Working Paper Series
Congressional Budget Office
Washington, D.C.

Quantifying the Uncertainty of Long-Term Economic Projections

U. Devrim Demirel
Congressional Budget Office
devrim.demirel@cbo.gov

James Otterson
(formerly of the Congressional Budget Office)

Working Paper 2022-07

April 2022

To enhance the transparency of the work of the Congressional Budget Office and to encourage external review of that work, CBO's working paper series includes papers that provide technical descriptions of official CBO analyses as well as papers that represent independent research by CBO analysts. Papers in this series are available at http://go.usa.gov/xUzd7.

The authors are grateful to Bruce E. Hansen of the University of Wisconsin, Lars Hansen of the University of Chicago, Ulrich Müller of Princeton University, James Stock of Harvard University, and David Wilcox of the Peterson Institute for International Economics for valuable comments on this paper. The authors also benefited greatly from comments by Molly Dahl, Mark Doms, Wendy Edelberg (formerly of CBO), John Kitchen, Jeffrey Kling, Julie Topoleski, Jeffrey Werling, and other members of CBO's staff. The authors gratefully acknowledge the editorial assistance provided by Rebecca Lanning.

Abstract

This paper presents a practical method for assessing the uncertainty of long-term economic projections. Economic variables play a central role in the Congressional Budget Office's analysis of federal spending and revenues, and the uncertainty of economic projections is a key driver of the uncertainty about the agency's budget projections. The presented method quantifies the uncertainty of economic variables by using simulations from a multivariate statistical model in which variables are formulated as sums of unobserved stationary and nonstationary components. Experiments on artificial data demonstrate that the method performs fairly well compared with alternative methods in terms of long-term predictive accuracy and coverage.

Keywords: long-term uncertainty, prediction interval, unobserved components, state-space models

JEL Classification: C32, C53, E17

Contents

1 Introduction 1
2 The Method 3
3 Estimation and Prediction Intervals 5
3.1 The State-Space Form 5
3.2 Priors 6
3.3 Estimating the Model 7
3.4 Constructing Prediction Intervals 8
3.5 Long-Term Correlations 12
4 Assessing the Predictive Performance of the UC Method 14
4.1 Data Generating Processes 14
4.2 Monte Carlo Analysis 16
4.3 Results 18
5 Conclusion 19
References 20

1 Introduction

Each year, the Congressional Budget Office publishes 30-year projections for federal revenues, spending, and debt in its annual *Long-Term Budget Outlook* report. Those projections are based on the assumption that current laws governing federal spending and revenues generally remain unchanged. But even if current laws remained unchanged, budgetary outcomes would differ from those in CBO's projections, primarily because of unexpected changes in economic and demographic factors. To quantify the uncertainty of budget projections arising from those factors, CBO uses a series of simulations for spending and revenues, each reflecting an alternative path for a set of key economic and demographic variables, and examines how federal deficits and debt would evolve under each path. Those simulations produce a range of budgetary outcomes and thereby illustrate how the uncertainty of economic and demographic variables translates into budgetary uncertainty.[1]

This paper presents a practical method for assessing the long-term uncertainty of the economic variables that underpin budget projections. For the purposes of CBO's analysis, long-term uncertainty is defined as the uncertainty of variables' average values over a long period of time—typically, several decades. (For example, across the 30-year periods that occurred since early 1950s, the average growth rate of total factor productivity, or TFP, varied by about 1 percentage point; the long-term uncertainty of TFP growth arises from that variability.) A significant part of the uncertainty of budgetary outcomes in the long term stems from persistent changes in economic trends rather than transitory economic fluctuations, or business cycles (which tend to average closer to zero over longer periods). Examining the variability of those trends is crucial for assessing the risks to the long-term sustainability of federal debt and for designing policies that can help to mitigate the budgetary effects of unfavorable economic developments. Analyzing the likely range of long-term economic outcomes can also help lawmakers better evaluate the size and timing of the policy changes that they may choose to implement to address the long-term budget imbalance.

CBO's analysis of economic uncertainty is based on simulations from a multivariate unobserved components (UC) model that represent a range of potential future paths for economic variables, such as the rate of productivity growth and the interest rate on federal debt. In the model, variables are specified as sums of individually unobserved stationary and nonstationary components. Distinguishing between those components is important for accurately measuring the uncertainty of variables over the long term. That is because, although the short-term uncertainty of the variables arises from the variability of both components, the long-term uncertainty stems, in large part, from the variability of the nonstationary component.

---

[1] Other important sources of the uncertainty about CBO's budget projections include the uncertainty of future government policies and of the models the agency uses to project future economic and budgetary outcomes.

Uncertainty in long-term economic projections can be expressed by constructing prediction intervals for long-horizon averages of economic time series. A prediction interval is a range in which a future outcome is expected to fall with a specified probability. For example, in *The 2019 Long-Term Budget Outlook*, CBO estimated that there was roughly a two-thirds chance that the average annual growth rate of TFP over the next three decades would be in the range of 0.6 percent to 1.6 percent. CBO also estimated long-term prediction intervals for other major economic variables, including the real interest rate and the unemployment rate. Changes in those variables have important effects on the federal budget: Faster TFP growth and a lower unemployment rate would mean higher taxable income and revenues and, therefore, lower deficits and debt. By contrast, higher interest rates would increase the government's interest payments, causing deficits and debt to be larger than they would be otherwise.

Estimating prediction intervals for long-term averages of economic time series presents significant challenges. First, there is relatively limited information about the variability of long-term averages in the available sample data for most economic variables. Estimating that variability is particularly difficult when the prediction horizon is long relative to the sample size, resulting in very few observations of long-term averages. For example, since the end of World War II, there have not yet been three nonoverlapping 30-year periods, whereas there have been 25 nonoverlapping 3-year periods. In addition, long-term statistical properties of a variable depend on the exact form of persistence that the variable displays. For example, as discussed in Müller and Watson (2018), random walks have different long-term properties than fractionally integrated or serially correlated stationary variables. But there is often limited information in the sample data to precisely distinguish between different forms of persistence.

The recent research offers new insights into long-term prediction and inference. Müller and Watson (2016) develop methods for conducting inference about the long-term variability of economic time series and construct prediction sets for long-term averages. Those methods provide theoretically valid prediction intervals under a wide range of stochastic processes exhibiting different forms of persistence, including local-level, local-to-unity, and fractionally integrated forms. Zhou et al. (2010) propose an alternative approach to constructing prediction intervals based on estimated long-term standard deviations and sample quantiles. Chudy at al. (2020) propose adjustments that improve the predictive performance of the method of Zhou et al. (2010) over long horizons. In a related branch of the literature, Stock (1996) and Phillips (1998) show that standard methods of estimating long-term prediction intervals and impulse response functions (which describe how economic variables respond to various shocks over time) produce biased estimates if variables exhibit a high degree of persistence, and Pesavento and Rossi (2006) propose an approach to constructing confidence intervals for impulse responses when variables are highly persistent.

