

# Comparing Measures of Potential Output

Amy Y. Guisinger, [Michael T. Owyang](#), and Hannah G. Shell

One of the goals of stabilization policy is to reduce the output gap—the difference between potential and actual output—during downturns. Potential output, however, is an unobserved variable whose definition can vary. For example, some view potential output as the level of output that can be produced when employment is at the natural rate. Others use trend measures of output to measure potential. We survey some of these measures using both full-sample data (all of the data that would be available through June 2017) and real-time data (the actual data that would have been available at different points in the sample). We construct six different measures of potential: a linear trend, a quadratic trend, the Congressional Budget Office measure, and three filtered trends. We compare these measures across methods and across time. We also use the measures to compute the monetary policy prescription in a standard interest rate rule and find very little difference across methods. (JEL C22, E32)

Federal Reserve Bank of St. Louis *Review*, Fourth Quarter 2018, 100(4), pp. 297-316.  
<https://doi.org/10.20955/r.100.297-316>

## INTRODUCTION

One of the goals of stabilization policy is closing the output gap—that is, the difference between actual and potential output. Obviously, then, a key component of stabilization policy and its timing is the proper measurement of potential output. Potential output is an unobserved measure of the economy's ability to generate output. Unfortunately, many decades of academic research have failed to converge on a single measure (or even definition) of potential output, leaving one of policymakers' key objectives—at least, in part—out of focus.

As an example, one of the early works on measuring potential output was by Okun (1962) and led to the development of Okun's law, an empirical relationship between output and unemployment. Since that paper, a large literature has arisen around the measurement of potential output in various countries. Orphanides and van Norden (2002) show that conven-

Amy Y. Guisinger is an assistant professor of economics at Lafayette College. Michael T. Owyang is an assistant vice president and economist and Hannah G. Shell is a senior research associate in the Research Division at Federal Reserve Bank of St. Louis. The authors thank Kevin Kliesen and Jeremy Piger for helpful discussion.

© 2018, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

tional statistical measures produce unreliable estimates of the U.S. output gap in real time because they are subject to large revisions and the unreliability of end-of-sample estimates; however, Edge and Rudd (2016) find that despite changes in productivity, the revision properties of output-gap estimates improved significantly in the late 1990s and early 2000s. Fernald (2014) closely examines the productivity slowdown in the 2000s, showing that declines in productivity preceded the Great Recession and that the current low-productivity state is not related to the downturn. This slowdown implies slower long-term potential growth (Fernald estimates around 2.1 percent) and a narrower output gap than that identified by the Congressional Budget Office (CBO).

Marcellino and Musso (2011) find that estimates of the output gap in the euro area are also highly uncertain. However, they attribute the uncertainty to parameter instability and model uncertainty instead of data revisions. Champagne, Poulin-Bellisle, and Sekkel (2016) extend the U.S.-based studies to Canada's output-gap estimates, concluding that the Bank of Canada's staff estimates have become more reliable over the past 30 years. Dungey, Jacobs, and Tian (2017) use an unobserved components (UC) filter to forecast one-step-ahead potential for the G-7 countries and find that allowing for correlated shocks is more valuable to the measurement of potential than allowing for structural breaks.<sup>1</sup>

One of the central issues in measuring potential output is that there is no single theoretical definition. Kiley (2014) highlights the importance of defining concepts related to various methodologies for estimating potential output. Some economists prefer to define potential output as the trend—sometimes linear or quadratic—in output. This definition allows the analyst to use univariate (or sometimes multivariate) econometric techniques to measure potential output without much theory. However, even these techniques must rely on often implicit theoretical assumptions (Basu and Fernald, 2009). These techniques generally consist of filters that separate the trend from cyclical or low-frequency fluctuations. There are, however, a variety of different filters that are commonly used, and their popularity often depends on their ease of use or availability of executable code rather than the reliability or sensibility of the resulting potential output series. Other techniques, like those used by the CBO, define potential as the level of output possible if all resources are employed at their full potential, defined as trend growth in productive capacity. The theoretical assumptions underlying these techniques directly influence the calculation of potential; however, there are many unknowns associated with these models, including structural (parameter) shifts and even the appropriateness of the model itself to describe the economy (Basu and Fernald, 2009).

In this article, we explore some of the myriad of methods to measure potential output. We start by examining a measure produced by the CBO, which uses empirical relationships among output, the unemployment rate, and inflation, among other variables. We then consider a few common methods used to extract a trend-version of potential output: (i) a linear trend, (ii) a quadratic trend, (iii) the Hodrick-Prescott (1997) filtered trend, (iv) a univariate version of the UC model (Harvey, 1989, Clark, 1987, and Watson, 1986), and (v) a multivariate version of the UC model that includes inflation (Basistha and Startz, 2008).

Evaluating these methods, however, is a difficult task. Unlike standard out-of-sample forecast experiments with observed data, potential output is latent, so there is no “truth” against

which we can compare our estimates. However, one way to examine the measures is to view the implications of their use through a policy lens. We consider the use of the six different measures (five trend measures and the CBO measure) in a standard version of the Taylor rule, asking whether different methods for estimating potential output would lead to dramatically different conclusions about how policy should be implemented at different points. We first conduct this experiment *ex post*, using all of the data in the sample. We then consider how these conclusions would have varied in “real time,” using only data that were available at the time.

We find small differences across the measures of potential output. These differences are magnified at the beginning and end of the sample around turning points and when using real-time data. However, we find that these differences do not typically affect the policy prescription resulting from a standard Taylor-type interest rate rule.

The balance of the article is laid out as follows: Section 1 describes four of the six measures of potential output that we consider. Section 2 describes the data and the full-sample estimates of potential output using linear and quadratic trends and each of the four measures described in Section 1. Section 3 recomputes the estimates of the measures in real time, while Section 4 shows the implications for monetary policy using a calibrated Taylor rule. Section 5 concludes.

## METHODS OF COMPUTING POTENTIAL OUTPUT

Potential output is an unobserved construct often thought to be synonymous with the maximum level of sustainable output. Because it is unobserved, potential output is generally constructed from observed data such as gross domestic product (GDP). Some definitions of potential output are based on theoretical or empirical relationships that are imposed on the data. For example, one might believe in both a Phillips curve-type relationship (between unemployment and inflation) and an Okun’s law-type relationship (between output and unemployment) and construct potential output from a three-variable New Keynesian model. In this section, we discuss a few common methods of estimating potential output and how they differ, both theoretically and empirically. Obviously, the methods here describe only a small subset of the literature; they are, however, relatively commonly used measures.

