoods under the various unobserved components models. These expressions are then evaluated using band and sparse matrix routines. The technical details are given in Appendix B.

## 4 Empirical Results

In this section we compare the performance of the various unobserved components models discussed in Section 2 in fitting US real GDP. The main goal of this exercise is to establish the types of model features that are useful in trend-cycle decompositions. For example, does allowing for correlation between permanent and transitory shocks substantially improve model fit? Or is it more important to allow for a break in GDP growth? If yes, when is the break date?

We use US quarterly real GDP from 1947Q1 to 2014Q4 for our analysis, which is sourced from the Federal Reserve Bank of St. Louis economic database. The data are then transformed by taking the logs and multiplying by 100. We first report the model comparison results in Section 4.1. Trend-cycle decompositions and variance decompositions for selected models are reported in Section 4.2. Then, in Section 4.3 we present the trend output growth estimates and investigate the robustness of these estimates by considering additional bivariate models.

### 4.1 Model Comparison Results

All the models are estimated using the new sampling approach based on band matrix routines discussed in Section 2.2 and Appendix A. The marginal likelihoods are computed using the improved cross-entropy method of Chan and Eisenstat (2015), which is outlined in Section 3. Each set of results is based on 100000 posterior draws after a burn-in period of 10000. For computing each marginal likelihood value, we use 50000 importance sampling draws.

We first address the timing of a break in GDP growth. Motivated by the results in Perron and Wada (2009) and Luo and Startz (2014), we consider two classes of models with a break, namely, UCUR- $t_0$  and DT- $t_0$ . For each class of models, we consider 10 possible break dates: every first quarter from 1971Q1 to 1975Q1, and from 2005Q1 to 2009Q1. The latter dates are chosen so that we can determine whether trend GDP growth has changed following the Great Recession. We only consider a break in the first quarter of each year, given that the difference between models with breaks in consecutive quarters is expected to be small. Hence, it might be more useful to think of the break occurring in that year than in that particular quarter.

The model comparison results are reported in Table 1. By comparing UCUR- $t_0$  and DT- $t_0$  for each break date, we conclude that UCUR- $t_0$  uniformly outperforms DT- $t_0$  in fitting the US data. For example, the Bayes factor in favor of UCUR-07 against DT-07 is about 28—if we assume both models are equally likely *a priori*, the former becomes 28 times more likely given the data—indicating strong evidence in favor of UCUR-07. This indicates that the GDP trend is better modeled as a stochastic process rather than a deterministic one. In addition, for both classes of models, a break date in 2007 is most favored by the data among the 10 possible break dates. This is similar to the break date of 2006Q1 identified in Luo and Startz (2014). These results suggest that there seems to be a structural break in trend GDP growth in the early stages of the Great Recession.

Table 1: Log marginal likelihoods of the UCUR- $t_0$  and DT- $t_0$  models with various break dates. Numerical standard errors are in parentheses.

DT-71	DT-72	DT-73	DT-74	DT-75
-368.96	-368.52	-367.95	-367.50	-368.20
(0.008)	(0.010)	(0.005)	(0.009)	(0.019)
DT-05	DT-06	DT-07	DT-08	DT-09
-367.43	-367.55	<b>-367.37</b>	-367.62	-369.86
(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
UCUR-71	UCUR-72	UCUR-73	UCUR-74	UCUR-75
-366.25	-366.03	-365.69	-365.39	-365.99
(0.046)	(0.077)	(0.046)	(0.090)	(0.049)
UCUR-05	UCUR-06	UCUR-07	UCUR-08	UCUR-09
-364.45	-364.29	<b>-364.04</b>	-364.29	-365.60
(0.045)	(0.065)	(0.039)	(0.036)	(0.050)

The above results show that *if* a break is assumed, the break is most likely to have occurred in 2007. Next, we address the question of whether a break is needed by comparing models with and without a break. A related question is whether one break is sufficient. To answer the latter question, we also include a model with two breaks in 1973Q1 and 2007Q1, given the large literature around a break in 1973, which we call UCUR-(73,07). The model comparison results are reported in Table 2.

Among all the models we consider, the overall best model is UCUR-07. For example,

compared to DT-73—the preferred model in Perron and Wada (2009)—the Bayes factor in favor of UCUR-07 is about 50, showing strong support for the latter model. UCUR-07 also compares favorably against the more general UCUR-(73,07), indicating that in the presence of a break in 2007, an additional break in 1973 is not needed. This also gives an example of the built-in penalty for model complexity within the Bayes factor: it prefers the more restricted UCUR-07 when an additional break in 1973 does not substantially improve the model fit.

Table 2: Log marginal likelihoods of competing models with and without a break. Numerical standard errors are in parentheses.

DT	UC0	UCUR	DT-07	UCUR-07	UCUR-(73,07)
-370.63	-370.54	-365.02	-367.37	<b>-364.04</b>	-365.30
(0.004)	(0.030)	(0.026)	(0.007)	(0.039)	(0.087)

Next, we investigate whether allowing for a nonzero correlation between the permanent and transitory shocks substantially improves model fit. This can be done by directly comparing UC0 and UCUR. Indeed, the Bayes factor in favor of UCUR against UC0 is about 250, indicating overwhelming support for the former. This is consistent with the findings in Morley et al. (2003) and Oh, Zivot, and Creal (2008). The former reports the maximum likelihood estimate of the correlation  $\rho$  to be about  $-0.9$  with a relatively small standard error; the latter rejects the null hypothesis that  $\rho = 0$  using the likelihood ratio test. In contrast, using information criteria to compare models, Morley and Piger (2012) find mixed results: Akaike information criterion weakly prefers UCUR but Bayesian information criterion slightly favors UC0.

To conclude, the model comparison results show that it is useful to allow for a nonzero correlation between the permanent and transitory shocks, and a break in trend GDP growth is likely to have occurred in the early stages of the Great Recession.

## 4.2 Trend-Cycle Estimates and Variance Decomposition

In this section we first report the trend and cycle estimates under three models: DT-73, the deterministic trend model with a break in trend growth in 1973; UCUR, the correlated unobserved components model of Morley et al. (2003); and UCUR-07, an extension of UCUR with a break in growth in 2007. Then, we consider whether permanent or transitory shocks are more important under these models. In the next section, we investigate the breaks in trend output growth and address the question of whether trend output growth has substantially slowed in the past decade.

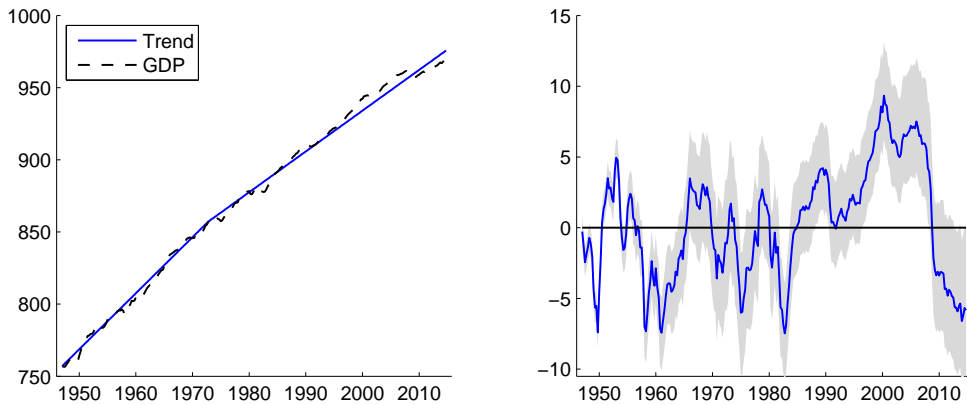


Figure 1: Estimates of trend (left panel) and cycle (right panel) under DT-73. The shaded region represents the 10% and 90% quantiles.

Figure 1 plots the trend and cycle estimates under DT-73.<sup>2</sup> It can be seen in the left panel that there is a kink in the trend in 1973, reflecting a slower estimated trend growth rate after the break date. Under this deterministic trend model, all variation in output is attributed to the cyclical component. Consequently, the cycles are large and persistent. For example, output started to outpace its trend from the early 1990s, and the cyclical component reached a peak in the new millennium. Output was substantially above trend until the Great Recession—since then it has dropped below trend. Interestingly, the output gap seems to have widened since 2010, reflecting the slower growth in GDP compared to the historical trend.

We report the trend and cycle estimates of UCUR and UCUR-07 in Figures 2 and 3. Compared to those of DT-73, the cycle estimates are much smaller in magnitude and less persistent. This highlights the sensitivity of trend and cycle estimates to model specification, and hence the importance of model comparison. The cycle estimates also suggest that output is above trend in the run-up to the Great Recession, even though not to the same magnitude as in DT-73.

