

Finding a Role for Slack in Real-Time Inflation Forecasting*

N. Kundan Kishor^a and Evan F. Koenig^b

^aUniversity of Wisconsin-Milwaukee

^bFederal Reserve Bank of Dallas

Real-time forecasting of PCE inflation is most successful when headline inflation is stripped of high-frequency noise and core inflation's trend and cycle are separately forecasted. It proves helpful, additionally, to allow cyclical inflation to respond to labor market slack, to allow for a late-1990s break in the behavior of trend inflation, and to explicitly model revisions to headline inflation. We do all of this within the context of an unobserved-common-components model of inflation and slack. The model's real-time inflation forecasts are significantly more accurate than those generated by benchmark models. That outperformance and the finding that cyclical inflation responds to slack are robust to an alternative measure of slack, an alternative model of trend inflation, and an alternative treatment of data revisions.

JEL Codes: E31, E37.

1. Introduction

Forecasting inflation is of great importance to policymakers, households, and businesses, and the literature on forecasting inflation is

*The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Dallas or the Federal Reserve System. We have benefited from comments received at the 2015 CIRANO Realtime Workshop (Montreal, October 2015), Midwest Econometrics Group Meeting (St. Louis, October 2015), the "day-ahead" meeting of the Federal Reserve System Committee on Macroeconomics (Cleveland, November 2015), and the SNDE Symposium (Tuscaloosa, March 2016). Author contact: Kishor: Professor, Department of Economics, Box 413, Bolton Hall 822, University of Wisconsin-Milwaukee, Milwaukee, WI 53201. E-mail: kishor@uwm.edu. Koenig: Senior Vice President (retired), Federal Reserve Bank of Dallas, 2200 N. Pearl Street, Dallas, TX 75201. E-mail: evan.f.koenig@dal.frb.org.

vast. Nevertheless, there is no consensus on the relative usefulness of different inflation-forecasting methodologies. This lack of agreement can be partly attributed to instability in U.S. inflation dynamics over time. Notably, Stock and Watson (2007) have shown that inflation in the United States became much less variable during the post-1983 “Great Moderation” period, and that changes in inflation simultaneously became much harder to predict: Mean-squared forecast errors from a variety of standard inflation models shrink during the Great Moderation, but it is very difficult to improve on the forecasts generated by a parsimonious autoregressive or random-walk model.¹

One strand of the inflation-forecasting literature attributes the apparent instability of the inflation process to changes in the behavior of long-run trend inflation, which is often linked to or assumed to be reflected in long-horizon inflation expectations. Studies that include time-series models of trend inflation include, for example, Kozicki and Tinsley (2001, 2005); Cogley, Primiceri, and Sargent (2010); Stock and Watson (2010); and Mertens (2016). In contrast, Koenig and Atkinson (2012), Faust and Wright (2013), and Ball and Mazumder (2019) equate inflation’s trend to the long-horizon expectations of professional forecasters. Clark and Doh (2014) assess alternative measures of trend inflation on the basis of their helpfulness in forecasting, and find that survey expectations do about equally as well as time-series models.

Two other prominent issues in the inflation-forecasting literature are the usefulness of stripping transitory variation from headline inflation and the usefulness of controlling for economic slack. The intuitive argument for using lagged values of a “core” inflation measure on the right-hand side of the inflation-forecasting equation, rather than lagged headline inflation, is that inflation components with mostly high-frequency variation are unlikely to have predictive power at the horizons of interest to policymakers: Core inflation ideally removes from headline inflation those components that have no signal for future headline inflation and leaves in those components

¹Similarly, Rossi and Sekhposyan (2010) conclude that “the predictive power of the Phillips curve disappeared at around the time of the Great Moderation.”

that do have signaling power.² However, the existing literature suggests that there is, in practice, little or no gain from using any of a variety of core inflation measures to forecast headline inflation (Crone et al. 2013). There is also a theoretical argument for using a stripped-down inflation measure on the *left*-hand side of the forecasting equation during estimation, even if it is headline inflation that is of ultimate interest: Removing unforecastable noise from the left-hand-side variable improves the precision of coefficient estimates.³ In this regard, Koenig and Atkinson’s (2012) finding that forecasts of trimmed-mean personal consumption expenditure (PCE) inflation also perform well as forecasts of headline PCE inflation is encouraging.⁴

