Are all Output Gap Estimates Unstable in Real Time?

Alessandro Barbarino*
Federal Reserve Board

Travis J. Berge
Federal Reserve Board

Han Chen
Federal Reserve Board

Andrea Stella
Federal Reserve Board

March 1, 2019

Preliminary and incomplete, please do not redistribute

Abstract

The output gap estimate produced by the Federal Reserve Staff is known to be more reliably estimated in real time than univariate de-trending models used in the literature (Edge and Rudd, 2016). The purpose of this paper is to understand why. We estimate several multivariate unobserved-component (UC) models of the economy and show that the real-time stability of the Federal Reserve estimates is likely due to the use of labor market data to inform the estimation of the cyclical state of the economy. We find that a simple two equation UC model that estimates the output gap using output and the unemployment rate produces real-time output gap estimates with time-series and revision properties very similar to the Federal Reserve staff's judgmental estimate. We also investigate the usefulness of the output gap estimates when forecasting inflation.

JEL classifications: L11, E24

Keywords: Output gap estimation; real-time data; unobserved components model; trend-cycle decomposition.

*The views expressed here should not be interpreted as reflecting the views of the Federal Reserve Board of Governors or any other person associated with the Federal Reserve System.

1 Introduction

The output gap is a fundamental economic concept for economic policymakers.[1] However, the literature started by Orphanides and van Norden (2002) carries a fundamentally negative message relative to the use of the output gap for policy purposes: the "standard" detrending techniques used by practitioneers to produce this central concept are extremely unreliable, with ex post revisions of the gap in the United States of the same order of magnitude as the estimated gap itself, jeopardizing any serious attempt to use it by policy makers.[2]

Further, when an explicit link with inflation is hard-wired into the detrending model via a Phillips curve, such as in the model by Kuttner (1994), real-time stability does not improve, rather it worsens sligthly and Orphanides and van Norden (2005) show that the resulting unstable real-time output gap does not help at all in forecasting inflation in real-time. Not only, detrending models aside, they also show that the Federal Reserve's staff estimates of the output gap publicly available at the time of their analysis (from the 1980s to the 1990s) do not escape such instability.[3] Such pessimistic message has been tempered over the years.

A few researchers before us investigated the reasons for the poor real-time performance of the detrending methods considered by Orphanides and van Norden (2002) and proposed alternatives that can overcome such unreliability. Garratt et al. (2008) suggest that the detrending statistical models considered by Orphanides and Van Norden are fraught with end-point estimation problems and propose to catch two birds with a stone by using forecasts from a small VAR model to fix those problems and jointly fold in information from data revisions. Notice that although Garratt et al. (2008) propose the use of a VAR, their detrending methodology is in a sense "univariate" being squarely focused on different GDP releases and no other information, such as inflation or labor

[1] See, for example, the January 19, 2017 speech of Janet Yellen "The Economic Outlook and the Conduct of Monetary Policy").

[2] Orphanides and van Norden (2002), in addition to deterministic detrending techniques, considered a battery of statistical detrending models available at the time including the filter by Hodrick and Prescott (1997) (HP filter), the univariate unobserved components approaches of Watson (1986), Harvey (1985) and Clark (1987) and the bivariate unobserved component approaches of Kuttner (1994) and Gerlach and Smets (1997) that pair a Phillips curve to the detrending model for GDP.

[3] Federal Reserve staff forecasts and estimates are contained in a document called Tealbook (once called Greenbook) and become available to the public with a 5 years lag.

market variables.

