

A Service of



Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Crump, Richard K.; Gospodinov, Nikolaj; Wieman, Hunter

Working Paper Sparse trend estimation

Staff Reports, No. 1049

Provided in Cooperation with: Federal Reserve Bank of New York

Suggested Citation: Crump, Richard K.; Gospodinov, Nikolaj; Wieman, Hunter (2023) : Sparse trend estimation, Staff Reports, No. 1049, Federal Reserve Bank of New York, New York, NY

This Version is available at: https://hdl.handle.net/10419/272862

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU



NO. 1049 FEBRUARY 2023

Sparse Trend Estimation

Richard K. Crump | Nikolay Gospodinov | Hunter Wieman

FEDERAL RESERVE BANK of NEW YORK

Sparse Trend Estimation

Richard K. Crump, Nikolay Gospodinov, and Hunter Wieman *Federal Reserve Bank of New York Staff Reports*, no. 1049 February 2023 JEL classification: C13, C30, C33, E27, E32

Abstract

The low-frequency movements of many economic variables play a prominent role in policy analysis and decision-making. We develop a robust estimation approach for these slow-moving trend processes, which is guided by a judicious choice of priors and is characterized by sparsity. We present some novel stylized facts from longer-run survey expectations that inform the structure of the estimation procedure. The general version of the proposed Bayesian estimator with a slab-and-spike prior accounts explicitly for cyclical dynamics. The practical implementation of the method is discussed in detail, and we show that it performs well in simulations against some relevant benchmarks. We report empirical estimates of trend growth for U.S. output (and its components), productivity, and annual mean temperature. These estimates allow policymakers to assess shortfalls and overshoots in these variables from their economic and ecological targets.

Key words: slow-moving trends, sparsity, Bayesian inference, latent variable models, trend output growth

Crump: Federal Reserve Bank of New York (email: richard.crump@ny.frb.org). Gospodinov: Federal Reserve Bank of Atlanta (email: niko-lay.gospodinov@atl.frb.org). Wieman: Princeton University (email: hw8318@princeton.edu). This paper was previously circulated under the title "Elusive 'Stars': Robust Trend Estimation." The authors are grateful to Serena Ng for detailed and constructive discussions on this project. They also thank Martín Almuzara, Marco Del Negro, Alex Thorp, and seminar participants at the 2021 NBER Summer Institute for helpful comments and discussions.

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in this paper are those of the author(s) and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve Bank of Atlanta, or the Federal Reserve System. Any errors or omissions are the responsibility of the author(s).

1 Introduction

Quantifying and characterizing the low-frequency behavior of time series processes has a long tradition in economics. The appeal of constructing accurate and robust estimates for the slow-moving trend component of an economic series for informing policy analysis cannot be overstated. However, as the true low-frequency component is inherently unobserved, its extraction and validation for guiding economic policy is fraught with difficulties arising from substantial underlying uncertainty.

The commonly employed state-space approach essentially assumes a particular parametric structure about the deterministic and stochastic trend components but it tends to exhibit fragilities to potential misspecification and strong identification requirements. For example, it is often the case that even a seemingly innocuous misspecification in the trend component could induce a severe distortion in the estimated trend-cycle decomposition by erroneously attributing some of the lowfrequency persistence to the cyclical dynamics. On the other hand, adopting a more agnostic approach by starting from an unrestricted set of nonparametric estimators may produce very wide (highly uncertain) estimates that are of little practical relevance. This suggests that to achieve more informative inference, one may need to impose discipline on the low-frequency movements via prior information or restrictions that ensure sparsity.

To inform our approach to trend estimation, we present novel survey evidence on "longer-run" forecasts – which offer a convenient way to define the low-frequency or trend component – of main variables of interest in the U.S. In addition to the "slow-moving" trend that is embedded in these forecasts, we also provide a new stylized fact, summarized by histograms for the first and second difference of the forecast, that individual forecasters appear to change their long-run forecasts only rarely. To accommodate these features – slowly-evolving and infrequently-changing low-frequency component – we rely on sparsity and shrinkage through our choice of a slab-and-spike prior. This prior information that we impose on the unobserved trend process is intuitive and aligns with the beliefs of professional forecasters. To allow for high-frequency cyclical fluctuations, we employ a stationary autoregressive process. We implement these ideas by resorting to the Bayesian counterpart of generalized LASSO estimation with separate penalties on the long-run and short-run components. Our method thus generalizes some existing procedures for robust trend estimation (Kim, Koh, Boyd, and Gorinevsky (2009), Tibshirani (2014)), and Roualdes (2015)) with the aim of capturing the salient features of economic data.

Prominent contributions to the study of the low-frequency behavior of economic series include Hodrick and Prescott (1980), Beveridge and Nelson (1981), Baxter and King (1999), Morley, Nelson, and Zivot (2003), among many others. A primary motivation of this literature is to isolate the lowfrequency, slow-moving component of a series – which reflects the secular and structural factors that underlie its dynamic behavior – from its high-frequency, possibly cyclical, variations. Potential output, natural rate of unemployment, neutral real rate of interest, inflation expectations, and common variation in real activity are only a few examples of such slow-moving latent processes that are often denoted by and referred to as "stars." There are numerous papers that have implemented these methodologies in wide applications across economics and finance (see, for example, Laubach and Williams (2003), Holston, Laubach, and Williams (2017), Del Negro, Giannone, Giannoni, and Tambalotti (2018), Crump, Eusepi, Giannoni, and Şahin (2019) among many others). The theoretical literature continues to evolve with recent work such as Müller and Watson (2008), Grant and Chan (2017), Hamilton (2018), Phillips and Shi (2020), and Lee, Liao, Seo, and Shin (2021) among others (see Hodrick (2020) and Canova (2020) for comprehensive discussion).

The point of departure of this paper from the existing literature is to operationalize the "slowmoving" trend assumption through an imposition of sparsity that is informed by survey evidence of financial market participants. In particular, our contributions can be summarized as follows. First, we provide novel evidence on individual survey forecasts of the trend component of key economic variables. We utilize non-public data from the Survey of Primary Dealers which is the only survey of professional forecasters which explicitly solicits longer-run forecasts on a consistent basis. Our matched panel data allows us to characterize the distribution of adjustments to longer-run forecasts over time. On the methodological side, we contribute to the literature on Bayesian sparse trend estimation utilizing slab-and-spike priors and a serially correlated cycle component. The embedded sparsity of our estimator offers an alternative statistical characterization of a "slow-moving" trend. Furthermore, our Bayesian setup allows to incorporate uncertainty around our estimates in an internally consistent manner.

Our main empirical application focuses on the estimation of the low-frequency trend in real GDP growth. This is a perennial question in empirical macroeconomics with important implications for monetary and fiscal policy (for recent contributions, see, e.g., Fernald, Hall, Stock, and Watson (2017), Coibion, Gorodnichenko, and Ulate (2018) and Müller, Stock, and Watson (2020)). Our estimates for 1947–2019 suggest that trend real GDP growth in the U.S. has been falling since the 1960s and ends the sample at around 2%. Importantly, our method produces slow-moving trend estimates which do not systematically co-vary with the cycle. We contrast our estimates to those based on recently proposed estimators of Müller and Watson (2008), Grant and Chan (2017), and Phillips and Shi (2020). We further provide a detailed analysis of the underlying trend in the contributions to real GDP growth from its constituent components. Despite the fact that these components generally exhibit different dynamic properties, we find that the individual trends estimated for the GDP contributions can be aggregated and match closely the direct trend estimate from headline real GDP growth. We also explore two other empirical applications: post-war total factor productivity (TFP) growth and mean temperature changes in the U.S. over the last 140 years. These estimates facilitate policy makers' efforts to assess shortfalls and overshoots in these variables from their economic and ecological targets.

This paper is organized as follows. In Section 2, we provide heuristic motivation, based on long-run survey forecasts, for our general estimation approach which is introduced in Section 3. We first discuss the estimation methodology for a white noise cycle in Section 3.1 and its extension to a general cycle component in Section 3.2. Our main empirical application on the trend in U.S. real GDP growth is provided in Section 4. Section 5 assesses the finite-sample properties of the method in a small simulation experiment. Section 6 concludes. Finally, the Appendix provides full details on the implementation of our proposed methodology.

2 Evidence on Survey-Based Trends

In this section, we present new stylized facts about longer-run survey forecasts of professional forecasters. This empirical evidence will serve as a motivation for the methodological approach that we introduce later in the paper. To cement ideas, it is useful to discuss a convenient mathematical formulation of an underlying trend (Beveridge and Nelson, 1981) of a variable y_t as $\lim_{h\to\infty} \mathbb{E}[y_{t+h} | \mathcal{F}_t]$ where \mathcal{F}_t is the information set available at time t. Thus, this may be interpreted as a very long-run forecast of the variable y_t which aligns directly with our survey data.

We use non-public data from the Federal Reserve Bank of New York's Survey of Primary Dealers (SPD).¹ The SPD is conducted by the Trading Desk of the New York Fed one to two weeks before each regularly scheduled Federal Open Market Committee (FOMC) meeting and survey respondents are primary dealers (at the time of the survey) to the Federal Reserve Bank of New York.² Starting in July 2012 the survey began asking respondents about their "longer-run" forecasts for relevant economic variables. To our knowledge, the SPD is the only survey that solicits forecasts from private-sector respondents on explicitly "longer-run" values of economic variables.³ This stands in contrast to a number of surveys which request farther in the future forecasts for economic variables with specific horizon. The SPD collects longer-run forecasts of real GDP growth, the unemployment rate, the federal funds rate, and PCE inflation. We obtain an unbalanced panel of these forecasts for the entire available sample starting in mid-2012 which comprises 77 survey observations. Our sample includes 25 different primary dealers of which 16 primary dealers have forecasts for at least 70 of the 77 survey dates.⁴

Figure 1 presents the time series of different quantiles of the cross-sectional distribution of forecasts for real GDP growth, the unemployment rate, the federal funds rate, and the real federal funds rate. The latter is obtained by subtracting the longer-run PCE inflation forecast from the nominal federal funds rate forecast. First, there appears to be clear commonality in the movement of these longer-run forecasts with only modest disagreement across respondents. All four variables exhibit a downward trend over our sample period reflecting the perceived decline in potential output, the natural rate of unemployment, and the natural rate of interest (e.g., Crump, Eusepi, and Moench (2018), Holston, Laubach, and Williams (2017), Del Negro, Giannone, Giannoni, and Tambalotti (2018)). Notably, the longer-run forecasts appear remarkably stable even after the onset of the COVID-19 pandemic.

Importantly, across all four variables, we can observe two key properties of these longer-run forecasts that will later serve to motivate our theoretical approach. First, the longer-run values of these economic variables are perceived to move over time but in a slow and deliberate fashion. Second, there are distinct periods of unchanged forecasts along with periods where forecasts change

¹See https://www.newyorkfed.org/markets/primarydealer_survey_questions.

²Further information is available here: https://www.newyorkfed.org/markets/primarydealers.

³The Survey of Economic Projections (SEP), conducted by the FOMC, also reports longer-run values for these variables which is then mirrored by the SPD. In the SEP, "[l]onger-run projections represent each participants assessment of the rate to which each variable would be expected to converge under appropriate monetary policy and in the absence of further shocks to the economy." See, for example, https://www.federalreserve.gov/monetarypolicy/files/fomcprojtabl20220921.pdf.