The connection between slack and inflation—which has been studied going back at least to Phillips (1958)—remains controversial. In a widely cited paper, Atkeson and Ohanian (2001) show that slack adds nothing to the forecasting power of a simple random-walk inflation model. Similarly, Ang, Bekaert, and Wei (2007) find that a variety of real-activity and slack measures have no predictive power for headline and ex-food-and-energy inflation once the influence of lagged inflation has been taken into account.⁵ Researchers

²Our concern, here, is whether certain measures of “core” inflation are useful when forecasting headline inflation. However, some core inflation concepts may be useful for other purposes. For example, on theoretical grounds, a central bank might want to stabilize sticky-price inflation rather than headline inflation. Then, a core inflation measure that excludes flexible-price goods and services would be of independent interest. Similarly, in some models monetary policy should react differently to supply-side shocks than to demand-side shocks. So, a core inflation measure that strips out the effects of supply shocks could be useful. Relatedly, Stock and Watson (2019) construct a core inflation measure that weights PCE subcomponents according to their sensitivity to cyclical variation in real activity. For an early discussion of alternative core-inflation concepts, see Bryan and Cecchetti (1993).

³Based on this reasoning, Koenig, Dolmas, and Piger (2003) demonstrate that it is desirable to put first-release data on the left-hand side of forecasting equations when data revisions are mostly “news” rather than “noise.” Also see Faust and Wright (2013).

⁴Dolmas (2005) describes the procedures used to calculate the trimmed-mean PCE inflation measure used by Koenig and Atkinson. Data are available on the Federal Reserve Bank of Dallas website.

⁵Rossi and Sekhposyan (2010) find a handful of slack and real-activity variables that are useful for predicting 12-month CPI inflation—but the usefulness of those variables evaporates after the early 1980s.

who focus on *medium*-frequency inflation movements, though, have had some success in identifying a Phillips-curve relationship. Examples of such studies include Stock and Watson (2010), which looks at the impact of slack on ex-food-and-energy PCE inflation during and immediately following recessions; Koenig and Atkinson (2012), which examines the link between the unemployment rate and deviations of trimmed-mean PCE inflation from long-run inflation expectations; and Ball and Mazumder (2019), which is similar to Koenig and Atkinson (2012) but uses median CPI inflation in place of trimmed-mean PCE inflation and short-term unemployment in place of headline unemployment.

Our analysis explores whether inflation is usefully separated into trend, cyclical, and noise components, where those components are inferred from multiple published inflation series using an unobserved-common-components (UC) model. Observable inputs include real-time headline PCE inflation, ex-food-and-energy (“core”) consumer price index (CPI) inflation, and survey measures of long-horizon inflation expectations. We allow for a break in the dynamics of trend inflation in the late 1990s.⁶ When slack is introduced into the analysis, it too is treated as a variable to be inferred. That inference is informed by the difference between the current unemployment rate and a survey measure of expected average unemployment 7 to 11 years in the future.⁷

By doing our UC inference over an expanding sample, at each point using only then-available data, we are able to construct real-time estimates of trend and cyclical PCE inflation from 1992 onward. Ex-food-and-energy and Federal Reserve Bank of Dallas trimmed-mean PCE inflation data, in contrast, are available in real time only back to 1996 and 2005, respectively. Partly for that reason, most of

⁶See Ball and Mazumder (2011, 2019), Koenig and Atkinson (2012), and Nalewaik (2015) for evidence of a 1990s shift to an anchored inflation process. As a robustness exercise, we also consider a simple alternative specification in which trend inflation is obtained by applying a one-sided Hodrick-Prescott (HP) filter to long-run inflation expectations as captured by survey data.

⁷Our approach is similar to that taken by Chan, Clark, and Koop (2015), and is in the spirit of Basistha and Nelson (2007) and Basistha and Startz (2008), who show that multivariate unobserved-component models provide precise and economically meaningful estimates of the output gap and natural rate of unemployment.

the existing inflation-forecasting literature has used latest-vintage data.⁸ Unlike those studies, throughout our analysis we restrict ourselves to data available when the inflation forecast would have been prepared.

Key findings are as follows: First, we confirm the Atkeson and Ohanian (2001) and Ang, Bekaert, and Wei (2007) result that adding slack to autoregressive models of headline and ex-food-and-energy inflation fails to improve forecasting performance, and the Crone et al. (2013) result that using ex-food-and-energy inflation to predict headline inflation produces little improvement in forecasting performance. Second, nevertheless, headline-inflation forecasts from a UC model of inflation and slack substantially and significantly outperform forecasts from autoregressive models of headline or ex-food-and-energy inflation.