Fleischman and Roberts (2011) approach the output gap estimation problem from a truly multivariate perspective and find that the enlargement of the information set on which a UC model feeds with the inclusion of labor market variables (above all the unemployment rate) is effective in abating the instability of gaps estimates from UC models.[4] However, their exercise is quasi real-time, they do not disentangle the effects of the unemployment rate and inflation, they do not study the inflation forecasting properties of their gap and above all they do not relate their gap to the Federal Reserve's official estimate, the main exercise of this paper. More recently Hamilton (2018) seemingly deals the last blow to the HP filter by reviewing previously known defects and unveiling new ones. In addition, Hamilton comes up with a simple to implement alternative univariate detrending technique and he proves it has desirable statistical properties. Hamilton's approach, in contrast to Garratt et al. (2008), is backward looking and does not require forecasts. We show that Hamilton's detrending procedure is very effective in conferring real-time stability to the output gap and the intuition of this result is that such procedure does not suffer from an end-point problem by construction. This said, we also find that relation with the Federal Reserve judgmental gap is tenuous and the resulting gap does not seem informative in order to conduct monetary policy. The same considerations apply to the gap obtained with the comprehensive environment proposed by Mueller and Watson (2017) in which to detrend and do inference in the long-run: that is the resulting gap is fairly stable in real-time however it does not seem very useful for policy purposes. On another front, Edge and Rudd (2016) reconsider the Federal Reserve's staff estimates of the output gap and show that they have become much more stable in real-time after the 1990s. They also show that the resulting measure of slack has comparable performance in forecasting inflation in real-time to alternative models, a reassuring property given the first leg of the Federal Reserve mandate.[5]

This paper first updates and confirms the results by Edge and Rudd (2016) including data

[4]Trimbur (2009) also finds that a bivariate UC model that includes manufacturing capacity utilization (CU), an eminently cyclical series, has also excellent real-time properties. Trimbur does not impose a common cycle yet he estimates a very high correlation between GDP and CU cycles of 0.9.

[5]Since Edge and Rudd work with the Federal Reserve staff estimate of the output gap, which is much more stable in real-time than the gaps considered by Orphanides and van Norden (2005), they are able to exclude that real-time instability of the gap is at the root of the poor real-time performance of gap-based inflation forecasting models. Instead, they view their results as reflecting the general decline in the forecastability of inflation documented by Stock and Watson (2007, 2009).

through 2012, then it addresses the reasons why the Federal Reserve's official estimates of the output gap have become more stable in real-time.[6] We speculate that the real-time stability of the Federal Reserve gap estimates have been affected by a combination of factors, specifically the inclusion of labor market information in the judgmental estimate of the gap and the adoption of newer detrending techniques in the background preliminary work that leads to the judgmental estimate. It is a well-known fact that the Federal Reserve pays quite a bit of attention to labor market outcomes, also for the assessment of slack (see last speech of chairman Powell for instance). However, we explicitly show that only multivariate specifications of an unobserved component (UC) model that include the unemployment rate are able to jointly i) reproduce output gaps that have the same degree of real-time stability as the Federal Reserve estimate, and ii) are most closely related to the Federal Reserve's staff estimates of the output gap.

A second contribution of the paper is showing that the alternative and more recent detrending techniques reviewed above, although produce output gaps that are equally stable in real-time, the measures of slack that they produce are only loosely related to the Federal Reserve estimate of the output gap and do not appear "smooth" enough to guarantee an effectively implementable monetary policy.

The rest of the paper is organized as follows. We set the stage in Section 2 by reviewing and updating the real-time stability properties of the Federal Reserve staff's estimate of the gap; we also inspect some of its dynamic revision properties not highlighted by prior research. In Section 3 we show that the inclusion of the unemployment rate in UC detrending models coupled with a "sufficiently tight" Okun's law is a key ingredient in order to confer real-time stability to the gap estimates. We explore alternative detrending methods in Section 3.3. Finally in Section 4 we evaluate the properties of the gaps from the various detrending methods that we consider. We find that the gap from a UC model that folds in unemployment information is the closest model-based match to the Federal Reserve judgmental estimate in the sense of being the most correlated with it, inducing the closest monetary policy actions, having similarly good real-time stability properties and having

[6] Including data through 2012 allows to fully take into account the great recession in output gap estimates, with the non-trivial consequence of offering the arguably most severe available test of the real-time stability of output gap estimates since the great moderation.

a similar inflation forecasting performance (superior to the other methods).

