Congressional Budget Office
Nonpartisan Analysis for the U.S. Congress

Quantifying the Uncertainty of Long-Term Economic Projections

June 17, 2022

Presentation at the

28th International Conference of The Society for Computational Economics

U. Devrim Demirel and James Otterson

For more information about the conference, see https://comp-econ.org/CEF_2022/index.htm. For further discussion of the results in this presentation, see U. Devrim Demirel and James Otterson, *Quantifying the Uncertainty of Long-Term Economic Projections,* Working Paper 2022-07 (Congressional Budget Office, April 2022), www.cbo.gov/publication/57711.

Introduction

The Uncertainty of Long-Term Economic and Budget Projections

The Congressional Budget Office (CBO) publishes 30-year projections for federal revenues, spending, and debt in its annual *Long-Term Budget Outlook* report.

CBO projects that, if current laws governing taxes and spending generally remained unchanged, federal debt as a percentage of gross domestic product (GDP) would surpass its highest level in history within a decade and continue to rise over the following several decades. But those projections are subject to substantial uncertainty.

A significant part of the uncertainty of long-term budgetary outcomes stems from persistent changes in economic trends rather than transitory economic fluctuations.

Quantifying Long-Term Uncertainty

The paper upon which this presentation is based proposes a practical method for assessing the long-term uncertainty of the economic variables that underpin budget projections. The method is based on simulations from a multivariate unobserved components (UC) model.

Variables are specified as sums of individually unobserved stationary and nonstationary components. Distinguishing between those components is important because the long-term uncertainty stems, in large part, from the variability of the nonstationary components.

We quantify the uncertainty of long-term economic projections by constructing prediction intervals for long-horizon averages of the variables.

Challenges of Long-Term Prediction Intervals

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.

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, local-to-unity, and fractionally integrated processes have different long-term properties.

But there is often limited information in the sample data to precisely distinguish between different forms of persistence.

The Method

Multivariate Unobserved Components Model

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 (X_t) and nonstationary (μ_t) components.

$$Y_t = \mu_t + X_t + w_t$$

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

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

The framework generalizes a stationary vector autoregression model. Furthermore, restrictions on the elements of the autoregressive and covariance matrices deliver models similar to those of Laubach and Williams (2003), Holston et al. (2017), and Lewis and Vazquez-Grande (2019).

The State-Space Form

The system of equations can be expressed in the state-space form:

$$\mathbf{S}_{t+1} = \mathbf{F} \cdot \mathbf{S}_t + \mathbf{v}_{t+1}$$

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

where

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

Estimation

We use Bayesian methods to estimate the parameters of the UC model and the historical paths of the nonstationary (μ_t) and stationary (X_t and w_t) components of the variables.

The variability of the nonstationary, or trend, components plays an important role in our analysis. For the standard deviations of the trend shocks, we choose an inverse-gamma (IG) prior.

The positive domain of the IG prior counters the tendency of maximum-likelihood estimates of the standard deviations to be biased toward zero—a phenomenon known as the pileup problem (Stock, 1994).

In addition, we incorporate the prior belief that the autoregressive parameters of the stationary component are independent and normally distributed.

Estimation (Continued)