⁴All of our empirical results are robust to excluding the 9 primary dealers who have shorter reporting periods.

at a constant rate. This provides suggestive evidence that differences or second differences of forecasts for longer-run variables regularly take on values of zero.

Figure 1. Time Series Properties of the Cross-Sectional Distribution of Forecasts. This figure shows the cross-sectional median (solid line) and first and third quartiles (dashed lines) for the longer-run forecasts from the Survey of Primary Dealers (SPD). The real federal funds rate longer-run forecast is constructed by subtracting the individual longer-run PCE inflation forecast from the corresponding longer-run federal funds rate forecast. The sample period is 2012m7–2022m3.



We can demonstrate these properties by examining the empirical distribution of changes in longer-run forecasts. Figure 2 shows histograms of the first- and second-differences of longer-run forecasts from the SPD. Each chart presents two histograms. The first histogram pools all forecasts across survey respondents and time periods. The second histogram is an optimal aggregator of the histograms for each individual survey respondent. The weights are obtained by minimizing the Hellinger distance between the pooled histogram and the optimal-weighted aggregator.⁵ The most striking feature of Figure 2 is the disproportionate peak in each histogram at the bin centered at zero. This pattern holds across all variables and for both first and second differences.⁶ Importantly, this behavior is driven by exact zero values rather than very small values. In particular, all four variables have between 80% and 85% of their first differences equal to zero and 66% to 74% of second differences equal to zero. As we move away from zero, we tend to observe a relatively wide but

⁵For more details and an application of this aggregation approach to asset-pricing models, see Gospodinov and Maasoumi (2021).

⁶In unreported results, we confirm that the same patterns hold at the individual forecaster level. These results are omitted to comply with data confidentiality restrictions of the SPD.

flat spread of observations. These histograms do not appear to be consistent with the behavior of a standard continuous distribution like the Gaussian distribution. Instead, the disparate behavior for small values relative to larger values (in magnitude) appears more consistent with a mixture distribution. This is precisely the methodological approach we undertake in the next section.

Figure 2. Changes in Longer-Run Survey Forecasts. This figure summarizes some of the properties of the changes in longer-run forecasts for real GDP growth, the unemployment rate, the federal funds rate and the real federal funds rate from the Survey of Primary Dealers. The top (bottom) panel displays the histogram of (the change in) one-period changes in longer-run forecasts. A purple hue denotes overlap in the pooled histogram and the optimally weighted histogram. The sample period is 2012m7–2022m3.



3 Sparse Trend Methodology

3.1 Trend plus white noise

We begin with a description of the setup adopted throughout the paper. Suppose that we observe the time series $\{y_t\}_{t=1}^T$. We will assume that y_t may be decomposed as

$$y_t = g_t + \epsilon_t, \qquad t = 1, \dots, T,\tag{1}$$

for some "trend," g_t and corresponding deviation from trend ϵ_t . Without further assumptions we cannot make any progress in separating g_t from its deviation. To see this, note that we can set $\hat{g}_t = y_t$ and fit the data with no error. To avoid such an outcome, a popular class of estimators of the trend is of the form:

$$\min_{g_1,\dots,g_T} \sum_{t=1}^T \left(y_t - g_t \right)^2 + \lambda_g \sum_{t=1}^T \ell\left(\Delta^k g_t\right),\tag{2}$$

where Δ^k denotes the k-th order difference operator for $k \ge 1$ and $\ell(\cdot)$ is a penalty function with penalty parameter λ_g . When $\ell(z) = z^2$ and k = 1, then this is the likelihood function of the local level model, where $\Delta g_t = v_t$ for some white noise process $\{v_t\}_{t=1}^T$. Similarly, when $\ell(z) = z^2$ and k = 2, we obtain the HP filter (Hodrick and Prescott 1980). Finally, for $\ell(z) = |z|$ and a fixed k we obtain the trend filter of Tibshirani (2014).

We appeal to the common assumption of a "slow moving" trend. This assumption is implemented, in practice, almost exclusively by modeling the trend component as non-stationary with a "small" innovation variance. In contrast, our approach is to instead induce the desired lowfrequency behavior through sparsity in changes in the estimated trend component. As a convenient way to model the slow-moving trend and its associated estimation uncertainty, we develop our methods within a Bayesian framework.

A natural starting place would be the trend filter as the frequentist version induces sparsity since the estimation setup may be nested as a generalized LASSO problem. From a Bayesian perspective, the trend filter can be motivated following Park and Casella (2008). Park and Casella (2008) study the standard linear regression setting, where $Y|X, \beta, \sigma^2 \sim \mathcal{N}(X\beta, \sigma^2 I_T)$ and

$$p(\beta_1, \dots, \beta_p | \sigma^2) = \prod_{j=1}^p \frac{1}{2\sqrt{\sigma^2}} \exp\left\{-\lambda \frac{|\beta_j|}{\sqrt{\sigma^2}}\right\}.$$
(3)

In its hierarchical form,⁷

$$\beta_1, \dots, \beta_p \left| \sigma^2, \varkappa_1, \dots, \varkappa_p \sim \mathcal{N} \left(0, \sigma^2 \cdot \operatorname{diag} \left(\varkappa_1, \dots, \varkappa_p \right) \right),$$

$$\tag{4}$$

⁷The hierarchical form uses the representation of the Laplace distribution as a scale mixture of normals (with an exponential mixing density). See Park and Casella (2008) for further discussion.

$$p(\varkappa_1, \dots, \varkappa_p | \sigma^2, \lambda) = \prod_{i=1}^p \left(\frac{\lambda^2}{2}\right) \exp\left\{\frac{-\lambda^2 \varkappa_i}{2}\right\}$$
(5)

with

$$p(\sigma^2) = 1/\sigma^2. \tag{6}$$

The Bayesian trend filter can then be obtained setting p = T, $X = I_T$, $g_t = \beta_t$ and the priors applied to each of the $\Delta^k g_t$ (Roualdes 2015). A drawback to this approach is that frequentist sparsity does not translate to Bayesian sparsity so that any individual draw of g_t will have nonzero changes with probability one. Consequently, so does the posterior median or other common summary objects of interest. To alleviate this issue, we instead utilize a spike-and-slab prior so that individual draws possess the desired sparsity.

As a simple illustrative example for the spike-and-slab approach, consider the process as in equation (1) with $\Delta g_t = v_t$, where v_t are *i.i.d.* which follow a mixture distribution with $v_t \sim \mathcal{N}(0, \varpi)$ with probability ξ and $v_t = 0$ with probability $1 - \xi$. Intuitively, this mixture distribution weakens the link between small movements and large movements in the random variable. This accommodates slow-moving dynamics without compromising some occasional bursts of movement in the series (as seen in Figure 1).

To transform this motivating example to a more general setup, suppose that the observed data $y = (y_1, \ldots, y_T)'$ satisfies

$$y \mid g, \sigma^2 \sim \mathcal{N}(g, \sigma^2 I_T) \tag{7}$$

and we propose the following prior (and hyperprior) distributions on the parameters. First, we assume that $g = (g_1, \ldots, g_T)'$ satisfies

$$g \mid \sigma^2, \omega \sim \mathcal{N}(0, \sigma^2 \Sigma_g), \qquad \Sigma_g^{-1} = D'_k \operatorname{diag}(\omega_1, \dots, \omega_{T-k})^{-1} D_k,$$
(8)

where D_k is the $(T-k) \times T$, k-difference matrix which maps a vector $(x_1, \ldots, x_T)'$ to $(\Delta^k x_{k+1}, \ldots, \Delta^k x_T)'$ and ω_j are the mixing variables for the Laplace distribution. These mixing weights for g, $\{\omega_j : j = 1, \ldots, T-k\}$, are conditionally independent with corresponding density

$$p(\omega_j \mid \lambda_g, \lambda_0, \theta_j) = (1 - \theta_j) \frac{\lambda_0^2}{2} \exp\left\{\frac{-\lambda_0^2 \omega_j}{2}\right\} + \theta_j \frac{\lambda_g^2}{2} \exp\left\{\frac{-\lambda_g^2 \omega_j}{2}\right\},\tag{9}$$

where λ_0 is fixed at a large number such that $\lambda_0 >> \lambda_g$. This corresponds to the continuous spikeand-slab setup (see Ročková and George (2018)), where θ_j can take on the value very close to 0 – the "spike" – or the value of 1 – the "slab", governed by λ_g . The parameter λ_g has the interpretation of representing a signal-to-noise ratio, where the "signal" corresponds to the underlying trend, g_t .

Further, it is assumed that θ_j are distributed as

$$\theta_j \sim \text{Bernoulli}(\xi)$$
 (10)

with common parameter ξ , where

$$\xi \sim \mathcal{B}\mathrm{eta}(a,b). \tag{11}$$

Here, a and b are hyperparameters. In our implementation, we choose a = b = 1 which corresponds to a uniform prior.⁸ Next, as in Park and Casella (2008), we place an uniformative prior on σ^2 as

$$p(\sigma^2) = 1/\sigma^2. \tag{12}$$

Finally, λ_q is distributed as

$$\lambda_g \sim \Gamma(r_g, \delta_g) \tag{13}$$

with hyperparameters r_g and δ_g , where r_g is the *shape* parameter, and δ_g is the *rate* parameter. In principle, we could estimate λ_g via an empirical Bayes approach; however, we have found that a diffuse gamma prior gives similar posterior median estimates but is more computationally efficient and also allows us to accommodate the uncertainty from this parameter.

It is instructive to compare the prior assumptions underpinning the Bayesian LASSO (equations 3–6) to our slab-and-spike formulation. In the Bayesian LASSO, a single Laplace density is used to draw from. In contrast, Figure 3 provides an example of the two Laplace densities that comprise the the continuous spike-and-slab setup.⁹ The blue line represents the "spike" and the red line represents the "slab". We can observe the clear similarities in the implied mixture distribution of Figure 3 and the survey data presented in Figure 2. In Appendix A.1, we fully operationalize our model, choice of priors, and the individual steps of the Gibbs sampling procedure.

Figure 3. Spike and Slab Densities. This figure illustrates the roles of the two densities in the continuous spike and slab setup.



⁸In some situations, one may want to place much more informative priors on ξ .

⁹In practice, when we implement the "slab and spike" formulation, the difference between the two densities is even more extreme but we diminish the differences in Figure 3 for presentation purposes.

To gain further intuition for our approach, Figure 4 shows the trend estimate g of U.S. total factor productivity (TFP). Quarterly data for the period 1947:Q2–2019:Q4 are obtained from the Federal Reserve Bank of San Francisco¹⁰ for business sector TFP, defined as output growth less the contribution of capital and labor (see, for example, Fernald (2014)). The TFP series is constructed as a percent change at an annual rate. The left-hand plot presents estimates using k = 1 and k = 2along with the observed data series whereas the right-hand plot shows only the trend estimates. We choose r_g and δ_g to produce a diffuse prior on λ_g (see Appendix A.1). Despite the diffuse prior on ξ , we observe that the posterior median of g evolves smoothly with a corresponding posterior coverage interval which is relatively tight. We also observe only small differences between the two choices of k.

In the right graph of Figure 4 we show the trend estimate for both choices of k along with the 95% pointwise posterior coverage intervals. The median estimate of trend TFP growth hovers around 2% through the 1960s before a steep decline, falling to about 1.0% in 1980. In the 1980s and 1990s trend TFP growth partially reversed this fall before declining again in the last part of the sample. At the end of 2019, the estimated trend is somewhere between 0.5% and 0.7% although the uncertainty band is fairly wide.