2 Properties of The Federal Reserve's Output Gap Estimate

Before delving into model-based estimates of the output gap, we update through 2012 the real-time evaluation of the Federal Reserve's judgmental output gap estimate originally performed by Orphanides and van Norden (2002) and subsequently by Edge and Rudd (2016).

2.1 Real-time Stability Results

Coming very soon.

3 UC Models and the Role of Labor Market Information

The business cycle, C, is assumed to follow a stationary AR(2) process

$$C_t = \phi_1 C_{t-1} + \phi_2 C_{t-2} + \varepsilon_t^C \tag{1}$$

so that it responds to shocks in the standard hump-shaped manner.

The model decomposes the log of real GDP per capita (y) into a trend (y^*), cycle (C), and error (e_y):

$$y_t = y_t^* + C_t + e_{yt}. \tag{2}$$

The term e_{yt} captures measurement error and transitory shocks to the level of GDP. Trend output is modeled as a random walk with drift, τ, which itself follows a random walk:

$$y_t^* = y_{t-1}^* + \tau_{t-1}^* + \varepsilon_{y^*t} \tag{3}$$

$$\tau_t = \tau_{t-1} + \varepsilon_{\tau t} \tag{4}$$

The drift term captures slow-moving changes to the growth rate of trend output, posibly due to changing productivity growth and demographic changes to the labor force.

The unemployment rate, u, is the sum of trend (u^*), a common cycle, and measurement error.

$$u_t = u_t^* + \alpha_0 C_t + \alpha_1 C_{t-1} + e_{ut}. \tag{5}$$

The parameters α_0 and α_1 are the Okun's law coefficients; they relate deviations in the unemployment rate from its natural rate to the output gap. We model the unemployment rate trend as a random walk, as Staiger et al. (1997):

$$u_t^* = u_{t-1}^* + \varepsilon_{u^*t}. \tag{6}$$

Inflation, π, is the sum of its trend (π^*), the common cycle, and measurement error. Again, trend inflation follows a random walk:

$$\pi_t = \pi_t^* + \beta_0 C_t + \beta_1 C_{t-1} + e_t^\pi \tag{7}$$

$$\pi_t^* = \pi_{t-1}^* + \varepsilon_{\pi^*t} \tag{8}$$

Trend inflation can be thought of representing inflation expectations, so that the inflation specification is akin to a New-Keynesian Phillips curve.[7] The β parameters are the Phillips curve parameters that describe inflation's response to slack.

Group the observables into the vector $Y_t = (y_t, u_t, \pi_t)'$ and the unobserved states into the vector $s_t = (y_t^*, \tau_t^*, u_t^*, \pi_t^*, C_t, C_{t-1})'$. The model is then compactly written in state-space form:

$$Y_t = F s_t + e_t; \quad e_t \sim \mathrm{N}(0, \Sigma) \tag{9}$$

$$s_t = G s_{t-1} + \varepsilon_t; \quad \varepsilon_t \sim \mathrm{N}(0, \Omega) \tag{10}$$

For simplicity, we assume the measurement errors are uncorrelated with each other as well as with the shocks to the states.

We will consider several nested versions of the model described above. The first is the model of Clark (1987) with $Y_t = (y_t)'$, which we denote UC1. We then gradually enlarge the information

[7](Stock and Watson 2007, Clark and Doh 2014).

set to include the unemployment rate and inflation, that is, a UC2 model with $Y_t = (y_t, u_t)'$) and a UC2 model with $Y_t = (y_t, \pi_t)'$. Considering results from a bivariate UC model with GDP and the unemployment rate allows us to single out the individual contributions of labor markets and price information and their complementarity. Finally, we include all three indicators in the UC3 model, $Y_t = (y_t, u_t, \pi_t)'$.