We achieve a further improvement in forecasting performance by incorporating a simple model of PCE inflation revisions into our UC model. Mixing first-release and lightly revised data with data that have undergone multiple rounds of revision is inappropriate and can substantially impair forecasting performance (Koenig, Dolmas, and Piger 2003; Kishor and Koenig 2012, 2014). In the current context we demonstrate that explicit modeling of revisions to headline inflation yields real-time inflation forecasts that are significantly more accurate, at most forecast horizons, than those produced when end-of-sample-vintage data are taken at face value.

In principle, an advantage of UC analysis is that it allows one to bring a variety of indicators of inflation, slack, and inflation's trend to bear when identifying inflation's components and their relation to slack. In practice, UC analysis quickly becomes computationally challenging as the number of variables expands, absent strong a priori assumptions on the variance/covariance matrix of disturbance terms. Assumptions must be made, too, about the dynamics of unobserved state variables, and how observed and unobserved variables are related to each other. The particular choices and assumptions

⁸Koenig and Atkinson (2012) is a notable exception. However, the limited availability of real-time trimmed-mean PCE inflation data casts some doubt on the robustness of that study's findings. Stock and Watson's (2010) analysis is "pseudo real-time," which means that they conduct a recursive forecasting exercise using latest-vintage data.

that we make in our baseline specification are, we feel, reasonable, and the resultant model is significantly more successful in real-time forecasting than benchmark alternatives. Moreover, we are able to demonstrate that including labor market slack in our model contributes to its forecasting success. That said, there certainly might be equally reasonable modeling choices that would produce forecasts as good as, or better than, those we obtain. Recognizing that possibility, we undertake a sensitivity analysis that alters several of our modeling assumptions, one at a time. We find that our main results are robust to a different treatment of data revisions, a more-flexible model of inflation's long-run trend, and an alternative indicator of labor market slack.

A particular concern is that our UC model may produce forecasts that, while better *on average* than those of an autoregression, fail to *consistently* outperform and which may even underperform over some subperiods (Giacomini and Rossi 2010; Rossi 2019). We examine those stability issues, and while our short forecast period (65 quarters) limits our ability to reach strong conclusions, forecast-comparison results appear to be consistent over subperiods. Notably, the performance advantage of our UC model is most strongly evident in subperiods that include the recovery from the 2008–09 Great Recession—a period during which Federal Reserve officials have claimed labor market slack became less useful as a real-time predictor of inflation and, to that extent, less relevant to the formulation of policy.⁹ Our results cast doubt on the validity of that argument.

To shed light on *why* we find a role for slack in inflation forecasting when many other analysts have not, we undertake two exercises that consider non-marginal changes to our forecasting approach. In the first exercise, we detrend headline inflation exactly as in our baseline UC analysis, but make no effort to strip inflation of high-frequency variation. Labor market slack then proves to be unhelpful in real-time forecasting. In the second exercise, we filter out inflation's high-frequency variation exactly as in our baseline analysis,

⁹Clarida (2020), for example, observes that over the past decade “price inflation seems less responsive to resource slack, and also, . . . estimates of resource slack based on historically estimated price Phillips curve relationships are less reliable and subject to more material revision than was once commonly believed.” He uses that observation as motivation for a monetary policy reaction function that ignores real-time measures of labor market tightness.

but make no effort to control for inflation's long-run trend. Again, labor market slack proves to be unhelpful in forecasting. In combination with our baseline results, these exercises strongly suggest that the key to finding a role for slack in real-time inflation forecasting is to strip inflation of high-frequency noise, identify its long-run trend, and then separately forecast trend and cycle, using slack to help predict inflation's cyclical component. That result is intuitive and consistent with studies by Stock and Watson (2010), Koenig and Atkinson (2012), and Ball and Mazumder (2019).

Our work is independent of, but broadly similar to, a study by Hasenzagl et al. (2018). Like us, Hasenzagl et al. decompose headline inflation into trend, cycle, and noise using an unobserved-common-components model. However, they eschew the information about inflation and unemployment trends that is available in surveys of professional forecasters' long-horizon expectations. Their model includes, instead, survey measures of households' and professional forecasters' *near*-term inflation expectations.¹⁰ To identify plausible cycles and trends, they find that they need to use Bayesian estimation methods and to impose strong orthogonality assumptions. The data-revisions issue is ignored.

The remainder of the paper is organized as follows: Section 2 presents our real-time, unobserved-common-components model of inflation and describes our data set; Section 3 presents empirical results, comparing the forecasting performance of our baseline UC model to the performance of alternative models; Section 4 takes a closer look at *why* our model performs so well; and Section 5 concludes.