Figure 4. Trend Estimates of Total Factor Productivity Growth. This figure presents trend estimates of total factor productivity (TFP). The left chart displays TFP growth along with the estimated trend for k = 1 and k = 2 based on the methodology introduced in Section 3.1. The right chart displays the estimated trend for k = 1 and k = 2 along with the corresponding 95% posterior coverage interval. The sample period is 1947Q2–2019Q4.



In the above example, the deviation from trend appears to be well approximated by a white noise process. In many economic examples, we observe more persistent cycles (e.g., Morley, Nelson, and Zivot (2003), and so in the next section we augment our approach to accommodate such behavior.

3.2 General Setup

In this section, we will assume that y_t may be decomposed as

$$y_t = g_t + c_t, \qquad t = 1, \dots, T, \tag{14}$$

¹⁰https://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/

where c_t now represents the "cycle" and is presumed to be serially correlated over time with degree of persistence determined by a parameter vector γ . For parsimony, we assume $c_t = y_t - g_t$ follows an $AR(\mathbf{q})$ with

$$c_t = \gamma_1 c_{t-1} + \dots + \gamma_q c_{t-q} + \epsilon_t, \qquad t = 1, \dots, T,$$
(15)

where $\gamma = (\gamma_1, \ldots, \gamma_q)'$. In principle, one could use a more general ARMA specification but we do not pursue this further. This approach for the cycle term leads to a more general form of equation (2),

$$\min_{g_1,\dots,g_T} \sum_{t=1}^T \epsilon_t^2 + \lambda_g \sum_{t=1}^T \ell\left(\Delta^k g_t\right) + \lambda_\gamma \sum_{i=1}^q h_i(\gamma),\tag{16}$$

where $h_i(\gamma)$ represents a transformation of the autoregressive parameters and λ_{γ} is the associated penalty parameter. Two natural choices for $h_i(\cdot)$ are $h_i(\gamma) = \gamma_i$ and $h_i(\gamma) = \varphi_i$, where φ_i is the *i*-th partial autocorrelation coefficient (see Schmidt and Makalic (2013) for regularization of partial autocorrelation coefficients when g_t is constant over time). In this more general setup, λ_g again represents the signal-to-noise ratio but comparing standard deviation of the trend innovation to the cycle.

We now assume that the observed data $y = (y_1, \ldots, y_T)'$ satisfies

$$y \mid g, \gamma, \sigma^2 \sim \mathcal{N}(g, \sigma^2 \cdot \mathcal{V}_{\gamma}), \tag{17}$$

where \mathcal{V}_{γ} denotes the variance-covariance matrix of a strictly stationary $AR(\mathbf{q})$ process with $\sigma^2 = 1$. The trend, $g = (g_1, \ldots, g_T)'$, satisfies

$$g \mid \sigma^2, \omega, \gamma \sim \mathcal{N}(0, \sigma^2 \mathcal{V}_{\gamma, 11} \cdot \Sigma_g), \tag{18}$$

where $\mathcal{V}_{\gamma,11}$ is the (1, 1) element of \mathcal{V}_{γ} . This additional factor ensures that our approach is scale invariant to y_t .¹¹ The mixing weights for g, $\{\omega_j : j = 1, \ldots, T - k\}$, remain conditionally independent with corresponding density as in equation (9). Moreover, the remaining parameters, σ^2 , θ_j , ξ , and λ_g have priors as in the previous section.

For the cycle component, $c = (c_1, \ldots, c_T)'$, we assume that $c \sim \mathcal{N}(0, \sigma^2 \mathcal{V}_{\gamma})$ and partition the density as,

$$p(c|\gamma,\sigma^2) = p(c_{q+1:T} \mid c_{1:q},\gamma,\sigma^2)p(c_{1:q} \mid \gamma,\sigma^2),$$
(19)

where $c_{1:q}$ are the first q observations of the cycle and $c_{q+1:T}$ are the last T-q observations and

$$p(c_{q+1:T} \mid c_{1:q}, \gamma, \sigma^2) \propto \frac{1}{\sigma^{(T-q)}} \exp\left\{-\frac{(c - B_{c,q}\gamma)'(c - B_{c,q}\gamma)}{2\sigma^2}\right\}$$
(20)

¹¹In some contexts, an alternative that does not scale Σ_g may be preferred (Moran, Ročková, and George (2019)). In our empirical applications, we have found that the results are robust to the choice of scaling.

with $B_{c,q}$ denoting a $(T-q) \times q$ matrix with *i*-th row equal to $(c_{i+q-1}, c_{i+q-2}, \ldots, c_i)$. For the initial conditions we have,

$$p(c_{1:q} \mid \gamma, \sigma^2) \propto \frac{1}{\det(\sigma^2 \mathcal{V}_{\gamma})^{1/2}} \exp\left\{-\frac{c'_{1:q} \mathcal{V}_{\gamma}^{-1} c_{1:q}}{2\sigma^2}\right\}.$$
(21)

Next, following a similar approach as for the trend, we impose¹²

$$\gamma \mid \tau \sim \mathcal{N}(0, \Sigma_{\gamma}) \tag{22}$$

where

$$\Sigma_{\gamma} = \operatorname{diag}(\tau_1, \dots, \tau_q) \tag{23}$$

and the mixing weights τ_j are conditionally independent for $j = 1, \ldots, q$ with density

$$p(\tau_j \mid \lambda_\gamma) = \frac{(\lambda_\gamma)_j^2}{2} \exp\left\{\frac{-(\lambda_\gamma)_j^2 \tau_j}{2}\right\}.$$
(24)

Finally, the additional penalty parameter λ_{γ} is distributed as,

$$\lambda_{\gamma} \sim \Gamma(r_{\gamma}, \delta_{\gamma}), \tag{25}$$

with hyperparameters r_{γ} and δ_{γ} , where r_{γ} is the shape parameter, and δ_{γ} is the rate parameter. In our empirical implementation, we choose r_{γ} and δ_{γ} that result in a relatively tight prior distribution for λ_{γ} around a small value. As we have discussed, our motivation is a cyclical component around a slow-moving trend which corresponds to little to no prior penalization on this term. Appendix A.2 and Appendix A.3 provide operational details about the implementation of proposed method.

To illustrate the implementation of the method from this section, we apply our estimator with and without a cycle to data on the annual mean temperature change in the United States obtained from NASA, which cover the sample period 1880-2019.¹³ The data are expressed as a deviation from the average temperature (in Celsius) over 1951–1980. Figure 5 presents these results. The top row shows the estimated underlying trend for both k = 1 and k = 2 when a cycle is included and when it is omitted. When k = 1, we observe that the underlying trend displays more pronounced local movements whereas for k = 2, the results are much smoother. Nevertheless, both trend estimates display a clear and strong upward trend since around 1980. The second row presents again the estimated trends along with 95% posterior coverage intervals. We can observe that the estimated trend is precisely estimated and we can comfortably reject the hypothesis that current levels of the trend are consistent with the data observed up to 1980. Finally, we highlight that the inclusion of a cycle term has little effect on all of these results. In the next section, we will provide

¹²The formulation using partial autocorrelation coefficients is omitted to conserve space and is available upon request from the authors.

¹³Data are available at https://data.giss.nasa.gov/gistemp/graphs_v4/.

empirical examples where the role of the cycle term is more prominent.

Figure 5. Trend Estimates of the Annual Mean Temperature Change in the United States. This figure presents trend estimates of the annual mean temperature change in the United States. The charts in the top row displays the annual mean temperature change along with the estimated trend for k = 1 and k = 2 based on the methodology introduced in Section 3. The bottom row displays the estimated trend for k = 1 and k = 2, with and without a possibly serially correlated cycle, along with the corresponding 95% posterior coverage interval. The sample period is 1880–2019.



4 Empirical Application: Real GDP Growth

As our main empirical application, we study the underlying trend in real GDP growth in the United States in the post-war era. The estimate of the unobserved trend component in real GDP growth, which is stripped of cyclical fluctuations, can inform policymakers of the long-run potential of the U.S. economy. The data is quarterly (annualized) real GDP growth over the pre-pandemic period, 1947:Q2–2019:Q4, although we also use the observations from 2020:Q1 to 2022:Q2 to quantify the output shortfall from potential during the pandemic. As an additional exercise, we also construct a "bottom-up" estimate of the trend using the individual component contributions to real GDP growth.¹⁴

¹⁴All series are available from https://fred.stlouisfed.org/: Personal Consumption Expenditures (DPCERY2Q224SBEA); Change in Inventories (A014RY2Q224SBEA); Residential Investment (A011RY2Q224SBEA); Busi-

The top row of Figure 6 presents the time series of real GDP growth along with the estimated low frequency component for k = 1 (left column) and k = 2 (right column). As it is well known, the first half of the sample period is characterized by more frequent recessions (denoted by the shaded areas) and higher output growth volatility. Each chart compares estimates which allow for a serially correlated cycle to those which do not. We can see immediately that for real GDP growth, there is a clear role for a cyclical component. In the top left plot, for k = 1, there is a large difference between the estimated trend with and without a cyclical component. This stands in stark contrast to the results presented in Section 3.2 for annual temperature change which were largely unchanged with the addition of a cyclical component. For k = 2, the estimated trend is quite similar with and without a cyclical component.

In the bottom left chart of Figure 6, we report results only for the estimated trends with a cyclical component, along with the pointwise 95% posterior coverage intervals. We can observe that the estimated trend and coverage intervals are broadly similar for each choice of k. The estimated trend has been declining over the sample and both choices produce an estimated trend of about 2.25% at the end of the sample. This is somewhat higher than the corresponding estimate available from surveys of professional forecasters, the CBO or the FOMC which are all closer to 1.9%. The corresponding uncertainty around these estimates is slightly higher for k = 2, reflecting the modestly higher variability in the trend estimate. Furthermore, there are more pronounced boundary effects when k = 2 which is reflected in the widening of the confidence intervals on either end of the sample. This was also the case for TFP in Figure 4.

ness Fixed Investment (A008RY2Q224SBEA); Federal Government (A823RY2Q224SBEA); State & Local Government (A829RY2Q224SBEA); Net Exports (A019RY2Q224SBEA).

Figure 6. Trend Estimates of Real GDP Growth. This figure presents trend estimates of real GDP growth. The charts in the top row display real GDP growth along with the estimated trend for k = 1 and k = 2, with and without a serially correlated cycle, based on the methodology introduced in Section 3. The left chart in the middle row displays the estimated trend for k = 1 and k = 2 with a serially correlated cycle, along with the corresponding 95% posterior coverage interval. The right chart in the middle row displays the estimated trend for k = 1 and k = 2, with a serially correlated cycle, along with the estimated based on Grant and Chan (2017). The charts in the bottom row display the estimated trend for k = 1 and k = 2 with a serially correlated cycle, along with the estimated based on Grant and Chan (2017). The charts in the bottom row display the estimated trend for k = 1 and k = 2 with a serially correlated cycle, along with the estimated based on Phillips and Shi (2020) and Müller and Watson (2008). The sample period is 1947Q1-2019Q4.



















k = 2

For comparison, we consider several alternative approaches commonly used in the economic literature for decomposing economic series into trend and cycle. The workhorse models for estimating the trend and the cyclical component are the unobserved components (UC) models. The closest UC model to our approach is that of Grant and Chan (2017) which specifies a random walk trend in the growth rate of real GDP along with an additive ARMA cyclical component. The middle right chart presents our trend estimates for both choices of k along with the UC model of Grant and Chan (2017) with a choice of an AR(2) cycle. Although all three estimated trends broadly co-move, we can observe that the trend estimate from the UC model is more volatile and displays residual cyclical behavior, tending to achieve local minima in NBER recessions. This demonstrates the advantages of our sparse modeling framework which produces trend estimates that do not systematically co-vary with the cycle.