Compared with some of the recently developed methods, including those of Müller and Watson (2016) and Chudy et al. (2020), the simulation-based UC approach presented in this paper can more easily accommodate a multivariate framework. Assessing the interactions of multiple variables is important for CBO's analysis because correlations between economic variables matter for budgetary outcomes. For example, if slower output growth is associated with lower interest rates, a reduction in revenues caused by slower growth in TFP would be paired with lower interest payments on debt, which could offset the effect on deficits from slower TFP growth. Müller and Watson (2018) develop a method for assessing the long-term covariability of economic variables. But implementing that method in a setting with more than two variables carries a heavy computational burden.

The simulation-based UC method offers a practical solution to assessing the long-term variability of more than two economic variables. The method is computationally manageable and relatively simple to implement, and it offers a seamless analysis of both short- and long-term uncertainty. It is, however, less robust under alternative forms of persistence that economic variables may display than are some of the recently developed, state-of-the-art methods (including those of Müller and Watson, 2018). Nevertheless, a comparison of the predictive performances of different methods based on Monte Carlo experiments indicates that, despite being less robust than some of the recently developed methods, the UC method fares reasonably well in terms of long-term predictive accuracy and coverage. Ultimately, however, all methods are subject to the fundamental limitation that there are very few observations of long-term averages in any limited sample period. Therefore, the amount of low-frequency information that can be gleaned from the available data and our ability to precisely estimate long-term prediction intervals are limited.

Section 2 of this paper lays out the UC model and describes how it is used to decompose variables into stationary and nonstationary components. Section 3 discusses the estimation of model parameters and presents the estimated prediction intervals for TFP growth, the unemployment rate, and the real interest rate, which are based on a series of simulations each reflecting an alternative future path for those variables. Section 4 conducts Monte Carlo experiments on simulated data produced by artificial processes that mimic the historical behavior of a set of key macroeconomic variables and also compares the predictive performance of the UC method with that of alternative methods. Section 5 provides concluding remarks.

2 The Method

CBO assesses the relationship between economic factors and budgetary outcomes by examining how federal spending and revenues would change if key economic variables differed from the agency's baseline projections (see CBO, 2016). Previously, those assessments included variables such as the size of the labor force, the growth rate of TFP, the interest rate on federal debt, and the growth rate of federal spending per beneficiary for Medicare and Medicaid. One set of simulations shows the range of budgetary outcomes that would occur if a single economic variable differed from CBO's baseline. A second set of simulations shows the range of outcomes that would occur if multiple variables differed from baseline projections simultaneously. Those simulations illustrate the sensitivity of budgetary outcomes to economic factors but shed no light on the probabilities of different economic scenarios. The UC approach to conducting multivariate simulations offers a basis for probabilistic assessments of future economic developments.

To distinguish between transitory and permanent movements (underpinning the short- and long-term uncertainty of the variables), the UC approach formulates each variable as a sum of individually unobserved stationary and nonstationary components by using the following specification:

$$Y_t = \mu_t + X_t + w_t \quad (1)$$

$$\mu_t = \mu_{t-1} + \varepsilon_t$$

$$X_t = A_1 X_{t-1} + \cdots + A_q X_{t-q} + u_t$$

The vector Y_t collects the variables of interest, including the growth rate of TFP, the unemployment rate, and the real interest rate on 10-year Treasury securities (calculated by subtracting inflation as measured by the consumer price index from the nominal interest rate on those securities). We use annual data for all three variables, and our sample period runs from 1953 to 2021 (see Table 1).

Table 1.

Statistical Summaries of the Variables

Percentage Points

	Mean	Median	SD	17th Percentile	83rd Percentile
TFP Growth	1.3	1.1	1.6	0.1	2.9
Unemployment Rate	5.9	5.6	1.6	4.4	7.4
Real Interest Rate	2.1	2.2	2.5	0.3	3.9

Data source: Congressional Budget Office.

The sample includes annual data from 1953 to 2021. The unemployment rate is the number of unemployed people as a percentage of the civilian labor force. The real interest rate in each year is calculated by subtracting inflation as measured by the percentage change in the consumer price index for all urban consumers over 12 months from the nominal interest rate on 10-year Treasury notes. Numbers are rounded to the nearest tenth of a percentage point.

SD = standard deviation; TFP = total factor productivity.

The vector random walk process, μ_t, captures the nonstationary components of the variables. The vector autoregression (VAR) process, X_t, represents the stationary components, which may themselves be highly persistent. The third component, w_t, is a white noise process, which may represent one-off shocks or measurement error. Vectors w_t, ε_t, and u_t are normally distributed with zero means and covariance matrices Ω_w, Ω_ε, and Ω_u, respectively.

The UC framework (1) nests a number of relevant special cases. The framework generalizes a stationary VAR model, which corresponds to the case $\Omega_\varepsilon = 0$. Furthermore, restrictions on the elements of $A_1, \ldots, A_q$ and Ω_u deliver models similar to those of Laubach and Williams (2003), Holston et al. (2017), and Lewis and Vazquez-Grande (2019), which have been used to estimate the natural rate of interest.

The UC model provides a useful and simple framework for analyzing both the short- and the long-term variation of economic variables, but it is subject to limitations. One limitation stems from the assumption of homoskedasticity—the specification that the variances of the error terms ε_t and u_t are constant over time. Also, ε_t, u_t, and w_t are assumed to be independent of one another (which is a commonly adopted specification in the literature that uses UC models). In addition, the elements of $A_1, \ldots, A_q$ are time invariant. Relaxing all or a subset of those assumptions can result in a more general framework but also increases the number of parameters to be estimated. With small samples, increasing the number of estimated parameters by incorporating additional structure into the model (for example, by introducing time-varying parameters or cross-correlations across all elements of ε_t, u_t, and w_t) tends to result in greater estimation uncertainty and less accurate out-of-sample predictions. The relatively parsimonious specification of model (1) helps to reduce estimation uncertainty while incorporating sufficient detail to account for both the stationary and the nonstationary components of the variables.

3 Estimation and Prediction Intervals

In the first step of our analysis, we use Bayesian methods to estimate the parameters of the UC model (1) and the historical paths of the nonstationary (μ_t) and stationary (X_t and w_t) components of the variables. We then construct prediction intervals by conducting stochastic simulations (also known as Monte Carlo analysis). Those simulations yield ranges (determined with a specified probability) in which future values of variables are expected to fall.

