Reconciling Output Gaps: Unobserved Components Model and Hodrick-Prescott Filter*

Angelia L. Grant
Research School of Economics,
Australian National University

Joshua C.C. Chan
Research School of Economics,
Australian National University

December 2016

Abstract

This paper reconciles two widely used trend-cycle decompositions of GDP that give markedly different estimates: the correlated unobserved components model yields output gaps that are small in amplitude, whereas the Hodrick-Prescott (HP) filter generates large and persistent cycles. By embedding the HP filter in an unobserved components model, we show that this difference arises due to differences in the way the stochastic trend is modeled. Moreover, the HP filter implies that the cyclical components are serially independent—an assumption that is decidedly rejected by the data. By relaxing this restrictive assumption, the augmented HP filter provides comparable model fit relative to the standard correlated unobserved components model.

Keywords: trend-cycle decomposition, HP filter, structural break

JEL classification: C11, C52, E32

*We thank seminar participants at Australian National University, University of Queensland and University of New South Wales. In particular, this paper has benefited from the discussion with James Morley and Benjamin Wong. Joshua Chan would also like to acknowledge financial support by the Australian Research Council via a Discovery Early Career Researcher Award (DE150100795).

1 Introduction

The estimation of the output gap—the deviation of the output of an economy from its potential or trend output—is an important problem for both academics and policymakers. A particularly important task is to reconcile the differences between estimation methods, given that different approaches provide vastly different results. Morley, Nelson, and Zivot (2003) make an important contribution to such reconciliations. They show that the difference between two widely used trend-cycle decompositions—the Beveridge and Nelson (1981) decomposition and the unobserved components (UC) model of Watson (1986)—is entirely due to one restriction imposed in the UC model: the correlation between the innovations to the trend and cycle is assumed to be zero. When this restriction is relaxed, they find that the two trend-cycle decompositions are identical. In particular, both methods yield output gaps that are noisy and small in amplitude.

However, output gap estimates produced by the popular Hodrick and Prescott (1980, 1997) decomposition are often large and highly persistent. For example, the Hodrick-Prescott (HP) filter gives an estimate of the US output gap that is as large as -3% at the trough of the Great Recession. In contrast, the corresponding estimate from the correlated UC model of Morley et al. (2003) is close to zero. To reconcile the differences, Kamber, Morley, and Wong (2016) investigate how one can generate large cyclical components using the Beveridge-Nelson decomposition, which typically gives small and noisy cycles. They find that by setting the noise-to-signal ratio to be large, instead of estimating it from the data, the cycles obtained are large and the timing of troughs matches the chronology dated by the National Bureau of Economic Research (NBER).

We take an alternative modeling approach to reconciling the small versus large output gaps generated by different methods. In particular, we address the question of how best to model the trend output and the implications of different specifications on the estimates of the output gap and potential output growth. We start by embedding the HP filter in an UC model similar to that in Morley et al. (2003)—but instead of a random walk output trend, the trend follows a second-order Markov process.[1] By formulating an econometric model for the HP filter and comparing it to other popular UC models, we aim to identify the source of differences in the estimated output gaps.

[1]The idea of embedding the HP filter in an unobserved components model can be traced back to Harvey and Jaeger (1993). However, they consider an unobserved components model where all the innovations are independent. In contrast, the main purpose of this paper is to reconcile the output gaps obtained from the HP filter and the correlated UC model of Morley et al. (2003).

Under the UC framework, we show that the HP filter implicitly makes three assumptions: 1) shocks to the trend growth and cyclical components are uncorrelated; 2) the noise-to-signal ratio is fixed; and 3) the cyclical components are serially uncorrelated. We then construct a series of "augmented HP filters"—UC models that progressively relax these assumptions. By formally comparing these models, we can identify which of these modeling assumptions are not supported by the data.

The aim of this paper is similar to Perron and Wada (2009) and Grant and Chan (2016) in that we assess the support of different trend output specifications. However, while they focus on random walk trend output models with a deterministic break in the growth rate, we consider models with both stochastic trend output and stochastic growth. Within this broader framework, we assess the support of the level shocks to trend output in the presence of a stochastic growth rate.

The estimation results, using US real GDP data, show that the output gap estimated under the UC model with a second-order Markov trend specification is substantially larger in magnitude than that obtained from the correlated UC model of Morley et al. (2003). Given that the HP filter can be embedded into this model, this suggests that the difference between the output gaps obtained from the correlated UC model and the HP filter arises due to differences in the way trend output is modeled. Furthermore, by comparing the augmented HP filters, we show the implicit assumption that the cyclical components are serially uncorrelated is decidedly rejected by the data.

Unlike the correlated UC model, which assumes a constant trend growth rate, the second-order Markov trend specification allows for time-varying trend growth. In the application, we show that there is substantial time variation in the trend output growth rate. Specifically, the annualized trend growth rate fluctuates between 3.5% and 4% from 1950 to 1970; remains stable at about 3% from the mid-1970s to 2000; and starts a gradual decline to about 1.7% in the middle of the Great Recession. These estimates are comparable to those reported in Perron and Wada (2009) and Grant and Chan (2016)—all obtained using UC models with breaks in the trend growth rate.

In contrast, the annualized trend growth rate implied by the HP filter displays conspicuous cyclical patterns: e.g., it drops from about 4% in 1985 to about 2.4% in 1990, only to rebound again to 4% in 1997. Due to the implicit assumption of the HP filter that the cyclical components are serially independent, part of the business cycle variation is absorbed by the time-varying trend growth, resulting in its counter-intuitive large time-variation.

In this paper we focus on the trend-cycle decomposition of real GDP, but our new modeling approach has broader implications. The HP filter continues to be popular, despite concerns that using it to remove low-frequency movements in the data may lead to poor model fit and forecasts (see, e.g., Morley and Piger, 2012; Baştürk, Çakmakli, Ceyhan, and Van Dijk, 2014; Canova, 2014, among many others). It is therefore useful to provide an alternative model-based approach that generalizes the HP filter to accommodate salient features of the data. In addition, using a model-based approach also allows us to formally assess the model fit and compare it with other models.

2 UC Model with a Smoother Trend

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, namely, a nonstationary component that represents the trend and a transitory deviation from the trend. Based on the correlated unobserved components model of Morley et al. (2003), consider the decomposition of the log real GDP y_t:

$$y_t = \tau_t + c_t, \tag{1}$$

where τ_t is the trend and c_t is the stationary, cyclical component. The nonstationary trend τ_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:

$$c_t = \phi_1 c_{t-1} + \cdots + \phi_p c_{t-p} + u_t^c, \tag{2}$$

$$\tau_t = \mu + \tau_{t-1} + \widetilde{u}_t^\tau, \tag{3}$$

where the innovations u_t^c and $\widetilde{u}_t^\tau$ are jointly normal

$$\begin{pmatrix} u_t^c \\ \widetilde{u}_t^\tau \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_c^2 & \rho\sigma_c\widetilde{\sigma}_\tau \\ \rho\sigma_c\widetilde{\sigma}_\tau & \widetilde{\sigma}_\tau^2 \end{pmatrix}\right). \tag{4}$$

Note that the innovation $\widetilde{u}_t^\tau$ impacts the level of the trend output τ_t. Following Morley et al. (2003), we refer to this model as UCUR. The drift μ can be interpreted as the average growth rate of trend output. Morley et al. (2003) show that the trend-cycle decomposition from this correlated unobserved components model is equivalent to the one obtained from the Beveridge and Nelson (1981) decomposition using an unrestricted

ARIMA model—both attribute most of the variance in output to the variation in trend and the cyclical component is small in amplitude.