The literature has highlighted some fragilities in the estimation of UC models and, as a result, flexible trend estimation approaches may be preferred due to their ostensible robustness properties. By far, the most common technique for estimating low-frequency trends in economic time series is the HP filter (Hodrick and Prescott (1980)).¹⁵ Recently, Phillips and Shi (2020) have refined the original HP filter to produce a data-driven implementation with desirable theoretical properties. In the bottom left chart of Figure 6, we present the estimated trend for real GDP growth rate, based on an application of the original HP filter and the boosted HP filter of Phillips and Shi (2020). We apply these procedures to the level of log real GDP as is the standard approach in practice. We can see immediately that the estimated trend growth rate is far more volatile than that of our procedure. Moreover, the recession shading illuminates that the variability of the two trend estimates appear to be governed by the state of the business cycle. In particular, trend real GDP growth is adjudged to be at its local minimum around recessions throughout the sample. Furthermore, during the Great Recession, the HP-boosted estimate of trend growth falls below zero which would have suggested far less slack in the economy. In contrast, our trend estimates do not exhibit any cyclical behavior which is more akin to an HP filter with a very large choice of the penalty parameter (see, for example, Coibion, Gorodnichenko, and Ulate (2018)). Finally, the bottom right graph provides a comparison to another flexible approach to estimating the low-frequency component, introduced by Müller and Watson (2008) (see also Müller and Watson (2020)). Similar to the HP filter, this approach also produces an estimate which is more variable and cyclical than our method.

There is a natural commonality in all of the approaches to low frequency estimation as they can accommodate different behavior with different choices for the tuning parameter. All of these approaches, for a specific choice of tuning parameter, would produce trend estimates that are similar to our method. However, it is important to emphasize that our tuning parameter selection is conducted in a data-driven way and obviates the need for "rules of thumb" based on the frequency of the data.

Our discussion thus far has focused on the trend estimate but the estimates for the cycle component are also of interest. For our implementation, we choose q = 4, allowing for up to 4

¹⁵Grant and Chan (2017) show that the HP filter can be obtained as the posterior mean of their estimator under the assumption of uncorrelated trend and white-noise cycle components, and the penalty parameter equal to the inverse of the signal-to-noise ratio.

lags in the cyclical component of real GDP growth. This appears to be more than sufficient as the posterior distribution of the γ parameters is characterized by a cycle variable with only a moderate degree of serial correlation. For example, the posterior median of the sum of the γ coefficients is about 0.3 which suggests weak time-series dependence. The standard deviation of the innovation to the cycle component has a 95% posterior coverage interval of 3.2% to 3.8%, which is consistent with the slow-moving trend estimates presented in Figure 6.

The recent COVID pandemic coincided with a brief but dramatic decline in real output. Given the rapid recovery, a natural question to ask is whether there still remains a shortfall relative to the pre-pandemic baseline. The left chart of Figure 7 presents the cumulative shortfall during the pandemic from our k = 2 estimate of log potential output as of 2022Q3. The sharp rebound in economic activity in the latter half of 2020 and through 2021 had almost closed the gap relative to the trend. In contrast, the right chart shows the corresponding path of cumulative real GDP growth after the Great Financial Crisis (GFC). The pseudo real-time estimate of the trend of real GDP growth using k = 2 was 2.7% prior to the GFC. The differential behavior of output growth is striking. Whereas after the pandemic, the real GDP shortfall was nearly closed, no such convergence occurred over a similar time span after GFC. By 2013, the gap was actually getting larger.

Figure 7. Real GDP Shortfalls. This figure presents log real GDP before and after the onset of the COVID pandemic (left chart) and Great Financial Crisis (right chart) along with the counterfactual path based on the estimated trend for k = 2 (red dashed line) and using the methodology introduced in Section 3.



Although trend estimates of real output are relatively common in the literature, the underlying components generally receive scant attention. A bottom-up approach to trend estimation may reveal differential properties then using aggregate output as the target series. In Figure 8, we apply our methodology to the underlying contributions to real GDP growth from personal consumption expenditures, residential investment, business fixed investment, federal government expenditures, state and local government expenditures, net exports and the change in inventories. The charts in Figure 8 show the realized series along with the trend estimates using k = 1 and k = 2.

Figure 8. Trend Estimates of Component Contributions to Real GDP Growth. This figure presents trend estimates of component contributions to real GDP growth: personal consumption expenditures, residential investment, business fixed investment, federal government expenditures, state and local government expenditures, net exports and the change in inventories. All charts display the estimated trend for k = 1 and k = 2, with a serially correlated cycle, based on the methodology introduced in Section 3. The sample period is 1947Q2–2019Q4.



1950 1960 1970 1980 1990 2000 2010 2020

-3 -4

> -5 1940





First, several features of the data are worth noting. A few series display outsized volatility at the beginning of the sample, likely representing the transition into a post-war economy. The so-called Great Moderation – a decline in the variability in real GDP growth starting in the mid 1980s – is only reflected in the underlying behavior of a subset of series. In particular, residential investment, federal government expenditures, and the change in inventories exhibit pronounced reductions in volatility in the second half of the sample. Finally, the different sub-components of real GDP growth appear to display different serial correlation properties around a time-varying central tendency.

Figure 8 shows that across most of the components, we estimate relatively slow-moving trends for both k = 1 and k = 2. For example, the trend growth rate of personal consumption expenditures (which represents the dominant component of U.S. GDP) largely follows similar dynamics as for real GDP growth with a global peak in the mid-1960s at about 2.5% (k = 2) and a sharp drop from 2.3% in the late 1990s to a local trough of 1.3% in 2011. At the end of the sample, both k = 1 and k = 2 suggest a trend growth of 1.6%.

Residential investment and federal government expenditures are the two sub-components that display the largest differences between the trend estimates using k = 1 versus k = 2. For residential investment, we observe modestly more volatility in the trend estimate using k = 1 relative to that of k = 2; however, the trend estimate for k = 1 is still slow-moving with relative variability far less than the actual series. In contrast, there is a marked difference in the trend estimate for federal government expenditures using k = 1 versus k = 2. This is primarily driven by the very large contributions to real GDP growth arising from elevated military spending related to the Korean and Vietnam Wars. This results in heterogenous dynamics in the series such as different volatility regimes, changes in persistence, or more pronounced non-Gaussian features. When k = 1, we are necessarily assuming less smoothness in the trend and so the estimator may be more sensitive to these types of features of the data. On the other hand, the trend estimate for k = 2 is robust to extreme observations and highly persistent cyclical components. In general, we would recommend reporting both k = 1 and k = 2 as differences can be informative about the underlying trend behavior and the time series properties of the variables.

Figure 9 compares the "top-down" and "bottom-up" trend estimates using our methodology for k = 1 and k = 2. By "top-down" we refer to trend estimation based on aggregate real GDP growth as reported in Figure 6 and "bottom-up" refers to adding up the trend estimates of the sub-components reported in Figure 8. For k = 2 (right plot), we can observe that there is essentially no difference between the two approaches with an average absolute deviation of less than 5 basis points. While for k = 1 (left plot) there is substantial deviation in the first half of the sample – driven exclusively by the estimated trend in federal government expenditures, this difference is largely eliminated by the mid 1970s with an average absolute deviation of 13 basis points.

Using real GDP growth as a benchmark application, we show both the benefits and the flexibility of our new methodology. We use the underlying contributions to real GDP as a pseudo "out-of-sample" metric to investigate the robustness properties of our approach. In this respect, we exploit the richness of the GDP data to "stress-test" trend estimation procedures. The underFigure 9. Real GDP Growth Trend Estimates: Top-Down and Bottom-Up. This figure presents trend estimates of real GDP growth based on the aggregate series ("Top-Down") and the component contributions ("Bottom-Up"). The trend estimates are for k = 1 and k = 2 with a serially correlated cycle. The sample period is 1947Q2–2019Q4.



lying contributions display heterogeneous dynamics and other features which allow us to assess the robustness properties of our method. We find that when the data are homogenous, the choice of k = 1 is preferred and performs well as shown by Figure 6; however, in the presence of outliers, volatility clustering, and other characteristics, k = 2 produces a more robust slow-moving trend estimate. Indeed, we saw in Figure 9, that the bottom-up and top-down estimates were nearly indistinguishable for k = 2.

5 Simulation Evidence

In this section, we present results of a small simulation experiment that evaluates the performance of our proposed estimator and compares it to the HP-boosting estimator (Phillips and Shi (2020)) from the previous section. The HP-boosting estimator has a data-driven implementation based on an information-criterion rule. This makes it a natural comparison to our methodology for a simulation experiment.

For our data-generating process, we purposefully simulate data with a slow-moving trend and allow for both a serially uncorrelated and serially correlated cycle component. For the trend, $y_{1,t}$, we choose

$$y_{1,t} = y_{1,t-1} + \eta_t, \tag{26}$$

where $\eta_t \sim \mathcal{N}(0, 0.05^2)$ with probability p_1 and $\eta_t \sim \mathcal{N}(0, 0.05^2 + 0.15^2)$ with probability $1 - p_1$. We set p_1 to 5% to obtain a generally slow-moving trend which features infrequent sharp moves to align with the empirical evidence presented in Section 2. For the cycle component, $y_{2,t}$, we employ a first-order autoregression,

$$y_{2,t} = \rho \cdot y_{2,t-1} + \sqrt{(1-\rho^2)}\nu_t \tag{27}$$

where $\nu_t \sim \mathcal{N}(0, 1)$. The observed data is then $y_t = y_{1,t} + y_{2,t}$. We consider two choices of ρ , $\rho = 0$ and $\rho = 0.4$. The simulation results are based on a sample size of T = 150 and S = 500 replications. We simulate the trend component S times only and then generate the cycle component for each of $\rho = 0$ and $\rho = 0.4$. This ensures comparability of the results across the two data-generating processes. To provide supporting evidence that we have calibrated our DGP to the type of data encountered in previous sections, Figure 10 presents a representative realization for each choice of ρ . Despite the simple data generating process, it is evident from this figure that the simulated data mimic the key features of the data sets we have used throughout the paper.

Figure 10. Simulated Data. This figure presents two draws from the data generating process described by equations (26) and (27) for T = 150. The left and right charts represent simulated data for the series (y_t) and the underlying trend $(y_{1,t})$ when $\rho = 0$ or $\rho = 0.4$, respectively.



Tables 1 and 2 present the results for five different estimators: our sparse trend estimators for k = 1 and k = 2, with and without a serially-correlated cycle component, and the HP boosting estimator.¹⁶ To evaluate the performance, we use the root mean-square error (RMSE) and mean absolute error (MAE) between the estimate and the true trend. In particular, each table reports the minimum and maximum along with the 10th, 25th, 50th, 75th, and 90th quantiles of the realized RMSE and MAE across simulations. To implement the procedures introduced in Section 3, we follow the same approach as in the empirical applications (summarized in Appendices A.1 and A.2). For specifications with a serially correlated cycle component, we choose q = 4. Relative to the true DGP, we are comfortably overparameterized, but this allows us to assess what the costs of this flexibility are in practice.