3.1 Overview of Estimation

This model is a linear Gaussian state-space model that we estimate using a Gibbs sampler. The intuition of the sampler is straightforward: conditional on the data and states, equations 9 and 10 are a set of independent linear, Gaussian regressions for which parameters can be sampled from their known posterior distribution. Conditional on the model's parameters, the states can be drawn using the Kalman filter. Thus, we estimate the model using a Gibbs sampler alternates between draws of model parameters that are conditioned on the states and draws of the states that are conditioned on model parameters.

Each step of the sampler is standard, so we leave the details to the appendix. The algorithm consists the following five steps:

1. Sample F. Conditional on the data, states and Σ, draw the Okun's law and Phillips curve parameters (α and β).

2. Sample Σ. Conditional on the data, states, and F, draw the variances of the measurement errors.

3. Sample G. Conditional on states and Ω, draw the AR parameters of the cycle. We impose stationarity.

4. Sample Ω. Conditional on the data (y^T), states (s^T) and G, draw the variances of the shocks to the states.

5. Sample the states. Conditional on the data and parameters, we draw the states using the Carter and Kohn (1994) algorithm.

We produce 11,000 draws from this sampler, ignoring the first 1,000 and using a thinning factor of 5 for a total of 2,000 draws from each conditional posterior distribution.

3.1.1 Priors

We use the H-P filter to calibrate the prior variance of the shocks to the unobserved components.

3.2 Real-Time Stability Results for UC Models

The four UC model specifications allow us to evaluate the merits of different information sets in conferring additional real-time stability to a simple univariate model. Table 1 summarizes a few key statistics of the revision (revision being "final minus real-time") properties of the gaps from the different UC models we estimate.

Table 1: UC Models Revision Statistics: Final Versus Real-Time Estimates

Method	*MEAN*	*SD*	*RMS*	*MIN*	*MAX*	*AR*
Hodrick-Prescott	0.15	1.58	1.58	-3.28	3.36	0.93
Hamilton (2018)	-0.79	1.46	1.66	-6.42	4.29	0.73
UC1 (gdp)	-0.02	0.96	0.95	-3.20	2.21	0.82
UC2 (gdp, π)	0.13	1.3	1.31	-5.67	4.58	0.85
UC2 (gdp, u)	0.10	0.66	0.66	-1.09	2.98	0.94
UC3 (gdp, π, u)	0.16	0.66	0.68	-0.89	3.16	0.86

Note — *SD* is the standard deviation of the revision; *RMS* denotes the mean square of the revision series shown, and *AR* the first-order serial correlation of the series. All statistics are for 1965:4–2017:4

Table 2 collects the handful of reliability statistics proposed by Orphanides and van Norden (2002) and that have become pretty standard in real-time evaluation of output gap measures since then. The column denoted "COR" is the simple correlation of real-time and final output measures. The column denoted "NS" reports the ratio of the standard deviation of the revision to that of the final estimate of the gap: the revision should not be much more variable than the inherent variability of an output gap estimate from a given model, therefore a high NS statistic denotes real-time instability. The column denoted "NSR" reports a measure of real-time instability somewhat similar to NS: it is the ratio of the root mean square of the revision to the standard deviation of the final estimate of the gap and again a high NRS statistics is a symptom of instability. Finally one desirable stability property of a given model is not suggesting a change in monetary policy prescriptions between real-time and final output gap estimates and one rather coarse yet simple way of measuring it is the probability that real-time and final gaps have the opposite signs (measured here by the relative frequency that real-time and final gaps have opposite signs); here again a low OPSIGN is desirable.