A limitation of the model is that the trend growth rate μ is a constant, which is not supported by the data. For example, using US real GDP from 1947Q1 to 1998Q2, Perron and Wada (2009) find a break in trend growth at 1973Q1. Using more recent data, Luo and Startz (2014) and Grant and Chan (2016) find a break at 2006Q1 and 2007Q1, respectively. One way to relax this restrictive assumption is to allow for a time-varying growth rate, as considered in Harvey (1985) and Clark (1987), replacing (3) by

$$\tau_t = \mu_t + \tau_{t-1} + v_t^\tau, \tag{5}$$

$$\mu_t = \mu_{t-1} + v_t^\mu. \tag{6}$$

Under this model, the innovation v_t^τ affects the level of the trend output τ_t—it plays the role of $\widetilde{u}_t^\tau$ in the UCUR model—whereas v_t^μ changes the trend growth μ_t. While this variant allows for a time-varying growth rate, Perron and Wada (2009, p. 751) conclude that "this generalization leaves the trend-cycle decompositions virtually unchanged". That is, if the innovations are correlated, this extension is unable to generate output gaps as large as those from the Hodrick-Prescott (HP) filter.

2.1 An Alternative Specification for the Trend Component

Here we consider an alternative unobserved components model that has a different trend specification, with the goal of providing a formal modeling framework for the HP filter. To keep our specification as similar as possible to Morley et al. (2003), we only modify the state equation for τ_t. In particular, instead of the random walk process with drift in (3), we consider the following second-order Markov process:[2]

$$\Delta\tau_t = \Delta\tau_{t-1} + u_t^\tau, \tag{7}$$

[2]Note that (7) can be written as $\tau_t = 2\tau_{t-1} - \tau_{t-2} + u_t^\tau$, hence it is a second-order Markov process. Although formulated differently, this trend specification can be shown to be equivalent to the one in Harvey, Trimbur, and Van Dijk (2007). The cyclical component in their model, however, is different from ours and the associated cycle estimates from both models are drastically different. For example, their estimated trend growth rate of US GDP does not exhibit any discernible long-term trend over the period from 1950s to 2006, whereas ours shows a gradual downward trend.

where Δ is the first difference operator such that $\Delta x_t = x_t - x_{t-1}$. As before we assume the innovations u_t^c and u_t^τ are jointly normal

$$\begin{pmatrix} u_t^c \\ u_t^\tau \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_c^2 & \rho\sigma_c\sigma_\tau \\ \rho\sigma_c\sigma_\tau & \sigma_\tau^2 \end{pmatrix}\right). \tag{8}$$

We refer to this model as UCUR-2M. Since $\Delta\tau_t$ can be interpreted as the trend output growth at time t, the specification in (7) implies that the growth rate follows a random walk. As such, by construction the UCUR-2M model incorporates a stochastic trend growth rate, in contrast to the constant trend growth rate implied by (3). The former specification is more flexible and can accommodate breaks in trend output growth found in some recent studies such as Perron and Wada (2009) and Luo and Startz (2014).

We emphasize that UCUR-2M differs from the correlated unobserved components model in Morley et al. (2003) in only the state equation for τ_t. That is, the proposed model is defined by (1), (2), (7) and (8). In addition, we follow Morley et al. (2003) and set $p = 2$. Furthermore, the initial values τ_{-1} and τ_0 are treated as parameters to be estimated, and for simplicity we assume that $c_{-1} = c_0 = 0$.

It is instructive to compare the second-order Markov transition for the trend in (7) to the specification in (5)-(6), which also allows for a time-varying growth rate. After some algebra, one can show that the latter specification implies

$$\Delta\tau_t - \Delta\tau_{t-1} = v_t^\mu + v_t^\tau - v_{t-1}^\tau,$$

i.e., the change in trend growth follows a moving average process. The specification in (7) can therefore be viewed as a limiting case when the variance of v_t^τ goes to zero. That is, the UCUR-2M model classifies all permanent shocks as shocks to the trend growth.

2.2 Augmented HP Filters

The Hodrick-Prescott decomposition is based on the smoothing problem initiated by Bohlmann (1899) but often credited to Whittaker (1923). Given a positive integer q and a smoothing parameter λ, the trend is the solution of the following minimization problem:

$$\underset{\boldsymbol{\tau}}{\operatorname{argmin}} \left[\sum_{t=1}^{T}(y_t - \tau_t)^2 + \lambda\sum_{t=1}^{T}(\Delta^q\tau_t)^2\right].$$

When $q = 2$, the solution to this smoothing problem is the HP trend, which we denote as $\widehat{\boldsymbol{\tau}}_{\text{HP}}$. Here λ is a fixed constant that penalizes variability in the trend component. The larger the value of λ, the smoother is the associated HP trend $\widehat{\boldsymbol{\tau}}_{\text{HP}}$. Hodrick and Prescott (1980, 1997) highlight that λ may be viewed as the noise-to-signal ratio under certain restrictive conditions—namely, that "the cyclical components and the second differences of the [trend] components were identically and independently distributed." They then suggest setting $\lambda = 1600$ for US quarterly data.

In Appendix B we show that the UCUR-2M model nests the HP filter as a special case by fixing the parameters to certain values. We summarize this result in the following proposition.

Proposition 1. *Consider the unobserved components model with a second-order Markov transition defined by* (1), (2), (4) *and* (7). *The Hodrick-Prescott trend* $\widehat{\boldsymbol{\tau}}_{\text{HP}}$ *is the posterior mean of* $\boldsymbol{\tau}$ *by fixing* $\rho = 0$, $\boldsymbol{\phi} = \mathbf{0}$ *and* $\lambda = \sigma_c^2/\sigma_\tau^2$.

To evaluate the support from data for each of the three assumptions underlying the HP filter, we construct a series of nested augmented HP filters as follows. Starting from the HP filter with $\lambda = 1600$, we first allow the AR coefficients $\boldsymbol{\phi}$ to be estimated rather than setting to zero—it amounts to allowing the cyclical components to be serially correlated. We call this version HP-AR. Then, we further allow λ to be estimated, and this version is called UC-2M. Finally, we have the fully flexible UCUR-2M where ρ, $\boldsymbol{\phi}$, σ_c^2 and σ_τ^2 are all estimated from the data. Results from these models are presented in the next section.

3 Empirical Results

In this section we report the cycle estimates and other parameters of interest under the unobserved components model with a second-order Markov trend transition (UCUR-2M). We compare these estimates with those obtained under standard benchmarks to assess the impact of different trend specifications. In particular, we present results from the correlated unobserved components (UCUR) model of Morley et al. (2003). Moreover, given the findings in Perron and Wada (2009), Luo and Startz (2014) and Grant and Chan (2016), we also consider a version in which the trend growth rate has a break at t_0. More precisely, we replace (3) with

$$\tau_t = \mu(t < t_0) + \gamma 1(t \geqslant t_0) + \tau_{t-1} + \widetilde{u}_t^\tau,$$

where $1(A)$ is the indicator function that takes the value 1 if the condition A is true and 0 otherwise. In other words, the stochastic trend τ_t has a growth rate of μ before the break t_0 and a growth rate of γ after the break. We set the break date at 2007Q1 and refer to this model as UCUR-07.