The results in both tables confirm the appealing properties of our trend estimator demonstrated in the empirical applications. When $\rho = 0$, shown in Table 1, all four of our trend estimators outperform uniformly the HP boosting procedure across simulations. For example, the median RMSE/MAE of HP boosting exceeds by more than 20% the RMSE/MAE of the k = 2 with a serially correlated cycle. We should expect the method using k = 1 without a cycle to perform the

¹⁶Following Phillips and Shi (2020), we implement the HP boosting estimator using an initial value of $\lambda = 1600$ and a maximum of 30 boosting iterations.

best within our group of estimators and this is indeed the case. However, for either k = 1 and k = 2 with a serially-correlated cycle, there is very little deterioration in performance. This suggests that the cost of being more agnostic about the true DGP is relatively low for our method and aligns with the empirical evidence presented earlier. Moreover, there is little difference in performance between choosing k = 1 and k = 2. This is because the data-generating process produces relatively homogenous dynamics even though it features time-varying volatility.

Table 1. White-Noise Cycle. This table reports results from data generated as in equations (26) and (27) with $\rho = 0$ for T = 150. Results are based on 500 simulations.

		1					
	min	10%	25%	50%	75%	90%	max
k = 1, no cycle	0.0860	0.1288	0.1517	0.1765	0.2081	0.2406	0.3264
k = 2, no cycle	0.0919	0.1325	0.1572	0.1849	0.2132	0.2455	0.3365
k = 1, w/cycle	0.0904	0.1296	0.1519	0.1791	0.2098	0.2417	0.3244
k = 2, w/cycle	0.0870	0.1332	0.1564	0.1859	0.2147	0.2459	0.3305
HP Boosting	0.1103	0.1697	0.1925	0.2240	0.2579	0.2931	0.3833

Quantiles of Root Mean-Square error (RMSE)

Quantiles of Mean Absolute Error (MAE)

	\min	10%	25%	50%	75%	90%	\max
k = 1, no cycle	0.0698	0.1036	0.1214	0.1430	0.1701	0.1985	0.2860
k = 2, no cycle	0.0744	0.1069	0.1257	0.1500	0.1755	0.2031	0.2905
k = 1, w/cycle	0.0744	0.1042	0.1216	0.1453	0.1721	0.2001	0.2776
k = 2, w/cycle	0.0700	0.1075	0.1263	0.1497	0.1759	0.2035	0.2863
HP Boosting	0.0883	0.1347	0.1541	0.1799	0.2097	0.2405	0.3101

In Table 2 we report our performance metrics for the DGP with $\rho = 0.4$. The presence of a serially correlated cycle presents a more challenging environment for differentiating the trend component. Consistent with this observation, we observe that both MSE and MAE are larger than their counterparts in Table 1. Perhaps unsurprisingly, the sparse trend estimators without a serially correlated cycle, uniformly underperform their counterparts that feature a dependent cycle component. In this specification the estimators can attribute movements in the cycle to the trend as we saw in the case of k = 1 in Figure 6. Combined with the results in Table 1, it becomes clear that a simple and robust approach is to always utilize the estimators with a dependent cycle as there is little cost when the errors are white noise while there are substantial benefits otherwise. These latter estimators uniformly outperform the HP boosting procedure across simulations. For example, the median RMSE/MAE of HP boosting exceeds by more than 30% the RMSE/MAE of the k = 2 with a serially correlated cycle. Finally, in unreported results, we also calculated RMSE and MAE for both Table 1 and 2 by omitting the first 10 and last 10 observations to ensure that boundary effects were not disproportionately affecting the results. We draw exactly the same conclusions from the trimmed results as from those above.

Table 2. Serially-Correlated Cycle. This table reports results from data generated as in equations (26) and (27) with $\rho = 0.4$ for T = 150. Results are based on 500 simulations.

min	10%	25%	50%	75%	90%	\max
0.1332	0.2108	0.2648	0.3499	0.4608	0.5576	0.8849
0.1012	0.1785	0.2126	0.2623	0.3245	0.3904	0.6031
0.1007	0.1520	0.1791	0.2251	0.2742	0.3272	0.5389
0.0939	0.1591	0.1949	0.2405	0.2883	0.3411	0.4695
0.1388	0.2414	0.2742	0.3259	0.3815	0.4387	0.5780
	min 0.1332 0.1012 0.1007 0.0939 0.1388	min 10% 0.1332 0.2108 0.1012 0.1785 0.1007 0.1520 0.0939 0.1591 0.1388 0.2414	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	min10%25%50%0.13320.21080.26480.34990.10120.17850.21260.26230.10070.15200.17910.22510.09390.15910.19490.24050.13880.24140.27420.3259	min 10% 25% 50% 75% 0.1332 0.2108 0.2648 0.3499 0.4608 0.1012 0.1785 0.2126 0.2623 0.3245 0.1007 0.1520 0.1791 0.2251 0.2742 0.0939 0.1591 0.1949 0.2405 0.2883 0.1388 0.2414 0.2742 0.3259 0.3815	min10%25%50%75%90%0.13320.21080.26480.34990.46080.55760.10120.17850.21260.26230.32450.39040.10070.15200.17910.22510.27420.32720.09390.15910.19490.24050.28830.34110.13880.24140.27420.32590.38150.4387

Quantiles of Root Mean-Square error (RMSE)

Quantiles of Mean Absolute Error (MAE)

					· /		
	\min	10%	25%	50%	75%	90%	max
k = 1, no cycle	0.1031	0.1702	0.2149	0.2832	0.3727	0.4470	0.7080
k = 2, no cycle	0.0802	0.1429	0.1714	0.2138	0.2643	0.3259	0.4782
k = 1, w/cycle	0.0839	0.1245	0.1467	0.1861	0.2299	0.2725	0.4523
k = 2, w/cycle	0.0743	0.1287	0.1582	0.1943	0.2357	0.2844	0.4112
HP Boosting	0.1122	0.1924	0.2229	0.2640	0.3137	0.3592	0.4684

6 Conclusion

Many variables of interest in economics and finance are obtained by decomposing observable processes into unobserved components. This is innately a challenging problem as assumptions are necessary to reliably separate the contributions from each component. In this paper, we adopt an approach motivated by the frequently used notion that the economic trend of interest is "slowmoving." Using novel survey data, we show that perceptions of slow-moving trends are characterized by sparse adjustments. We incorporate this observation to inform the priors in a general Bayesian framework for trend estimation. We combine "slab and spike" priors on changes in the underlying trend with a setup where the deviation from trend is allowed to be serially correlated. This ensures that our estimated trend features sparse adjustments and that it can capture the dynamics of commonly used series in economics. We illustrate in simulations that our method performs well.

We apply our method in three different empirical settings: post-war total factor productivity (TFP) growth, mean temperature changes in the U.S. over the last 140 years, and post-war real GDP growth and its sub-components. For TFP growth, our trend estimates confirm the suspected slowdown in trend productivity growth since the early 2000s. At the eve of the COVID pandemic, our estimates suggest that trend TFP growth is between 0.5% and 0.7%. In our climate application, we estimate a clearly upward underlying trend with an inflection point around 1980. Since then, the underlying trend in temperature has risen by about 1.0° C (Celsius) with a 95% posterior coverage interval of $(0.7, 1.4)^{\circ}$ C. The trend estimates at the end of our sample are consistent with a rise in temperatures of between 0.2 and 0.3° C over the next ten years. In our primary application, we estimate that trend real GDP growth in the U.S. has been falling since the 1960s and now stands

slightly above 2%. While we show that trend estimates from other popular procedures appear to inherit some residual business cycle cyclicality, our method produces slow-moving trend estimates which do not systematically co-vary with the cycle. We further analyze the underlying trend in the contributions to real GDP growth from its constituent components which generally exhibit very different dynamic properties. Despite that, we find that the individual trends estimated by these components can be aggregated to match closely the direct trend estimate from headline real GDP growth.

Our method can be readily extended to a multivariate setting. We saw in the SPD data that the longer-run forecasts appear to co-move across different economic variables, suggesting that they could be modeled jointly. Moreover, this could be beneficial as using external information or a multivariate model to impose economic restrictions (e.g., Phillips curve, Okun's law, etc.) is likely to sharpen identification and inference (e.g., see Müller, Stock, and Watson (2020)). Our method is also applicable in settings – for example, impulse responses in local projection models – that may require flexible de-trending of the data prior to the analysis. These extensions are currently under investigation by the authors.

References

- BAXTER, M., AND R. G. KING (1999): "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series," *Review of Economics and Statistics*, 81(4), 575–593.
- BEVERIDGE, S., AND C. NELSON (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, 7(2), 151–174.
- CANOVA, F. (2020): "FAQ: How do I measure the Output gap?," Press Discussion Paper 14943, CEPR.
- COIBION, O., Y. GORODNICHENKO, AND M. ULATE (2018): "The Cyclical Sensitivity in Estimates of Potential Output," *Brookings Papers on Economic Activity*, 49(2), 343–411.
- CRUMP, R. K., S. EUSEPI, M. GIANNONI, AND A. ŞAHIN (2019): "A Unified Approach to Measuring u^{*}," *Brookings Papers on Economic Activity*, 50(1), 143–214.
- CRUMP, R. K., S. EUSEPI, AND E. MOENCH (2018): "The Term Structure of Expectations and Bond Yields," Staff Report 775, Federal Reserve Bank of New York.
- DEL NEGRO, M., D. GIANNONE, M. P. GIANNONI, AND A. TAMBALOTTI (2018): "Safety, Liquidity, and the Natural Rate of Interest," *Brookings Papers on Economic Activity*, 49(1), 235–94.
- (2019): "Global trends in Interest Rates," Journal of International Economics, 118(1), 248–262.
- FERNALD, J. G. (2014): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," Working Paper 2012-19, Federal Reserve Bank of San Francisco.
- FERNALD, J. G., R. E. HALL, J. H. STOCK, AND M. W. WATSON (2017): "The Disappointing Recovery of Output after 2009," *Brookings Papers on Economic Activity*, 48(2).
- GOSPODINOV, N., AND E. MAASOUMI (2021): "Generalized aggregation of misspecified models: With an application to asset pricing," *Journal of Econometrics*, 222(1(B)), 451–467.
- GRANT, A. L., AND J. C. CHAN (2017): "Reconciling Output Gaps: Unobserved Components Model and Hodrick-Prescott filter," Journal of Economic Dynamics and Control, 75(1), 114–121.
- HAMILTON, J. D. (2018): "Why You Should Never Use the Hodrick-Prescott Filter," *Review of Economics and Statistics*, 100(5), 831–843.
- HODRICK, R. J. (2020): "An Exploration of Trend-Cycle Decomposition Methodologies in Simulated Data," Working Paper 26750, National Bureau of Economic Research.
- HODRICK, R. J., AND E. C. PRESCOTT (1980): "Postwar US Business Cycles: An Empirical Investigation," Discussion paper, Carnegie-Mellon University.
- HOLSTON, K., T. LAUBACH, AND J. C. WILLIAMS (2017): "Measuring the Natural Rate of Interest: International Trends and Determinants," *Journal of International Economics*, 108(1), S59–S75.
- KIM, S.-J., K. KOH, S. BOYD, AND D. GORINEVSKY (2009): "L1 Trend Filtering," *SIAM Review*, 51(2), 339–360.
- LAUBACH, T., AND J. C. WILLIAMS (2003): "Measuring the Natural Rate of Interest," *Review of Economics and Statistics*, 85(4), 1063–1070.
- LEE, S., Y. LIAO, M. H. SEO, AND Y. SHIN (2021): "Sparse HP filter: Finding kinks in the COVID-19 contact rate," *Journal of Econometrics*, 220(1), 158–180.