2. An Unobserved-Common-Components Model of Inflation

Because Federal Reserve policymakers have defined price stability in terms of headline PCE inflation and include that measure in their quarterly forecasting exercises (as described in the Federal Reserve's Summary of Economic Projections), our primary focus

¹⁰Oil prices are also included in the analysis. Oil-price movements help to explain why the one-year inflation expectations of households and professional forecasters diverge.

is on forecasting headline PCE inflation. However, we hypothesize that there is information useful for forecasting headline PCE inflation in several series besides headline PCE inflation itself, including ex-food-and-energy CPI inflation and long-run inflation expectations from the quarterly Survey of Professional Forecasters (SPF) and the semi-annual Blue Chip survey of long-range forecasts. Specifically, we assume that these series share a common trend, and that the two realized short-term inflation measures share a common cyclical component. We extract trend and cycle using state-space methods.¹¹ This approach has several advantages. First, using a multivariate model improves the precision of our estimates of PCE inflation's trend and cycle.¹² Second, we can tailor our definition of core PCE inflation (headline PCE inflation, stripped of unforecastable noise) to our purpose (forecasting headline inflation).¹³ Finally, our approach overcomes constraints on the availability of real-time data that apply to off-the-shelf measures of core PCE inflation like trimmed-mean or ex-food-and-energy PCE inflation.

Our baseline model has five observation equations and five transition equations. The observed variables are first-release headline PCE inflation (π_t^{FR}); ex-food-and-energy "core" CPI inflation (π_t^{CPI}); the average of nine-year, one-year-forward SPF CPI inflation expectations and five-year, six-year-forward Blue Chip CPI inflation expectations (π_t^{LR}); the gap between the current unemployment rate (U_t) and the Blue Chip five-year, six-year-forward expected unemployment rate (U_t^*); and the revision to last-quarter's headline PCE inflation rate ($\varepsilon_{t-1,t}$). The observation equations for those five variables are as follows:

¹¹Related efforts include Stock and Watson (2016), which uses disaggregated data on sectoral inflation to estimate trend PCE inflation, and Kiley (2008), which uses a bivariate common-trend framework to examine the role of food and energy prices in the dynamics of trend inflation for both PCE and CPI inflation. Another effort to infer trend inflation using multivariate models is Mertens (2016). Notably, Mertens finds that trimmed-mean PCE inflation, because it so effectively filters out short-term noise, is especially useful. Trimmed-mean PCE inflation is unsatisfactory for our purposes, however, because real-time estimates only become available in 2005.

¹²The multivariate approach has been shown to be useful in a state-space setting by Clark (1989), Basistha and Nelson (2007), and Basistha and Startz (2008).

¹³Cf. footnote 2, above.

$$\pi_t^{FR} = \mu^{PCE} + \tau_t + c_t - \varepsilon_{t,t+1} + \eta_t^{PCE} \tag{1}$$

$$\pi_t^{CPI} = \mu^{CPI} + \tau_t + c_t + \eta_t^{CPI} \tag{2}$$

$$\pi_t^{LR} = \tau_t + c_t^{LR} + \eta_t^{LR} \tag{3}$$

$$U_t - U_t^* = S_t + \eta_t^U \tag{4}$$

$$\varepsilon_{t-1,t} = \varepsilon_{t-1,t}. \tag{5}$$

Headline PCE and core CPI inflation share a common trend (τ_t) with long-run inflation expectations. They also share a cyclical component (c_t). Long-forward inflation expectations differ from trend inflation by two error terms, one of which (c_t^{LR}) is allowed to be serially correlated. The unemployment gap is a noisy measure of labor market slack (S_t). The revision to last-quarter’s headline inflation rate ($\varepsilon_{t-1,t} \equiv \pi_{t-1,t}^{SR} - \pi_{t-1}^{FR}$, where $\pi_{t-1,t}^{SR}$ is the quarter- t official estimate of the prior quarter’s PCE inflation rate) is observed perfectly. All other equations include noise terms (the η_t^i terms) that are joint-normally distributed with mean zero.¹⁴ There are no restrictions on the contemporaneous variance/covariance matrix of those terms. However, we assume that $cov(\eta_t^i, \eta_t^j) = 0$ for all i, j when $s \neq t$.

Figures 1 and 2 plot the observable variables that feed into our UC analysis: π_t^{FR} , π_t^{CPI} , π_t^{LR} , and $\varepsilon_{t-1,t}$ are shown in Figure 1, while Figure 2 plots U_t and U_t^* .