Table 2: Summary of UC Models Reliability Indicators

Method	*COR*	*NS*	*NSR*	*OPSIGN*
HP Filter	0.51	1.04	1.04	0.41
Hamilton (2018)	0.89	0.45	0.51	0.12
UC1 (gdp)	0.73	0.72	0.72	0.25
UC2 (gdp,π)	0.56	0.86	0.86	0.24
UC2 (gdp,u)	0.98	0.22	0.22	0.06
UC3 (gdp,u,π)	0.98	0.23	0.23	0.05

Note — In the table *COR* denotes the correlation of the real-time and final estimates. *NS* denotes the ratio of the standard deviation of the revision to that of the final estimate of the gap. *NSR* denotes the ratio of the root mean square of the revision to the standard deviation of the final estimate of the gap. *OPSIGN* denotes the frequency with which the real-time and final gap estimates have opposite signs. All statistics are for 1965:4–2017:4

In terms of the real-time stability of the output gap measures that come out of our experiments, we note that UC1 and UC2 with inflation are the worst performers. The addition of price information in the latter model surprisingly does not appear to carry any gains of sort, contrary to what was found by Planas and Alessandro (2004) on a shorter sample. To the contrary, although the correlation between final and real-time gaps are similar (see column denoted "COR"), UC1 appears to be more stable in real-time when stability is measured by all other statistics: i) UC2 with inflation is slightly worse considering the ratio of the standard deviation of the revision to that of the final estimate (column denoted "NS"), ii) it is worse considering the ratio of the root mean square of the revision to the standard deviation of the final estimate of the gap ("NSR" column) and iii) the probability that real-time and final output gaps have opposite signs is also slightly higher (column "OPSIGN"). By contrast folding in information on the unemployment rate as in the UC2 with unemployment offers a drastic improvement on all accounts. Finally, UC3 performs very similarly to the UC2 with unemployment.

3.3 Alternative Detrending Techniques

Although the HP filter has been widely used in macro-econometric analysis for a long time, aside from the critiques by Hamilton (2018), the paper by Orphanides and van Norden (2002) already showcased its shortcomings in extracting a measure of the output gap. We confirm in Table 2 those earlier results also with data through 2018. By contrast, we are the first, to our knowledge, to test the real-time stability of Hamilton's alternative detrending technique and we find that it performs impressively well, above all considering its implementation simplicity and the fact that it uses only information on GDP. Table 2 highlights that Hamilton's model is not too far from the excellent performance of the UC2 with unemployment and UC3.

4 Evaluating Output Gaps Beyond Real-Time Stability

Results so far suggest that in terms of real-time stability of the extracted output gap, the HP filter appears a sub-optimal solution and UC models greatly benefit from the inclusion of the unemployment rate in their information set. This said, aside UC models we also found that Hamilton's methods confer similar real-time stability to the respective output gaps. However, the improved real-time stability of the gap alone does not imply that it can be a useful policy tool. Therefore we assess the capability of gaps from the models we estimate in mimicking the gap estimated by the FED, which for the moment we assume to be a proxy for a useful policy gap.

4.1 Relationship of Model-based Output Gaps to the FED Output Gap

In this section we show that the output gaps from the UC models that contain the unemployment rate are much closer to the judgmental output gap from the Federal Reserve than other model-based gaps. In order to evaluate such distance and to preserve uniformity with previous sections we use the same metrics that we have used to evaluate the stability in real-time with appropriate modifications. That is, we focus on the real-time version of the model-based gaps and we study its closeness with the final version of the Federal Reserve gap by computing the correlation of the two gaps in the first column of Table 3. In other words we take the FED's final estimate of the output gap as "truth" instead of the model's final estimate. As evident the UC models gaps are much more correlated with the output gap estimate produced by the FED staff. We also summarize the distance between real-time model-based and FED gaps in the second column denoted *NS* (reporting the relative standard deviation of the distance between the two, in which the denominator is the final FED staff gap) or in the column *NSR* that reports the root mean square error of the relative distance between the two (again relative to the FED staff estimate).[8] Again UC models that exploit the unemployment rate are closer to the FED judgmental gap. Moreover as shown in the last column of Table 3 gaps from UC models with the unemployment rate are also less likely to differ from the FED gap in sign.