3.1 The State-Space Form

We first express the system of equations (1) in the following state-space form:

$$\mathbf{S}_{t+1} = \mathbf{F} \cdot \mathbf{S}_t + \mathbf{v}_{t+1} \qquad (2)$$

$$\mathbf{Y}_t = \mathbf{M} + \mathbf{H} \cdot \mathbf{S}_t + w_t$$

where the state vector, $\mathbf{S}_t$, and the vector of innovations, $\mathbf{v}_t$, are defined as

$$\mathbf{S}_t = \begin{bmatrix} \mu_t \\ X_t \\ X_{t-1} \\ \vdots \\ X_{t-q+1} \end{bmatrix} \quad \mathbf{v}_t = \begin{bmatrix} \varepsilon_t \\ u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and the matrices **F**, **H**, and **M** have the following forms:

$$\mathbf{F} = \begin{bmatrix} I_{n\times n} & \mathbf{0}_{n\times n} & \cdots & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & A_1 & \cdots & A_q \\ \mathbf{0}_{n\times n} & I_{n\times n} & \cdots & \mathbf{0}_{n\times n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n\times n} & \cdots & I_{n\times n} & \mathbf{0}_{n\times n} \end{bmatrix} \quad \mathbf{H} = [I_{n\times n} \;\; I_{n\times n} \; \mathbf{0}_{n\times n} \; \cdots \; \mathbf{0}_{n\times n}] \quad \mathbf{M} = \mathbf{0}_{n\times 1},$$

where n denotes the number of variables in Y_t, and $I_{n\times n}$ is an n-by-n identity matrix. We set the parameter q (the lag length of the VAR component) to 2 and specify the covariance matrix of $\mathbf{v}_t$ as

$$\begin{bmatrix} \Omega_\varepsilon & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n(q-1)} \\ \mathbf{0}_{n\times n} & \Omega_u & \mathbf{0}_{n\times n(q-1)} \\ \mathbf{0}_{n(q-1)\times n} & \mathbf{0}_{n(q-1)\times n} & \mathbf{0}_{n(q-1)\times n(q-1)} \end{bmatrix},$$

where the term $\mathbf{0}_{x\times y}$ represents an x-by-y matrix of all zeros, and

$$\Omega_\varepsilon = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \end{bmatrix}. \tag{3}$$

As the covariance matrix (3) suggests, we allow each variable to have a nonstationary component and also allow those components to be correlated with each other.

3.2 Priors

In the next step, we set the prior distributions of the parameters. The variability of the nonstationary, or trend, components of Y_t (elements of μ_t) plays an important role in our analysis. For the standard deviations of the trend shocks (elements of ε_t), we choose an inverse-gamma (IG) distribution. The domain of the IG distribution is the set of positive real numbers and therefore excludes zero. By incorporating the prior belief that trend shocks have strictly positive variance, the IG specification counters the tendency of maximum-likelihood estimates of that variance to be biased toward zero—a phenomenon known as the pileup problem (see Stock, 1994). We set the means of the standard deviations of the trend shocks to reflect the prior belief that the elements of X_t drive about one quarter of the variation of each element of Y_t that is generated by cycles with periodicities longer than 12 years (and estimated by using a band-pass filter).

In addition, we incorporate the prior belief that the autoregressive parameters of the stationary component, X_t, are independent and normally distributed with the following properties:

$$E\left[(A_s)_{jk}\right] = 0 \text{ for all } s, j, \text{ and } k.$$

$$Var\left[(A_s)_{jk}\right] = \begin{cases} \frac{\lambda^2}{s^2} & \text{if } j = k \\ \omega \frac{\lambda^2}{s^2}\frac{\sigma_j^2}{\sigma_k^2}, & \text{otherwise.} \end{cases}$$

The term $(A_s)_{jk}$ represents the j^{th} row and k^{th} column element of the matrix A_s (where s denotes lag length), and $E[x]$ and $Var[x]$ denote the expected value and variance of x, respectively. The parameter λ controls the tightness of the prior information, ω measures the extent to which the tightness of the priors on a variable's own lags differs from the tightness of those on other variables' lags, and the ratio σ_j^2/σ_k^2 corrects for the differences in the scales of the variables j and k. The prior distribution of the autoregressive parameters of X_t are centered around zero because the stationary components of the variables, once separated from the random walk components, are expected to display a substantial degree of mean reversion. The specification of variances (which is based on a set of beliefs known as the Minnesota prior) reflects the view that each variable's own lagged values provide better prior information about that variable's dynamics than do the lags of other variables, and distant lags of a variable are less important drivers of its variation than are the variable's more recent lags (see Litterman, 1986). We set the tightness parameter, λ^2, to 0.64, which implies that there is roughly a two-thirds prior probability that the elements of A_1 are between −0.8 and 0.8, and follow Doan (1990) in setting w to 0.5. Following Litterman (1986), we set σ_j (for j = 1, 2, ..., n) by using the standard deviations of the residuals from regressions of the elements of Y_t on a constant and q of their own lagged values. We use diffuse priors for the remaining parameters.

3.3 Estimating the Model

Estimated parameters include the elements of **F**, **M**, and **H** and the covariance matrices of $\mathbf{v}_t$ and w_t. We calculate the log of the posterior density of the parameters (up to a normalizing constant) by summing the log of the sample likelihood, $\log L(Y_t|\,\Theta)$, and the log of the prior density of the parameters, $\log \pi\,(\Theta)$. That is, we calculate

$$Q(\Theta|Y_1, Y_2, \ldots, Y_T) = \sum_{t=1}^{T} \log L\,(Y_t|\,\Theta) + \log \pi\,(\Theta) \tag{4}$$

where the vector Θ contains the parameters of the model and T denotes the size of the sample. We compute the sample likelihood by starting with an initial guess for the state vector, $\mathbf{S}_t$, and recursively applying the Kalman filter (see Hamilton, 1994).[2] We simulate the parameters' posterior distribution by using a Markov chain Monte Carlo (MCMC) procedure.[3] The MCMC method produces a joint distribution that approximates the true posterior distribution of the parameters.

Figure 1 shows the historical paths of the nonstationary components of the growth rate of TFP, the unemployment rate, and the real interest rate estimated by using the mode of the posterior distribution described by equation (4). The nonstationary component of the real interest rate is estimated to be more volatile than those of TFP growth and the unemployment rate. In addition, the trend component of TFP growth is, on average, greater in the first half of the sample than in the second half, and it shows a noticeable drop that starts in early 2000s. The nonstationary component of the real interest rate also exhibits a downward drift that seems to have started in the mid-1980s and accelerated after the early 2000s. The properties of the estimated trend components of TFP growth and the real interest rate are consistent with the empirical regularities highlighted in the literature examining the causes and consequences of the so-called productivity slowdown (see, for example, Gordon, 2014; Fernald, 2015) and secular stagnation (see Laubach and Williams, 2003; Holston et al., 2017). Those patterns are insensitive to specifying the unemployment rate as a stationary variable (based on the finding that standard univariate tests reject the null hypothesis of the unit root at conventional levels of significance) by setting all corresponding elements of the covariance matrix (3) (that is, σ_{12}, σ_{23}, and σ_2^2) to zero. However, the simulated distribution of the parameters is sensitive to the assumptions about the number of variables that incorporate a nonstationary component.