We also present results from the unobserved components model with a random walk growth rate as in Clark (1987), which we denote as UC-LS for unobserved components model with a local slope. Finally, we include also the augmented HP filters HP-AR (HP filter with correlated cycles) and UC-2M (HP filter with correlated cycles and free λ).

We use US quarterly real GDP from 1947Q1 to 2014Q4 for our analysis. The data are sourced from the Federal Reserve Bank of St. Louis economic database, and the series is then transformed by taking the logs and multiplying by 100. Each set of results below is based on 100000 posterior draws after a burn-in period of 10000.

3.1 Parameter Estimates and Model Comparison

Figure 1 presents the contour plot of the bivariate posterior density $p(\phi_1, \phi_2 \,|\, \mathbf{y})$ under the UCUR-2M model.[3] It is evident that most of the mass of this density is concentrated around $(1.3, -0.4)$ and there is no mass near the origin.

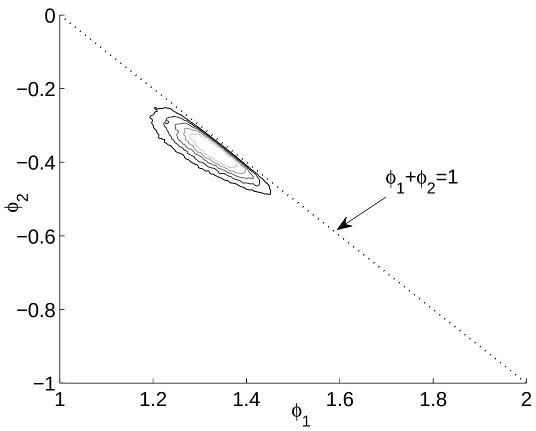

Figure 1: Contour plot of the bivariate posterior density $p(\phi_1, \phi_2 \,|\, \mathbf{y})$ under UCUR-2M.

This can be viewed as evidence that the assumption $\boldsymbol{\phi} = \mathbf{0}$ implied by the HP filter is

[3] Posterior draws of $\boldsymbol{\phi} = (\phi_1, \phi_2)'$ are first obtained using the MCMC sampler. These draws are then used to compute the density using the kernel density estimator of Botev, Grotowski, and Kroese (2010).

strongly rejected by the data. (Below we also perform a formal model comparison.) It is also worth noting that one of the stationarity restrictions, namely, $\phi_1 + \phi_2 < 1$, appears to be binding, reflecting the fact that the cyclical components are highly persistent.

Next, we report in Table 1 the parameter estimates under the UCUR-2M model, as well as estimates from the other six models for comparison. Consistent with the contour plot in Figure 1, the posterior means of ϕ_1 and ϕ_2 under UCUR-2M are estimated to be 1.31 and -0.37 respectively with small standard errors. Moreover, the estimates of ϕ_1 and ϕ_2 are similar across all models that have a serially correlated cyclical component, highlighting that the cycles are highly persistent regardless of how the trend is modeled.

Table 1: Estimated posterior means for selected models. Numerical standard errors are in parentheses.

	UCUR	URUR-07	UC-LS	UCUR-2M	UC-2M	HP-AR	HP
μ	0.78 (0.08)	0.84 (0.08)	–	–		–	–
γ	–	0.37 (0.20)	–	–		–	–
ϕ_1	0.95 (0.34)	1.10 (0.36)	1.44 (0.11)	1.31 (0.07)	1.32 (0.06)	1.33 (0.06)	–
ϕ_2	-0.36 (0.18)	-0.44 (0.18)	-0.50 (0.11)	-0.37 (0.06)	-0.37 (0.06)	-0.37 (0.06)	–
σ_c^2	1.12 (0.55)	0.90 (0.49)	0.50 (0.18)	0.76 (0.07)	0.76 (0.07)	0.77 (0.07)	2.30 (0.15)
$\widetilde{\sigma}_\tau^2$	1.85 (0.49)	1.42 (0.59)	0.22 (0.15)	–	–	–	–
σ_τ^2	–	–		0.0028 (0.002)	0.0018 (0.002)	–	–
σ_μ^2			0.0018 (0.002)			–	–
ρ	-0.87 (0.07)	-0.76 (0.25)		-0.01 (0.55)	–	–	–

Under UCUR-2M the estimate of σ_τ^2 is 0.0028. This means that the estimated standard deviation of the shock to the trend growth is about 0.052. Given the Gaussian state equation, this implies that the difference $\Delta\tau_t - \Delta\tau_{t-1}$ is within $(-0.1, 0.1)$ with about probability 0.95. That is, the annualized quarterly change in trend growth is between -0.4% and 0.4% with about probability 0.95.

In addition, the estimate of σ_c^2 is 0.76, giving an estimate for the noise-to-signal ratio of about 271, which is smaller than the typical value of $\lambda = 1600$ used in the HP filter. However, the standard error associated with the estimate of σ_τ^2 is relatively large, indicating substantial parameter certainty and it is not immediately apparent that λ is statistically different from 1600. Below we formally test the hypothesis that $\lambda = 1600$. Finally, we note that under UCUR-2M, the estimated ρ is close to 0 with a large standard error, indicating that the innovations to the cyclical components and the trend growth rate are uncorrelated. Indeed, as the plot of the prior and posterior densities of ρ in Appendix C shows, even though the posterior of ρ is relatively flat, it has more mass around 0 compared to the prior.

We now compare the seven models in fitting US real GDP in a formal Bayesian model comparison using the marginal likelihood. The marginal likelihood can be interpreted as a density forecast of the data under the model evaluated at the actual observed data. Hence, if the observed data are likely under the model, the associated marginal likelihood would be "large". For a more detailed discussion, see, e.g., Koop (2003) and Geweke and Amisano (2011). The marginal likelihoods are computed using the adaptive importance sampling method proposed in Chan and Eisenstat (2015). The results are reported in Table 2.

Table 2: Log marginal likelihoods of competing models. Numerical standard errors are in parentheses.

HP	HP-AR	UC-2M	UCUR-2M	UC-LS	UCUR	UCUR-07
-601.1	-368.4	-369.8	-369.8	-370.1	-365.0	-364.0
(0.37)	(0.01)	(0.01)	(0.02)	(0.04)	(0.03)	(0.04)

According to the marginal likelihood, the HP filter is by far the worst model. This supports the conclusion in Morley and Piger (2012, p.214) that the HP filter is "strongly at odds with the data." By simply relaxing the assumption $\boldsymbol{\phi} = \mathbf{0}$, the HP-AR model substantially improves the model-fit over the HP filter. This is consistent with the estimation results above, which indicate that the cyclical component is highly persistent.

Beyond allowing for a serially correlated cyclical component, there is little gain in relaxing the assumptions $\lambda = 1600$ or $\rho = 0$. Specifically, HP-AR is nested within UC-2M with $\lambda = 1600$, but the marginal likelihood of the latter is in fact worse than that of the former. This is an example where a more flexible model does worse—this shows that the marginal likelihood has a built-in penalty for model complexity. Similarly, UC-2M is nested within UCUR-2M with $\rho = 0$, but the marginal likelihoods for the two models are the same, suggesting equal evidence in support of and against the hypothesis $\rho = 0$.