Not surprisingly, the cycle estimates of UCUR and UCUR-07 are fairly similar. The only noticeable difference occurs after the Great Recession: the former model suggests that output has been at trend since 2009 and remains there, whereas the latter model indicates that output has been slightly below trend since 2009. Figure 4 plots the trend estimates of the three models from 2005Q1 to 2014Q4.

---

<sup>2</sup>The trend estimates reported in this section are *smoothed* values obtained conditional on the whole sample. More specifically, the point estimates are the posterior means  $\mathbb{E}(\boldsymbol{\tau} | \mathbf{y})$ . The cycle estimates are then given by  $\mathbf{y} - \mathbb{E}(\boldsymbol{\tau} | \mathbf{y})$ .

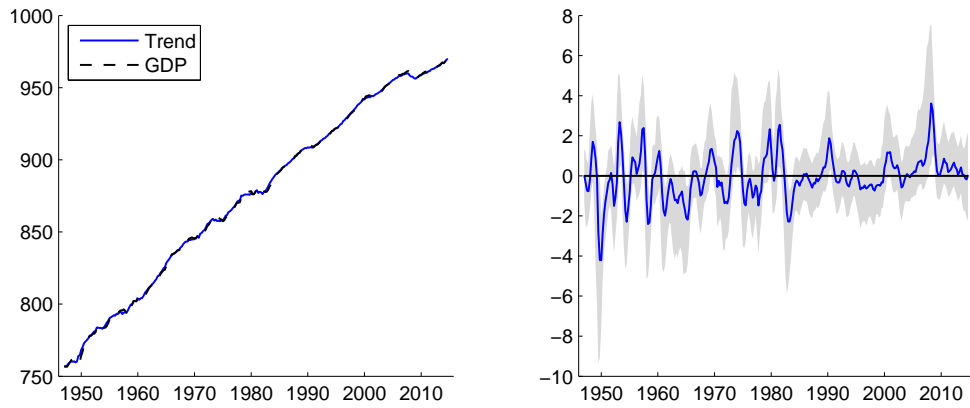


Figure 2: Estimates of trend (left panel) and cycle (right panel) under UCUR. The shaded region represents the 10% and 90% quantiles.

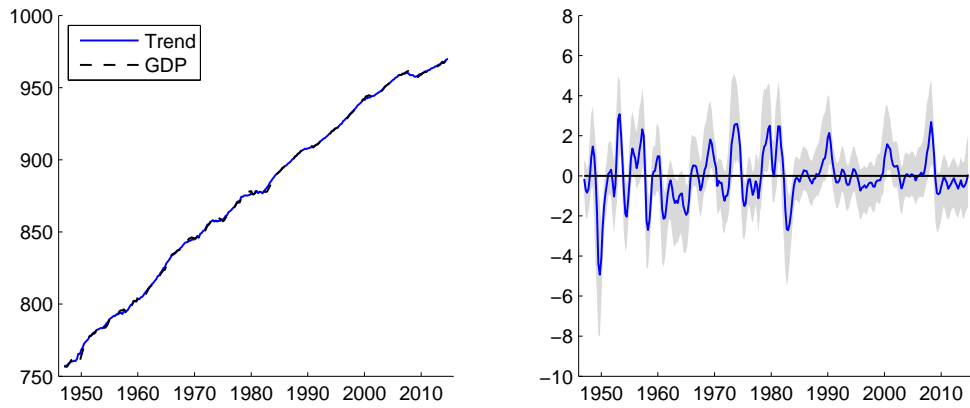


Figure 3: Estimates of trend (left panel) and cycle (right panel) under UCUR-07. The shaded region represents the 10% and 90% quantiles.

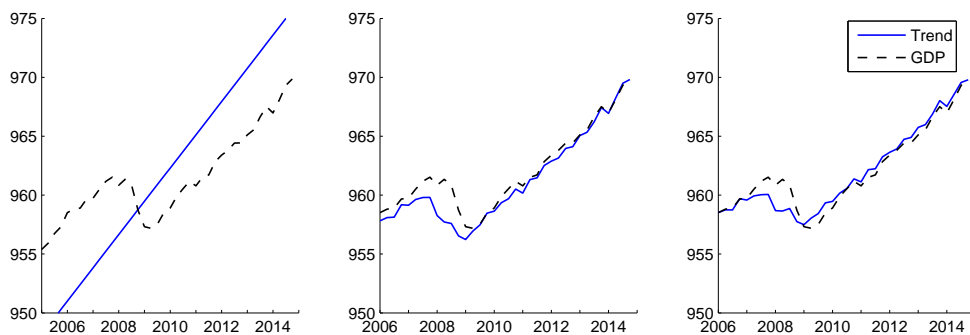


Figure 4: Estimates of trend from 2005Q1 to 2014Q4 under the DT-73 model (left panel), UCUR model (middle panel) and UCUR-07 model (right panel).

In addition to the different trend-cycle decompositions, these models also differ in their conclusions regarding the relative importance of permanent and transitory shocks. As mentioned above, DT-73 has a deterministic trend and therefore all variation of output is due to the transitory shocks. In contrast, both UCUR and UCUR-07 allow us to decompose the variance of output into the portions contributed by the permanent and transitory shocks. Table 3 reports the parameter estimates of the three models. For comparison, we also include results from UCUR-(73,07), a model with two breaks in 1973Q1 and 2007Q1.

Under UCUR-07 the estimates of  $\sigma_\tau^2$  and  $\sigma_c^2$  are 1.42 and 0.9 respectively, giving a ratio of about 1.5. This indicates that permanent shocks are relatively more important compared to transitory shocks; similar conclusions can be drawn from the results of UCUR and UCUR-(73,07). These results are in line with the conclusion in Morley et al. (2003). However, since the variance parameters are not precisely estimated, we need to take account of parameter uncertainty.

Table 3: Estimated posterior means under DT-73, UCUR, UCUR-07 and UCUR-(73,07). Numerical standard errors are in parentheses.

	DT-73	UCUR	UCUR-07	UCUR-(73,07)
$\phi_1$	1.34 (0.057)	0.95 (0.343)	1.10 (0.361)	1.17 (0.338)
$\phi_2$	-0.37 (0.057)	-0.36 (0.184)	-0.44 (0.180)	-0.46 (0.169)
$\sigma_c^2$	0.79 (0.069)	1.12 (0.553)	0.90 (0.486)	0.84 (0.478)
$\sigma_\tau^2$	–	1.85 (0.494)	1.42 (0.593)	1.17 (0.695)
$\rho$	–	-0.87 (0.071)	-0.76 (0.246)	-0.61 (0.417)
$\mathbb{P}(\sigma_\tau^2 > \sigma_c^2)$	–	0.92	0.83	0.71

In Figure 5 we plot the posterior densities of the variance ratio  $\sigma_\tau^2/\sigma_c^2$  under the three models.<sup>3</sup> The majority of the mass for the densities is in regions that are larger than unity. In fact, the posterior probabilities  $\mathbb{P}(\sigma_\tau^2 > \sigma_c^2 | \mathbf{y})$  are 0.92, 0.83 and 0.71 for UCUR, UCUR-07 and UCUR-(73,07), respectively. These results show that despite the high parameter uncertainty, the three models conclude that real shocks are relatively more important in explaining the variation in output.

<sup>3</sup>For each model, posterior draws of the ratio  $\sigma_\tau^2/\sigma_c^2$  are first obtained using the MCMC sampler described in Appendix A. These draws are then used to compute the density using the kernel density estimator of Botev, Grotowski, and Kroese (2010).

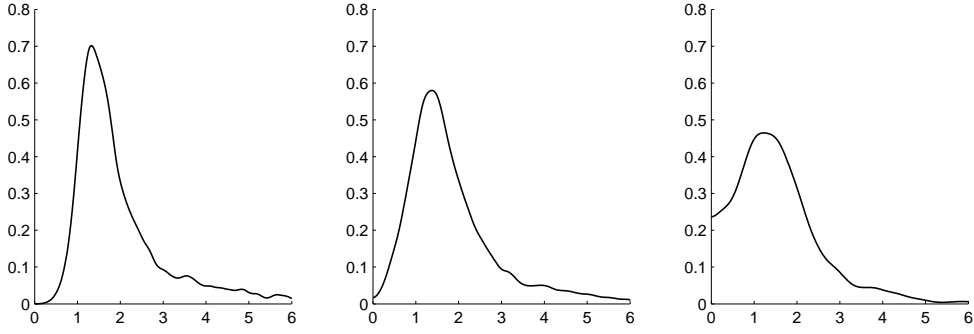


Figure 5: Posterior densities of the variance ratio  $\sigma_\tau^2/\sigma_c^2$  under UCUR (left panel), UCUR-07 (middle panel) and UCUR-(73,07) (right panel).