### *The CBO Measure of Potential Output*

The CBO defines potential output as the trend growth in the productive capacity of the economy. To estimate potential output, it uses a model that attributes real GDP growth to the growth in three factor inputs: capital, labor, and technological progress. The CBO method divides GDP into five sectors: nonfarm business, government, farm, households and non-profits, and housing.<sup>2</sup> For each sector, the CBO estimates a production function based on labor, capital accumulation, and total factor productivity (TFP) using a “potential value” for each factor.<sup>3</sup>

Output in sector  $i$  is assumed to be generated from a Cobb-Douglas production function:

$$(1) \quad Y_{i,t} = A_{i,t} L_{i,t}^{(1-\alpha)} K_{i,t}^{\alpha},$$

where  $Y_{i,t}$  is period- $t$  output in sector  $i$ ,  $L_{i,t}$  is hours worked in sector  $i$  during period  $t$ ,  $K_{i,t}$  is the level of the capital stock in sector  $i$  at time  $t$ , and  $A_{i,t}$  is period- $t$  TFP for sector  $i$  (CBO, 2001). The parameter  $\alpha$  represents the capital share—capital’s contribution to output growth.<sup>4</sup> Aggregate output is assumed to be the sum of output across the five sectors. The CBO defines potential output for each sector as the output obtained when all of the factors ( $A$ ,  $L$ , and  $K$ ) are at their potential levels.

The potential level of the hours worked factor,  $L_{i,t}^*$ , is computed as the product of average weekly hours,  $\bar{L}_{i,t}$ , and potential employment, which in turn is determined by the potential labor force,  $N_{i,t}^*$ :

$$L_{i,t}^* = \bar{L}_{i,t} (1 - U_{i,t}^*) N_{i,t}^*,$$

where  $U^*$  is the natural rate of unemployment.<sup>5</sup> The size of the labor force is assumed to be a function of the unemployment gap (i.e., the deviation of the unemployment rate from the natural rate) and the state of the business cycle,  $S_t$ :

$$(2) \quad \log(N_{i,t}) = \beta_i (U_{it} - U_{it}^*) + f_i(S_t) + e_{i,t},$$

where  $S_t$  is a matrix of business cycle dummy variables, with each column representing one business cycle, and  $e_{i,t}$  is an iid normal error term. A representative element of  $S_t$  is set to 1 during the expansion associated with that column. The CBO uses the fitted values from this regression, setting  $U_{it} + U_{it}^*$ , to obtain an estimate for the potential size of the labor force,  $N_{i,t}^*$ .

Because different types of capital have different potential productivities, the CBO uses a capital services index that assumes that the marginal productivities of different types of capital are proportional to capital cost shares.<sup>6</sup> The types of capital in the CBO index are computers, software, communications equipment, other equipment, nonresidential structures, inventories, and land. The index is then the average of capital growth rates weighted by cost share:

$$\log\left(\frac{K_{i,t}}{K_{i,t-1}}\right) = \sum_j \omega_{ij,t} \log\left(\frac{K_{ij,t}}{K_{ij,t-1}}\right),$$

where  $j$  denotes the different types of capital. The weights  $\omega_{ij,t}$  are a two-period average of the relative cost share of each type of capital  $j$  (CBO, 2001) for sector  $i$ . The resulting index does not have to be adjusted to potential because the productivities already represent the potential contribution of capital to output.

The final component of the nonfarm business sector is TFP, which is estimated by using the residual historical growth values not attributed to capital and labor. These values are then cyclically adjusted similar to the way the labor factor uses separate dummy time trends for each business cycle.

While the sectors of the CBO model generally follow the same format, there is some variation. For example, the government sector separates output into local, state, and federal levels and then calculates each individual component by using potential compensation of employees and capital cost allowance. Compensation is modeled as a function of potential employment

and potential productivity. Output for the farm sector is a function of potential farm employment and potential output per employee, and output from the households and nonprofit sector is calculated using potential hours worked and potential productivity in the same manner outlined above. The housing sector is a bit different and relies on projections of the residential housing stock based on residential investment and projections of the productivity of that stock using the rate of residential capital depreciation.

### **HP Filter**

Another method for computing potential output that is commonly used by both the media and academics defines potential as “trend output.” While this method can be as simple as computing a deterministic (say, linear or quadratic) trend from output data, the literature has typically adopted slightly more complicated methods of extracting the trend. One popular method for extracting a trend from a single series of data is the Hodrick-Prescott (1997, henceforth HP) filter. The objective of the HP filter is to separate the high-frequency or cyclical fluctuations from the low-frequency or trend movements.

Suppose that  $Y_t$  is the series in question (in our case, output) and  $T$  is the length of the data series. The HP-filtered trend is obtained from

$$\min \left( \sum_{t=1}^T (y_t - y_t^*)^2 + \lambda \sum_{t=2}^{T-1} [(y_{t+1}^* - y_t^*) - (y_t^* - y_{t-1}^*)]^2 \right),$$

where  $y_t = \log Y_t$  and  $y_t^*$  is the value of the potential output (measured as the trend) at time  $t$ . The parameter  $\lambda$  is a smoothing parameter that is discussed below. The two terms of the equation lead to a trade-off between allowing the trend to follow the data (the first term) and minimizing movements in the trend (the second term). The smoothing parameter increases the weight on the latter: The larger the smoothing parameter, the more weight is placed on minimizing the higher-frequency fluctuations. Thus, as the data frequency increases, it is desirable to set the smoothing parameter to larger values (see Ravn and Uhlig, 2002, for a discussion of the values commonly used for various data frequencies).

While the HP filter’s relative ease of use is an advantage, the method does have a number of drawbacks. Aside from choosing the “correct” value of the smoothing parameter, the HP filter makes a number of assumptions about the data to be examined, including that the series is  $I(2)$ . If the series is not  $I(2)$ , the HP filter can introduce false fluctuations in the trend.<sup>2</sup>

### **Unobserved Components**

An alternative to the HP filter that may have better statistical characteristics is the UC model first introduced in Harvey (1989). Similar to the HP filter, the UC model assumes that the data series  $y_t$  (in our case, log output) can be written as the sum of a unit root trend  $y_t^*$  and stationary cycle  $c_t$ :

$$y_t = y_t^* + c_t,$$

where the trend is a random walk with drift  $\tau$ ,

$$y_t^* = \tau + y_{t-1}^* + \varepsilon_t,$$

and the cycle exhibits stationary AR( $P$ ) dynamics,

$$c_t = A(L)c_{t-1} + v_t,$$

where  $\varepsilon_t$  and  $v_t$  are iid normal error terms and  $P \geq 2$  is a common identification assumption.

A typical assumption in the UC literature is that the variance-covariance of the vector of errors is diagonal—that is,  $\varepsilon_t$  and  $v_t$  are orthogonal. However, Morley, Nelson, and Zivot (2003) showed that the assumption of zero correlation between the trend and cycle innovations can be relaxed. Moreover, imposing  $\sigma_{\varepsilon v} \approx -1$  yields another common filter, the Beveridge-Nelson decomposition. For our application, we will consider the uncorrelated UC model but refer the reader to Morley, Nelson, and Zivot (2003) for the application with correlation between the shocks to the trend and cycle.

### **Multivariate Unobserved Components**

One issue with the potential output measures computed in the previous two sections is that they are simply trends extracted from output without any economic theory. We can add an element of economic theory by incorporating other variables that might fluctuate when the difference between output and potential output changes. For example, if one believed in Okun’s law and a Phillips curve (as in a New Keynesian model), then there should exist a relationship between output’s deviation from potential and inflation. We can capture this relationship by estimating a multivariate version of the UC model, where we allow some correlation between the cross-series innovations in the trends and cycles.