3.4 Constructing Prediction Intervals

We quantify uncertainty by constructing prediction intervals, or ranges, that contain the future values of the variables of interest with a specified probability. Our approach to constructing prediction intervals captures both the uncertainty stemming from the randomness of shocks (that is, the uncertainty of the future values of the elements of $\mathbf{v}_t$ and w_t) and parameter uncertainty. The former reflects forecast uncertainty about what the future will hold. The latter arises because the model parameters (the elements of **F**, **M**, and **H** and the covariance matrices of $\mathbf{v}_t$ and w_t) are

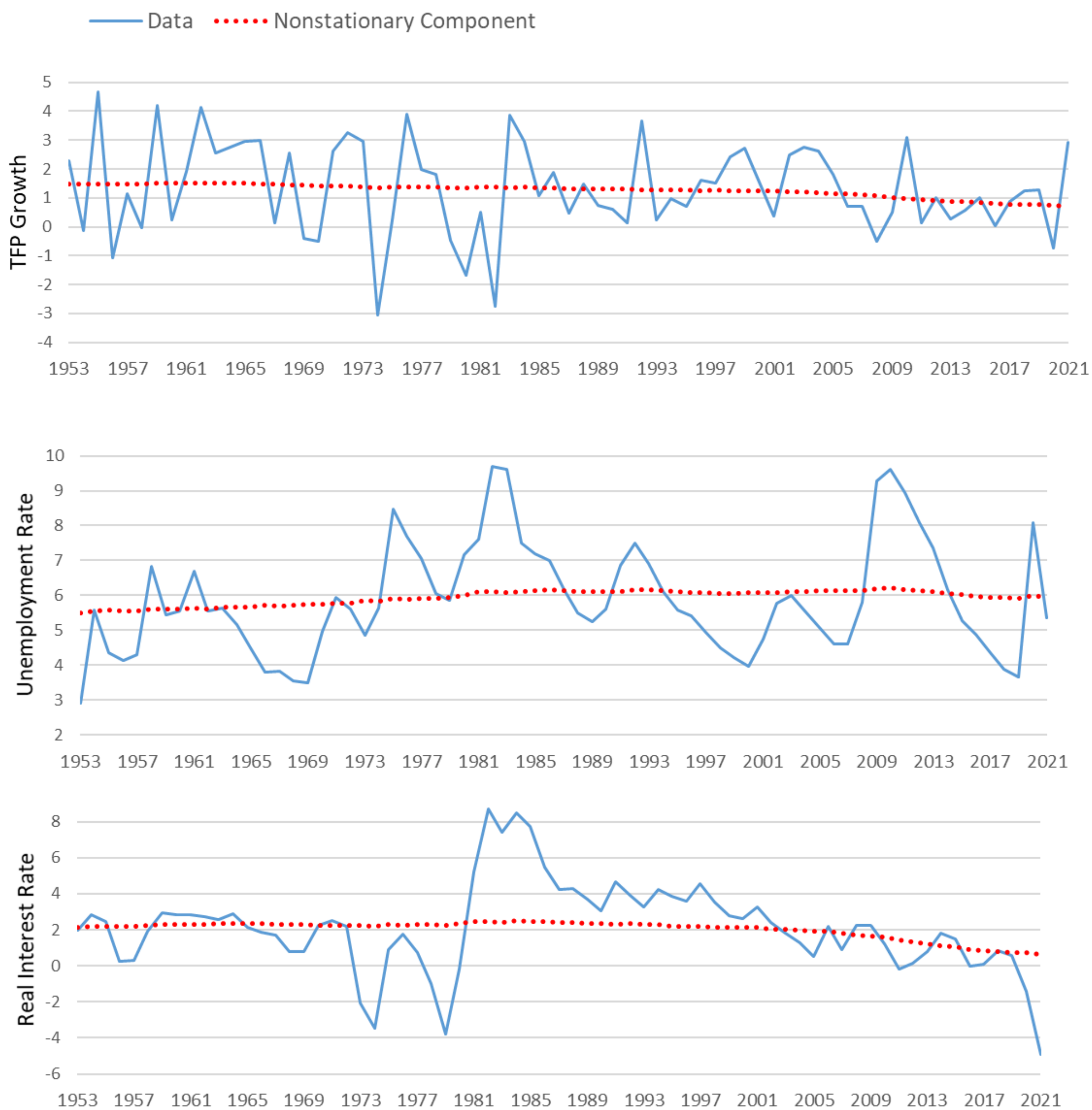
[2] To reduce the sensitivity of the results to the initial state vector, we assign large values to the diagonal elements of the initial state covariance matrix.

[3] See Chib (2001) for a detailed discussion of MCMC methods.

Figure 1.

Estimated Nonstationary Components of the Variables

Percentage Points



Data source: Congressional Budget Office.

All variables are defined in percentage points. The unemployment rate is the number of unemployed people as a percentage of the civilian labor force. The real interest rate is calculated by subtracting inflation as measured by the consumer price index for all urban consumers from the nominal interest rate on 10-year Treasury notes.

TFP = total factor productivity.

unknown and their estimates are based on sample data and, therefore, are subject to sampling uncertainty.

We capture parameter uncertainty in our simulations by drawing sets of parameter values from the posterior distribution produced by the MCMC method. To capture forecast uncertainty, for each draw from the posterior distribution of the parameters, we draw a large number of sequences—each the same length as the data sample—for the elements of $\mathbf{v}_t$ and w_t. Then, for each sequence, we iterate the system of equations (2) forward and construct 30-year series, each corresponding to an alternative future path for the variables in Y_t. For each series, we then calculate 30-year moving averages and construct prediction intervals containing the middle two-thirds of the simulated distribution (that is, the range of values between the 17th and 83rd percentiles) of the averages of each variable.

CBO does not use the simulation-based UC method to estimate the long-term averages of economic variables. Those estimates are based on CBO's full forecasting framework and detailed analysis. The agency uses the UC method only to quantify the uncertainty of those averages. To ensure that the ranges produced by the UC method are consistent with CBO's central projections, when conducting simulations to construct prediction intervals, we calculate variables' deviations from the unconditional forecasts produced by the UC model in each simulation and apply those deviations to CBO's central projections.