Next, we consider the estimates of  $\rho$  under the three correlated unobserved components models. The estimates of  $\rho$  are negative and large in magnitude for all three models. In particular,  $\rho$  is estimated to be  $-0.87$ ,  $-0.76$  and  $-0.61$  for UCUR, UCUR-07 and UCUR-(73,07), respectively. In contrast, Luo and Startz (2014) report an estimate of  $0.18$  under their preferred model of UCUR-06.<sup>4</sup> The posterior densities of  $\rho$  are plotted in Figure 6. The posterior modes of the three densities are near  $-0.9$ , which is in line with the results in Morley et al. (2003). In addition, all densities have little mass near  $0$ , showing the empirical relevance of allowing for nonzero correlation between the permanent and transitory shocks.

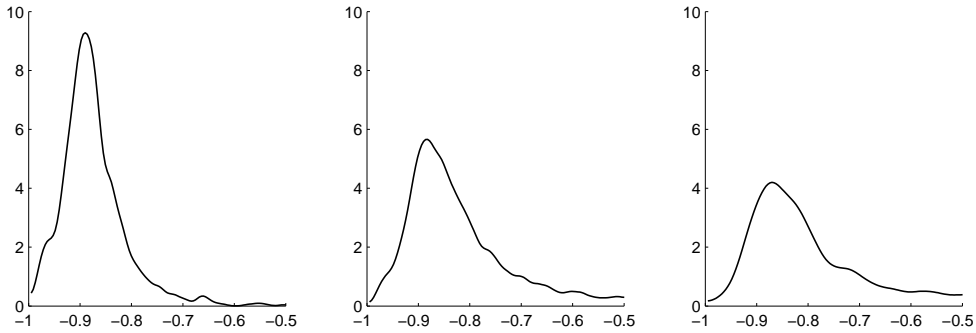


Figure 6: Posterior densities of  $\rho$  under UCUR (left panel), UCUR-07 (middle panel) and UCUR-(73,07) (right panel).

<sup>4</sup>The difference in our results might be due to different parameterizations and priors. Specifically, instead of using the parameterization in terms of  $(\sigma_c^2, \sigma_\tau^2, \rho)$ , Luo and Startz (2014) transform the parameters to  $(\sigma_1^2, \sigma_2^2, b)$  via the transformation

$$\begin{pmatrix} \sigma_\tau^2 & \rho\sigma_c\sigma_\tau \\ \rho\sigma_c\sigma_\tau & \sigma_c^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}.$$

Then, they assume priors on the transformed parameters  $(\sigma_1^2, \sigma_2^2, b)$ . It is unclear what the implied prior for  $\rho$  is. Instead, we parameterize the model in terms of  $(\sigma_c^2, \sigma_\tau^2, \rho)$  and assume a uniform prior on  $(-1, 1)$  for  $\rho$ .



### 4.3 Breaks in Trend Output Growth

In this section we first report the trend output growth rates for the four univariate models: DT-73, UCUR, UCUR-07 and UCUR-(73,07). Our main conclusion is that the trend output growth rate has substantially slowed since 2007. We then check the robustness of this result by considering bivariate correlated unobserved components models of GDP and inflation, and GDP and the unemployment rate.

Table 4 presents the trend output growth rates for the four univariate models. For ease of comparison, each row reports the growth rates for a specific period. For instance, under the UCUR model, the *annualized* trend growth rate from 1947 to 2014 is 3.12% ( $0.78\% \times 4$ ). As the table shows, there is substantial variation across different periods.

Table 4: Estimated trend output growth rates under DT-73, UCUR, UCUR-07 and UCUR-(73,07). Numerical standard errors are in parentheses.

	DT-73	UCUR	UCUR-07	UCUR-(73,07)
$\mu$ : 1947-2014	–	0.78 (0.082)	–	–
$\mu$ : 1947-2006	–	–	0.84 (0.077)	–
$\mu$ : 1947-1972	0.97 (0.039)	–	–	0.93 (0.106)
$\mu$ : 1973-2014	0.70 (0.039)	–	–	–
$\mu$ : 1973-2006	–	–	–	0.77 (0.093)
$\mu$ : 2007-2014	–	–	0.37 (0.199)	0.32 (0.197)

First, recall the model comparison results in Section 4.1 that show between the two deterministic trend models DT-73 and DT, the data favor the former. The parameter estimates of the trend growth under DT-73 support this conclusion. In particular, the annualized trend growth rate drops from 3.88% before 1973 to 2.80% after 1973. The left panel of Figure 7 plots the posterior density of the difference in trend growth rates, which has virtually no mass below zero.

Also recall that the overall best model is UCUR-07, which allows for a break in 2007. Under UCUR-07, the annualized trend growth rate more than halves before and after 2007—dropping from 3.36% to 1.48%. UCUR-(73,07) gives similar estimates: the annualized trend growth rate drops from 3.08% between 1973-2006 to 1.28% after 2007. The low estimates of the trend growth rates after 2007 for UCUR-07 and UCUR-(73,07) are partly due to the influence of the Great Recession—we only have eight years of data after the 2007 break. Even so, our results support the view that growth has slowed after the Great Recession.

The posterior densities of the difference in trend growth for UCUR-07 and UCUR-(73,07) in Figure 7 show that there is more parameter uncertainty for both models compared to DT-73, reflecting the difficulty in estimating the growth of a stochastic trend as opposed to a deterministic one. Nevertheless, both densities have little mass below zero, reinforcing the conclusion that trend growth has slowed after 2007.

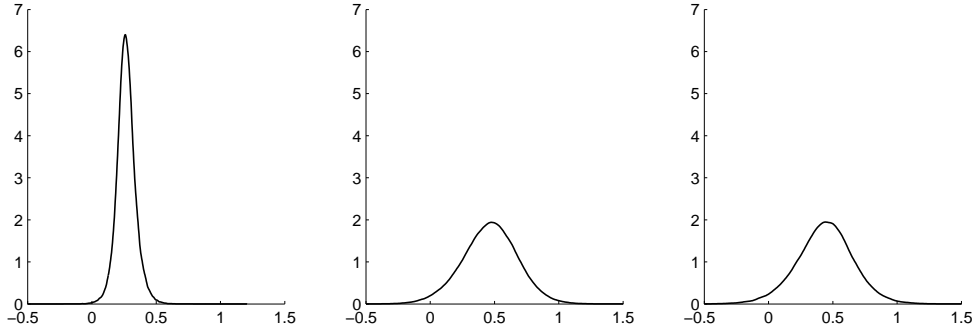


Figure 7: Posterior densities of the difference in trend growth between 1947-1972 and 1973-2014 under DT-73 (left panel); between 1947-2006 and 2007-2014 under UCUR-07 (middle panel); between 1973-2006 and 2007-2014 under UCUR-(73,07) (right panel).

Next, we investigate the robustness of our results by considering various bivariate models. A few recent papers have extended univariate unobserved components models to multivariate settings. For example, Basistha (2007) and Basistha and Nelson (2007) introduce bivariate unobserved components models for GDP and inflation; Sinclair (2009) proposes a bivariate correlated unobserved components model for GDP and unemployment; and Berger and Kempa (2011) consider a trivariate system that includes GDP, inflation and the exchange rate.

Following Sinclair (2009), we consider the following bivariate correlated unobserved components model:

$$y_{it} = \tau_{it} + c_{it} \quad (10)$$

for  $i = 1, 2$ , where the trend and cyclical components are modeled as:

$$\tau_{it} = \mu_{i1}1(t < t_0) + \mu_{i2}1(t \geq t_0) + \tau_{i,t-1} + u_{it}^\tau, \quad (11)$$

$$c_{it} = \phi_{i1}c_{i,t-1} + \phi_{ip}c_{i2} + u_{it}^c. \quad (12)$$

We further assume that the innovations  $\mathbf{u}_t = (u_{1t}^c, u_{2t}^c, u_{1t}^\tau, u_{2t}^\tau)'$  are jointly normal, i.e.,  $\mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$ . We fit this model with two sets of data: GDP and inflation, and GDP and the unemployment rate.<sup>5</sup> As before, we also consider two types of break dates: a break at 2007Q1, and two breaks at 1973Q1 and 2007Q1.

<sup>5</sup>The inflation rate is computed from the consumer price index and is available from 1947Q2 to 2014Q4. The sample for the unemployment rate is from 1948Q1 to 2014Q4. Both series are sourced from the Federal Reserve Bank of St. Louis economic database.