For this model, we assume that both output and inflation can be written as the simple sum of a trend and a cycle:

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} y_t^* \\ \pi_t^* \end{bmatrix} + \begin{bmatrix} c_t^y \\ c_t^\pi \end{bmatrix},$$

where the trends in each variable are unit roots with separate drifts:

$$\begin{aligned} \pi_t^* &= \tau^\pi + \pi_{t-1}^* + \varepsilon_t^\pi, \\ y_t^* &= \tau^y + y_{t-1}^* + \varepsilon_t^y. \end{aligned}$$

The cycles  $\mathbf{c}_t = [c_t^y, c_t^\pi]'$  evolve as a VAR( $P$ ):

$$\mathbf{c}_t = A(L)\mathbf{c}_{t-1} + \mathbf{v}_t,$$

where the covariance matrix is block diagonal,

$$\Omega = \begin{bmatrix} \omega_{yy}^{\tau} & \omega_{y\pi}^{\tau} & 0 & 0 \\ \omega_{y\pi}^{\tau} & \omega_{\pi\pi}^{\tau} & 0 & 0 \\ 0 & 0 & \omega_{yy}^c & \omega_{y\pi}^c \\ 0 & 0 & \omega_{y\pi}^c & \omega_{\pi\pi}^c \end{bmatrix}.$$

This formulation allows an outside variable—in this case, inflation—to influence the estimate of potential output through the contemporaneous correlation in  $\Omega$  and the lagged correlation in the VAR.<sup>8</sup> The block diagonal structure allows contemporaneous correlation across the variables in the cycles or trends but does not allow the trend and cycle of a variable to be contemporaneously correlated and does not allow the trend of one variable to be correlated with the cycle of the other.

## COMPARING SIX MEASURES OF POTENTIAL OUTPUT

### Data

The standard measure for output is real GDP, which is available at a quarterly frequency from the Bureau of Economic Analysis (BEA). The BEA releases three estimates of output: an advance estimate available one month after the end of a quarter, a preliminary estimate available two months after the end of a quarter, and a final estimate available three months after the end of a quarter. The final numbers are continually revised. As is typical in these types of analyses, we present the revised final estimates, seasonally adjusted and annualized.

In a later section, we consider the computation of real-time estimates of potential output using only the final estimates of real GDP that would have been available at the time. Moreover, we use the data vintage from each period, ignoring subsequent revisions to the data. Thus, in 1990:Q1, we would compute potential output using only data available at that point in time and using only the revisions that were available up to that date.

Inflation data are the yearly percent change in quarterly, seasonally adjusted personal consumption expenditures compiled by the BEA.

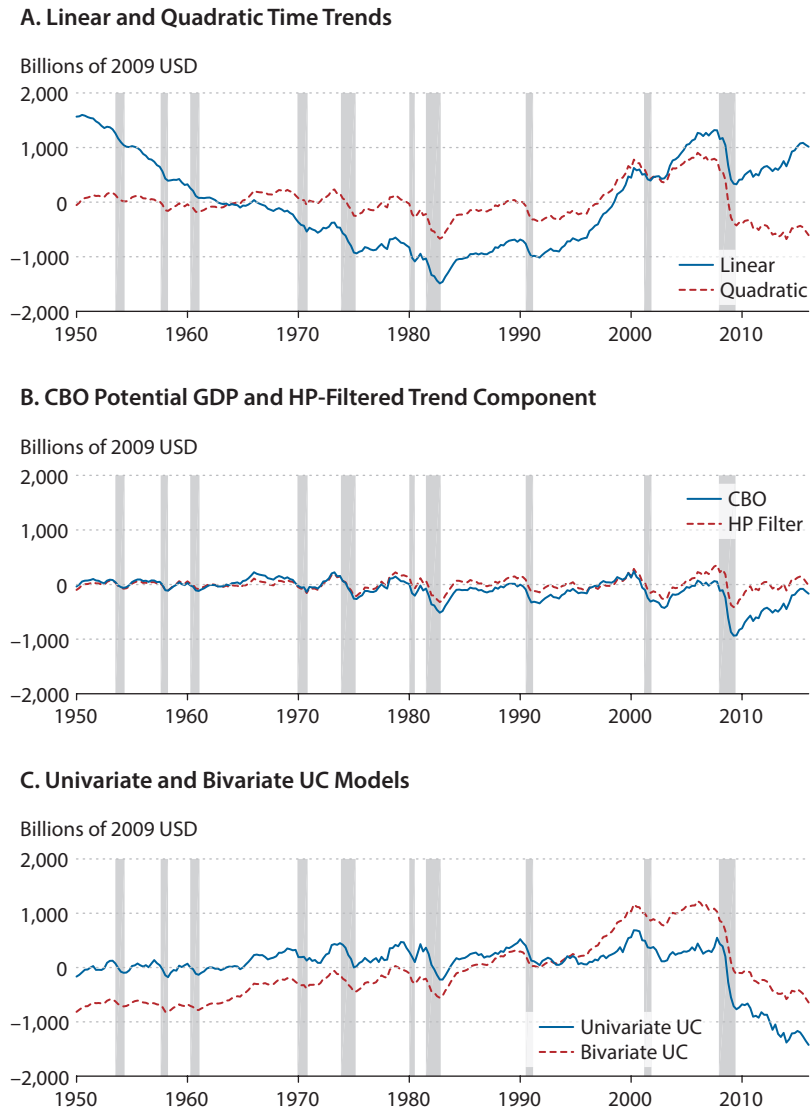
These two data series, along with the CBO's potential output, are accessed from the Federal Reserve Bank of St. Louis FRED<sup>®</sup> database. We use data from 1950:Q1 through 2015:Q2 for the full-sample analysis. The real-time measures of output, inflation, and the CBO's potential output come from the FRED<sup>®</sup> vintage data service, ALFRED<sup>®</sup>. For the output measures, we use the first vintage available in January 1992; for inflation, we use the first vintage in January 1996.

### Full-Sample Results

Figure 1 presents output gaps—the difference between actual output and potential output—for six alternative measures of potential GDP. Panel A shows the linear and quadratic time trends, Panel B shows the CBO measure and the trend component of the HP filter, and Panel C shows the two UC models. The linear time trend assumes that the growth rate in

**Figure 1**

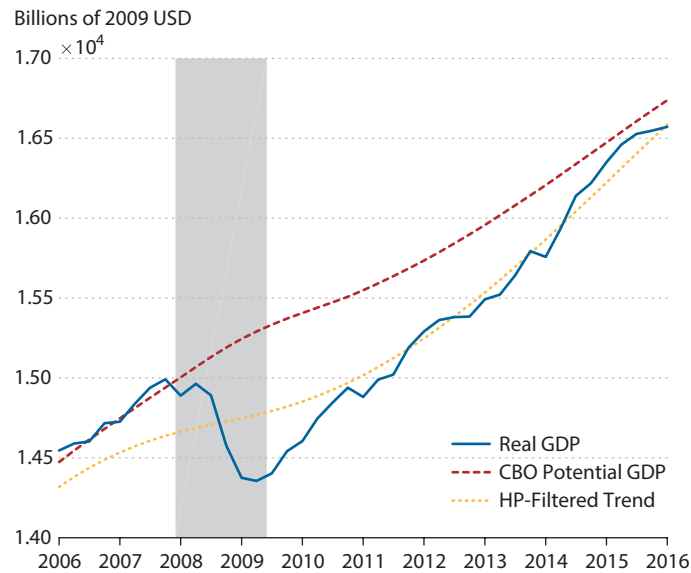
**Output Gaps for the Six Alternative Measures of Potential Output**



NOTE: Gray bars indicate recessions as determined by the National Bureau of Economic Research (NBER).  
SOURCE: BEA, CBO, FRED®, and authors' calculations.

potential output is constant over time. A prominent feature of the data is that GDP growth appears slower in the first half of the sample than in the second half. Because the linear time trend assumes that the growth rate of potential output is constant over the whole sample, the output gap is negative for much of the middle years of the sample. Additionally, the output gap is large and positive (actual output is greater than potential) at the beginning and end of the sample.