Recall that UC-LS includes UC-2M as a limiting case when the variance of the level shock to the trend becomes zero. The marginal likelihood indicates that the more restricted UC-2M does slightly better, showing that the data favor the second-Markov trend specification over the more general specification where the trend has a random walk growth rate.

Finally, the UCUR-07 model with a break in 2007Q1 is the best model, followed by UCUR and HP-AR. The data generally prefer UCUR or UCUR-07 over HP-AR, but the evidence is not overwhelming. As we show in the next section, UCUR-07 and UCUR give small output gaps, whereas those from HP-AR are much larger in magnitude. If the researcher prefers to have larger output gaps, the HP-AR model seems to be the best choice—it produces a more persistent output gap at only a slight cost of model fit.

3.2 Output Gap Estimates

Figure 2 plots the estimates of the output gap c_t under the UCUR and UCUR-07 models. It is evident that the estimates from both models are remarkably similar until the mid-2000s, when the UCUR-07 model allows for a break in trend growth in 2007Q1.

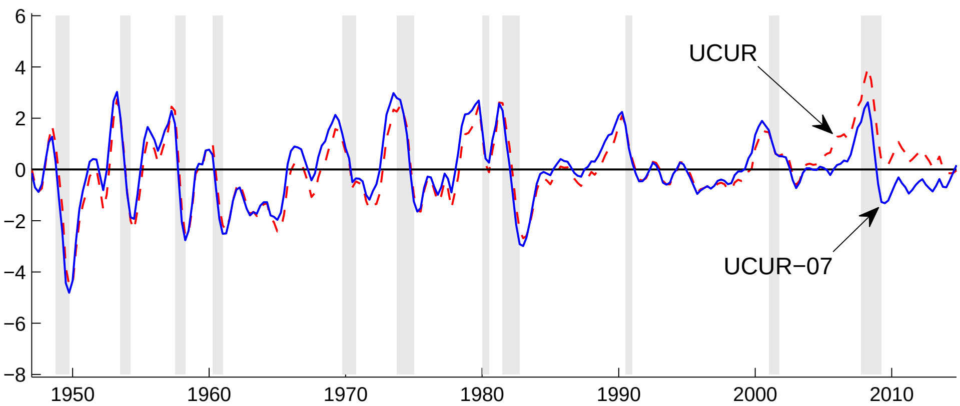

Figure 2: Estimates of the output gap c_t under UCUR and UCUR-07. The shaded regions are the NBER recession dates.

This finding contrasts with the results in Perron and Wada (2009), who find that when a break is allowed for, the estimates from correlated UC models with and without a break differ substantially. One source of difference is that their output gap estimates are computed given the maximum likelihood estimates, thus ignoring parameter uncertainty.

In contrast, ours take into account parameter uncertainty by averaging the output gap estimates over posterior draws of the model parameters.

The output gap estimates in Figure 2 generally coincide with the NBER recession dates, but they are relatively small. For example, the trough of the 1981-1982 recession is only about -3%; during the Great Recession the output gap estimates are about -1.3% for the UCUR-07 model and close to 0 for the UCUR.

The corresponding estimates under the UCUR-2M model are reported in Figure 3. It is clear that the plot of the output gap has a similar shape as that in Figure 2; it traces the NBER recession dates closely as well. The main difference is that the new output gap estimates are noticeably larger. For instance, the trough of the 1981-1982 recession is now at about -7% and the trough in the Great recession is about -4%.

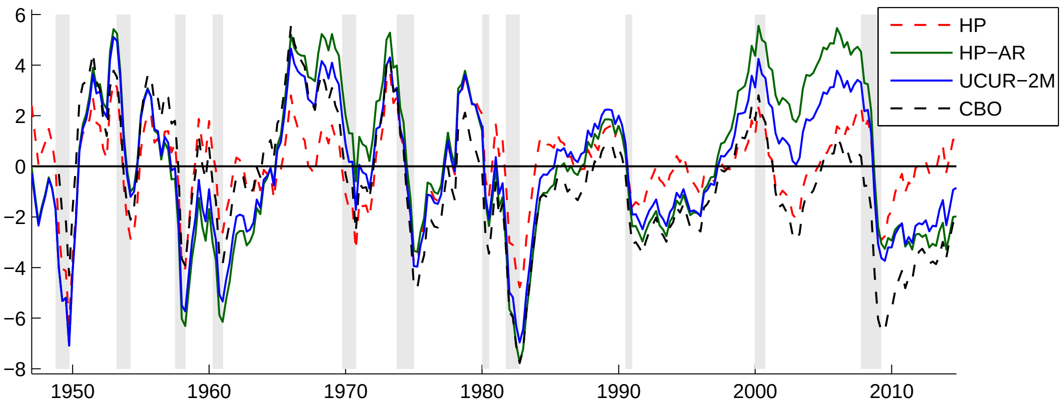

Figure 3: Estimates of the output gap c_t under UCUR-2M, HP-AR and the HP filter, as well as the estimates from the Congressional Budget Office. The shaded regions are the NBER recession dates.

For comparison, we also plot the output gap estimates under the HP-AR model and estimates from the Congressional Budget Office (CBO). The three estimates are remarkably similar before late 1990s. For example, at the trough of the 1980s recession, all three estimates are between -7% and -8%. However, they diverge in the late 1990s—the two UC models give sizable positive output gaps throughout the 2000s whereas the estimates from the CBO are much smaller.

Compared to the HP filter, the output gap estimates under the UCUR-2M model are typically larger and more persistent. For instance, the HP filter estimates drop to 0 in 2011-2012, whereas the UCUR-2M model shows a persistent output gap—it remains around -0.9% at 2014Q4. This difference is due to the assumption that the cyclical components are independent under the HP filter, i.e., $\boldsymbol{\phi} = \mathbf{0}$. In contrast, when we allow $\boldsymbol{\phi}$ to be estimated as in the UCUR-2M model, the estimates are far from zero. This in turn makes the output gap estimate under the UCUR-2M model more persistent.

3.3 Trend Output Growth Estimates

Next, we present the estimated annualized trend output growth rate under UCUR-2M—the posterior means of $4\Delta\tau_t = 4(\tau_t - \tau_{t-1})$—in Figure 4. It is evident that there is substantial time variation in trend growth over the past six decades. In particular, the annualized trend growth rate fluctuates between 3.5% and 4% from the beginning of the sample until 1970. It then begins a steady decline and reaches about 3% in the mid-1970s. These estimates are similar to those in Berger, Everaert, and Vierke (2016), who find substantial time variation in trend output growth using a trivariate unobserved components model of output, inflation and unemployment rate.