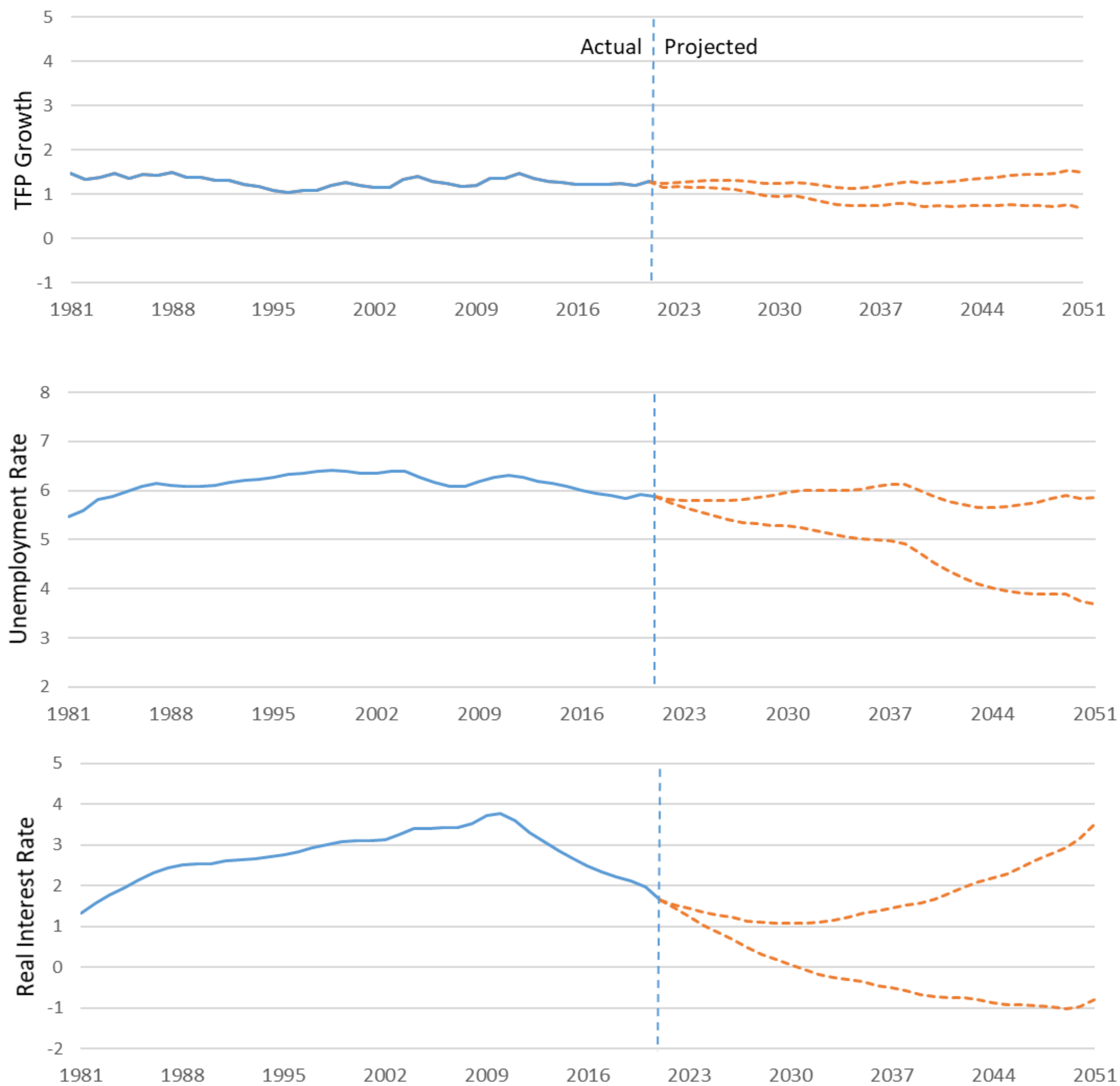
An important statistical property of the macroeconomic variables we examine is the skewness in the distribution of the unemployment rate: The right tail of that distribution extends farther out than its left tail (see McGrane, 2022, and Dupraz et al., 2020). We capture that property in our simulations by including the log of the unemployment rate in the vector Y_t rather than the unemployment rate itself. Under the linear structure of the UC model, shocks to the stationary and nonstationary components of Y_t induce symmetric movements in the log of the unemployment rate. Because the log function is a concave transformation, symmetric increases and decreases in the log of the unemployment rate translate into increases in the unemployment rate that are larger in magnitude than decreases, thereby inducing a distribution for the unemployment rate that is skewed toward the right tail.

Figure 2 shows the prediction intervals for the 30-year moving averages of the variables. We also construct prediction intervals for averages over the next 15 and 30 years (see Table 2). Our results indicate that there is a two-thirds chance that the average annual growth rate of TFP will be between roughly 0.7 percent and 1.5 percent over the next three decades (when rounded to the nearest tenth of a percentage point). That range is somewhat narrower than the 67 percent prediction interval estimated in Müller and Watson (2016) for the average TFP growth rate over 25 years. Although we do not report it here, we also find that the intervals for the average TFP

Figure 2.

Prediction Intervals for 30-Year Moving Averages

Percentage Points



Data source: Congressional Budget Office.

All variables are defined in percentage points. The upper and lower bounds indicated by the dashed lines correspond to the 17th and 83rd percentiles of the 30-year moving averages of each variable. The unemployment rate is the number of unemployed people as a percentage of the civilian labor force. The real interest rate is calculated by subtracting inflation as measured by the consumer price index for all urban consumers from the nominal interest rate on 10-year Treasury notes.

TFP = total factor productivity.

growth over the next few years are wider than those for longer-term averages. That is because the transitory component of TFP growth—which drives a relatively large share of that variable's overall variance in the sample period—averages closer to zero over longer periods.

The range that covers two-thirds of simulated outcomes is −0.8 percent to 3.5 percent for the 30-year average real interest rate and 3.7 percent to 5.9 percent for the unemployment rate. Those ranges are considerably wider than the ranges for the growth rate of TFP because the low-frequency components of those series (especially those of the real interest rate) are more variable than the low-frequency components of the TFP growth rate.

Co-movements of variables' low-frequency components are a key factor underpinning the range of long-term outcomes. The estimates for Ω_ε and Ω_u indicate that shocks to the nonstationary components of the variables (elements of ε_t) are correlated, as are the shocks to the stationary components of the variables (elements of u_t). Therefore, the factors driving, for example, the long-term uncertainty of TFP growth also affect the long-term uncertainty of the unemployment and real interest rates.

Table 2.

Prediction Intervals for Long-Term Averages

Percentage Points

	17th Percentile	83rd Percentile
15-Year Horizon		
TFP growth rate	0.7	1.5
Unemployment rate	3.6	5.8
Real interest rate	-1.3	2.4
30-Year Horizon		
TFP growth rate	0.7	1.5
Unemployment rate	3.7	5.9
Real interest rate	-0.8	3.5

Data source: Congressional Budget Office.

The unemployment rate is the number of unemployed people as a percentage of the civilian labor force. The real interest rate is calculated by subtracting inflation as measured by the consumer price index for all urban consumers from the nominal interest rate on 10-year Treasury notes. Numbers are rounded to the nearest tenth of a percentage point.

TFP = total factor productivity.

3.5 Long-Term Correlations

The correlations among economic variables matter for budgetary outcomes, in both the short and the long term. For example, if the growth rate of TFP is positively correlated with the real interest rate, then faster economic growth resulting in higher-than-projected federal revenues could also mean higher interest rates and, therefore, increased interest payments on federal debt, which could offset the effect of higher revenues on deficits. If TFP growth correlates negatively with the unemployment rate, then higher revenues and lower deficits resulting from faster growth of taxable income would be paired with reduced federal spending for programs such as unemployment insurance, which would further reduce deficits.

A useful summary of the long-term statistical relationships between economic variables is provided by long-term correlations, which are defined as the correlations between long-horizon averages of variables. They differ from conventional correlations when variables exhibit serial dependence—a property that permeates virtually all of the economic variables we examine. The simulation-based approach to constructing prediction intervals outlined in the previous section can also be used to assess the long-term correlations between economic variables. Each simulated sequence of the vector Y_t—produced under a given parameterization of the model (2)—yields a value for the 30-year average of each macroeconomic variable. We estimate the long-term correlations between variables by calculating the correlations among those 30-year averages. We then estimate the distribution of each long-term correlation (and construct confidence intervals) by repeatedly drawing from the simulated chain of model parameters and calculating the correlations produced under each draw.

The median estimate of the long-term correlation between the growth rate of TFP and the unemployment rate is −0.43 (see Table 3). An increase in the average TFP growth rate is, therefore, associated with a drop in the average unemployment rate over the long term. That result is consistent with the findings of Müller and Watson (2018) and Staiger et al. (2001): Using different estimation methods, both studies find a negative long-term correlation between TFP growth and the unemployment rate. The estimated standard error is large (primarily reflecting parameter uncertainty), which indicates the inexactness of estimating long-term correlations in relatively short samples.

Table 3.