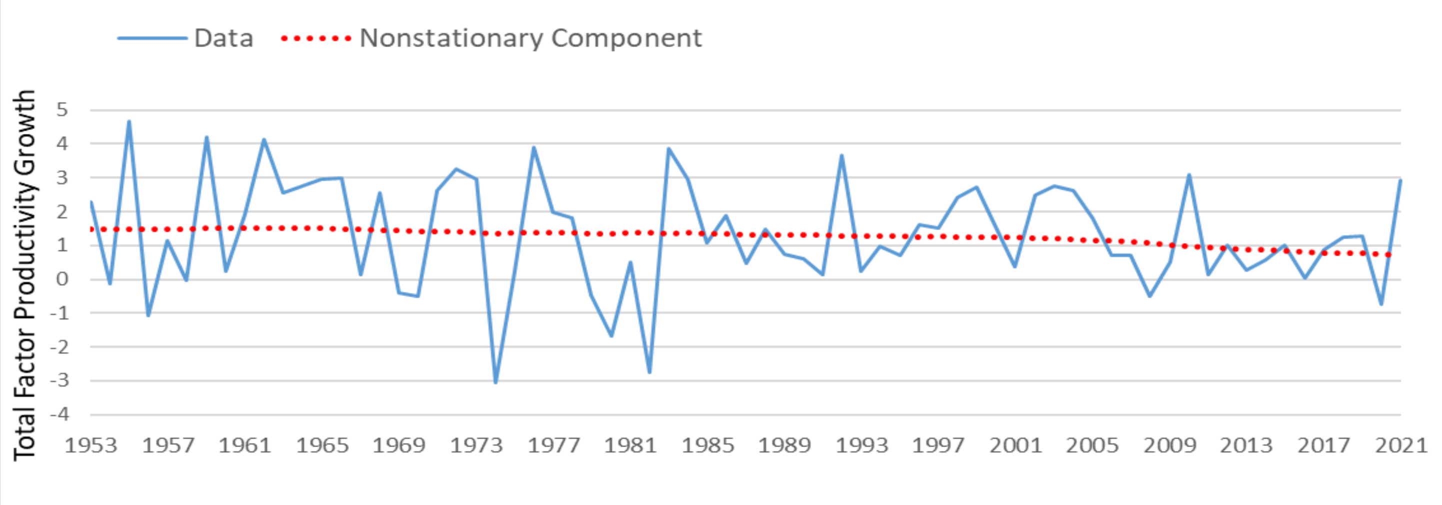
The mean of the prior distribution of the autoregressive parameters is zero.

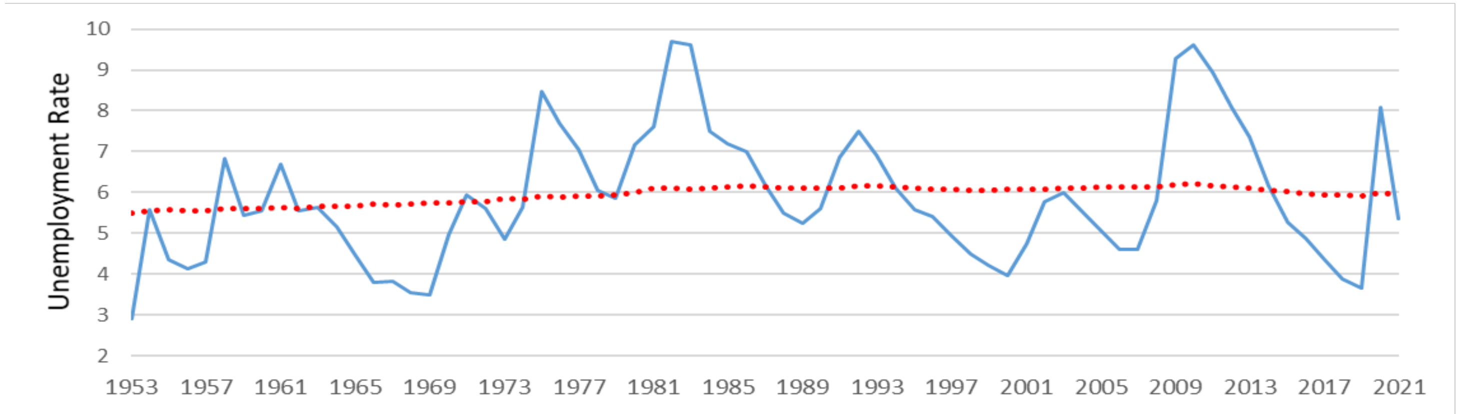
The specification of variances (which is based on the Minnesota prior) reflects the view that each variable's own lagged values provide better prior information about its dynamics than do the lags of other variables, and distant lags of a variable are less important drivers of its variation than are the variable's more recent lags.