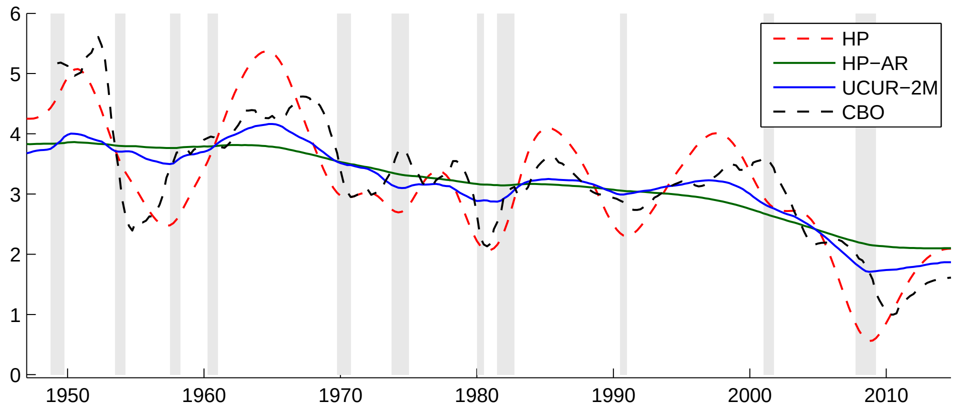

Figure 4: Estimates of the annualized growth in trend $4\Delta\tau_t$ under the UCUR-2M model and HP filter. The shaded regions are the NBER recession dates.

The estimated trend growth rate remains stable at about 3% from the mid-1970s until 2000, when it starts another gradual decline to about 1.7% in the middle of the Great Recession. These results are similar to the estimated trend growth rates reported in Grant and Chan (2016)—they find that the trend growth rate drops from 3.4% before 2007 to about 1.5% afterward. Our estimate is also in line with the forecast of US potential output growth of 1.5% to 1.55% from 2007 to 2032 given in Gordon (2014).

The estimates under HP-AR are also similar, although they tend to be smoother than those under UCUR-2M.

For comparison we also plot the annualized trend growth rate implied by the HP filter and the estimates from the CBO in Figure 4. The most prominent feature of the HP filter estimates is the apparent cyclical pattern, where the fluctuations are much more pronounced than the two UC models as well as the CBO estimates. For instance, the growth rate under the HP filter drops from about 4% in 1985 to about 2.4% in 1990, only to rebound again to 4% in 1997.

This large variation in trend growth might not be surprising as the HP trend preserves any movements whose corresponding frequency is below a certain level. When the cyclical components are very persistent, some of the variation becomes part of the trend component by the definition of the HP filter. Hence, this result is quite natural from the view point of spectral analysis, as summarized and documented in King and Rebelo (1993) and Harvey and Jaeger (1993). In contrast, when we allow the cyclical components to be correlated via an AR(2) process in the UCUR-2M model, the estimated trend growth rate is much smoother and picks up only some very low-frequency movement.

4 Concluding Remarks and Future Research

We formulate a new correlated unobserved components model with a second-order Markov process and show one can recover the HP trend as a special case. Using this model-based approach, we directly compare the HP filter with other popular unobserved components models and shed light on the source of differences in the cycle estimates. We show that by relaxing the implicit assumption that the cyclical components are independent under the HP filter, the new model provides comparable model fit relative to the standard correlated unobserved components model.

Many recent papers, including Carriero, Kapetanios, and Marcellino (2009), Banbura, Giannone, and Reichlin (2010) and Koop (2013), have demonstrated the gains of incorporating the information content in a large number of macroeconomic variables. For future research, it would be interesting to embed the proposed UC model in a large Bayesian vector autoregression to assess if other macroeconomic variables provide additional information about the output gap.

Online Appendix

Appendix A: Estimation Details

This appendix discusses the priors and provides the estimation details of the unobserved components model with a second-order Markov trend specification. The sampler discussed is based on the posterior simulator developed in Grant and Chan (2016) for fitting the correlated unobserved components model in Morley et al. (2003). A key novel feature of our approach is that it draws on recent advances in band matrix algorithms developed in Chan and Jeliazkov (2009), McCausland, Miller, and Pelletier (2011) and Chan (2013), which are shown to be more efficient than the conventional Kalman filter-based algorithms.

We assume proper but relatively noninformative priors for the model parameters $\boldsymbol{\phi} = (\phi_1, \phi_2)', \sigma_c^2, \sigma_\tau^2, \rho, \tau_0$ and τ_{-1}. In particular, we assume independent priors for $\boldsymbol{\phi}$, τ_0 and τ_{-1}:

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

where $\mathbf{R}$ is the stationarity region. The prior on the AR coefficients $\boldsymbol{\phi}$ affects how persistent the cyclical components are. We assume relatively large prior variance $\mathbf{V}_{\boldsymbol{\phi}} = \mathbf{I}_2$ so that *a priori* $\boldsymbol{\phi}$ can take on a wide range of values. The prior mean is assumed to be $\boldsymbol{\phi}_0 = (1.3, -0.7)'$, which implies that the AR(2) process has two complex roots. The prior mean is similar to the values reported in, e.g., Morley et al. (2003). On the other hand, the hyperparameters V_τ and τ_{00} are relatively unimportant provided that V_τ is sufficiently large, as they only affect the initial values. We set $V_\tau = 100$ and $\tau_{00} = 750$.

Next, we assume the priors on σ_c^2, σ_τ^2 and ρ to be uniform:

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

Since the correlation coefficient ρ is bounded between -1 and 1, the uniform prior on $(-1, 1)$ is a natural choice. The priors on σ_c^2 and σ_τ^2 affect the size of the shocks to the cycle and the trend growth rate, respectively. From estimates of similar parameters reported in Morley et al. (2003), the upper bounds $b_c = 3$ and $b_\tau = 0.01$ seem to be sufficiently large.

Next, we provide the details of the posterior sampler. To that end, stack $\mathbf{y} = (y_1, \ldots, y_T)'$ and $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_T)'$. Then, posterior draws can be obtained by sequentially sampling from the following densities: 1. $p(\boldsymbol{\tau} \,|\, \mathbf{y}, \boldsymbol{\phi}, \sigma^2_c, \sigma^2_\tau, \rho, \tau_0, \tau_{-1})$; 2. $p(\boldsymbol{\phi} \,|\, \mathbf{y}, \boldsymbol{\tau}, \sigma^2_c, \sigma^2_\tau, \rho, \tau_0, \tau_{-1})$; 3. $p(\sigma^2_c \,|\, \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma^2_\tau, \rho, \tau_0, \tau_{-1})$; 4. $p(\sigma^2_\tau \,|\, \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma^2_c, \rho, \tau_0, \tau_{-1})$; 5. $p(\rho \,|\, \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma^2_c, \sigma^2_\tau, \tau_0, \tau_{-1})$; 6. $p(\tau_0, \tau_{-1} \,|\, \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma^2_c, \sigma^2_\tau, \rho)$.