Long-Term Correlations of the Growth Rate of TFP With the Unemployment Rate and Real Interest Rates

	Long-Term Correlations With TFP Growth Rate	
	Unemployment Rate	Real Interest Rate
Median	-0.43	0.18
Standard Error	0.37	0.36

Data source: Congressional Budget Office.

The unemployment rate is the number of unemployed people as a percentage of the civilian labor force. The real interest rate is calculated by subtracting inflation as measured by the consumer price index for all urban consumers from the nominal interest rate on 10-year Treasury notes.

TFP = total factor productivity.

The median estimate of the long-term correlation between the growth rate of TFP and the real interest rate is 0.18 (but the standard error is large, as is that of the long-term correlation between TFP growth and the unemployment rate). A positive long-term correlation between TFP growth and the real interest rate is consistent with the view that one contributor to the steady decline in real interest rates over the past four decades is the slowdown in the trend growth rate of the economy. Holston et al. (2017) find evidence that the natural rate of interest (defined as the real interest rate consistent with an economy operating at full employment) and the trend growth rate of output display a strong co-movement. Likewise, Del Negro et al. (2017) find that one of the main factors driving the persistent decline in the natural rate of interest has been slow economic growth.[4] A positive correlation between the long-term growth rate of the economy and the real interest rate also emerges in a broad class of dynamic general equilibrium models. For example, in the standard neoclassical growth model, an increase in the trend growth rate of the economy has the same effect on the real interest rate as a drop in households' time discount factor, which boosts the real rate by reducing the desire to save.[5]

4 Assessing the Predictive Performance of the UC Method

In this section, we evaluate the predictive performance of the UC method by using Monte Carlo experiments, whereby we apply the method to artificially generated data with known statistical properties. The data generating processes (DGPs) we use in these experiments mimic the key characteristics of U.S. macroeconomic time series, in particular, the growth rate of TFP and the real interest rate. We also compare the accuracy of our method with those of alternative approaches.

4.1 Data Generating Processes

Assessments of a method's long-term predictive accuracy that are based on actual macroeconomic time series are often uninformative because the period over which variables are averaged is long relative to the length of the available data sample, resulting in very few observations that can be used to assess a method's performance. For example, the available sample for TFP growth rates does not yet include three nonoverlapping 30-year subsamples, thereby offering very few actual observations for the 30-year average TFP growth rate. Because of that limitation, we test the predictive performance of the UC method on simulated data produced by artificial DGPs with specified properties.

[4] By using a sample going back to the early 19th century, Hamilton et al. (2016) find that the correlation between the growth rate of the economy and the real interest rate has been positive after World War II. But they also find that the correlation is weak in the full sample. See Gamber (2020) for a review of the literature examining the relationship between the slowdown in trend growth and the decline in real interest rates.

[5] CBO's long-term economic projections also incorporate a positive relationship between the projected growth rate of TFP and the real interest rate. See Congressional Budget Office (2021).

To generate data that closely mimic the statistical properties of macroeconomic variables such as the TFP growth rate and the real interest rate, we adopt DGPs that produce stationary and nonstationary variation as well as level and volatility breaks. We conduct two sets of experiments: First, we generate data for three variables by adopting the following independent processes used in Müller and Watson's (2016) Monte Carlo analysis:

DGP 1: $x_{1t} = \mu_{1t} + r_{1t}$ where $\mu_{1t} = \mu_{1t-1} + s_{1t}\delta_{1t}$ and $r_{1t} \sim N(0, \omega_1^2)$,

DGP 2: $x_{2t} = \mu_{2t} + r_{2t}$ where $\mu_{2t} = \mu_{2t-1} + s_{2t}\delta_{2t}$, $r_{2t} = \rho r_{2t-1} + \epsilon_t$, and $\epsilon_t \sim N(0, \omega_2^2)$,

DGP 3: $x_{3t} = \mu_{3t} + \sigma_t r_{3t}$ where $r_{3t} = r_{3t-1} + \epsilon_t$, $\log(\sigma_t) = \log(\sigma_{t-1}) + s_{3t}\delta_{3t}$, and $\mu_{3t} \sim N(0,1)$, $\epsilon_t \sim N(0, \omega_3^2)$.

The three DGPs cover various key properties of U.S. macroeconomic time series in the post–World War II period. DGP 1 mimics the behavior of TFP growth: The zero-mean stationary component, r_{1t}, gives the process the appearance of a serially independent series. But the random walk component, μ_{1t}, displays permanent level shifts. Those shifts occur sporadically, and their timing is determined by the indicator variable s_{1t}, which follows an independently and identically distributed Bernoulli process with success probability $P(s_{1t} = 1) = p$. The level shifts are of size $|\delta_{1t}| = \delta$, where positive and negative signs have the same chance of occurring.[6] DGP 2 mimics the behavior of the real interest rate, exhibiting a high degree of serial correlation and persistent drifts. The main feature distinguishing DGP 2 from DGP 1 is serial correlation in the error term r_{2t}, which now follows an autoregressive process.[7] Finally, DGP 3 involves shifts in volatility; that process is motivated by the properties of the macroeconomic data, such as the persistent reduction in the volatility of the growth rates of real gross domestic product (GDP) and various other measures of real economic activity that is estimated to have started in mid-1980s and ended with the 2008 financial crisis. The three DGPs together provide a useful testing framework because they feature a more general set of dynamic properties (including time-varying volatility) than those the UC framework is designed to capture.

In the second set of experiments, we allow the nonstationary components of the DGPs to be correlated with each other. Specifically, we set $s_{it} = s_t$ for all i's (where the Bernoulli process s_t has the success probability p), and we specify the joint distribution of the shocks to the nonstationary components as

---

[6] The distributions of s_{it} are independent and identical.

[7] Müller and Watson (2016) consider alternative values for p and in their Monte Carlo analysis. In the first set of experiments, we generate data by using Müller and Watson's larger values (which they denote p_{large} and d_{large}) and set p so that permanent shifts occur, on average, once in a decade for all three DGPs, and we set δ to 0.375, 1, and 0.1875 for DGP 1, DGP 2, and DGP 3, respectively. Also following Müller and Watson, we set the autoregressive parameter, ρ, to 0.98. We set the standard deviations w_1, w_2, and w_3 to 1.5, 0.46, and 0.707, respectively.

$$\begin{bmatrix}\delta_{1t}\\ \delta_{2t}\\ \delta_{3t}\end{bmatrix} \sim N\left(\mathbf{0}_{3\times 1}, \begin{bmatrix}\sigma_1^2 & \sigma_{12} & \sigma_{13}\\ \sigma_{12} & \sigma_2^2 & \sigma_{23}\\ \sigma_{13} & \sigma_{23} & \sigma_3^2\end{bmatrix}\right)$$

where both the diagonal and nondiagonal elements of the covariance matrix are different from zero.[8]

4.2 Monte Carlo Analysis

To test the predictive performance of our method, we use the DGPs to produce a large number of data sequences—each of length T—for a vector process including three variables. We then split each sequence into two parts, one containing the first T_1 observations in the series and the other containing the remaining $T_2 = T - T_1$ observations. We treat the first part of each sequence as actual sample data and use each method to construct prediction intervals for the average values of the variables over the next T_2 years. Then, for each variable, we examine whether the prediction interval contains the average value of the variable in the second part of the sequence. We calculate the coverage probability for each variable by assessing the frequency with which the estimated prediction intervals contain the actual average values (that is, by measuring the rate at which actual T_2-year averages fall within the prediction intervals). In all experiments, we set the averaging horizon, T_2, to 30 years. To facilitate comparison with the results of Müller and Watson (2016), we set T_1 to 65 years, which matches the length of the samples they simulate in their Monte Carlo analysis.[9]

Coverage probability provides a useful criterion for assessing the accuracy of a method's prediction intervals. If one seeks to construct intervals that contain, for example, two-thirds or roughly 67 percent of the potential average values of a variable, then one prefers the coverage probability to be as close as possible to the nominal value of 67 percent. Values smaller or greater than 67 percent indicate undercoverage or overcoverage, which might occur, for example, if the method in question fails to accurately capture the true composition of stationary and nonstationary components of the variables.