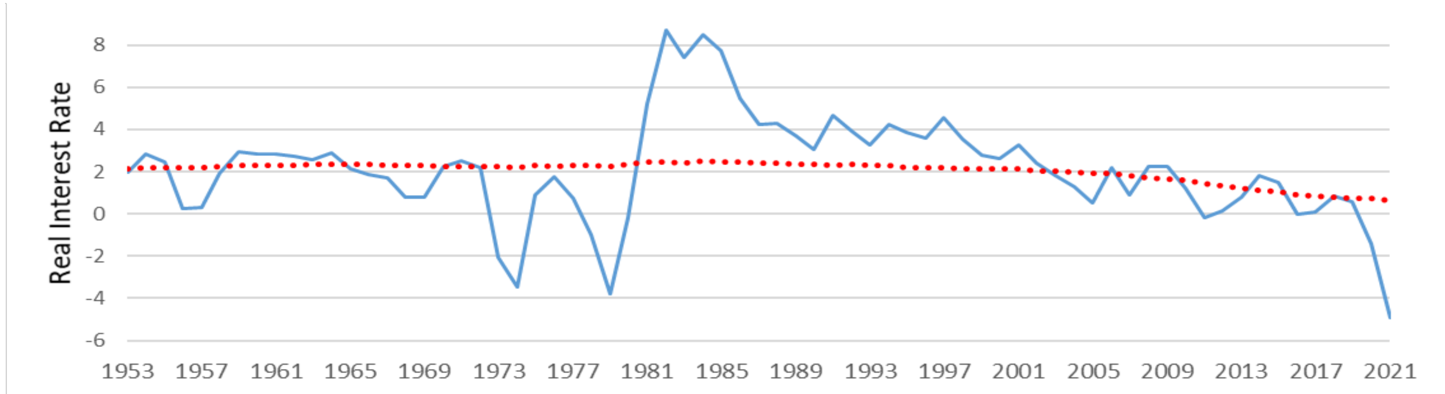
$$E\left[(A_s)_{jk}\right] = 0 \text{ for all } s, j, \text{ and } k.$$