To implement Step 1, let $\mathbf{c} = (c_1, \ldots, c_T)'$, and similarly define $\mathbf{u}^c$ and $\mathbf{u}^\tau$. Then, we rewrite the system (1), (2) and (7) in the following matrix form:

$$\begin{aligned} \mathbf{y} &= \boldsymbol{\tau} + \mathbf{c}, \\ \mathbf{H}_{\boldsymbol{\phi}}\mathbf{c} &= \mathbf{u}^c, \\ \mathbf{H}_2\boldsymbol{\tau} &= \widetilde{\boldsymbol{\alpha}} + \mathbf{u}^\tau, \end{aligned}$$

where $\widetilde{\boldsymbol{\alpha}} = (2\tau_0 - \tau_{-1}, -\tau_0, 0, \ldots, 0)'$ and

$$\mathbf{H}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}, \quad \mathbf{H}_{\boldsymbol{\phi}} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\phi_1 & 1 & 0 & 0 & \cdots & 0 \\ -\phi_2 & -\phi_1 & 1 & 0 & \cdots & 0 \\ 0 & -\phi_2 & -\phi_1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\phi_2 & -\phi_1 & 1 \end{pmatrix}.$$

Note that both $\mathbf{H}_2$ and $\mathbf{H}_{\boldsymbol{\phi}}$ are band matrices with only a few nonzero elements arranged around the main diagonal. Further, since both are square matrices with unit determinant, they are invertible. Hence, given the parameters $\boldsymbol{\phi}$, σ^2_c, σ^2_τ, ρ, τ_0 and τ_{-1}, it follows from (4) that we have

$$\begin{pmatrix} \mathbf{c} \\ \boldsymbol{\tau} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0} \\ \boldsymbol{\alpha} \end{pmatrix}, \begin{pmatrix} \sigma^2_c(\mathbf{H}_{\boldsymbol{\phi}}'\mathbf{H}_{\boldsymbol{\phi}})^{-1} & \rho\sigma_c\sigma_\tau(\mathbf{H}_2'\mathbf{H}_{\boldsymbol{\phi}})^{-1} \\ \rho\sigma_c\sigma_\tau(\mathbf{H}_{\boldsymbol{\phi}}'\mathbf{H}_2)^{-1} & \sigma^2_\tau(\mathbf{H}_2'\mathbf{H}_2)^{-1} \end{pmatrix} \right),$$

where $\boldsymbol{\alpha} = \mathbf{H}_2^{-1}\widetilde{\boldsymbol{\alpha}}$. 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

$$(\boldsymbol{\tau} \,|\, \sigma^2_\tau, \tau_0, \tau_{-1}) \sim \mathcal{N}(\boldsymbol{\alpha}, \sigma^2_\tau(\mathbf{H}_2'\mathbf{H}_2)^{-1}),$$

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

$$(\mathbf{y} \,|\, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma^2_c, \sigma^2_\tau, \rho, \tau_0, \tau_{-1}) \sim \mathcal{N}\left(\mathbf{H}_{\boldsymbol{\phi}}^{-1}\mathbf{a} + \mathbf{H}_{\boldsymbol{\phi}}^{-1}\mathbf{B}\boldsymbol{\tau}, (1-\rho^2)\sigma^2_c(\mathbf{H}_{\boldsymbol{\phi}}'\mathbf{H}_{\boldsymbol{\phi}})^{-1}\right),$$

where

$$\mathbf{a} = -\frac{\rho\sigma_c}{\sigma_\tau}\mathbf{H}_2\boldsymbol{\alpha}, \quad \mathbf{B} = \mathbf{H}_{\boldsymbol{\phi}} + \frac{\rho\sigma_c}{\sigma_\tau}\mathbf{H}_2.$$

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

$$p(\boldsymbol{\tau} \,|\, \sigma_\tau^2, \tau_0, \tau_{-1}) = (2\pi\sigma_\tau^2)^{-\frac{T}{2}}\mathrm{e}^{-\frac{1}{2\sigma_\tau^2}(\boldsymbol{\tau}-\boldsymbol{\alpha})'\mathbf{H}_2'\mathbf{H}_2(\boldsymbol{\tau}-\boldsymbol{\alpha})} \tag{9}$$

$$p(\mathbf{y} \,|\, \boldsymbol{\tau}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \tau_0, \tau_{-1}) = (2\pi\sigma_y^2(1-\rho^2))^{-\frac{T}{2}}\mathrm{e}^{-\frac{1}{2(1-\rho^2)\sigma_c^2}(\mathbf{H}_{\boldsymbol{\phi}}\mathbf{y}-\mathbf{a}-\mathbf{B}\boldsymbol{\tau})'(\mathbf{H}_{\boldsymbol{\phi}}\mathbf{y}-\mathbf{a}-\mathbf{B}\boldsymbol{\tau})}. \tag{10}$$

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

$$(\boldsymbol{\tau} \,|\, \mathbf{y}, \boldsymbol{\phi}, \sigma_c^2, \sigma_\tau^2, \rho, \tau_0, \tau_{-1}) \sim \mathcal{N}(\widehat{\boldsymbol{\tau}}, \mathbf{K}_{\boldsymbol{\tau}}^{-1}), \tag{11}$$

where

$$\mathbf{K}_{\boldsymbol{\tau}} = \frac{1}{\sigma_\tau^2}\mathbf{H}_2'\mathbf{H}_2 + \frac{1}{(1-\rho^2)\sigma_c^2}\mathbf{B}'\mathbf{B}, \quad \widehat{\boldsymbol{\tau}} = \mathbf{K}_{\boldsymbol{\tau}}^{-1}\left(\frac{1}{\sigma_\tau^2}\mathbf{H}_2'\mathbf{H}_2\boldsymbol{\alpha} + \frac{1}{(1-\rho^2)\sigma_c^2}\mathbf{B}'(\mathbf{H}_{\boldsymbol{\phi}}\mathbf{y}-\mathbf{a})\right).$$

Since $\mathbf{H}_2$, $\mathbf{H}_{\boldsymbol{\phi}}$ and $\mathbf{B}$ are all band matrices, so is the precision matrix $\mathbf{K}_{\boldsymbol{\tau}}$. As such, the precision sampler of Chan and Jeliazkov (2009) can be used to sample $\boldsymbol{\tau}$ efficiently.

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

$$\begin{pmatrix}\mathbf{u}^c\\ \boldsymbol{\tau}\end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix}\mathbf{0}\\ \boldsymbol{\alpha}\end{pmatrix}, \begin{pmatrix}\sigma_c^2\mathbf{I}_T & \rho\sigma_c\sigma_\tau(\mathbf{H}_2')^{-1}\\ \rho\sigma_c\sigma_\tau\mathbf{H}_2^{-1} & \sigma_\tau^2(\mathbf{H}_2'\mathbf{H}_2)^{-1}\end{pmatrix}\right), \tag{12}$$

where $\boldsymbol{\alpha} = \mathbf{H}_2^{-1}\widetilde{\boldsymbol{\alpha}}$ with $\widetilde{\boldsymbol{\alpha}} = (2\tau_0 - \tau_{-1}, -\tau_0, 0, \ldots, 0)'$. Hence, the conditional distribution of $\mathbf{u}^c$ given $\boldsymbol{\tau}$ and other parameters is

$$(\mathbf{u}^c \,|\, \boldsymbol{\tau}, \sigma_c^2, \sigma_\tau^2, \rho, \tau_0, \tau_{-1}) \sim \mathcal{N}\left(\frac{\rho\sigma_c}{\sigma_\tau}\mathbf{H}_2(\boldsymbol{\tau}-\boldsymbol{\alpha}), (1-\rho^2)\sigma_c^2\mathbf{I}_T\right).$$