- MORAN, G. E., V. ROČKOVÁ, AND E. I. GEORGE (2019): "Variance Prior Forms for High-Dimensional Bayesian Variable Selection," *Bayesian Analysis*, 14(4), 1091–1119.
- MORLEY, J. C., C. R. NELSON, AND E. ZIVOT (2003): "Why Are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP so Different?," *Review of Economics and Statistics*, 85(2), 235–243.
- MÜLLER, U. K., J. H. STOCK, AND M. W. WATSON (2020): "An Econometric Model of International Growth Dynamics for Long-horizon Forecasting," *Review of Economics and Statistics*, forthcoming.
- MÜLLER, U. K., AND M. W. WATSON (2008): "Testing Models of Low-Frequency Variability," *Econometrica*, 76(5), 979–1016.
- (2020): "Low-Frequency Analysis of Economic Time Series," in *Handbook of Econometrics*, ed. by S. N. Durlauf, L. P. Hansen, J. J. Heckman, and R. L. Matzkin, vol. 7. Elsevier.
- PARK, T., AND G. CASELLA (2008): "Journal of the American Statistical Association,," Journal of the American Statistical Association, 103(482), 681–686.
- PHILLIPS, P. C. B., AND Z. SHI (2020): "Boosting: Why You Can Use the HP Filter," *International Economic Review*, forthcoming.
- ROUALDES, E. A. (2015): "Bayesian Trend Filtering," Working paper.
- ROČKOVÁ, V., AND E. I. GEORGE (2018): "The Spike-and-Slab LASSO," Journal of the American Statistical Association, 113(521), 431–444.
- SCHMIDT, D. F., AND E. MAKALIC (2013): "Estimation of stationary autoregressive models with the Bayesian LASSO," *Journal of Time Series Analysis*, 34, 517–531.
- TIBSHIRANI, R. J. (2014): "Adaptive Piecewise Polynomial Estimation Via Trend Filtering," Annals of Statistics, 42(1), 285–323.

Appendix A: Description of the Gibbs Sampling Algorithm

This Appendix provides detailed steps on how to implement the estimation procedure introduced in Section 3. Section A.1 provides steps for the case of a white-noise cycle whereas Section A.2 generalizes to the case of a serially-correlated cycle. For further details underlying each individual step, see Appendix A.3.

A.1 Trend Plus Noise

Throughout, we use the same notation as introduced in Section 3.1. In addition, define $S := \sum_{j=1}^{T-k} \theta_j$. The steps of the Gibbs sampler are as follows:

- 1. Draw $(\lambda_g \mid g, \theta) \sim \Gamma\left(S + r_g, \ \delta_g + \sum \frac{\theta_j |(D_k g)_j|}{\sqrt{\sigma^2}}\right)$.
- 2. Draw $(\sigma^2 \mid y, g, \gamma) \sim \text{INVERSE-GAMMA}\left(T \frac{k}{2}, \frac{1}{2}\left((y g)'(y g) + g'\Sigma_g^{-1}g\right)\right).$
- 3. Draw $(g \mid y, \sigma^2, \gamma, \omega) \sim \mathcal{N}\left(\left(I_T + \Sigma_g^{-1}\right)^{-1} y, \sigma^2 \left(I_T^{-1} + \Sigma_g^{-1}\right)^{-1}\right).$
- 4. Draw $(\xi \mid \theta) \sim \mathcal{B}eta(S+a, T-k-S+b).$
- 5. Draw each $(\theta_j \mid g, \sigma^2, \lambda_g, \lambda_0, \xi)$ as an independent Bernoulli random variable with

$$P(\theta_{j}=0) = \frac{(1-\xi)\lambda_{0}e^{-\sqrt{\lambda_{0}^{2}(D_{k}g)_{j}^{2}/\sigma^{2}}}}{(1-\xi)\lambda_{0}e^{-\sqrt{\lambda_{0}^{2}(D_{k}g)_{j}^{2}/\sigma^{2}}} + \xi\lambda_{g}e^{-\sqrt{\lambda_{g}^{2}(D_{k}g)_{j}^{2}/\sigma^{2}}}},$$
$$P(\theta_{j}=1) = \frac{\xi\lambda_{g}e^{-\sqrt{\lambda_{0}^{2}(D_{k}g)_{j}^{2}/\sigma^{2}}}}{(1-\xi)\lambda_{0}e^{-\sqrt{\lambda_{0}^{2}(D_{k}g)_{j}^{2}/\sigma^{2}}} + \xi\lambda_{g}e^{-\sqrt{\lambda_{g}^{2}(D_{k}g)_{j}^{2}/\sigma^{2}}}},$$
$$j = 1, \dots, T-k.$$

6. Draw each $(\frac{1}{\omega_j} \mid g, \theta_j, \lambda_g) \sim \text{INVERSE-GAUSSIAN}\left(\sqrt{\frac{((1-\theta_j)\lambda_0^2 + \theta_j\lambda_g^2)\sigma^2}{|(D_kg)_j|^2}}, (1-\theta_j)\lambda_0^2 + \theta_j\lambda_g^2\right)$ independently for $j = 1, \ldots, T-k$.

In our empirical implementations, we choose $\lambda_0 = 100$, $(r_g, \delta_g) = (0.25, 2.5)$ for k = 1, and $\lambda_0 = 4,000$, $(r_g, \delta_g) = (0.0075, 2.25)$ for k = 2.

A.2 Trend Plus Cycle

for

Throughout, we use the same notation as introduced in Section 3.2. In addition, define $v_{sc}(\gamma) = \sigma^2 \cdot \mathcal{V}_{\gamma,11}$. Note that the Metropolis-Hastings step (Step 8 below) is required to draw γ in the Gibbs sampler because (1) we assume a strictly stationary autoregressive process and (2) the prior variance of the trend g is a function of \mathcal{V}_{γ} . However, we follow the approach of Del Negro, Giannone, Giannoni, and Tambalotti (2019) but tailored to our setting, and use the conditional distribution of the LASSO estimator (conditional on the q initial conditions of c = y - g) as our candidate distribution. This results in a simple accept-reject rule which performs well in practice. Consequently, our Gibbs sampler remains computationally efficient with standard software implementation.¹⁷ The steps of the Gibbs sampler are as follows:

¹⁷In our simulations and empirical exercises, we use MATLAB.

1. Draw
$$(\lambda_g \mid g, \gamma, \sigma^2, \theta, r_g, \delta_g) \sim \Gamma\left(S + r_g, \ \delta_g + \sum_{j=1}^{T-k} \frac{\theta_j |(D_k g)_j|}{\sqrt{v_{sc}(\gamma)}}\right)$$

- 2. Draw $(\sigma^2 \mid y, g, \gamma) \sim \text{INVERSE-GAMMA}\left(T \frac{k}{2}, \frac{1}{2}\left((y-g)'\mathcal{V}_{\gamma}^{-1}(y-g) + \frac{1}{\mathcal{V}_{\gamma,11}}g'\Sigma_g^{-1}g\right)\right).$
- 3. Draw $(\xi \mid \theta) \sim \mathcal{B}eta(S+a, T-k-S+b).$
- 4. Draw each $(\theta_j \mid g, \sigma^2, \lambda_g, \lambda_0, \xi)$ as an independent Bernoulli random variable with

$$P(\theta_j = 0) = \frac{(1 - \xi)\lambda_0 e^{-\sqrt{\lambda_0^2(D_k g)_j^2/\sigma^2}}}{(1 - \xi)\lambda_0 e^{-\sqrt{\lambda_0^2(D_k g)_j^2/\sigma^2}} + \xi\lambda_g e^{-\sqrt{\lambda_g^2(D_k g)_j^2/\sigma^2}}},$$
$$P(\theta_j = 1) = \frac{\xi\lambda_g e^{-\sqrt{\lambda_g^2(D_k g)_j^2/\sigma^2}}}{(1 - \xi)\lambda_0 e^{-\sqrt{\lambda_0^2(D_k g)_j^2/\sigma^2}} + \xi\lambda_g e^{-\sqrt{\lambda_g^2(D_k g)_j^2/\sigma^2}}}.$$

for j = 1, ..., T - k.

- 5. Draw each $(\frac{1}{\omega_j} \mid g, \theta_j, \lambda_g) \sim \text{INVERSE-GAUSSIAN}\left(\sqrt{\frac{\left((1-\theta_j)\lambda_0^2 + \theta_j\lambda_g^2\right)\sigma^2}{|(D_kg)_j|^2}}, (1-\theta_j)\lambda_0^2 + \theta_j\lambda_g^2\right)$ independently for $j = 1, \ldots, T-k$.
- 6. Draw $(g \mid y, \sigma^2, \gamma, \omega) \sim \mathcal{N}\left(\left(\mathcal{V}_{\gamma}^{-1} + \frac{\sigma^2}{v_{sc}(\gamma)}\Sigma_g^{-1}\right)^{-1}\mathcal{V}_{\gamma}^{-1}y, \sigma^2\left(\mathcal{V}_{\gamma}^{-1} + \frac{\sigma^2}{v_{sc}(\gamma)}\Sigma_g^{-1}\right)^{-1}\right).$

7. Draw
$$(\lambda_{\gamma} \mid \gamma, r_{\gamma}, \delta_{\gamma}) \sim \Gamma\left(q + r_{\gamma}, \delta_{\gamma} + \sum_{j=1}^{q} |\gamma_j|\right).$$

- 8. Given the values in Steps 1–7, jointly draw $(\gamma, \Sigma_{\gamma})$ through the following steps:
 - (a) Draw each $(\frac{1}{\tau_j^{\rm o}}) \sim \text{INVERSE-GAUSSIAN}\left(\left|\frac{\lambda_{\gamma}}{\gamma_j}\right|, \lambda_{\gamma}^2\right)$ independently for $j = 1, \dots, q$. Set $\Sigma_{\gamma}^{\rm o} = \text{diag}(\tau_1^{\rm o}, \dots, \tau_q^{\rm o}).$
 - (b) Draw $\gamma^{o} \sim \mathcal{N}\left(CB'_{c,q}c_{q+1:T}, \sigma^{2}C\right)$, where $C = (B'_{c,q}B_{c,q} + \sigma^{2}(\Sigma_{\gamma}^{o})^{-1})^{-1}$ and $c_{q+1:T}$ is the vector comprised of the last T q elements of c.
 - (c) Let $\gamma^{(t)}$ denote the previous Gibbs draw of γ . If γ^{o} implies a stationary cycle component, set $(\gamma, \Sigma_{\gamma}) = (\gamma^{o}, \Sigma_{\gamma}^{o})$ with probability

$$\min\left\{ \left(\frac{v_{sc}(\gamma^{(t)})^{T-k} |\mathcal{V}_{\gamma^{(t)}}|}{v_{sc}(\gamma^{\mathrm{o}})^{T-k} |\mathcal{V}_{\gamma^{\mathrm{o}}}|} \right)^{1/2} \exp\left\{ \frac{c_{1:q}^{\prime} \left(\mathcal{V}_{\gamma^{(t)}}^{-1} - \mathcal{V}_{\gamma^{\mathrm{o}}}^{-1} \right) c_{1:q}}{2\sigma^2} + \frac{g' \Sigma_g^{-1} g}{2\sigma^2} \left(\frac{1}{v_{sc}(\gamma^{(t)})} - \frac{1}{v_{sc}(\gamma^{\mathrm{o}})} \right) \right\}, \ 1 \right\}$$

and set $(\gamma, \Sigma_{\gamma}) = \left(\gamma^{(t)}, \Sigma_{\gamma}^{(t)} \right)$ otherwise.