The estimated trend output growth rates for these four specifications are reported in Table 5. For example, the second column reports the estimates from the bivariate model of GDP and inflation with a break in 2007. All four specifications give similar trend output growth estimates after 2007 compared to those obtained from univariate models. In particular, the annualized trend output growth rates are between 1.16% and 1.33%.

Table 5: Estimated trend output growth rates under four bivariate models. Numerical standard errors are in parentheses.

	2007 break		1973 + 2007 breaks	
	inflation	unemployment	inflation	unemployment
$\mu$ : 1947-2006	0.86 (0.045)	0.85 (0.087)	–	–
$\mu$ : 1947-1972	–	–	0.96 (0.060)	0.97 (0.126)
$\mu$ : 1973-2006	–	–	0.77 (0.053)	0.77 (0.111)
$\mu$ : 2007-2014	0.29 (0.144)	0.33 (0.221)	0.30 (0.132)	0.32 (0.212)

These results can be compared to the recent commentary on secular stagnation. Summers (2013, 2015) discusses a return of “secular stagnation” due to insufficient investment—the natural rate of interest is negative but the nominal interest rate cannot fall appreciably below zero, thus creating a saving glut and inadequate investment. This in turn raises the concern that actual output growth has slowed reflecting slower long-run potential growth.

In addition, Gordon (2012) argues that US labor productivity has already slowed markedly after 1972 compared to the previous eight decades, because the main ideas of the Second Industrial Revolution—the inventions of electricity, the internal combustion engine and running water with indoor plumbing—had mostly been implemented by then. In a follow-up paper, Gordon (2014) takes the realized productivity growth from 1972-2007 as the starting point, from which to subtract two “headwinds”: demographics and education. Together with a projected US population growth of 0.6%–0.65%, Gordon (2014) gives a forecast of 1.5%–1.55% of potential output growth from 2007 to 2032. While our estimates are not forecasts, our baseline model UCUR-07 shows that trend output growth from 2007 to 2014 has dropped to 1.48%. Results from the bivariate models give even slightly smaller estimates.

Next, Figure 8 plots the output gap estimates of the two bivariate models with a break in 2007. While these two models provide similar estimates of trend output growth, the estimates of the output gap are quite different: the output gap from the GDP and inflation model is substantially larger than that from the GDP and unemployment model.

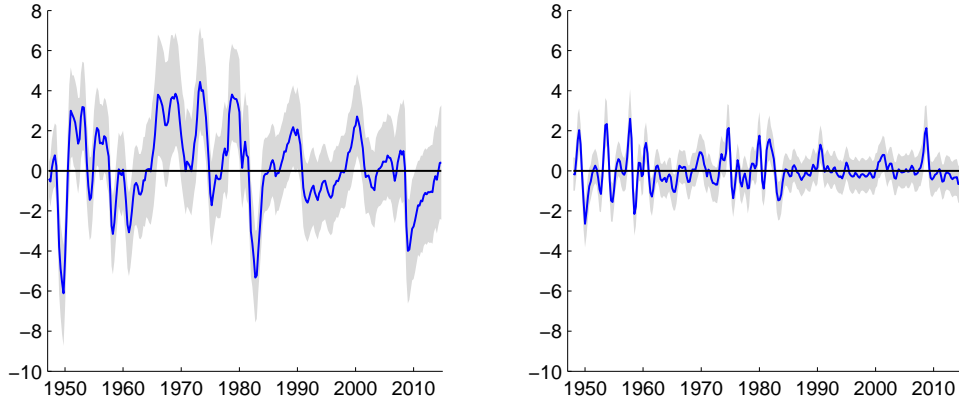


Figure 8: Estimates of the output gap for the two bivariate unobserved components models with a break in 2007: GDP and inflation (left panel) and GDP and unemployment (right panel).

Table 6: Selected parameter estimates for the GDP process under four bivariate models. Numerical standard errors are in parentheses.

	2007 break		1973 + 2007 breaks	
	inflation	unemployment	inflation	unemployment
variance of innovation to GDP cycle	0.51 (0.284)	0.75 (0.377)	0.47 (0.215)	0.74 (0.373)
variance of innovation to GDP trend	0.49 (0.285)	1.84 (0.413)	0.35 (0.173)	1.75 (0.446)
correlation between innovations	-0.18 (0.351)	-0.86 (0.043)	-0.04 (0.320)	-0.84 (0.155)

One factor that contributes to this difference is that under the model with inflation, the estimated variance of the innovation to the trend component of GDP becomes much smaller, which makes the relative contribution of the cyclical component larger (see Table 6). On the other hand, the estimates from the bivariate model with GDP and the unemployment rate are more similar to those from the univariate unobserved components models.<sup>6</sup> These results indicate that the estimated correlation structure of the innovations can be sensitive to the choice of variables included in the estimation.<sup>7</sup> This might reflect an inherent estimation problem: in a univariate model there is only one correlation

<sup>6</sup>Under the univariate models, we assume uniform priors for the variances and the correlation. Under the bivariate models, the covariance matrix  $\Sigma$  is assumed to have an inverse-Wishart prior centered at the identity matrix. Despite these two very different forms of priors, the parameter estimates from the bivariate model of GDP and unemployment are similar to those under the UCUR-07 model, indicating that priors do not play an important role in driving the results.

<sup>7</sup>Mertens (2014) raises the concern that the relationship between inflation and economic slack might not be stable over time. This might contribute to the difference in output gap estimates.

parameter; in a bivariate model, there are six. This makes it more difficult to precisely estimate these correlation parameters in the bivariate models.

Given the very different output gap estimates of the two bivariate models, it is of interest to perform a model comparison exercise to see which model fits the data better. However, this exercise is complicated by the fact that the two bivariate models fit different dependent variables—GDP and inflation versus GDP and unemployment—and therefore the marginal likelihoods from the two models are not directly comparable. To get around this problem, we compute the “conditional” marginal likelihood of GDP given the other variable.<sup>8</sup> This measure evaluates the fit of GDP given the information of the other variable. In addition, this measure is directly comparable to the marginal likelihood from the univariate models of GDP.

Table 7 reports the log marginal likelihoods of the two bivariate models with different types of breaks.<sup>9</sup> The bivariate model of GDP and unemployment fits the data substantially better than the model of GDP and inflation—the log Bayes factor in favor of the former model with a break in 2007 against the latter is 58. In fact, the bivariate model of GDP and unemployment does better than the univariate model of GDP alone, whereas the model of GDP and inflation does substantially worse. One can also conclude from the results of the bivariate model of GDP and unemployment that one break in 2007 is sufficient.

Table 7: Log marginal likelihoods of various univariate and bivariate models.

2007 break			1973 + 2007 breaks		
UCUR	inflation	unemployment	UCUR	inflation	unemployment
−359.61	−385.16	−327.39	−365.21	−384.48	−342.00
(0.053)	(0.388)	(0.205)	(0.092)	(0.360)	(0.605)

To conclude, these results suggest that the estimated correlation between the innovations in the GDP process can be sensitive to the choice of dependent variables: the estimated correlation is large and negative as in the univariate UCUR model if the unemployment rate is included, but it becomes close to zero if inflation is included instead. However, the model comparison results show that the bivariate model of GDP and inflation fits the data substantially worse compared to the univariate model as well as the bivariate model of GDP and unemployment. We take these results as evidence in favor of a strong negative correlation between the innovations to the trend and cycle components. In addition, all the results from the bivariate models support the conclusion that the trend output growth rate substantially slowed after 2007.

<sup>8</sup>Specifically, let  $\mathbf{y}_1$  denote log real GDP and let  $\mathbf{y}_2$  represent the other variable. The conditional version of the marginal likelihood under model  $M_k$  is defined as  $p(\mathbf{y}_1 | \mathbf{y}_2, M_k) = \int p(\mathbf{y}_1 | \mathbf{y}_2, \boldsymbol{\theta}_k, M_k) p(\boldsymbol{\theta}_k | M_k) d\boldsymbol{\theta}_k$ , where  $p(\mathbf{y}_1 | \mathbf{y}_2, \boldsymbol{\theta}_k, M_k)$  is the integrated likelihood of  $\mathbf{y}_1$  given  $\mathbf{y}_2$  and the parameter vector  $\boldsymbol{\theta}_k$ .

<sup>9</sup>Since data on the unemployment rate are only available after 1948Q1, the sample period for all the models in Table 7 is taken from 1948Q1 to 2014Q4. This ensures that the log marginal likelihoods are comparable across models.