**Figure 2****Real GDP, CBO Potential GDP, and the HP-Filtered Trend Component (2006:Q1–2016:Q1)**

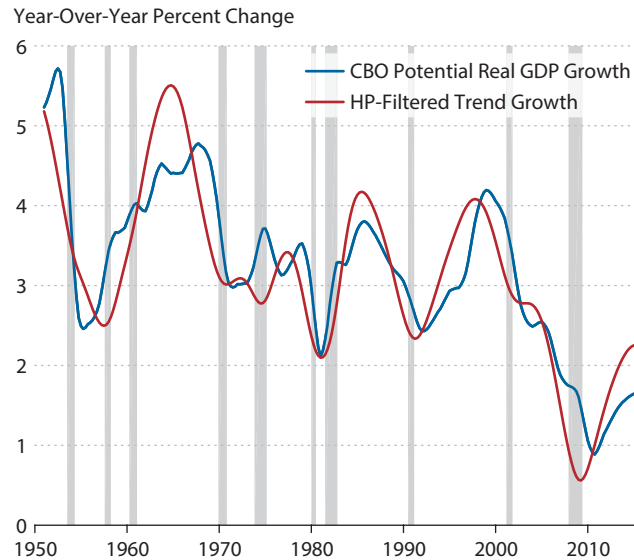
NOTE: Gray bar indicates recession as determined by the NBER.

SOURCE: BEA, CBO, FRED®, and authors' calculations.

The quadratic time trend introduces an additional term that allows the growth rate of potential output to change over time. The quadratic trend may be a more realistic estimate of potential GDP, yielding a negative output gap during economic downturns and a positive output gap during the height of some expansions. Based on the quadratic trend, GDP was above potential for most of the 2000s then fell below potential during the 2007-09 recession and has yet to recover. While simple to compute and easy to understand, these deterministic trends do not allow for possible structural change and are sensitive to the sample used to compute them.

Panel B of Figure 1 presents the CBO estimate and the trend component of the HP filter. Compared with the deterministic trends above, these two measures track real GDP more closely, usually resulting in smaller output gaps. For example, measured GDP was below both the HP-filtered and the CBO measures of potential output for short periods at the beginning of the 2000s. However, during the 2007-09 recession, the two measures diverged: Measured GDP was at the CBO measure of potential leading into the Great Recession but below the HP-filtered measure of potential.

These differences highlight one feature of the CBO model: It rarely places output above potential. Figure 2 highlights the last 10 years of data (2006:Q1–2016:Q1) for both the CBO measure and the HP filter. The CBO has the output gap much larger than the other measures during and after the Great Recession, with output still below potential into 2016.<sup>9</sup> The HP filter, on the other hand, has output above potential for the first half of the recession and shows a negative output gap only until mid-2011.

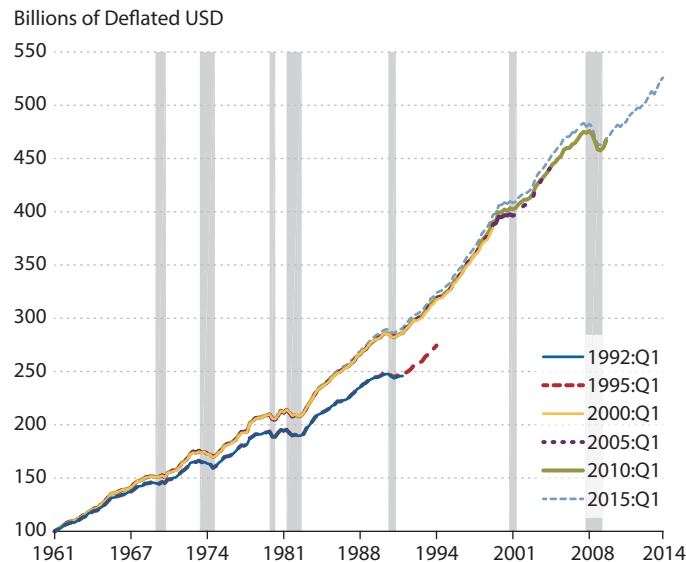
**Figure 3****CBO Potential Real GDP Plotted Against the HP-Filtered Trend Component**

NOTE: Gray bars indicate recessions as determined by the NBER.

SOURCE: BEA, CBO, FRED®, and authors' calculations.

Figure 3 compares the year-over-year percent change in the HP-filtered and the CBO potential output measures. The two series mostly move together, with some exceptions around the end of the sample, during the expansion between 1960 and 1970, and during the 1973 downturn. At the beginning of the sample in 1950, potential GDP growth was high in both measures, slightly above 5 percent. Potential output growth fluctuates over 5- to 10-year cycles for the next 50 years, averaging about 3 percent until the early 2000s when it falls to less than 1 percent before bottoming out. Figure 3 also shows that the HP filter's performance deteriorates around the ends of the sample. Because the filter has fewer data points on the ends, it is more difficult to identify the trend. Near the ends of the sample, the CBO potential growth rate is about 1.5 percent, while the HP filter's is higher and about 2.1 percent.<sup>10</sup>

Panel C of Figure 1 presents the trend component of GDP from the univariate and multivariate UC models. The univariate UC trend yields positive output gaps before the 1991, 2001, and 2007-09 recessions (albeit a much smaller positive gap leading into the 2007-09 recession). After the Great Recession, the univariate UC trend yields a large negative output gap. Unlike previous measures, the univariate UC output gap does not narrow during the Great Recession recovery. Instead, the gap widens. The bivariate UC trend, which allows for correlation between inflation and output in the estimation of potential, shows output below potential through the 1980s.<sup>11</sup> Output returns to potential for a while after the 1980s, until the mid-1990s, then goes above potential and remains there until the 2007-09 recession. Consistent with a narrative about unsustainable gains in productivity during the IT boom leading to

**Figure 4****Vintage GDP Used to Estimate Potential GDP in Figures 5 to 7**

NOTE: Gray bars indicate recessions as determined by the NBER.

SOURCE: BEA and ALFRED®.

growth above potential (Fernald, 2014), the bivariate UC model is the only measure for which measured output is above potential for all of the 2000s. After the 2007-09 recession, output falls below potential and the gap widens through the end of the sample. This model implies that GDP during the 2000s was growing above potential and that the 2007-09 recession actually served to adjust growth back down to potential levels.