We test the accuracy of the UC method by comparing its coverage probabilities with the nominal value of 67 percent. We then compare the method's coverage with those of two alternative

[8] We use the following values for the elements of the covariance matrix: $\sigma_1^2 = 0.10$ $\sigma_2^2 = \sigma_3^2 = 0.12, \sigma_{12} = 0.02, \sigma_{13} = 0.01, \sigma_{23} = -0.03$.

[9] We compute coverage probabilities by generating a total of 250,000 data sequences (or 500 sets with each set containing 500 sequences). Each sequence includes 95 years of simulated data for each variable.

methods. The first method is based on the estimated long-term, or asymptotic, covariance matrix of the variables.[10] The second one is the method developed in Müller and Watson (2016).

4.2.1 Asymptotic Covariance Method. Let $\bar{Y}_{T+1:T+n}$ denote the average of the vector Y_t between periods $T + 1$ and $T + n$. If the variables in Y_t are stationary, then the large-sample distribution of $\bar{Y}_{T+1:T+n}$ is $N(\mu, n^{-1}\Omega)$, where μ and Ω, respectively, denote the mean and the asymptotic variance of Y_t. However, the mean is unknown and, therefore, has to be estimated on the basis of sample data. Incorporating the additional uncertainty arising from the estimation of the mean, the distribution of $\bar{Y}_{T+1:T+n}$ can be found as

$$N\left[\bar{Y}_{1:T,}\ (T^{-1} + n^{-1})\Omega\right], \tag{5}$$

where $\overline{Y}_{1:T} = T^{-1}\Sigma_{t=1}^{T} Y_t$ denotes the sample average of Y_t. The asymptotic covariance matrix, Ω, is also unknown and has to be estimated. To that end, we use a well-established HAC (heteroskedasticity-and-autocorrelation-consistent) covariance matrix estimator, which takes the form

$$\widehat{\Omega} = \sum_{k=-(T-1)}^{T-1} g(k) \left(\frac{1}{T} \sum_{t=k+1}^{T} \widetilde{Y}_t \widetilde{Y}'_{t-k} \right) \text{ with } g(k) = \begin{cases} 1 - \frac{|k|}{r} \text{ if } |k| < r \\ 0 \text{ if otherwise} \end{cases},$$

where $\tilde{Y}_{1:T}$ denotes the difference between Y_t and the sample mean $\bar{Y}_{1:T}$, and the parameter r (known as the lag truncation number) is a positive integer.[11]

To construct prediction intervals for T_2-year averages of the variables, we generate a large number of draws from the distribution (5) after replacing Ω with its estimated counterpart and setting the averaging horizon n to T_2. We then compute the coverage probabilities in the same way as explained earlier.

4.2.2 Müller and Watson's Method. The method developed in Müller and Watson (2016) provides a basis for constructing prediction intervals for long-term averages of economic time series and yields asymptotically valid prediction sets under many different forms of long-term persistence, including local-level, local-to-unity, and fractionally integrated forms. Müller and Watson compute coverage probabilities for their prediction sets by using artificial data from five different data generating processes including DGPs 1, 2, and 3. They report two sets of coverage probabilities. The first set is based on a Bayesian approach and accounts for the uncertainty of the parameters governing variables' long-term persistence properties. The second set is based on a more robust approach to estimating prediction intervals and achieves frequentist coverage

---

[10] For an application of that approach to estimating the long-run correlation between productivity growth and real interest rates, see Hansen and Seshadri (2013).

[11] We set the value of that parameter by using the selection criterion proposed in Newey and West (1994).

across the space of parameters that determine the persistence properties of the variables. We use the coverage probabilities based on Müller and Watson's robust approach as a benchmark for evaluating the coverage performance of the UC method.

4.3 Results

Table 4 shows the coverage probabilities of the UC method and those produced by the alternative methods. Values that are significantly larger or smaller than the nominal value of 67 percent suggest overcoverage or undercoverage and indicate a mismatch between the estimated prediction intervals and the likely range of outcomes for the variable being examined.

Overall, the UC method performs reasonably well under the three DGPs. Although the method's coverage rates under DGP 1 (the process with serially independent transitory shocks) and DGP 2 (the process with nonstationary and serially correlated stationary components) are different from the nominal value, the rate is correct under DGP 3 (the process with stochastic volatility breaks) when DGPs have independent nonstationary components. Moreover, when the nonstationary components of the variables are correlated, the UC method exhibits slightly improved coverage performance under DGPs 1 and 2 compared with the case in which the nonstationary components are independent.

Table 4.

Coverage Probabilities

	DGP 1	DGP 2	DGP 3
With Independent Nonstationary Components			
UC method	0.71	0.61	0.67
Asymptotic covariance method	0.37	0.19	0.21
Müller and Watson's (2016) method	0.67	0.70	0.65
With Correlated Nonstationary Components			
UC method	0.70	0.63	0.69
Asymptotic covariance method	0.42	0.24	0.23
Müller and Watson's (2016) method	n.a.	n.a.	n.a.

Data sources: Congressional Budget Office; Müller and Watson (2016)

The coverage probabilities of Müller and Watson's (2016) method are reported on the eighth row of Table 3 in that work.

DGP = data generating process; UC = unobserved components; n.a.= not available.

Müller and Watson's (2016) method produces coverage rates that are slightly above the nominal value of 67 percent under DGP 2 and slightly below that value under DGP 3, and it achieves nominal coverage under DGP 1 when the nonstationary components of the variables are uncorrelated. Those produced by the asymptotic covariance method are much smaller than the nominal value under all DGPs in both sets of experiments. That method severely undercovers when variables contain nonstationary as well as stationary components because the estimated asymptotic covariance matrix underlying the method is based on the assumption that variables are stationary. When variables exhibit nonstationary behavior, the assumption of stationarity results in prediction intervals that are too narrow and, consequently, frequent occurrences in which the actual outcome falls outside of the estimated interval. By allowing the variables to incorporate nonstationary components, the simulation-based UC method exhibits improved coverage performance relative to the asymptotic covariance method when variables contain nonstationary components.