## 5 Concluding Remarks and Future Research

We have undertaken a formal Bayesian model comparison exercise to assess a number of models for decomposing US output into its trend and cyclical components. We find that it is empirically important to allow for correlation between permanent and transitory shocks. The correlated unobserved components model dominates any deterministic trend models with or without a break in growth. The overall best model is the correlated unobserved components model with a break in 2007. The annualized trend output growth decreases from about 3.4% to 1.2%-1.5% after the break. This model also indicates that permanent shocks are relatively more important in explaining the variation in output compared to transitory shocks. It would be interesting to see if these conclusions remain true if a broader set of nonlinear and asymmetric models are included, such as those in Morley and Piger (2008), Sinclair (2010) and Morley and Piger (2012).

Many recent papers, including Stock and Watson (2007), Chan (2013) and Clark and Doh (2014), have demonstrated the importance of allowing for stochastic volatility in modeling inflation using unobserved components models. More recently, Mertens (2014) has shown that stochastic volatility is also useful for decomposing output and unemployment. For future research, it would be worthwhile to add stochastic volatility to the bivariate unobserved components models for decomposing output. One particular focus would be on determining whether the addition of stochastic volatility reconciles the results between the GDP and inflation and GDP and unemployment rate models, and whether the models continue to give similar conclusions about trend output growth. More generally, it is of interest to develop parsimonious multivariate unobserved components models with stochastic volatility for decomposing output.

## Appendix A: Estimation Details

This appendix discusses the priors and provides the estimation details of the unobserved components models discussed in Section 2.

### Estimation of the UCUR Model

For the UCUR model in (1)–(3), the model parameters are  $\boldsymbol{\phi} = (\phi_1, \phi_2)'$ ,  $\sigma_c^2$ ,  $\sigma_\tau^2$ ,  $\rho$ ,  $\mu_1$  and  $\tau_0$ . We assume standard independent priors for  $\boldsymbol{\phi}$ ,  $\mu_1$  and  $\tau_0$ :

$$\boldsymbol{\phi} \sim \mathcal{N}(\boldsymbol{\phi}_0, \mathbf{V}_\phi)1(\boldsymbol{\phi} \in \mathbf{R}), \quad \mu_1 \sim \mathcal{N}(\mu_0, V_\mu), \quad \tau_0 \sim \mathcal{N}(\tau_{00}, V_\tau),$$

where  $\mathbf{R}$  is the stationarity region. We assume relatively large prior variances with  $\mathbf{V}_\phi = \mathbf{I}_2$ ,  $V_\mu = 1$  and  $V_\tau = 100$ . For the prior means, we set  $\boldsymbol{\phi}_0 = (1.3, -0.7)'$ ,  $\tau_{00} = 750$  and  $\mu_0 = 0.75$ . In particular, these values imply that the prior mean of the annualized growth rate is 3% and the AR(2) process of the transitory component has two complex roots. Next,  $\sigma_c^2$ ,  $\sigma_\tau^2$  and  $\rho$  have uniform priors:

$$\sigma_c^2 \sim \mathcal{U}(0, b_c), \quad \sigma_\tau^2 \sim \mathcal{U}(0, b_\tau), \quad \rho \sim \mathcal{U}(-1, 1),$$

where we set the upper bounds as  $b_c = b_\tau = 3$ .

Posterior draws are obtained by sequentially sampling from: 1.  $p(\boldsymbol{\tau} | \mathbf{y}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0)$ ; 2.  $p(\boldsymbol{\phi} | \mathbf{y}, \boldsymbol{\tau}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0)$ ; 3.  $p(\sigma_c^2 | \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_\tau^2, \rho, \mu_1, \tau_0)$ ; 4.  $p(\sigma_\tau^2 | \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \rho, \mu_1, \tau_0)$ ; 5.  $p(\rho | \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \mu_1, \tau_0)$ ; and 6.  $p(\tau_0, \mu_1 | \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho)$ . The implementation of Step 1 is discussed in Section 2.2. Here we provide the details of the other steps.

To sample  $\boldsymbol{\phi}$  in Step 2, recall that  $\mathbf{u}^c$  and  $\boldsymbol{\tau}$  are jointly normal:

$$\begin{pmatrix} \mathbf{u}^c \\ \boldsymbol{\tau} \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\alpha} \end{pmatrix}, \begin{pmatrix} \sigma_c^2 \mathbf{I}_T & \rho \sigma_c \sigma_\tau (\mathbf{H}')^{-1} \\ \rho \sigma_c \sigma_\tau \mathbf{H}^{-1} & \sigma_\tau^2 (\mathbf{H}' \mathbf{H})^{-1} \end{pmatrix} \right), \quad (13)$$

where  $\boldsymbol{\alpha} = \mathbf{H}^{-1} \tilde{\boldsymbol{\alpha}}$  with  $\tilde{\boldsymbol{\alpha}} = (\mu_1 + \tau_0, \mu_1, \dots, \mu_1)'$ . Hence, the conditional distribution of  $\mathbf{u}^c$  given  $\boldsymbol{\tau}$  and the other parameters is

$$(\mathbf{u}^c | \boldsymbol{\tau}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0) \sim \mathcal{N} \left( \frac{\rho \sigma_c}{\sigma_\tau} \mathbf{H} (\boldsymbol{\tau} - \boldsymbol{\alpha}), (1 - \rho^2) \sigma_c^2 \mathbf{I}_T \right).$$

Next, we write (3) as

$$\mathbf{c} = \mathbf{X}_\phi \boldsymbol{\phi} + \mathbf{u}^c,$$

where  $\mathbf{X}_\phi$  is a  $T \times 2$  matrix consisting of lagged values of  $c_t$ . Then, by standard regression results, we have

$$(\boldsymbol{\phi} | \mathbf{y}, \boldsymbol{\tau}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0) \sim \mathcal{N}(\hat{\boldsymbol{\phi}}, \mathbf{K}_\phi^{-1})1(\boldsymbol{\phi} \in \mathbf{R}),$$

where

$$\mathbf{K}_\phi = \mathbf{V}_\phi^{-1} + \frac{1}{(1-\rho^2)\sigma_c^2} \mathbf{X}'_\phi \mathbf{X}_\phi,$$

$$\widehat{\phi} = \mathbf{K}_\phi^{-1} \left( \mathbf{V}_\phi^{-1} \phi_0 + \frac{1}{(1-\rho^2)\sigma_c^2} \mathbf{X}'_\phi \left( \mathbf{c} - \frac{\rho\sigma_c}{\sigma_\tau} \mathbf{H}(\boldsymbol{\tau} - \boldsymbol{\alpha}) \right) \right).$$

A draw from this truncated normal distribution can be obtained by the acceptance-rejection method, i.e., keep sampling from  $\mathcal{N}(\widehat{\phi}, \mathbf{K}_\phi^{-1})$  until  $\phi \in \mathbf{R}$ .

To implement Steps 3 to 5, we first derive the joint density of  $\mathbf{u}^c$  and  $\mathbf{u}^\tau$ . To that end, note that given  $\sigma_c^2$ ,  $\sigma_\tau^2$  and  $\rho$ , we can factorize  $(u_t^c, u_t^\tau)$  as:

$$u_t^\tau \sim \mathcal{N}(0, \sigma_\tau^2), \quad (u_t^c | u_t^\tau) \sim \mathcal{N}\left(\frac{\rho\sigma_c}{\sigma_\tau} u_t^\tau, (1-\rho^2)\sigma_c^2\right).$$