Table 3: Distance of Model-Based to Federal Reserve Real-Time Output Gaps

Method	*COR*	*NS*	*NSR*	*OPSIGN*
Hodrick-Prescott	0.2	1.01	1.12	0.48
Hamilton (ar)	0.72	0.74	1.11	0.36
UC1 (gdp)	0.65	0.91	1	0.26
UC2 (gdp, π)	0.65	0.9	0.98	0.25
UC2 (gdp, u)	0.97	0.27	0.34	0.14
UC3 (gdp, π, u)	0.97	0.26	0.34	0.14

Note — In the table *COR* denotes the correlation of the real-time estimates of the models and the final estimate of the Federal Reserve. *NS* denotes the ratio of the standard deviation of the distance of model-based and FED gaps. *NSR* denotes the ratio of the root mean square of the distance of model-based and FED staff gaps. *OPSIGN* denotes the frequency with which the real-time model-based and final FED staff gaps have opposite signs. All statistics are for 1965:4–2012:4

Figure 1 shows this information visually. The picture plots, for seven different output gap estimates, the model's final estimate plotted against its real-time estimate (the horizontal axis). If a model produces a stable estimate of the output gap, the final and real-time estimates will lie very close to the 45 degree line. For ease of comparison, the Tealbook estimate is plotted in each figure,

[8]Notice that here all ratios are taken relative to the FED staff final estimate. In this case, since the denominators in the *NS* and *NSR* columns are all the same, the entries in Table 3 imply the same ranking for *absolute* distances as well.

alongside each of the other six estimates shown.

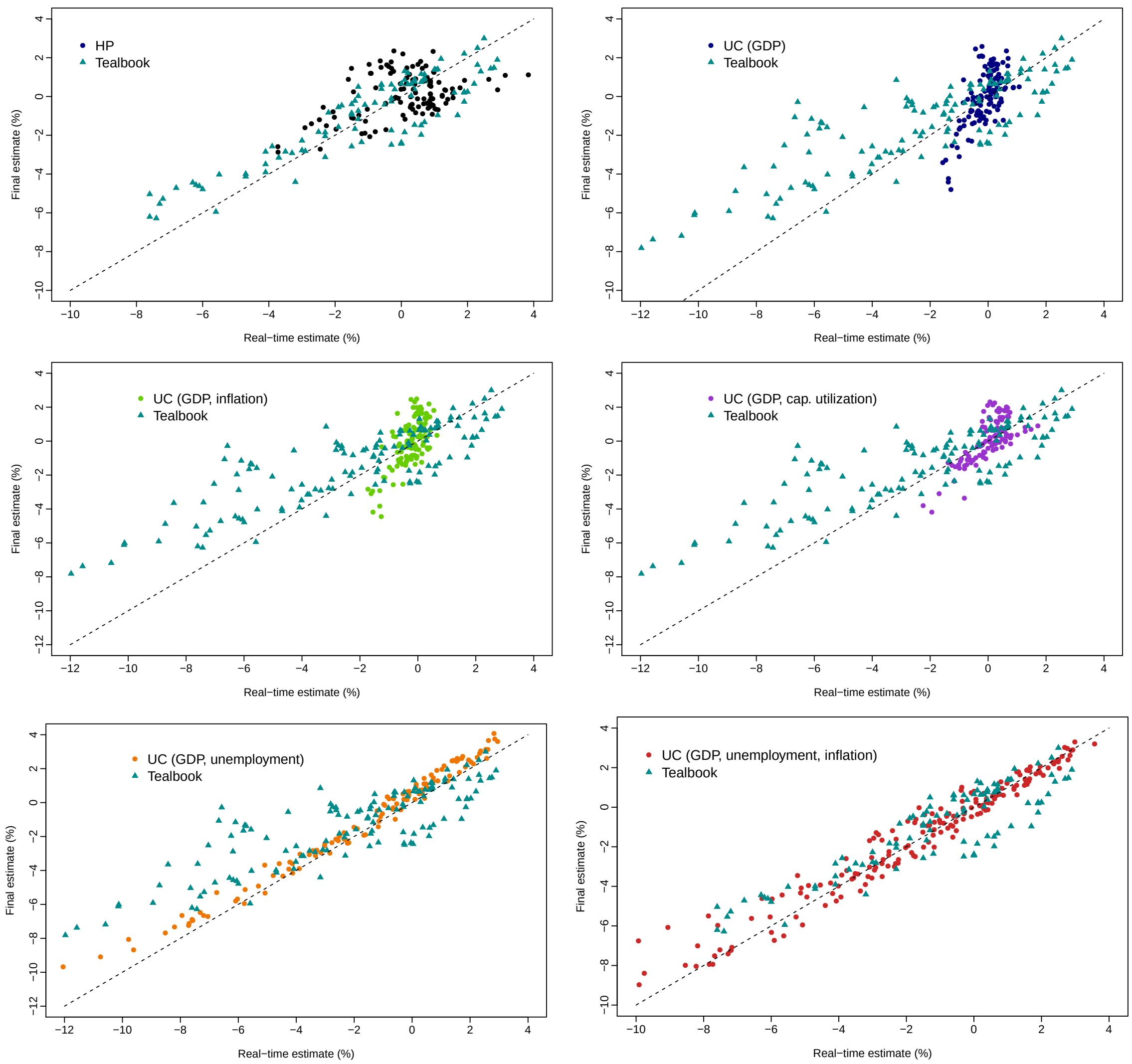
The top row of figure 1 plots two univariate output gap estimates—those from the HP filter (top left) and a univariate UC model (top right). In both cases, the real-time and final estimates appear only very weakly correlated with one another. In addition, it is clear that the models do not provide as "strong" a signal as the Tealbook output gap estimate: the variance of the output gap estimate is much smaller than the variability in the Tealbook estimate. The second row shows the results from two bivariate UC models: UC (GDP, inflation) on the left and UC (GDP, capacity utilization) in purple on the right. The GDP/inflation UC output gap estimate is broadly unchanged from the univariate UC estimate in the top right; inflation does little to stabilize the UC model's output gap estimate. In contrast, capacity utilization does appear to help stabilize the output gap estimate; the purple dots in the middle-right panel tend to be closer to the 45 degree line than the other models.