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

Estimated Nonstationary Components







Constructing Prediction Intervals

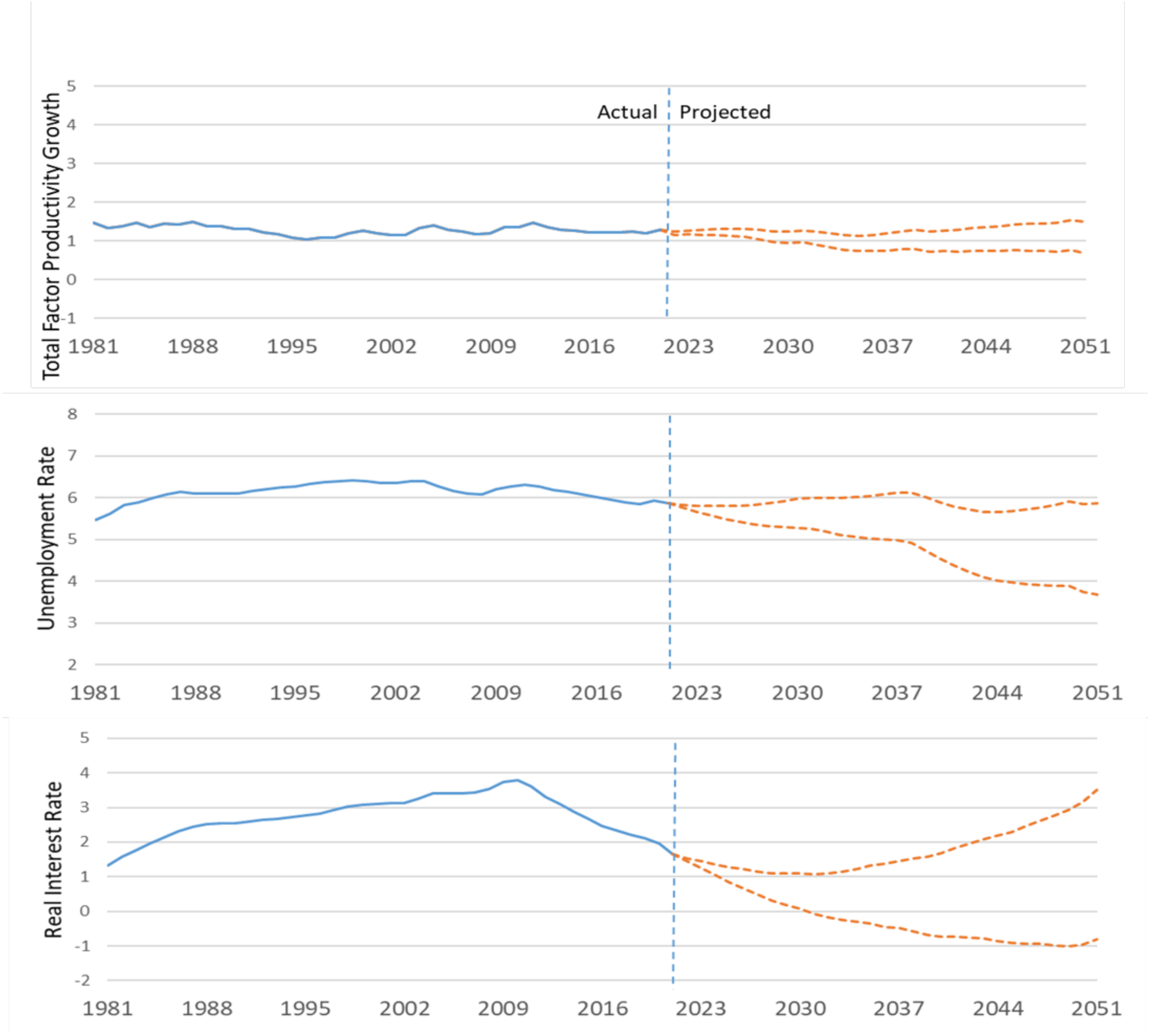
Our approach to constructing prediction intervals captures both parameter and forecast uncertainty.

We capture parameter uncertainty in our simulations by drawing sets of parameter values from the posterior distribution produced by the Markov chain Monte Carlo (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—and iterate the system of equations forward to construct 30-year series.

For each series, we then calculate 30-year moving averages and construct prediction intervals containing the middle two-thirds of the simulated distribution.

Prediction Intervals for 30-Year Moving Averages



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.

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.

Assessing the Performance of the UC Method

Data Generating Processes

We test the predictive performance of the UC method on simulated data produced by artificial data generating processes (DGPs) with specified properties.

We adopt three DGPs used in Müller and Watson's (2016) 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 $\mu_{3t} \sim N(0, 1)$, $\epsilon_t \sim N(0, \omega_3^2)$.

Monte Carlo Analysis

We calculate the coverage probability for each variable by assessing the frequency with which the estimated prediction intervals contain the actual average values.

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 methods.

- The first method, called the asymptotic covariance method, is based on the estimated long-run covariance matrix of the variables.
- The second method is the one developed in Müller and Watson (2016). That method yields asymptotically valid prediction sets under many different forms of long-term persistence.

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.

Conclusion

We propose a practical approach to quantifying economic uncertainty by using simulations from a multivariate UC model.

Compared with some recently developed state-of-the-art methods (including the methods of Müller and Watson, 2016, 2018), the UC approach is simpler to implement, can easily handle a multivariate framework, 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.

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.