Hence, the joint density of  $\mathbf{u}^c$  and  $\mathbf{u}^\tau$  is given by

$$p(\mathbf{u}^c, \mathbf{u}^\tau | \sigma_c^2, \sigma_\tau^2, \rho) \propto (\sigma_\tau^2)^{-\frac{T}{2}} e^{-\frac{1}{2\sigma_\tau^2} \sum_{t=1}^T (u_t^\tau)^2} \left( (1-\rho^2)\sigma_c^2 \right)^{-\frac{T}{2}} e^{-\frac{1}{2(1-\rho^2)\sigma_c^2} \sum_{t=1}^T (u_t^c - \frac{\rho\sigma_c}{\sigma_\tau} u_t^\tau)^2},$$

$$= \left( (1-\rho^2)\sigma_c^2 \sigma_\tau^2 \right)^{-\frac{T}{2}} e^{-\frac{1}{2\sigma_\tau^2} k_3 - \frac{1}{2(1-\rho^2)\sigma_c^2} \left( k_1 - \frac{2\rho\sigma_c}{\sigma_\tau} k_2 + \frac{\rho^2\sigma_c^2}{\sigma_\tau^2} k_3 \right)}, \quad (14)$$

where  $k_1 = \sum_{t=1}^T (u_t^c)^2$ ,  $k_2 = \sum_{t=1}^T u_t^c u_t^\tau$  and  $k_3 = \sum_{t=1}^T (u_t^\tau)^2$ . It follows from (14) that

$$p(\sigma_c^2 | \mathbf{y}, \boldsymbol{\tau}, \phi, \sigma_\tau^2, \rho, \mu_1, \tau_0) \propto p(\sigma_c^2) \times (\sigma_c^2)^{-\frac{T}{2}} e^{-\frac{1}{2(1-\rho^2)\sigma_c^2} \left( k_1 - \frac{2\rho\sigma_c}{\sigma_\tau} k_2 + \frac{\rho^2\sigma_c^2}{\sigma_\tau^2} k_3 \right)},$$

where  $p(\sigma_c^2)$  is the truncated normal prior specified above. This full conditional density of  $\sigma_c^2$  is not a standard density and we sample from it using a Griddy-Gibbs step. That is, we evaluate the full conditional density on a fine grid, and obtain a draw from the density using the inverse-transform method (see, e.g., Kroese, Taimre, and Botev, 2011, pp. 45–47). Steps 4 and 5 can be similarly implemented by noting that

$$p(\sigma_\tau^2 | \mathbf{y}, \boldsymbol{\tau}, \phi, \sigma_c^2, \rho, \mu_1, \tau_0) \propto p(\sigma_\tau^2) \times (\sigma_\tau^2)^{-\frac{T}{2}} e^{-\frac{1}{2\sigma_\tau^2} k_3 - \frac{1}{2(1-\rho^2)\sigma_c^2} \left( k_1 - \frac{2\rho\sigma_c}{\sigma_\tau} k_2 + \frac{\rho^2\sigma_c^2}{\sigma_\tau^2} k_3 \right)}$$

$$p(\rho | \mathbf{y}, \boldsymbol{\tau}, \phi, \sigma_c^2, \sigma_\tau^2, \mu_1, \tau_0) \propto p(\rho) \times (1-\rho^2)^{-\frac{T}{2}} e^{-\frac{1}{2(1-\rho^2)\sigma_c^2} \left( k_1 - \frac{2\rho\sigma_c}{\sigma_\tau} k_2 + \frac{\rho^2\sigma_c^2}{\sigma_\tau^2} k_3 \right)},$$

where  $p(\sigma_\tau^2)$  and  $p(\rho)$  are the priors for  $\sigma_\tau^2$  and  $\rho$  respectively.

Lastly, to jointly sample  $\tau_0$  and  $\mu_1$ , note that we can write  $\boldsymbol{\alpha} = \tau_0 \mathbf{1}_T + \mu_1 \mathbf{H}^{-1} \mathbf{1}_T = \mathbf{X}_\delta \boldsymbol{\delta}$ , where  $\mathbf{1}_T$  is a  $T \times 1$  column of ones,  $\mathbf{X}_\delta = (\mathbf{1}_T, \mathbf{H}^{-1} \mathbf{1}_T)$  and  $\boldsymbol{\delta} = (\tau_0, \mu_1)'$ . It follows from (13) that the conditional distribution of  $\boldsymbol{\tau}$  given  $\mathbf{u}^c$  and other parameters is

$$(\boldsymbol{\tau} | \mathbf{u}^c, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0) \sim \mathcal{N}\left(\mathbf{X}_\delta \boldsymbol{\delta} + \frac{\rho\sigma_\tau}{\sigma_c} \mathbf{H}^{-1} \mathbf{u}^c, (1-\rho^2)\sigma_\tau^2 (\mathbf{H}'\mathbf{H})^{-1}\right).$$

Then, by standard regression results, we have

$$(\tau_0, \mu_1 | \mathbf{y}, \boldsymbol{\tau}, \sigma_c^2, \sigma_\tau^2, \rho, \phi, \tau_0) \sim \mathcal{N}(\widehat{\boldsymbol{\delta}}, \mathbf{K}_\delta^{-1}),$$



where

$$\mathbf{K}_\delta = \mathbf{V}_\delta^{-1} + \frac{1}{(1 - \rho^2)\sigma_\tau^2} \mathbf{X}'_\delta \mathbf{H}' \mathbf{H} \mathbf{X}_\delta,$$

$$\hat{\boldsymbol{\delta}} = \mathbf{K}_\delta^{-1} \left( \mathbf{V}_\delta^{-1} \boldsymbol{\delta}_0 + \frac{1}{(1 - \rho^2)\sigma_\tau^2} \mathbf{X}'_\delta \mathbf{H}' \mathbf{H} \left( \boldsymbol{\tau} - \frac{\rho\sigma_\tau}{\sigma_c} \mathbf{H}^{-1} \mathbf{u}^c \right) \right),$$

where  $\mathbf{V}_\delta = \text{diag}(V_\tau, V_\mu)$  and  $\boldsymbol{\delta}_0 = (\tau_{00}, \mu_0)'$ .

## Estimation of the UCUR- $t_0$ Model

We now consider an extension of the UCUR model that allows for a break in the growth rate of the trend at time  $t_0$ . Specifically, we replace (4) with (5), which is reproduced below:

$$\tau_t = \mu_1 1(t < t_0) + \mu_2 1(t \geq t_0) + \tau_{t-1} + u_t^\tau,$$

where  $1(A)$  is the indicator function that takes the value 1 if the condition  $A$  is true and 0 otherwise. Compared to the UCUR model, the only additional parameter is  $\mu_2$ . Its prior is assumed to be the same as that of  $\mu_1$ , i.e.,  $\mu_2 \sim \mathcal{N}(\mu_0, V_\mu)$  with  $\mu_0 = 0.75$  and  $V_\mu = 1$ . For the common parameters, we assume exactly the same priors as in the UCUR model.

Only minor modifications of the sampler for the UCUR model are needed to fit this extension. For example, if we redefine

$$\tilde{\boldsymbol{\alpha}} = \left( \underbrace{\mu_1 + \tau_0, \mu_1, \dots, \mu_1}_{t_0-1}, \underbrace{\mu_2, \dots, \mu_2}_{T-t_0+1} \right)'$$

and  $\boldsymbol{\alpha} = \mathbf{H}^{-1} \tilde{\boldsymbol{\alpha}}$ , then Steps 1 and 2 can be implemented exactly as before. Similarly, if we compute  $\mathbf{u}^\tau$  using  $\mathbf{u}^\tau = \mathbf{H}\boldsymbol{\tau} - \tilde{\boldsymbol{\alpha}}$ , then Steps 3–5 remain the same as before. Lastly, to sample  $\tau_0$ ,  $\mu_1$  and  $\mu_2$  jointly, write  $\boldsymbol{\alpha} = \tau_0 \mathbf{1}_T + \mu_1 \mathbf{H}^{-1} \mathbf{d}_1 + \mu_2 \mathbf{H}^{-1} \mathbf{d}_2 = \mathbf{X}_\delta \boldsymbol{\delta}$ , where  $\mathbf{d}_1$  is a  $T \times 1$  vector of dummy variables where the first  $t_0 - 1$  elements are 1 and the rest are 0, and  $\mathbf{d}_2$  is defined so that  $\mathbf{d}_1 + \mathbf{d}_2 = \mathbf{1}_T$ . Note that  $\mathbf{X}_\delta$  and  $\boldsymbol{\delta}$  are redefined as  $\mathbf{X}_\delta = (\mathbf{1}_T, \mathbf{H}^{-1} \mathbf{d}_1, \mathbf{H}^{-1} \mathbf{d}_2)$  and  $\boldsymbol{\delta} = (\tau_0, \mu_1, \mu_2)'$ . Then, the last step is implemented as before.

## Estimation of the DT- $t_0$ Model

For the deterministic trend model with the trend

$$\tau_t = \mu_1 1(t < t_0) + \mu_2 1(t \geq t_0) + \tau_{t-1},$$

the model parameters are  $\boldsymbol{\phi}$ ,  $\sigma_c^2$ ,  $\mu_1$ ,  $\mu_2$  and  $\tau_0$ . We adopt the same priors as in the UCUR- $t_0$  model. Note that we can write this model as

$$\begin{aligned} \mathbf{y} &= \mathbf{X}_\delta \boldsymbol{\delta} + \mathbf{c}, \\ \mathbf{c} &= \mathbf{H}_\phi^{-1} \mathbf{u}^c, \end{aligned} \tag{15}$$

where  $\mathbf{X}_\delta = (\mathbf{1}_T, \mathbf{H}^{-1}\mathbf{d}_1, \mathbf{H}^{-1}\mathbf{d}_2)$  and  $\boldsymbol{\delta} = (\tau_0, \mu_1, \mu_2)'$ .