However, the difference once one adds the unemployment rate to the model is startling. The bivariate UC (GDP/unemployment) model shown in the bottom-left is very stable, as is the 3-variable UC model in the bottom-right. Further, these models are able to provide a strong output gap signal in real-time, in the sense that both the real-time and final estimates are quite variable.

Table 4 shows the distance of the real-time model gaps to the CBO staff output gap. As can be quickly glanced at in Table 4 the models maintain the same ranking relative to the CBO as they did relative to the FED gap. Contrary to the judgmental construction of the FED output gap, the methodology of the construction of the output gap by the CBO is well documented in Shackleton (2018), from which it is clear the important role played by the unemployment rate and labor market information. The CBO uses a complex, ad-hoc, disaggregated methodology to estimate the output gap, definitely not a UC model.[9] Despite this, Table 4 shows that the basic UC model that includes the unemployment rate delivers the gap closest to the CBO gap. Therefore we view Table 4 as indirect evidence that labor market information plays an important role in the construction of the FED gap.

[9]Although the CBO output gap does not come from a UC model, it includes an Okun law-type restrictions and Phillips curves.

Figure 1: Dynamic Properties of Output Gap Revisions



4.2 Forecasting Inflation with Slack

We now turn to the usefulness of the gaps from the various detrending models we reviewed in forecasting inflation compared to the gap produced by the FED Staff, arguably one of the most important use of the output gap by a monetary policy authority.

Coming soon!