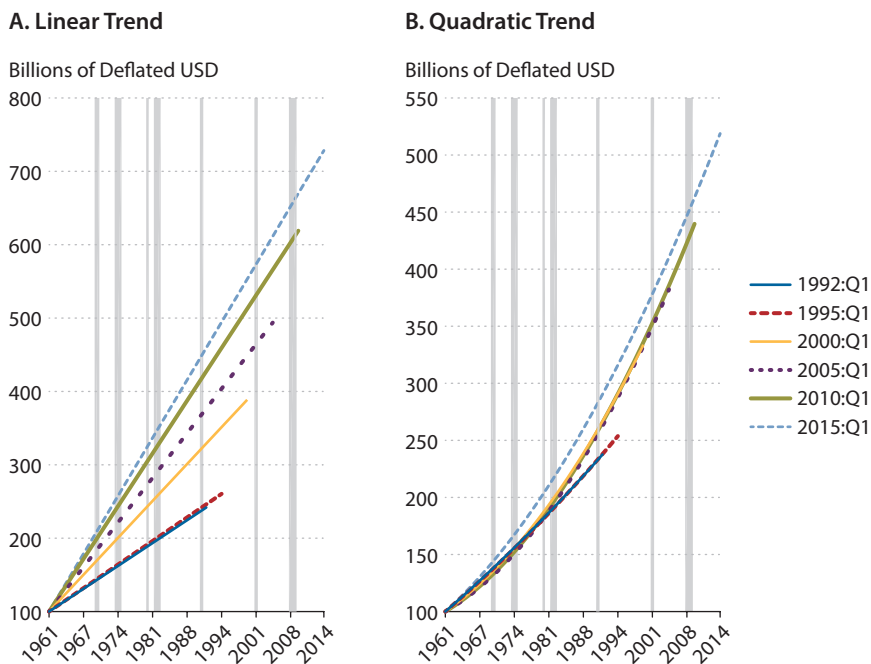
## REAL-TIME ESTIMATES OF POTENTIAL OUTPUT

The previous section demonstrated that differences in the definition of potential output for the methods used to construct it can alter our conclusions about the state of the economy. In that section, we used the data that were available in 2017:Q3. The data are, however, revised over time. For example, the 2001:Q1 observation of GDP might be different if observed in 2003:Q1 versus 2004:Q1. These revisions to the GDP data will obviously affect the calculation of potential.<sup>12</sup>

To see how the different measures of potential output change as the data are revised, we compute quasi-real-time estimates using different vintages of data. To compute a series for potential output at time  $\tau$ , we use the data that would be available at that time—that is, we use  $Y_1, \dots, Y_{\tau-1}$ . We consider a few different vintages of data: the first available vintage, 1992:Q1, and vintages for the first quarter of every fifth year, 1995:Q1, 2000:Q1, 2005:Q1, 2010:Q1, and 2015:Q1. Figure 4 shows the GDP vintages. The bivariate UC model is computed using

**Figure 5**

**Vintage Potential Output Estimates**



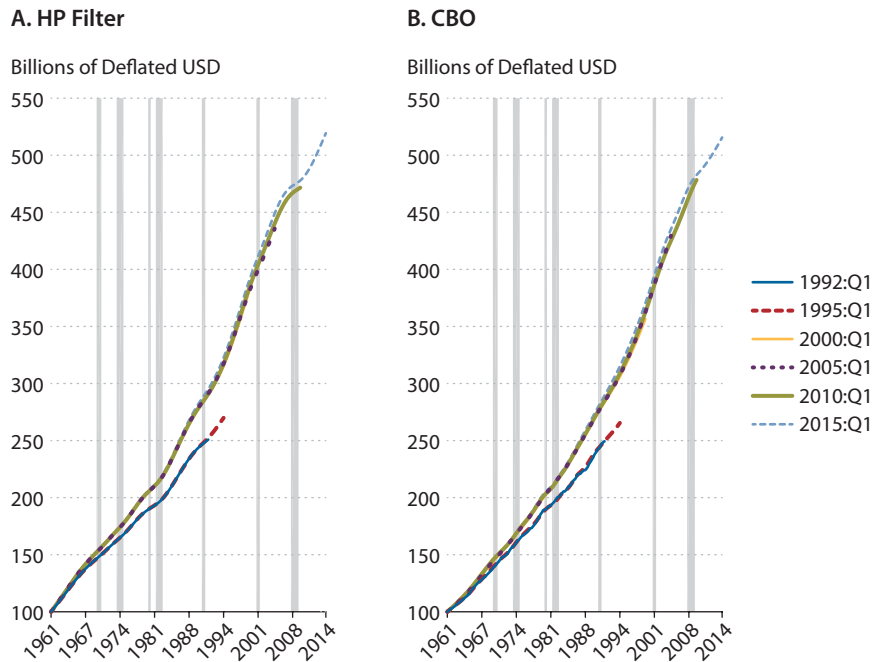
NOTE: Gray bars indicate recessions as determined by the NBER.

SOURCE: BEA, ALFRED®, and authors' calculations.

vintage data for both GDP and inflation. Because of limited availability of inflation vintages, the bivariate UC model is calculated starting in 1996:Q1 and then follows the same years as the univariate UC model (2000:Q1, 2005:Q1, and so forth). Figure 6 shows the six GDP vintages indexed to 100 at the beginning of the sample (January 1961 for the real-time analysis). A major revision between the 1995:Q1 and 2000:Q1 vintages shifts GDP up substantially. Another, less dramatic, upward shift occurs in the 2000s between the 2010:Q1 and 2015:Q1 vintages.

Figures 5-7 plot the vintage computations of potential output for the different methods outlined above. Notice that the real-time measure not only affects the last period of the vintage, but also can affect the estimates of potential output for all periods before. This effect can occur because the data are different (revised) or because the latest observation of the data affects the inference for all periods before. The GDP vintages used are reported in real terms using varying dollar indexes. To compare vintages across different dollar indexes, we index the potential computations to 100 in the first period.

Figure 5 shows real-time potential output for the linear time trend (Panel A) and quadratic time trend (Panel B). Because of the restriction in the linear trend that output be constant over time, as the sample size increases, the slope of the potential output curve becomes steeper. This steady increase reflects both GDP growth over time and revisions that shifted historical growth upward. The quadratic trend picks up on the same increasing growth over time. The

**Figure 6****Vintage Potential Output Estimates**

NOTE: Gray bars indicate recessions as determined by the NBER.

SOURCE: BEA, CBO, ALFRED®, and author's calculations.