Posterior draws can be obtained by sequentially sampling from: 1.  $p(\boldsymbol{\phi} | \mathbf{y}, \sigma_c^2, \mu_1, \mu_2, \tau_0)$ ; 2.  $p(\sigma_c^2 | \mathbf{y}, \boldsymbol{\phi}, \mu_1, \mu_2, \tau_0)$ ; and 3.  $p(\tau_0, \mu_1, \mu_2 | \mathbf{y}, \boldsymbol{\phi}, \sigma_c^2)$ . To implement Step 1, note that  $(\mathbf{u}^c | \sigma_c^2) \sim \mathcal{N}(\mathbf{0}, \sigma_c^2 \mathbf{I}_T)$ . Hence, we have

$$(\boldsymbol{\phi} | \mathbf{y}, \sigma_c^2, \mu_1, \mu_2, \tau_0) \sim \mathcal{N}(\widehat{\boldsymbol{\phi}}, \mathbf{K}_\phi^{-1}) \mathbf{1}(\boldsymbol{\phi} \in \mathbf{R}),$$

where

$$\mathbf{K}_\phi = \mathbf{V}_\phi^{-1} + \frac{1}{\sigma_c^2} \mathbf{X}'_\phi \mathbf{X}_\phi, \quad \widehat{\boldsymbol{\phi}} = \mathbf{K}_\phi^{-1} \left( \mathbf{V}_\phi^{-1} \boldsymbol{\phi}_0 + \frac{1}{\sigma_c^2} \mathbf{X}'_\phi \mathbf{c} \right).$$

A draw from this truncated normal distribution can be obtained by using the acceptance-rejection method.

Next, the full conditional density of  $\sigma_c^2$  is given by

$$p(\sigma_c^2 | \mathbf{y}, \boldsymbol{\phi}, \mu_1, \mu_2, \tau_0) \propto p(\sigma_c^2) \times (\sigma_c^2)^{-\frac{T}{2}} e^{-\frac{1}{2\sigma_c^2} \sum_{i=1}^T (u_i^c)^2},$$

where  $\mathbf{u}^c$  can be computed by  $\mathbf{u}^c = \mathbf{H}_\phi(\mathbf{y} - \mathbf{X}_\delta \boldsymbol{\delta})$ . As before, a draw of  $\sigma_c^2$  can be obtained using the Griddy-Gibbs step. Lastly, it follows from (15) that

$$(\tau_0, \mu_1, \mu_2 | \mathbf{y}, \sigma_c^2, \boldsymbol{\phi}) \sim \mathcal{N}(\widehat{\boldsymbol{\delta}}, \mathbf{K}_\delta^{-1}),$$

where

$$\mathbf{K}_\delta = \mathbf{V}_\delta^{-1} + \frac{1}{\sigma_c^2} \mathbf{X}'_\delta \mathbf{H}'_\phi \mathbf{H}_\phi \mathbf{X}_\delta, \quad \widehat{\boldsymbol{\delta}} = \mathbf{K}_\delta^{-1} \left( \mathbf{V}_\delta^{-1} \boldsymbol{\delta}_0 + \frac{1}{\sigma_c^2} \mathbf{X}'_\delta \mathbf{H}'_\phi \mathbf{H}_\phi \mathbf{y} \right).$$

## Appendix B: Integrated Likelihood Evaluation

This appendix derives analytical expressions for the integrated likelihoods under the unobserved components models discussed in Section 2. These integrated likelihoods can then be evaluated using band matrix routines, which are more efficient than using the conventional Kalman filter.

### The UCUR Model

Recall that the prior density of  $\boldsymbol{\tau}$  and the conditional likelihood under the UCUR model are given by

$$p(\boldsymbol{\tau} \mid \sigma_\tau^2, \mu_1, \tau_0) = (2\pi\sigma_\tau^2)^{-\frac{T}{2}} e^{-\frac{1}{2\sigma_\tau^2}(\boldsymbol{\tau}-\boldsymbol{\alpha})'\mathbf{H}'\mathbf{H}(\boldsymbol{\tau}-\boldsymbol{\alpha})}$$

$$p(\mathbf{y} \mid \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0) = (2\pi\sigma_c^2(1-\rho^2))^{-\frac{T}{2}} e^{-\frac{1}{2(1-\rho^2)\sigma_c^2}(\mathbf{H}_\phi\mathbf{y}-\mathbf{a}-\mathbf{B}\boldsymbol{\tau})'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a}-\mathbf{B}\boldsymbol{\tau})},$$

where

$$\mathbf{a} = -\frac{\rho\sigma_c}{\sigma_\tau}\mathbf{H}\boldsymbol{\alpha}, \quad \mathbf{B} = \mathbf{H}_\phi + \frac{\rho\sigma_c}{\sigma_\tau}\mathbf{H}.$$

Let  $k_4 = (2\pi)^{-T}((1-\rho^2)\sigma_c^2\sigma_\tau^2)^{-\frac{T}{2}}$ . Then, the integrated likelihood can be derived as follows:

$$\begin{aligned} p(\mathbf{y} \mid \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0) &= \int p(\mathbf{y} \mid \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \mu_1, \tau_0)p(\boldsymbol{\tau} \mid \sigma_\tau^2, \mu_1, \tau_0)d\boldsymbol{\tau} \\ &= k_4 \int e^{-\frac{1}{2(1-\rho^2)\sigma_c^2}(\mathbf{H}_\phi\mathbf{y}-\mathbf{a}-\mathbf{B}\boldsymbol{\tau})'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a}-\mathbf{B}\boldsymbol{\tau})} e^{-\frac{1}{2\sigma_\tau^2}(\boldsymbol{\tau}-\boldsymbol{\alpha})'\mathbf{H}'\mathbf{H}(\boldsymbol{\tau}-\boldsymbol{\alpha})} d\boldsymbol{\tau} \\ &= k_4 \int e^{-\frac{1}{2}\left(\frac{1}{(1-\rho^2)\sigma_c^2}((\mathbf{H}_\phi\mathbf{y}-\mathbf{a})'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})-2\boldsymbol{\tau}'\mathbf{B}'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})+\boldsymbol{\tau}'\mathbf{B}'\mathbf{B}\boldsymbol{\tau})+\frac{1}{\sigma_\tau^2}(\boldsymbol{\tau}'\mathbf{H}'\mathbf{H}\boldsymbol{\tau}-2\boldsymbol{\tau}'\mathbf{H}'\mathbf{H}\boldsymbol{\alpha}+\boldsymbol{\alpha}'\mathbf{H}'\mathbf{H}\boldsymbol{\alpha}))\right)} d\boldsymbol{\tau} \\ &= k_4 e^{-\frac{1}{2}\left(\frac{1}{(1-\rho^2)\sigma_c^2}(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})+\frac{1}{\sigma_\tau^2}\boldsymbol{\alpha}'\mathbf{H}'\mathbf{H}\boldsymbol{\alpha}\right)} \int e^{-\frac{1}{2}(\boldsymbol{\tau}'\mathbf{K}_\tau\boldsymbol{\tau}-2\boldsymbol{\tau}'\mathbf{d}_\tau)} d\boldsymbol{\tau} \\ &= k_4 e^{-\frac{1}{2}\left(\frac{1}{(1-\rho^2)\sigma_c^2}(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})+\frac{1}{\sigma_\tau^2}\boldsymbol{\alpha}'\mathbf{H}'\mathbf{H}\boldsymbol{\alpha}-\mathbf{d}_\tau'\mathbf{K}_\tau^{-1}\mathbf{d}_\tau\right)} \int e^{-\frac{1}{2}((\boldsymbol{\tau}-\mathbf{K}_\tau^{-1}\mathbf{d}_\tau)'\mathbf{K}_\tau(\boldsymbol{\tau}-\mathbf{K}_\tau^{-1}\mathbf{d}_\tau))} d\boldsymbol{\tau} \\ &= k_4 e^{-\frac{1}{2}\left(\frac{1}{(1-\rho^2)\sigma_c^2}(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})+\frac{1}{\sigma_\tau^2}\boldsymbol{\alpha}'\mathbf{H}'\mathbf{H}\boldsymbol{\alpha}-\mathbf{d}_\tau'\mathbf{K}_\tau^{-1}\mathbf{d}_\tau\right)} (2\pi)^{\frac{T}{2}} |\mathbf{K}_\tau|^{-\frac{1}{2}} \\ &= (2\pi(1-\rho^2)\sigma_c^2\sigma_\tau^2)^{-\frac{T}{2}} |\mathbf{K}_\tau|^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{1}{(1-\rho^2)\sigma_c^2}(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a})+\frac{1}{\sigma_\tau^2}\boldsymbol{\alpha}'\mathbf{H}'\mathbf{H}\boldsymbol{\alpha}-\mathbf{d}_\tau'\mathbf{K}_\tau^{-1}\mathbf{d}_\tau\right)}, \end{aligned}$$

where

$$\mathbf{K}_\tau = \frac{1}{\sigma_\tau^2}\mathbf{H}'\mathbf{H} + \frac{1}{(1-\rho^2)\sigma_c^2}\mathbf{B}'\mathbf{B}, \quad \mathbf{d}_\tau = \frac{1}{\sigma_\tau^2}\mathbf{H}'\mathbf{H}\boldsymbol{\alpha} + \frac{1}{(1-\rho^2)\sigma_c^2}\mathbf{B}'(\mathbf{H}_\phi\mathbf{y}-\mathbf{a}).$$

Since  $\mathbf{H}$ ,  $\mathbf{H}_\phi$  and  $\mathbf{K}_\tau$  are band matrices, this integrated likelihood can be evaluated quickly using the band matrix algorithms discussed Chan and Grant (2016).