In our empirical implementations, we choose $(r_{\gamma}, \delta_{\gamma}) = (0.5, 0.25)$.

A.3 Details of Gibbs Sampler Steps

In this section, we provide some additional detail on the derivation of the Gibbs sampler for the trend plus cycle specification (Appendix A.2). The results apply readily to the simpler trend plus noise specification (Appendix A.1).

Detail for Steps 1–3

Drawing λ_g : In the text, we utilize the representation of the Laplace distribution as a scale mixture of normals with an exponential mixing density as in Park and Casella (2008) (i.e., equations (9) and (18)). We may write the conditional density of $(D_k g)_i$ directly as,

$$p((D_kg)_i \mid \gamma, \sigma^2, \theta_i, \lambda_g, \lambda_0) = (1 - \theta_j) \frac{\lambda_0}{2\sqrt{v_{sc}(\gamma)}} \exp\left\{-\lambda_0 \frac{|(D_kg)_i|}{\sqrt{v_{sc}(\gamma)}}\right\} + \theta_j \frac{\lambda_g}{2\sqrt{v_{sc}(\gamma)}} \exp\left\{-\lambda_g \frac{|(D_kg)_i|}{\sqrt{v_{sc}(\gamma)}}\right\}.$$

The conditional distribution of λ_g is then

$$p\left(\lambda_{g} \mid g, \gamma, \sigma^{2}, \theta, r_{g}, \delta_{g}\right) \propto p\left(g \mid \gamma, \sigma^{2}, \lambda_{g}, \theta, r_{g}, \delta_{g}\right) \cdot p\left(\lambda_{g} \mid r_{g}, \delta_{g}\right)$$
$$\propto \left(\prod_{j\mid\theta_{j}=1} \frac{\lambda_{g}}{2\sqrt{v_{sc}(\gamma)}} \exp\left\{-\lambda_{g} \frac{\left|(D_{k}g)_{j}\right|}{\sqrt{v_{sc}(\gamma)}}\right\}\right) \cdot (\lambda_{g})^{r_{g}-1} \exp\{-\delta_{g}\lambda_{g}\}.$$

Thus, as in Step 1, we have

$$\left(\lambda_g \mid g, \gamma, \sigma^2, \theta, r_g, \delta_g\right) \sim \Gamma\left(S + r_g, \ \delta_g + \sum \frac{\theta_j |(D_k g)_j|}{\sqrt{v_{sc}(\gamma)}}\right).$$

Drawing σ^2 : Writing out the full joint density and using the prior for σ^2 in equation (12), we obtain

$$p(y,g,\omega,\lambda_g,\gamma,\tau,\lambda_\gamma,\sigma^2) \propto \frac{1}{\sigma^T} \exp\left\{-\frac{(y-g)'\mathcal{V}_{\gamma}^{-1}(y-g)}{2\sigma^2}\right\} \frac{1}{\sigma^{T-k}} \exp\left\{-\frac{g'\Sigma_g^{-1}g}{2v_{sc}(\gamma)}\right\} \frac{1}{\sigma^2}$$
$$\propto \frac{1}{\sigma^{2T+2-k}} \exp\left\{-\frac{(y-g)'\mathcal{V}_{\gamma}^{-1}(y-g) + \frac{1}{\mathcal{V}_{\gamma,11}}g'\Sigma_g^{-1}g}{2\sigma^2}\right\}.$$

Thus, as in Step 2, we obtain

$$\left(\sigma^2 \mid y, g, \gamma\right) \sim \text{Inverse-Gamma}\left(T - \frac{k}{2}, \frac{1}{2}\left((y - g)'\mathcal{V}_{\gamma}^{-1}(y - g) + \frac{1}{\mathcal{V}_{\gamma,11}}g'\Sigma_g^{-1}g\right)\right).$$

Drawing ξ : Using the priors introduced in equations (10) and (11) we obtain,

$$p(\xi \mid \theta, a, b) \propto p(\theta \mid \xi) \cdot p(\xi \mid a, b) \propto (1 - \xi)^{T - k - S + b - 1} \xi^{S + a - 1}.$$

Thus, as in Step 3, we obtain

$$(\xi \mid \theta, a, b) \sim \mathcal{B}eta(S+a, T-k-S+b).$$

Detail for Steps 4–5

Our method for sampling from the joint distribution of ω_j and θ_j is to first sample θ_j from its marginal distribution and then sample ω_j from its corresponding conditional distribution. We first derive the joint distribution of ω_j and θ_j :

$$\begin{split} p(\omega_j, \theta_j \mid g, \gamma, \sigma^2, \lambda_g, \lambda_0, \xi) &\propto p(g \mid \omega_j, \theta_j, \gamma, \sigma^2) \cdot p(\omega_j \mid \lambda_g, \lambda_0, \theta_j) \cdot p(\theta_j \mid \xi) \\ &\propto \left((1 - \theta_j)(1 - \xi) \frac{\lambda_0^2}{2} \exp\left\{\frac{-\lambda_0^2 \omega_j}{2}\right\} + \theta_j \xi \frac{\lambda_g^2}{2} \exp\left\{\frac{-\lambda_g^2 \omega_j}{2}\right\} \right) \\ &\qquad \times \frac{1}{\omega_j^{1/2}} \exp\left\{-\frac{\omega_j^{-1}(D_k g)_j^2}{2\sigma^2}\right\}. \end{split}$$

This last equation can be rewritten as:

$$p(\omega_{j},\theta_{j} \mid g,\gamma,\sigma^{2},\lambda_{g},\lambda_{0},\xi) \propto (1-\theta_{j}) \frac{(1-\xi)\lambda_{0}^{2}}{\omega_{j}^{1/2}} \exp\left\{-\frac{\omega_{j}^{-1}(D_{k}g)_{j}^{2}}{2\sigma^{2}} - \frac{\lambda_{0}^{2}\omega_{j}}{2}\right\} + \theta_{j} \frac{\xi\lambda_{g}^{2}}{\omega_{j}^{1/2}} \exp\left\{-\frac{\omega_{j}^{-1}(D_{k}g)_{j}^{2}}{2\sigma^{2}} - \frac{\lambda_{g}^{2}\omega_{j}}{2}\right\}.$$
(A.1)

Drawing θ_j : We first rewrite equation (A.1) as

$$p(\omega_j, \theta_j \mid g, \gamma, \sigma^2, \lambda_g, \lambda_0, \xi) \propto (1 - \theta_j) \frac{\mathcal{C}_0}{\omega_j^{1/2}} \exp\left\{-\mathcal{A}\omega_j^{-1} - \mathcal{B}_0\omega_j\right\} + \theta_j \frac{\mathcal{C}_1}{\omega_j^{1/2}} \exp\left\{-\mathcal{A}\omega_j^{-1} - \mathcal{B}_1\omega_j\right\}$$

with $\mathcal{A} = \frac{(D_k g)_j^2}{2\sigma^2}$, $\mathcal{B}_0 = \frac{\lambda_0^2}{2}$, $\mathcal{B}_1 = \frac{\lambda_g^2}{2}$, $\mathcal{C}_0 = (1 - \xi)\lambda_0^2$, $\mathcal{C}_1 = \xi\lambda_g^2$. We now can integrate over ω_j to get the marginal distribution of θ_j

$$p(\theta_j \mid g, \gamma, \sigma^2, \lambda_g, \lambda_0, \xi) = \int_0^\infty p(\omega_j, \theta_j \mid g, \gamma, \sigma^2, \lambda_g, \lambda_0, \xi) \, d\omega_j$$

$$\propto \int_0^\infty \left[(1 - \theta_j) \frac{\mathcal{C}_0}{\omega_j^{1/2}} \exp\left\{-\mathcal{A}\omega_j^{-1} - \mathcal{B}_0\omega_j\right\} + \theta_j \frac{\mathcal{C}_1}{\omega_j^{1/2}} \exp\left\{-\mathcal{A}\omega_j^{-1} - \mathcal{B}_1\omega_j\right\} \right] d\omega_j$$

$$\propto (1 - \theta_j)(1 - \xi)\lambda_0 e^{-\sqrt{\lambda_0^2(D_k g)_j^2/\sigma^2}} + \theta_j \xi \lambda_g e^{-\sqrt{\lambda_g^2(D_k g)_j^2/\sigma^2}}.$$

The last step follows since the antiderivative of

$$\int_0^\infty \frac{\mathcal{C}}{\omega_j^{1/2}} \exp\left\{-\frac{\mathcal{A}}{\omega_j} - \mathcal{B}\omega_j\right\} d\omega_j$$

is

$$\frac{\mathcal{C}\sqrt{\pi}}{2\sqrt{\mathcal{B}}} \left(e^{2\sqrt{\mathcal{AB}}} \operatorname{erf}\left(\frac{\sqrt{\mathcal{A}} + \sqrt{\mathcal{B}}\omega_j}{\sqrt{\omega_j}}\right) - e^{-2\sqrt{\mathcal{AB}}} \operatorname{erf}\left(\frac{\sqrt{\mathcal{A}} - \sqrt{\mathcal{B}}\omega_j}{\sqrt{\omega_j}}\right) \right) + \frac{1}{2\sqrt{\mathcal{B}}} \left(e^{2\sqrt{\mathcal{AB}}} \operatorname{erf}\left(\frac{\sqrt{\mathcal{A}} - \sqrt{\mathcal{B}}\omega_j}{\sqrt{\omega_j}}\right) \right) + \frac{1}{2\sqrt{\mathcal{AB}}} \left(e^{2\sqrt{\mathcal{AB}}} \operatorname{erf}\left(\frac{\sqrt{\mathcal{A}} - \sqrt{\mathcal{AB}}\omega_j}{\sqrt{\omega_j}}\right) \right) + \frac{1}{2\sqrt{\mathcal{AB}}} \left(e^{2\sqrt{\mathcal{AB}}} \operatorname{erf}\left(\frac{\sqrt{\mathcal{A}} - \sqrt{\mathcal$$

where $\operatorname{erf}(x)$ is defined as $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$. Consequently,

$$\int_0^\infty \frac{\mathcal{C}}{\omega_j^{1/2}} \exp\left\{-\frac{\mathcal{A}}{\omega_j} - \mathcal{B}\omega_j\right\} \ d\omega_j = \frac{\mathcal{C}\sqrt{\pi}}{\sqrt{\mathcal{B}}} e^{-2\sqrt{\mathcal{A}\mathcal{B}}}.$$