Next, we write (2) as

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

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

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

where

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

$$\widehat{\boldsymbol{\phi}} = \mathbf{K}_{\boldsymbol{\phi}}^{-1}\left(\mathbf{V}_{\boldsymbol{\phi}}^{-1}\boldsymbol{\phi}_0 + \frac{1}{(1-\rho^2)\sigma_c^2}\mathbf{X}_{\boldsymbol{\phi}}'\left(\mathbf{c} - \frac{\rho\sigma_c}{\sigma_\tau}\mathbf{H}_2(\boldsymbol{\tau}-\boldsymbol{\alpha})\right)\right).$$

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

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 σ_c^2, σ_τ^2 and ρ, we can factor the joint distribution of (u_t^c, u_t^τ) 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

$$\begin{aligned} p(\mathbf{u}^c, \mathbf{u}^\tau \,|\, \sigma_c^2, \sigma_\tau^2, \rho) &\propto (\sigma_\tau^2)^{-\frac{T}{2}} \mathrm{e}^{-\frac{1}{2\sigma_\tau^2}\sum_{t=1}^T (u_t^\tau)^2}((1-\rho^2)\sigma_c^2)^{-\frac{T}{2}} \mathrm{e}^{-\frac{1}{2(1-\rho^2)\sigma_c^2}\sum_{t=1}^T \left(u_t^c - \frac{\rho\sigma_c}{\sigma_\tau}u_t^\tau\right)^2}, \\ &= ((1-\rho^2)\sigma_c^2\sigma_\tau^2)^{-\frac{T}{2}} \mathrm{e}^{-\frac{1}{2\sigma_\tau^2}k_3 - \frac{1}{2(1-\rho^2)\sigma_c^2}\left(k_1 - \frac{2\rho\sigma_c}{\sigma_\tau}k_2 + \frac{\rho^2\sigma_c^2}{\sigma_\tau^2}k_3\right)}, \end{aligned} \tag{13}$$

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 (13) that

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

where $p(\sigma_c^2)$ is the uniform prior specified above. This full conditional density of σ_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

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

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

Lastly, to jointly sample τ_0 and τ_{-1}, note that we can write $\boldsymbol{\alpha} = \mathbf{X}_{\boldsymbol{\delta}}\boldsymbol{\delta}$, where $\boldsymbol{\delta} = (\tau_0, \tau_{-1})'$

and

$$\mathbf{X}_{\boldsymbol{\delta}} = \begin{pmatrix} 2 & -1 \\ 3 & -2 \\ \vdots & \vdots \\ T+1 & -T \end{pmatrix}.$$

It follows from (12) that the conditional distribution of $\boldsymbol{\tau}$ given $\mathbf{u}^c$ and other parameters is

$$(\boldsymbol{\tau} \,|\, \mathbf{u}^c, \sigma_c^2, \sigma_\tau^2, \rho, \tau_0, \tau_{-1}) \sim \mathcal{N}\left(\mathbf{X}_{\boldsymbol{\delta}}\boldsymbol{\delta} + \frac{\rho\sigma_\tau}{\sigma_y}\mathbf{H}_2^{-1}\mathbf{u}^c, (1-\rho^2)\sigma_\tau^2(\mathbf{H}_2'\mathbf{H}_2)^{-1}\right).$$

Then, by standard regression results, we have

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

where

$$\mathbf{K}_{\boldsymbol{\delta}} = \mathbf{V}_{\boldsymbol{\delta}}^{-1} + \frac{1}{(1-\rho^2)\sigma_\tau^2}\mathbf{X}_{\boldsymbol{\delta}}'\mathbf{H}_2'\mathbf{H}_2\mathbf{X}_{\boldsymbol{\delta}},$$
$$\widehat{\boldsymbol{\delta}} = \mathbf{K}_{\boldsymbol{\delta}}^{-1}\left(\mathbf{V}_{\boldsymbol{\delta}}^{-1}\boldsymbol{\delta}_0 + \frac{1}{(1-\rho^2)\sigma_\tau^2}\mathbf{X}_{\boldsymbol{\delta}}'\mathbf{H}_2'\mathbf{H}_2\left(\boldsymbol{\tau} - \frac{\rho\sigma_\tau}{\sigma_y}\mathbf{H}_2^{-1}\mathbf{u}^c\right)\right),$$

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

Appendix B: HP Filter as an UC Model

In this appendix we show that one can recover the HP trend $\widehat{\boldsymbol{\tau}}_{\mathrm{HP}}$ as a posterior mean of $\boldsymbol{\tau}$ under the model defined in (1), (2) and (7).[4] To that end, first note that

$$\sum_{t=1}^{T}(\Delta\tau_t - \Delta\tau_{t-1})^2 = (\mathbf{H}_2\boldsymbol{\tau} - \widetilde{\boldsymbol{\alpha}})'(\mathbf{H}_2\boldsymbol{\tau} - \widetilde{\boldsymbol{\alpha}}) = (\boldsymbol{\tau} - \boldsymbol{\alpha})'\mathbf{H}_2'\mathbf{H}_2(\boldsymbol{\tau} - \boldsymbol{\alpha}),$$

[4]It is unclear whether Hodrick and Prescott (1980, 1997) had a statistical model in mind when they proposed the Hodrick-Prescott decomposition. On the one hand, they realized that their minimization problem can be solved by the Kalman filter—hence, it has a state space form. On the other hand, they explicitly stated that "[o]ur statistical approach does not utilize standard time series analysis. Our prior knowledge concerning the process generating the data is not of the variety that permits us to specify a probability model as required for application of that analysis."

The Bayesian interpretation of the HP filter as a posterior mean under a certain prior for $\boldsymbol{\tau}$ has a long history. For example, papers by Kitagawa and Gersch (1984) and Gersch (1993) derive an explicit expression for a smoothing prior for $\boldsymbol{\tau}$ that gives the HP trend.

where $\boldsymbol{\alpha} = \mathbf{H}_2^{-1}\widetilde{\boldsymbol{\alpha}}$ and $\widetilde{\boldsymbol{\alpha}} = (2\tau_0 - \tau_{-1}, -\tau_0, 0, \ldots, 0)'$. Hence, the minimization problem can be rewritten as

$$\widehat{\boldsymbol{\tau}}_{\text{HP}} \equiv \underset{\boldsymbol{\tau}}{\text{argmin}} \left[(\mathbf{y} - \boldsymbol{\tau})'(\mathbf{y} - \boldsymbol{\tau}) + \lambda(\boldsymbol{\tau} - \boldsymbol{\alpha})'\mathbf{H}_2'\mathbf{H}_2(\boldsymbol{\tau} - \boldsymbol{\alpha}) \right].$$

By differentiating the objective function with respect to $\boldsymbol{\tau}$, we obtain the following first-order condition:

$$\frac{\partial}{\partial \boldsymbol{\tau}} \left[\boldsymbol{\tau}'\boldsymbol{\tau} - 2\boldsymbol{\tau}'\mathbf{y} + \lambda\boldsymbol{\tau}'\mathbf{H}_2'\mathbf{H}_2\boldsymbol{\tau} - 2\lambda\boldsymbol{\tau}'\mathbf{H}_2'\mathbf{H}_2\boldsymbol{\alpha}) \right] = 2(\mathbf{I}_T + \lambda\mathbf{H}_2'\mathbf{H}_2)\boldsymbol{\tau} - 2(\mathbf{y} + \lambda\mathbf{H}_2'\mathbf{H}_2\boldsymbol{\alpha}).$$

Setting the first-order condition to zero and solving for $\boldsymbol{\tau}$, we obtain

$$\widehat{\boldsymbol{\tau}}_{\text{HP}} = (\mathbf{I}_T + \lambda\mathbf{H}_2'\mathbf{H}_2)^{-1}(\mathbf{y} + \lambda\mathbf{H}_2'\mathbf{H}_2\boldsymbol{\alpha}). \tag{14}$$

Since the Hessian is $2(\mathbf{I}_T + \lambda\mathbf{H}_2'\mathbf{H}_2)$, which is always positive definite provided that $\lambda > 0$, $\widehat{\boldsymbol{\tau}}_{\text{HP}}$ is the unique minimizer. Also, the expression (14) provides a quick way to compute the HP trend as all $T \times T$ matrices are banded. See also Weinert (2007) for an alternative way to compute the HP trend.