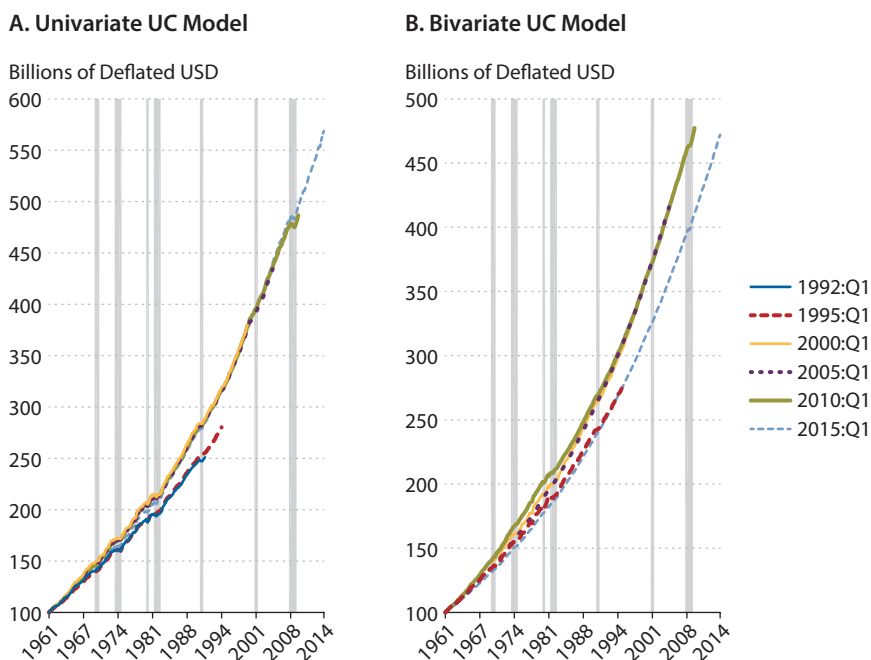
first two vintages, 1992:Q1 and 1995:Q1, move together, then the vintages between 2000:Q1 and 2010:Q1 have an upward shift with a more defined curve, and finally the last vintage in 2015:Q1 moves up again.

Figure 6 plots real-time vintage potential output for the HP-filtered (Panel A) and CBO measures (Panel B). These measures produce essentially the same potential output estimates for the 1992:Q1 and 1995:Q1 vintages. In 2000:Q1, the GDP revision causes an upward level shift in potential output. The real-time analysis further illustrates the problem of the HP filter being less accurate around the end points. As seen in the figure, around the Great Recession, the HP-filtered potential estimate increases notably from the 2010:Q1 vintage to the 2015:Q1 vintage. The CBO measure is more consistent through the recession, although the 2015:Q1 vintage shifts potential down slightly during the recession.

Figure 7 plots the vintage output estimates for the univariate UC (left) and bivariate UC (right) models. Potential obtained from the bivariate UC model has a different starting vintage to incorporate the personal consumption expenditures inflation data. The univariate UC vintages follow a similar pattern as those of the CBO and HP filter, reflecting an increase in potential output after 2000:Q1. Otherwise, the univariate model is mostly consistent despite the GDP revisions. The notable exception is around the 2007-09 recession, where the 2015:Q1 vintage estimates potential to be higher during the recession than the 2010:Q1 vintage.

**Figure 7**

**Vintage Potential Output Estimates**



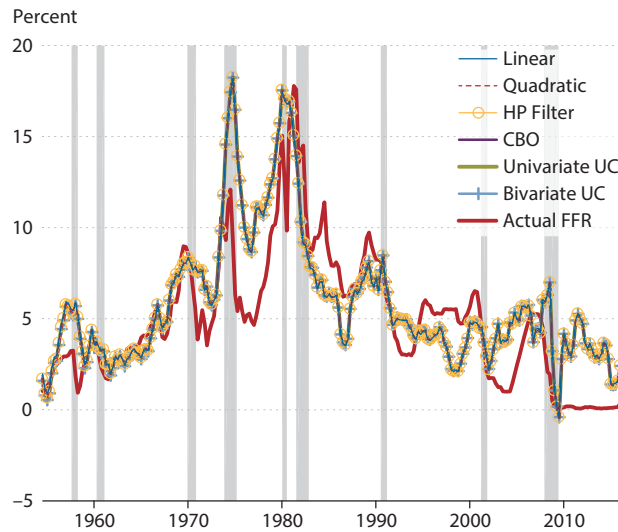
NOTE: Gray bars indicate recessions as determined by the NBER.

SOURCE: BEA, BLS, ALFRED®, and authors' calculations.

Data revisions for the bivariate UC model are substantively different from the other models. In the years following the Great Recession, data revisions for the other models generally lead to upward revisions in potential. For the bivariate UC model, however, the revised data show a downward adjustment of potential.

**IMPLICATIONS FOR POLICY**

One reason to care about measuring potential output is that it may be important for the conduct of monetary policy. While there appear to be differences in the measures of potential across methods and vintages, one might want to determine whether these differences are substantive enough to produce differences in policy. Although the Federal Reserve's dual mandate is to achieve full employment and stable prices, the conduct of monetary policy is often theorized to depend on output relative to its potential (which would be proportional to the deviation between the unemployment rate and the natural rate if Okun's law were assumed to hold). As an example, the Taylor (1993) rule, which is often characterized as an optimal rule for the stance of monetary policy, sets the policy rate,  $r_t$ , as a function of the deviation of the inflation rate,  $\pi_t$ , from a target rate,  $\pi^*$ , and the deviation of log output,  $Y_t$ , from log potential,  $Y_t^*$ :

**Figure 8****The Policy Rate Prescribed by the Taylor Rule, with the Output Gap Calculated Using the Six Alternative Measures of Potential Output**

NOTE: FFR, Federal funds rate. Gray bars indicate recessions as determined by the NBER.

SOURCE: FRED®, BEA, CBO, BLS, Board of Governors of the Federal Reserve System, and authors' calculations.

$$(3) \quad r_t = r_t^* + \pi_t + a_\pi (\pi_t - \pi^*) + a_y (Y_t - Y_t^*),$$

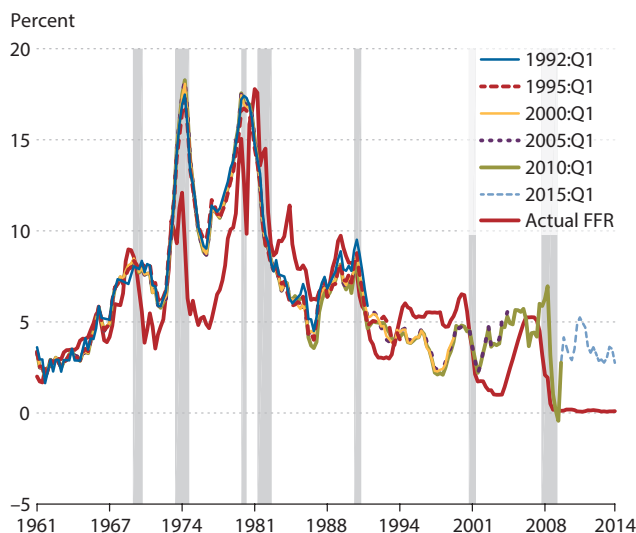
where  $r_t$  is the equilibrium interest rate. A common parameterization of the Taylor rule is to set  $a_\pi = a_y = 0.5$  and set the equilibrium interest rate  $r_t^* = 2.0$ .

The two coefficients  $a_\pi$  and  $a_y$  reflect the policymaker's responsiveness to inflation and output deviations, respectively. Larger coefficients imply larger movements in the policy instrument (usually the federal funds rate). However, the coefficients in the Taylor rule are not the only "free parameters." While the Fed has recently adopted a target band for inflation centered on 2 percent, how one computes potential output can affect the prescribed interest rate even with fixed parameters and a fixed inflation target.