## The UCUR- $t_0$ Model

For the extension of the UCUR model that allows for a break in the growth rate of the trend at time  $t_0$ , only minor modifications are needed. In particular, if we redefine

$$\tilde{\boldsymbol{\alpha}} = \left( \underbrace{\mu_1 + \tau_0, \mu_1, \dots, \mu_1}_{t_0-1}, \underbrace{\mu_2, \dots, \mu_2}_{T-t_0+1} \right)'$$

and  $\boldsymbol{\alpha} = \mathbf{H}^{-1}\tilde{\boldsymbol{\alpha}}$ , then the integrated likelihood of this generalization is exactly the same as that of the UCUR model.

## The DT- $t_0$ Model

For the deterministic trend model, recall that from (15) we have

$$\mathbf{y} = \mathbf{X}_\delta \boldsymbol{\delta} + \mathbf{H}_\phi^{-1} \mathbf{u}^c,$$

where  $\mathbf{u}^c \sim \mathcal{N}(\mathbf{0}, \sigma_c^2 \mathbf{I}_T)$ ,  $\mathbf{X}_\delta = (\mathbf{1}_T, \mathbf{H}^{-1} \mathbf{d}_1, \mathbf{H}^{-1} \mathbf{d}_2)$  and  $\boldsymbol{\delta} = (\tau_0, \mu_1, \mu_2)'$ . Hence, the likelihood is given by

$$p(\mathbf{y} | \boldsymbol{\phi}, \sigma_c^2, \mu_1, \mu_2, \tau_0) = (2\pi\sigma_c^2)^{-\frac{T}{2}} e^{-\frac{1}{2\sigma_c^2} (\mathbf{y} - \mathbf{X}_\delta \boldsymbol{\delta})' \mathbf{H}'_\phi \mathbf{H}_\phi (\mathbf{y} - \mathbf{X}_\delta \boldsymbol{\delta})}.$$

## References

- A. Basistha. Trend-cycle correlation, drift break and the estimation of trend and cycle in Canadian GDP. *Canadian Journal of Economics*, 40(2):584–606, 2007.
- A. Basistha and C. R. Nelson. New measures of the output gap based on the forward-looking new Keynesian Phillips curve. *Journal of Monetary Economics*, 54(2):498–511, 2007.
- T. Berger and B. Kempa. Bayesian estimation of the output gap for a small open economy: The case of Canada. *Economics Letters*, 112(1):107–112, 2011.
- S. Beveridge and C. R. Nelson. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘business cycle’. *Journal of Monetary Economics*, 7(2):151 – 174, 1981.
- Z. I. Botev, J. F. Grotowski, and D. P. Kroese. Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5):2916–2957, 10 2010.
- J. C. C. Chan. Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics*, 176(2):162–172, 2013.
- J. C. C. Chan and E. Eisenstat. Marginal likelihood estimation with the Cross-Entropy method. *Econometric Reviews*, 34(3):256–285, 2015.
- J. C. C. Chan and A. L. Grant. Fast computation of the deviance information criterion for latent variable models. *Computational Statistics and Data Analysis*, 100:847–859, 2016.
- J. C. C. Chan and I. Jeliazkov. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1-2):101–120, 2009.
- J. C. C. Chan and D. P. Kroese. Improved cross-entropy method for estimation. *Statistics and Computing*, 22(5):1031–1040, 2012.
- P. K. Clark. The cyclical component of US economic activity. *The Quarterly Journal of Economics*, 102(4):797–814, 1987.
- T. E. Clark and T. Doh. A Bayesian evaluation of alternative models of trend inflation. *International Journal of Forecasting*, 30(3):426–448, 2014.
- R. J. Gordon. Is US economic growth over? Faltering innovation confronts the six headwinds. Technical report, National Bureau of Economic Research, 2012.
- R. J. Gordon. The demise of US economic growth: Restatement, rebuttal, and reflections. Technical report, National Bureau of Economic Research, 2014.
- A. C. Harvey. Trends and cycles in macroeconomic time series. *Journal of Business and Economic Statistics*, 3(3):216–227, 1985.

- G. Koop. *Bayesian Econometrics*. Wiley & Sons, New York, 2003.
- G. Koop and D. Korobilis. Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics*, 3(4):267–358, 2010.
- G. Koop, D. J. Poirier, and J. L. Tobias. *Bayesian Econometric Methods*. Cambridge University Press, 2007.
- D. P. Kroese and J. C. C. Chan. *Statistical Modeling and Computation*. Springer, New York, 2014.
- D. P. Kroese, T. Taimre, and Z. I. Botev. *Handbook of Monte Carlo Methods*. John Wiley and Sons, New York, 2011.
- S. Luo and R. Startz. Is it one break or ongoing permanent shocks that explains US real GDP? *Journal of Monetary Economics*, 66:155–163, 2014.
- W. J. McCausland, S. Miller, and D. Pelletier. Simulation smoothing for state-space models: A computational efficiency analysis. *Computational Statistics and Data Analysis*, 55(1):199–212, 2011.
- E. Mertens. On the reliability of output gap estimates in real time. *Unpublished manuscript, Federal Reserve Board*, 2014.
- J. C. Morley. The two interpretations of the Beveridge–Nelson decomposition. *Macroeconomic Dynamics*, 15(3):419–439, 2011.
- J. C. Morley and J. Piger. Trend/cycle decomposition of regime-switching processes. *Journal of Econometrics*, 146(2):220 – 226, 2008.
- J. C. Morley and J. Piger. The asymmetric business cycle. *Review of Economics and Statistics*, 94(1):208–221, 2012.
- J. C. Morley, C. R. Nelson, and E. Zivot. Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different? *Review of Economics and Statistics*, 85(2):235–243, 2003.
- K. H. Oh, E. Zivot, and D. Creal. The relationship between the Beveridge-Nelson decomposition and other permanent-transitory decompositions that are popular in economics. *Journal of Econometrics*, 146(2):207–219, 2008.
- P. Perron and T. Wada. Let’s take a break: Trends and cycles in US real GDP. *Journal of Monetary Economics*, 56(6):749–765, 2009.
- R. Y. Rubinstein. Optimization of computer simulation models with rare events. *European Journal of Operational Research*, 99:89–112, 1997.
- R. Y. Rubinstein. The cross-entropy method for combinatorial and continuous optimization. *Methodology and Computing in Applied Probability*, 2:127–190, 1999.

- R. Y. Rubinstein and D. P. Kroese. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization Monte-Carlo Simulation, and Machine Learning*. Springer-Verlag, New York, 2004.
- T. M. Sinclair. The relationships between permanent and transitory movements in US output and the unemployment rate. *Journal of Money, Credit and Banking*, 41(2-3): 529–542, 2009.
- T. M. Sinclair. Asymmetry in the business cycle: Friedman’s plucking model with correlated innovations. *Studies in Nonlinear Dynamics and Econometrics*, 14(1):Article 3, 2010.
- J. H. Stock and M. W. Watson. Why has U.S. inflation become harder to forecast? *Journal of Money Credit and Banking*, 39(s1):3–33, 2007.
- L. Summers. IMF economic forum: Policy responses to crises, 2013. Speech delivered at the IMF Fourteenth Annual Research Conference, 9 November.
- L. Summers. Demand side secular stagnation. *American Economic Review: Papers & Proceedings*, 105(5), 2015.
- M. W. Watson. Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, 18(1):49–75, 1986.