Next, we show that $\widehat{\boldsymbol{\tau}}_{\text{HP}}$ is the mean of the conditional distribution of $\boldsymbol{\tau}$ given in (11). To see that, set $\rho = 0$ and $\boldsymbol{\phi} = \mathbf{0}$. Then, the posterior mean of $\boldsymbol{\tau}$ becomes

$$\begin{aligned}
\widehat{\boldsymbol{\tau}} &= \left(\frac{1}{\sigma_\tau^2}\mathbf{H}_2'\mathbf{H}_2 + \frac{1}{\sigma_c^2}\mathbf{I}_T \right)^{-1} \left(\frac{1}{\sigma_\tau^2}\mathbf{H}_2'\mathbf{H}_2\boldsymbol{\alpha} + \frac{1}{\sigma_c^2}\mathbf{y} \right) \\
&= \left(\frac{\sigma_c^2}{\sigma_\tau^2}\mathbf{H}_2'\mathbf{H}_2 + \mathbf{I}_T \right)^{-1} \left(\frac{\sigma_c^2}{\sigma_\tau^2}\mathbf{H}_2'\mathbf{H}_2\boldsymbol{\alpha} + \mathbf{y} \right) = \widehat{\boldsymbol{\tau}}_{\text{HP}}
\end{aligned}$$

with $\lambda = \sigma_c^2/\sigma_\tau^2$.

Appendix C: Additional Empirical Results

In this appendix we present additional estimation results. Specifically, Figure (5) plots the prior and posterior densities of ρ under UCUR-2M. Even though the posterior is relatively flat, it has more mass around 0 and less mass at the boundaries compared to the prior.

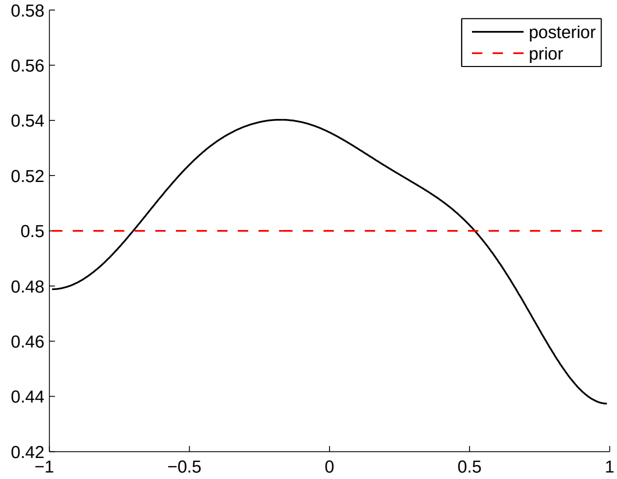

Figure 5: Prior and posterior densities of ρ under UCUR-2M.

References

N. Baştürk, C. Çakmakli, S. P. Ceyhan, and H. K. Van Dijk. Posterior-predictive evidence on US inflation using extended new Keynesian Phillips curve models with non-filtered data. *Journal of Applied Econometrics*, 29(7):1164–1182, 2014.

M. Banbura, D. Giannone, and L. Reichlin. Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92, 2010.

T. Berger, G. Everaert, and H. Vierke. Testing for time variation in an unobserved components model for the U.S. economy. *Journal of Economic Dynamics and Control*, 2016. Forthcoming.

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.

G. Bohlmann. Ein ausgleichungsproblem. *Nachrichten von der Gesellschaft der Wissenschaften zu Gottingen, Mathematisch-Physikalische Klasse*, pages 260–271, 1899.

Z. I. Botev, J. F. Grotowski, and D. P. Kroese. Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5):2916–2957, 2010.

F. Canova. Bridging DSGE models and the raw data. *Journal of Monetary Economics*, 67:1–15, 2014.

A. Carriero, G. Kapetanios, and M. Marcellino. Forecasting exchange rates with a large Bayesian VAR. *International Journal of Forecasting*, 25(2):400–417, 2009.

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 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.

P. K. Clark. The cyclical component of US economic activity. *The Quarterly Journal of Economics*, 102(4):797–814, 1987.

W. Gersch. Smoothness priors. In B. Brillinger et al., editor, *New Directions in Time Series Analysis Part II*, pages 113–146. Springer-Verlag, 1993.

J. Geweke and G. Amisano. Hierarchical Markov normal mixture models with applications to financial asset returns. *Journal of Applied Econometrics*, 26(1):1–29, 2011.

R. J. Gordon. The demise of US economic growth: Restatement, rebuttal, and reflections. Technical report, National Bureau of Economic Research, 2014.

A. L. Grant and J. C. C. Chan. A Bayesian model comparison for trend-cycle decompositions of output. *Journal of Money, Credit and Banking*, 2016. Forthcoming.

A. C. Harvey. Trends and cycles in macroeconomic time series. *Journal of Business and Economic Statistics*, 3(3):216–227, 1985.

A. C. Harvey and A. Jaeger. Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, 8(3):231–231, 1993.

A. C. Harvey, T. M. Trimbur, and H. K. Van Dijk. Trends and cycles in economic time series: A Bayesian approach. *Journal of Econometrics*, 140:618–649, 2007.

R. J. Hodrick and E. C. Prescott. Postwar US business cycles: An empirical investigation. *Carnegie Mellon University discussion paper, no. 451.*, 1980.

R. J. Hodrick and E. C. Prescott. Postwar US business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, pages 1–16, 1997.

G. Kamber, J. Morley, and B. Wong. Intuitive and reliable estimates of the output gap from a Beveridge-Nelson filter. *UNSW Business School Research Paper No. 2016-09*, 2016.

R. G. King and S. T. Rebelo. Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, 17(1-2):207–231, 1993.

G. Kitagawa and W. Gersch. A smoothness priors–state space modeling of time series with trend and seasonality. *Journal of the American Statistical Association*, 79(386): 378–389, 1984.

G. Koop. *Bayesian Econometrics.* Wiley & Sons, New York, 2003.

G. Koop. Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics*, 28(2):177–203, 2013.

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*. Wiley & 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.

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.

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.

M. W. Watson. Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, 18(1):49–75, 1986.

H. L. Weinert. Efficient computation for whittaker–henderson smoothing. *Computational Statistics and Data Analysis*, 52(2):959–974, 2007.

E. T. Whittaker. On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, 41:63–75, 1923.