Table 4: Distance of Model-Based to CBO Real-Time Output Gaps

Method	*COR*	*NS*	*NSR*	*OPSIGN*
Hodrick-Prescott	0.37	0.98	1.29	0.5
Hamilton	0.77	0.91	1.61	0.53
UC1 (gdp)	0.66	0.86	1.15	0.34
UC2 (gdp,π)	0.68	0.84	1.10	0.26
UC2 (gdp, u)	0.92	0.79	0.93	0.34
UC3 (gdp, π, u)	0.92	0.77	0.91	0.34

Note — In the table *COR* denotes the correlation of model-based estimates and CBO staff gaps. *NS* denotes the ratio of the standard deviation of the distance of model-based and CBO gaps. *NSR* denotes the ratio of the root mean square of the distance of model-based and CBO staff gaps. *OPSIGN* denotes the frequency with which the model-based and CBO staff gaps have opposite signs. All statistics are for 1965:4–2017:4

References

Carter, Chris K. and Robert Kohn, "On Gibbs Sampling for State Space Models," *Biometrika*, 1994, *81* (3), 541–53.

Clark, Peter K., "The Cyclical Component of U.S. Economic Activity," *Quarterly Journal of Economics*, 1987, *102* (4), 797–814.

Clark, Todd E. and Taeyoung Doh, "Evaluating Alternative Models of Trend Inflation," *International Journal of Forecasting*, 2014, *30* (3), 426–48.

Edge, Rochelle M. and Jeremy B. Rudd, "Real-time properties of the federal reserve's output gap," *The Review of Economics and Statistics*, 2016, *98* (4), 785–91.

Fleischman, Charles A. and John M. Roberts, "From Many Series, One Cycle: Improved Estimates of the Business Cycle from A Multivariate Unobserved Components Model," Finance and Economics Discussion Series 2011-46, Board of Governors of the Federal Reserve System, Washington 2011.

Garratt, Anthony, Kevin Lee, Emi Mise, and Kalvinder Shields, "Real-time Representations of the output gap," *The Review of Economics and Statistics*, 2008, *90* (4), 792–804.

Gerlach, Stefan and Frank Smets, "Output Gaps and Inflation: Unobservable-Components Estimates for the G-7 Countries," Technical Report, Bank for International Settlements, Basel 1997.

Hamilton, James D., "Why You Should Never Use the Hodrick-Prescott Filter," *The Review of Economics and Statistics*, 2018.

Harvey, Andrew C., "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 1985, *3*.

Hodrick, Robert J. and Edward C. Prescott, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 1997, *29*, 116–30.

Kuttner, Kenneth, "Estimating Potential Output as a Latent Variable," *Journal of Business and Economic Statistics*, 1994, *12* (3), 361–36.

Mueller, U. and M. Watson, "Low-Frequency Econometrics," *Advances in Economics and Econometrics: Eleventh World Congress of the Econometric Society*, 2017, *2*, 53–94.

Orphanides, Athanasios and Simon van Norden, "The Unreliability of Output-Gap Estimates in Real Time," *The Review of Economics and Statistics*, 2002, *84* (4), 569–83.

____ **and** ____ , "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time," *Journal of Money, Credit, and Banking*, 2005, *37*, 583–601.

Planas, Christophe and Rossi Alessandro, "Can inflation data improve the real-time reliability of output gap estimates?," *Journal of Applied Econometrics*, 2004, *19*, 121–33.

Shackleton, Robert, "Estimating and Projecting Potential Output Using CBO's Forecasting Growth Model," Technical Report, Congressional Budget Office, Washington 2018.

Staiger, Douglas, James H. Stock, and Mark W. Watson, "The NAIRU, Unemployment and Monetary Policy," *The Journal of Economic Perspectives*, 1997, *11* (1), 33–49.

Stock, James H. and Mark W. Watson, "Why Has U.S. Inflation Become Harder to Forecast?," *Journal of Money, Credit and Banking*, 2007, *Supplement to Vol. 39* (1), 3–33.

Watson, Mark W., "Univariate Detrending Methods with Stochastic Trends," *Journal of Monetary Economics*, 1986, *18*, 49–75.