5 Conclusion

CBO projects that, if current laws governing taxes and spending generally remained unchanged, federal debt as a percentage of GDP would surpass its highest level in history (reached shortly after World War II) within a decade and continue to rise over the following several decades. However, those projections are subject to substantial uncertainty, especially in the long term. A large part of the long-term uncertainty about budget projections arises from the uncertainty of economic and demographic variables. This paper outlines a practical approach to quantifying economic uncertainty by using simulations from a multivariate UC model (without quantifying the budgetary uncertainty that results from the uncertainty of economic and demographic variables). Monte Carlo experiments show that the simulation-based UC approach performs reasonably well when the variables incorporate both stationary and nonstationary components.

Assessing the long-term uncertainty of economic variables is challenging because of unpredictable structural breaks many variables undergo over long periods and scarcity of low-frequency information in the sample data (stemming from the small size of the available data sample relative to the prediction horizon). Compared with some recently developed state-of-the-art approaches to constructing long-term prediction intervals, including the methods of Müller and Watson (2016, 2018) and Chudy et al. (2020), the UC method is simpler to implement, can easily handle a multivariate framework—a key feature given the important role that interactions among economic variables play in shaping budgetary outcomes—and offers a unified analysis of both short- and long-term uncertainty. However, it is less robust under alternative forms of persistence that economic variables may exhibit. As discussed in Müller and Watson (2018), formulating a robust approach that seeks to account for all economically relevant forms of long-term persistence can result in very wide prediction intervals. But limiting the domain of the analysis to a narrower range of potential persistence patterns may lead to less reliable inference and reduced predictive accuracy. Results of our Monte Carlo analysis indicate that, despite being somewhat limited in scope and robustness, the UC method exhibits a fairly competitive overall performance in terms of predictive accuracy and coverage under some of the most commonly conjectured forms of long-term persistence.

References

Chib, S. 2001. "Markov Chain Monte Carlo Methods: Computation and Inference," in James J. Heckman and Edward Leamer, eds., *Handbook of Econometrics*, vol. 5 (Elsevier), pp. 3569–3649.

Chudy, Marek, Sayar Karmakar, and Wei Biao Wu. 2020. "Long-Term Prediction Intervals of Economic Time Series," *Empirical Economics*, vol. 58, no. 1 (January), pp. 191–222.

Congressional Budget Office. 2021. *The 2021 Long-Term Budget Outlook* (March), www.cbo.gov/publication/56977.

Congressional Budget Office. 2019. *The 2019 Long-Term Budget Outlook* (June), www.cbo.gov/publication/55331.

Congressional Budget Office. 2016. *The 2016 Long-Term Budget Outlook* (July), www.cbo.gov/publication/51580.

Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti. 2017. "Safety, Liquidity, and the Natural Rate of Interest," *Brookings Papers on Economic Activity* (Spring).

Doan, Thomas A. 1990. *RATS User's Manual* (VAR Econometrics, Suite 612, 1800 Sherman Avenue, Evanston, IL 60201).

Dupraz, Stéphane, Emi Nakamura, and Jón Steinsson. 2020. *A Plucking Model of Business Cycles*, Working Paper 748 (Banque de France, January).

Fernald, John. 2015. "Productivity and Potential Output Before, During, and After the Great Recession," in Jonathan A. Parker and Michael Woodford, eds., *NBER Macroeconomics Annual 2014*, vol. 29 (University of Chicago Press), pp. 1–51.

Gamber, Edward. 2020. *The Historical Decline in Real Interest Rates and Its Implications for CBO's Projections*, Working Paper 2020-09 (Congressional Budget Office, December), www.cbo.gov/publication/56891.

Gordon, Robert. 2014. "The Turtle's Progress: Secular Stagnation Meets the Headwinds," in Coen Teulings and Richard Baldwin, eds., *Secular Stagnation: Facts, Causes and Cures*. CEPR Press, 2014, pp. 47–60.

Hamilton, James. 1994. *Time Series Analysis* (Princeton University Press).

Hamilton, James, Ethan S. Harris, Jan Hatzius, and Kenneth D. West. 2016. “The Equilibrium Real Funds Rate: Past, Present, and Future,” *IMF Economic Review*, vol. 64, no. 4 (November), pp. 660–707.

Hansen, Bruce E., and Ananth Seshadri. 2013. *Uncovering the Relationship Between Real Interest Rates and Economic Growth*, Working Paper WP 2013-303 (University of Michigan Retirement Research Center).

Holston, Kathryn, Thomas Laubach, and John C. Williams. 2017. “Measuring the Natural Rate of Interest: International Trends and Determinants,” *Journal of International Economics*, vol. 108, no. S1 (May), pp. 559–575.

Laubach, Thomas, and John C. Williams. 2003. “Measuring the Natural Rate of Interest,” *Review of Economics and Statistics*, vol. 85, no. 4 (November), pp. 1063–1070.

Lewis, Kurt, and Francisco Vazquez-Grande. 2019. “Measuring the Natural Rate of Interest: A Note on Transitory Shocks,” *Journal of Applied Econometrics*, vol. 34, no. 3 (April), pp. 425–436.

Litterman, Robert B. 1986. “Forecasting With Bayesian Vector Autoregressions—Five Years of Experience,” *Journal of Business and Economic Statistics*, vol. 4, no. 1 (January), pp. 25–38.

McGrane, Michael. 2022. *A Markov-Switching Model of the Unemployment Rate*, Working Paper 2022-05 (Congressional Budget Office, March), www.cbo.gov/publication/57582.

Müller, Ulrich, and Mark Watson. 2018. “Long-Run Covariability,” *Econometrica*, vol. 86, no. 3 (May), pp. 775–804.

Müller, Ulrich, and Mark Watson. 2016. “Measuring Uncertainty About Long-Run Predictions,” *Review of Economic Studies*, vol. 83, no. 21 (January), pp. 1711–1740.

Newey, Whitney K., and Kenneth D. West. 1994. “Automatic Lag Selection in Covariance Matrix Estimation,” *Review of Economic Studies*, vol. 61 (October), pp. 631–653.

Pesavento, Elena, and Barbara Rossi. 2006. “Small-Sample Confidence Intervals for Multivariate Impulse Response Functions at Long Horizons,” *Journal of Applied Econometrics*, vol. 21, no. 8 (December), pp. 1135–1155.

Phillips, Peter C. 1998. “Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VARs,” *Journal of Econometrics*, vol. 83, no. 1-2 (March-April), pp. 21–56.

Staiger, Douglas, James H. Stock, and Mark W. Watson. 2001. “Prices, Wages, and the U.S. NAIRU in the 1990s,” in Alan B. Krueger and Robert M. Solow, eds., *The Roaring Nineties: Can Full Employment Be Sustained?* (Russell Sage Foundation), pp. 3–60.

Stock, James H. 1996. “VAR, Error Correction and Pretest Forecasts at Long Horizons,” *Oxford Bulletin of Economics and Statistics*, vol. 58, no. 4 (November), pp. 685–701.

Stock, James H. 1994. “Unit Roots, Structural Breaks and Trends,” in Robert F. Eagle and Daniel L. McFadden, eds., *Handbook of Econometrics*, vol. 4 (Elsevier), pp. 2739–2841.

Zhou, Zhou, Zhiwei Xu, and Wei Biao Wu. 2010. “Long-Term Prediction Intervals of Time Series,” *IEEE Transactions on Information Theory*, vol. 56, no. 3 (March), pp. 1436–1446.