Figure 8 shows the time series of the federal funds rate prescribed by the Taylor rule in Equation 3 using the six measures of potential output described above and the actual federal funds rate over the same period. Each of these series is computed using the full sample of data. Except for the beginning of the sample, the different measures of potential prescribe almost the same policy rate. In the beginning of the sample, the linear time trend and bivariate UC yield estimates that are slightly different from those of the other four measures because of large differences in each measure's estimated output gap. The linear trend prescribes a slightly higher policy rate in the early years of the sample because assuming a constant growth rate places potential below actual output for the first 10 years of data.

**Figure 9**

**Real-Time Policy Rate Prescriptions Calculated Using the Univariate UC Model as Potential Output in a Taylor Rule**



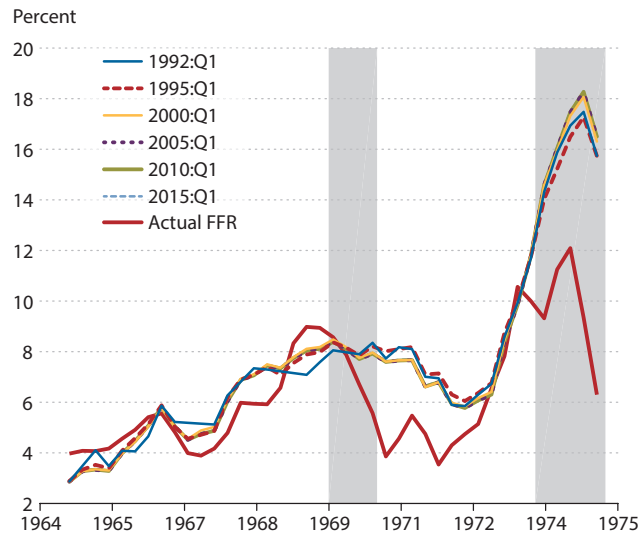
NOTE: FFR, Federal funds rate. Gray bars indicate recessions as determined by the NBER.

SOURCE: ALFRED®, BEA, BLS, Board of Governors of the Federal Reserve System, and authors' calculations.

One issue with the exercise in Figure 8 is that policymakers did not have access to all of the information for the full sample of data at every point in time. This availability issue is important because the HP filter is a two-sided filter and both UC methods use smoothers. The use of smoothers means that they use all of the data available to infer potential output—that is, data at time  $T$  at the end of the sample could influence the estimate of potential output at time  $t < T$ . Because the policymaker would not have this period- $T$  data available, his or her estimate of potential output could differ substantially and lead to a very different prescription for policy.

Figure 9 shows the real-time policy rates prescribed by the univariate UC model. The full-sample data above do not yield drastic differences for the policy rate; however, different vintages within the same measure do show important differences. We show only the vintages for the univariate UC model, but the real-time policy rates from the other measures yield similar results.<sup>13</sup> As more data become available and historical data are revised, estimated potential changes and so does the respective output gap. These updates to the data series lead us to



**Figure 10****Real-Time Taylor-Rule-Implied Policy Rates Using the Univariate UC Model for the Output Gap Calculation**

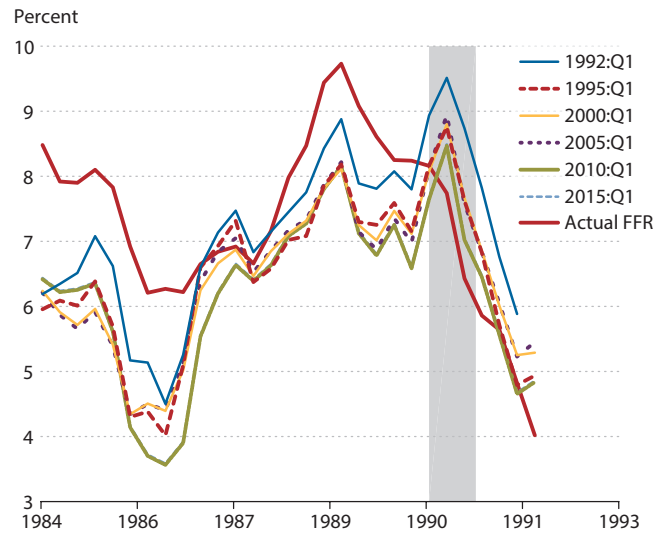
NOTE: The limited date range shows that, after the 1992:Q1 vintage, data revisions imply a later turning point in the 1960-1970 recession. FFR, Federal funds rate. Gray bars indicate recessions as determined by the NBER.

SOURCE: ALFRED®, BEA, BLS, Board of Governors of the Federal Reserve System, and authors' calculations.

reach different conclusions about policy rate levels around turning points. The real-time policy rule identifies an earlier turning point for the 1970 recession in the 1995:Q1 vintage than in the 1992:Q1 vintage and exhibits a lower max policy rate during the double-dip 1980s recession. After 1992:Q1, the revised data cause a downward level shift in the prescribed policy rate during the 1980s. Beyond a small blip in the late 1960s between the 1992:Q1 and 1995:Q1 vintages, the turning points occur on the same dates. However, the level of the policy rate at these turning points does vary. Figures 10 and 11 highlight level shifts in vintage policy rules around turning points. For example, the turning point before the 1990s recession was revised lower as new vintages were released. Figure 10 shows that after the 1992:Q1 vintage, data revisions imply a later turning point in the 1969-70 recession. Figure 11 highlights the downward shift in prescribed policy rates during the 1980s. As data revisions were released, policy rates shifted from a peak of around 9.5 percent from the 1992:Q1 vintage during the 1991 recession to around 8.5 percent in the 2010:Q1 and 2015:Q1 vintages.

**Figure 11**

**Real-Time Taylor-Rule-Implied Policy Rates Using the Univariate UC Model for the Output Gap Calculation**



NOTE: The limited date range shows that, as data were revised, Taylor rule policy rates were shifted lower through the 1980s. FFR, Federal funds rate. Gray bars indicate recessions as determined by the NBER.

SOURCE: ALFRED®, BEA, BLS, Board of Governors of the Federal Reserve System, and authors' calculations.