Consequently, θ_j is a Bernoulli Random variable with

$$P(\theta_j = 0) = \frac{(1 - \xi)\lambda_0 e^{-\sqrt{\lambda_0^2(D_k g)_j^2/\sigma^2}}}{(1 - \xi)\lambda_0 e^{-\sqrt{\lambda_0^2(D_k g)_j^2/\sigma^2}} + \xi\lambda_g e^{-\sqrt{\lambda_g^2(D_k g)_j^2/\sigma^2}}},$$
$$P(\theta_j = 1) = \frac{\xi\lambda_g e^{-\sqrt{\lambda_g^2(D_k g)_j^2/\sigma^2}}}{(1 - \xi)\lambda_0 e^{-\sqrt{\lambda_0^2(D_k g)_j^2/\sigma^2}} + \xi\lambda_g e^{-\sqrt{\lambda_g^2(D_k g)_j^2/\sigma^2}}}.$$

Drawing ω_j : We now derive the conditional distribution of ω_j given θ_j :

$$p(\omega_j \mid g, \gamma, \sigma^2, \lambda_g, \lambda_0, \theta_j) \propto p(g \mid \omega_j, \gamma, \sigma^2) \cdot p(\omega_j \mid \lambda_g, \lambda_0, \theta_j)$$

$$\propto \frac{1}{|\Sigma_g|^{1/2}} \exp\left\{-\frac{g'\Sigma_g^{-1}g}{2\sigma^2}\right\} \left((1-\theta_j)\frac{\lambda_0^2}{2}\exp\left\{\frac{-\lambda_0^2\omega_j}{2}\right\} + \theta_j\frac{\lambda_g^2}{2}\exp\left\{\frac{-\lambda_g^2\omega_j}{2}\right\}\right)$$

$$\propto \frac{1}{|\Sigma_g|^{1/2}} \exp\left\{-\frac{(D_kg)'\operatorname{diag}(\omega_1, \dots, \omega_{T-k})D_kg}{2\sigma^2}\right\} \exp\left\{\frac{-((1-\theta_j)\lambda_0^2 + \theta_j\lambda_g^2)\omega_j}{2}\right\}$$

The last step follows because θ_j only takes on the values of 0 or 1; in particular, only for these two values will the last two lines be proportional as a function of ω_j (their quotient is independent of ω_j , but dependent upon θ_j and λ_0, λ_g). Thus,

$$p(\omega_j \mid g, \gamma, \sigma^2, \lambda_g, \lambda_0, \theta_j) \propto \frac{1}{\omega_j^{1/2}} \exp\left\{-\frac{\sum_{i=1}^{T-k} \omega_i^{-1} (D_k g)_i^2}{2\sigma^2}\right\} \exp\left\{\frac{-\left((1-\theta_j)\lambda_0^2 + \theta_j \lambda_g^2\right)\omega_j}{2}\right\}.$$

Using the change of variables, $\eta_j = \frac{1}{\omega_j}$, we obtain

$$p(\eta_j \mid g, \gamma, \sigma^2, \lambda_g, \lambda_0, \theta_j) \propto \frac{1}{\eta_j^{3/2}} \exp\left\{\frac{-\eta_j (D_k g)_j^2 - \frac{\left((1-\theta_j)\lambda_0^2 + \theta_j \lambda_g^2\right)\sigma^2}{\eta_j}}{2\sigma^2}\right\}.$$

Thus, as in Step 5, we obtain,

$$1/\omega_j \sim \text{Inverse-Gaussian}\left(\sqrt{\frac{\left((1-\theta_j)\lambda_0^2+\theta_j\lambda_g^2\right)\sigma^2}{|(D_kg)_j|^2}}, \ (1-\theta_j)\lambda_0^2+\theta_j\lambda_g^2\right).$$

Detail for Steps 6–7

Drawing g: This follows directly from equations (17) and (18) as

$$p(g \mid y, \sigma^2, \gamma, \omega) \propto p(y \mid g, \sigma^2, \gamma) \cdot p(g \mid \sigma^2, \gamma, \omega) \propto \exp\left\{-\frac{(y-g)'\mathcal{V}_{\gamma}^{-1}(y-g)}{2\sigma^2}\right\} \exp\left\{-\frac{g'\Sigma_g^{-1}g}{2v_{sc}(\gamma)}\right\}$$

Thus, as in Step 6, we obtain that

$$\left(g \mid y, \sigma^2, \gamma, \omega\right) \sim \mathcal{N}\left(\left(\mathcal{V}_{\gamma}^{-1} + \frac{\sigma^2}{v_{sc}(\gamma)}\Sigma_g^{-1}\right)^{-1}\mathcal{V}_{\gamma}^{-1}y, \ \sigma^2\left(\mathcal{V}_{\gamma}^{-1} + \frac{\sigma^2}{v_{sc}(\gamma)}\Sigma_g^{-1}\right)^{-1}\right).$$

Drawing λ_{γ} : We will make use of equations (24) and (25) along with

$$p(\gamma, \Sigma_{\gamma}) \propto \mathbb{1}_{\gamma} \cdot \det(\Sigma_{\gamma})^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\gamma' \Sigma_{\gamma}^{-1} \gamma\right\} p(\Sigma_{\gamma}),$$
 (A.2)

where $\mathbb{1}_{\gamma} = 1$ indicates a strictly stationary cycle. Assuming $\mathbb{1}_{\gamma} = 1$, integrating out Σ_{γ} simplifies the prior for γ to

$$p(\gamma \mid \lambda_{\gamma}) = \prod_{j=1}^{q} \left(\frac{\lambda_{\gamma}}{2} \exp \left\{ -\lambda_{\gamma} |\gamma_{j}| \right\} \right),$$

so that

$$p(\lambda_{\gamma} \mid \gamma, r_{\gamma}, \delta_{\gamma}) \propto p(\gamma \mid \lambda_{\gamma}, r_{\gamma}, \delta_{\gamma}) \cdot p(\lambda_{\gamma} \mid r, \delta)$$
$$\propto \left(\prod_{j=1}^{q} \frac{\lambda_{\gamma}}{2} \exp\left\{-\lambda_{\gamma} |\gamma_{j}|\right\}\right) \cdot \lambda_{\gamma}^{r_{\gamma}-1} \exp\{-\delta_{\gamma} \lambda_{\gamma}\}.$$

Thus, as in Step 7, we obtain

$$(\lambda_{\gamma} \mid \gamma, r_{\gamma}, \delta_{\gamma}) \sim \Gamma\left(q + r_{\gamma}, \ \delta_{\gamma} + \sum_{j=1}^{q} |\gamma_j|\right).$$

Detail for Step 8

Drawing $\Sigma_{\gamma}^{o} \mid \gamma^{o}$: This step consists of jointly drawing all τ_{j} and follows by similar steps as for ω_{j} . We have,

$$p(\tau_j \mid \gamma) \propto \det(\Sigma_{\gamma})^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\gamma'\Sigma_{\gamma}^{-1}\gamma\right\} p(\Sigma_{\gamma}).$$

Using the change of variables, $\psi_j = \frac{1}{\tau_j}$, we obtain

$$p(\psi_j \mid \gamma) \propto \frac{1}{\psi_j^{3/2}} \exp\left\{-\frac{\psi_j \gamma_j^2 - (\lambda_\gamma)_j^2 \psi_j^{-1}}{2}\right\}.$$

Thus, we obtain

$$1/\tau_j^{\mathrm{o}} \sim \mathrm{Inverse-Gaussian}\left(\left| \frac{(\lambda_{\gamma})_j}{\gamma_j} \right|, \ (\lambda_{\gamma})_j^2 \right).$$

Drawing $\gamma^{o} \mid \Sigma_{\gamma}^{o}$: Equations (20) and (21) imply that

$$p(\gamma \mid c, g, \Sigma_{g}^{-1}, \Sigma_{\gamma}^{-1}, \sigma^{2}) \propto p(c \mid \gamma, g, \sigma^{2}) \cdot p(\gamma \mid g, \Sigma_{g}^{-1}, \Sigma_{\gamma}^{-1}, \sigma^{2}) \\ \propto p(c_{q+1:T} \mid c_{1:q}, \gamma, \sigma^{2}) \cdot p(c_{1:q} \mid \gamma, \sigma^{2}) \cdot p(g \mid \gamma, \Sigma_{g}^{-1}, \sigma^{2}) \cdot p(\gamma \mid \Sigma_{\gamma}^{-1}) \\ \propto \exp\left\{-\frac{(CB'_{c,q}c_{q+1:T} - \gamma)'C^{-1}(CB'_{c,q}c_{q+1:T} - \gamma)}{2\sigma^{2}}\right\} \\ \times \frac{1}{(\mathcal{V}_{\gamma,11})^{(T-k)/2}|\mathcal{V}_{\gamma}|^{1/2}} \exp\left\{-\frac{c'_{1:q}\mathcal{V}_{\gamma}^{-1}c_{1:q}}{2\sigma^{2}}\right\} \exp\left\{-\frac{g'\Sigma_{g}^{-1}g}{2v_{sc}(\gamma)}\right\}.$$
(A.3)

We cannot directly sample from equation (A.3). However, following Del Negro, Giannone, Giannoni, and Tambalotti (2019), we can use a Metropolis-Hastings step utilizing $\mathcal{N}\left(CB'_{c,q}c_{q+1:T}, \sigma^2 C\right)$ as the candidate distribution. Let $\tilde{p}(\gamma, c, g, \Sigma_g^{-1}, \Sigma_\gamma^{-1}, \sigma^2)$ be its corresponding pdf. Starting at $\gamma^{(t)}$, we draw a candidate γ° from the candidate distribution and calculate the acceptance probability

$$\begin{split} \alpha &= \min\left\{\frac{p\left(\gamma^{\mathrm{o}}|c^{(t+1)}, \dots, (\sigma^{2})^{(t+1)}\right)}{p\left(\gamma^{(t)}|c^{(t+1)}, \dots, (\sigma^{2})^{(t+1)}\right)} \frac{\tilde{p}\left(\gamma^{(t)}, c^{(t+1)}, \dots, (\sigma^{2})^{(t+1)}\right)}{\tilde{p}\left(\gamma^{\mathrm{o}}, c^{(t+1)}, \dots, (\sigma^{2})^{(t+1)}\right)}, 1\right\} \\ &= \min\left\{\left(\frac{v_{sc}(\gamma^{(t)})^{T-k}|\mathcal{V}_{\gamma^{(t)}}|}{v_{sc}(\gamma^{\mathrm{o}})^{T-k}|\mathcal{V}_{\gamma^{\mathrm{o}}}|}\right)^{1/2} \exp\left\{\frac{c_{1:q}^{\prime}\left(\mathcal{V}_{\gamma^{(t)}}^{-1} - \mathcal{V}_{\gamma^{\mathrm{o}}}^{-1}\right)c_{1:q}}{2\sigma^{2}} + \frac{g'\Sigma_{g}^{-1}g}{2\sigma^{2}}\left(\frac{1}{\left(v_{sc}(\gamma^{(t)})} - \frac{1}{v_{sc}(\gamma^{\mathrm{o}})}\right)\right)\right\}, 1\right\} \end{split}$$

Our second accept-reject step is to take our candidate draw of $(\gamma^{o}, \Sigma_{\gamma}^{o})$ (drawn using the above procedure) and accept with probability $\alpha = \mathbb{1}_{\gamma^{o}}$, where $\mathbb{1}_{\gamma^{o}}$ is an indicator for whether γ^{o} produces a stationary cycle.