## CONCLUSION

We considered a number of relatively common methods of measuring potential output. While the measures produce qualitatively similar results, they can also vary at important times: at the beginnings and ends of the sample periods and around turning points. Moreover, the measures can vary substantially in real time. However—and perhaps fortunately for the policy-maker—simple interest rate rules do not show wildly different policy prescriptions based on the different measures. ■

## NOTES

- <sup>1</sup> See also Kamber, Morley, and Wong (2018) for a survey.
- <sup>2</sup> The largest and most important sector in the CBO model is the nonfarm business sector, which accounts for approximately 70 percent of GDP.
- <sup>3</sup> For a more detailed discussion, see Arnold (2009).
- <sup>4</sup> The CBO sets  $\alpha = 0.3$  based on historical growth accounting data that measure payments to owners of capital as 30 percent of total income.
- <sup>5</sup> The CBO's NAIRU (non-accelerating inflation rate of unemployment) estimation relies on a Phillips curve assumption that indicates when the labor force is at the potential level.
- <sup>6</sup> This assumption is valid if firms minimize costs so that the cost share reflects productivity and if there are constant returns to scale between individual types of capital and the aggregate index (CBO, 2001).
- <sup>7</sup> De Jong and Sakarya (2016) formally outline the statistical characteristics of the filter. Hamilton (2017) argues against the use of the HP filter and proposes an alternative.
- <sup>8</sup> We allow correlation across the trends of the two variables and correlation across the cycles of the two variables. We maintain the assumption of independence across the trend and cycle within and across variables, consistent with our univariate UC model. Relaxation of these assumptions is straightforward to implement (see Morley, Nelson, and Zivot, 2003, or Sinclair, 2009).
- <sup>9</sup> For a full discussion of advantages and disadvantages related to the CBO's production-function-based approach, see Arnold (2009).
- <sup>10</sup> The CBO argues that filtered growth data is "trend output" rather than "potential output" because it does not tie directly to stable inflation (CBO, 2004).
- <sup>11</sup> Dungey, Jacobs, and Tian (2017) find that correlated innovations are more important than structural breaks for potential estimation, which suggests greater consideration should be given to multivariate models incorporating inflation into measures of potential.
- <sup>12</sup> Orphanides and van Norden (2002) have argued that estimates of the output gap computed in real time should be used cautiously. However, the policymaker does not have the luxury of knowing the magnitude of future data revisions: He or she must compute potential output in real time. Edge and Rudd (2016) show that the difficulty of estimating potential in real time may have improved in the later 2000s, after the period studied by Orphanides and van Norden.
- <sup>13</sup> Graphs presenting all vintages of the various potential measures are available upon request.

## REFERENCES

- Arnold, Robert W. "The Challenges of Estimating Potential Output in Real Time." *Federal Reserve Bank of St. Louis Review*, July/August 2009, 91(4), pp. 271-90; <https://doi.org/10.20955/r.91.271-290>.
- Basistha, Arabinda and Startz, Richard. "Measuring the NAIRU with Reduced Uncertainty: A Multiple-Indicator Common-Cycle Approach." *Review of Economics and Statistics*, November 2008, 90(4), pp. 805-811; <https://doi.org/10.1162/rest.90.4.805>.
- Basu, Susanto and Fernald, John G. "What Do We Know (and Not Know) About Potential Output?" *Federal Reserve Bank of St. Louis Review*, July/August 2009, 91(4), pp. 187-213; <https://doi.org/10.20955/r.91.187-214>.
- Champagne, Julien; Polin-Bellisle, Guillaume and Sekkel, Rodrigo. "The Real-Time Properties of the Bank of Canada's Staff Output Gap Estimates." *Bank of Canada Staff Working Paper No. 16-28*, 2016.
- Clark, Peter K. "The Cyclical Component of U.S. Economic Activity." *Quarterly Journal of Economics*, November 1987, 102(4), pp. 797-814; <https://doi.org/10.2307/1884282>.
- Congressional Budget Office. "CBO's Method for Estimating Potential Output: An Update." August 2001.

## Guisinger, Owyang, Shell

- Congressional Budget Office. "A Summary of Alternative Methods for Estimating Potential GDP." Background paper, March 2004.
- De Jong, Robert M. and Sakarya, Neslihan. "The Econometrics of the Hodrick-Prescott Filter." *Review of Economics and Statistics*, May 2016, 98(2), pp. 310-17; [https://doi.org/10.1162/REST\\_a\\_00523](https://doi.org/10.1162/REST_a_00523).
- Dungey, Mardi; Jacobs, Jan P.A.M. and Tian, Jing. "Forecasting Output Gaps in the G-7 Countries: The Role of Correlated Innovations and Structural Breaks." *Applied Economics*, 2017, 49(45), pp. 4554-66; <https://doi.org/10.1080/00036846.2017.1284998>.
- Edge, Rochelle M. and Rudd, Jeremy B. "Real-Time Properties of the Federal Reserve's Output Gap." *Review of Economics and Statistics*, October 2016, 98(4), pp. 785-91; [https://doi.org/10.1162/REST\\_a\\_00555](https://doi.org/10.1162/REST_a_00555).
- Fernald, John G. "Productivity and Potential Output Before and After the Great Recession." Working Paper Series 2014-15, Federal Reserve Bank of San Francisco, June 2014.
- Hamilton, James D. "Why You Should Never Use the Hodrick-Prescott Filter." Unpublished manuscript, 2017.
- Harvey, A.C. *Forecasting, Structural Time Series Models, and the Kalman Filter*. Cambridge, MA: Cambridge University Press, 1989.
- Hodrick, Robert and Prescott, Edward C. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit, and Banking*, 1997, 29(1), pp. 1-16; <https://doi.org/10.2307/2953682>.
- Kamber, Gunes; Morley, James and Wong, Benjamin. "Intuitive and Reliable Estimates of the Output Gap from a Beveridge-Nelson Filter." *Review of Economics and Statistics*, July 2018, 100(3), pp. 550-66; [https://doi.org/10.1162/rest\\_a\\_00691](https://doi.org/10.1162/rest_a_00691).
- Kiley, Michael T. "Output Gaps." *Journal of Macroeconomics*, 2013, 37, pp. 1-18; <https://doi.org/10.1016/j.jmacro.2013.04.002>.
- Marcellino, Massimiliano and Musso, Alberto. "The Reliability of Real-Time Estimates of the Euro Area Output Gap." *Economic Modeling*, 2011, 28, pp. 1842-56; <https://doi.org/10.1016/j.econmod.2011.03.014>.
- Morley, James C.; Nelson, Charles R. and Zivot, Eric. "Why Are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP So Different?" *Review of Economics and Statistics*, May 2003, 85(2), pp. 235-43; <https://doi.org/10.1162/003465303765299765>.
- Okun, Arthur M. "Potential GNP, Its Measurement and Significance." American Statistical Association, Proceedings of the Business and Economics Statistics Section, 1962, pp. 98-204.
- Orphanides, Athanasios and van Norden, Simon. "The Unreliability of Output Gap Estimates in Real Time." *Review of Economics and Statistics*, 2002, 84, pp. 569-83; <https://doi.org/10.1162/003465302760556422>.
- Ravn, Morten O. and Uhlig, Harald. "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations." *Review of Economics and Statistics*, May 2002, 84(2), pp. 371-76; <https://doi.org/10.1162/003465302317411604>.
- Sinclair, Tara M. "The Relationships between Permanent and Transitory Movements in U.S. Output and the Unemployment Rate." *Journal of Money, Credit, and Banking*, March-April 2009, 42(2-3), pp. 529-42; <https://doi.org/10.1111/j.1538-4616.2009.00220.x>.
- Taylor, John B. "Discretion Versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy*, December 1993, 39, pp. 195-214; [https://doi.org/10.1016/0167-2231\(93\)90009-L](https://doi.org/10.1016/0167-2231(93)90009-L).
- Watson, Mark W. "Univariate Detrending Methods with Stochastic Trends." *Journal of Monetary Economics*, 1986, 18, pp. 49-75; [https://doi.org/10.1016/0304-3932\(86\)90054-1](https://doi.org/10.1016/0304-3932(86)90054-1).