The five transition equations included in the baseline model are as follows:

$$\tau_t = \mu_1^\tau + \tau_{t-1} + u_{1,t}^\tau \text{ for } t \leq 1997:Q4 \text{ and}$$

$$\tau_t = \mu_2^\tau + \theta_2 \tau_{t-1} + u_{2,t}^\tau \text{ for } t > 1997:Q4, \text{ with } |\theta_2| < 1 \tag{6}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \sum_{i=1}^4 \delta_i S_{t-i} + u_t^c \tag{7}$$

$$S_t = \mu^S + \beta_1 S_{t-1} + \beta_2 S_{t-2} + u_t^S \tag{8}$$

$$c_t^{LR} = \gamma c_{t-1}^{LR} + u_t^{LR} \tag{9}$$

$$\varepsilon_{t,t+1} = \kappa \varepsilon_{t-1,t} + u_t^\varepsilon. \tag{10}$$

¹⁴See Morley, Nelson, and Zivot (2003) for a discussion of unobserved-components models with correlated error terms.

Figure 1. Inflation Measures and Revisions to Headline PCE Inflation

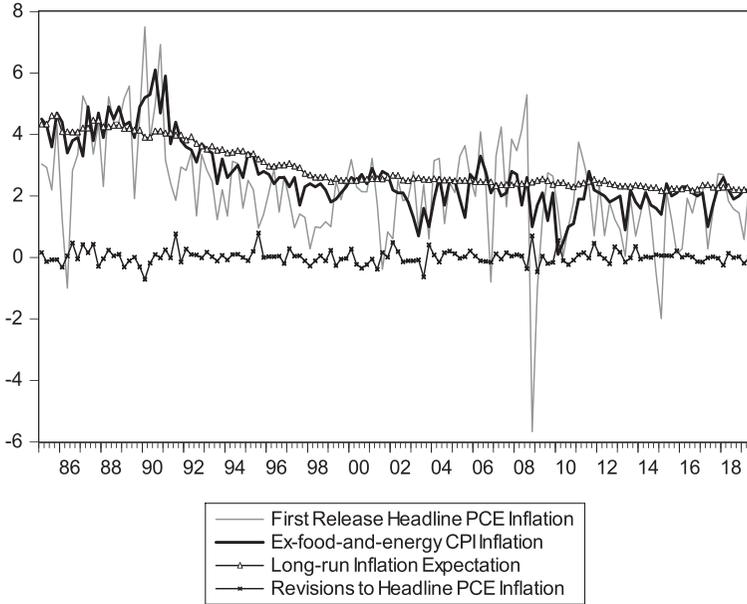
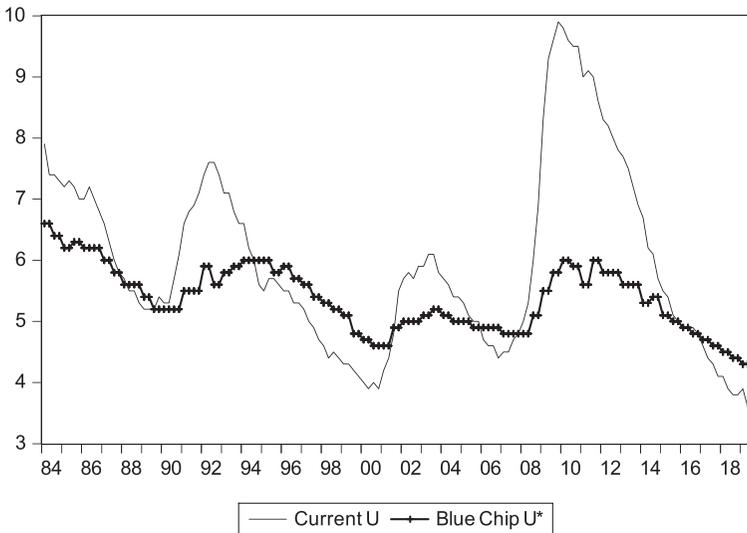


Figure 2. Current and Blue Chip Long-Run Unemployment Rates



An auxiliary equation defines core PCE inflation as (post-revision) headline PCE inflation stripped of noise:

$$\pi_t^{CORE} \equiv \mu^{PCE} + \tau_t + c_t. \quad (11)$$

This UC-filtered core inflation measure plays no active role in our analysis, but for assessing inflation pressures provides a potentially useful alternative to conventional core inflation measures such as ex-food-and-energy and the Federal Reserve Bank of Dallas trimmed-mean PCE inflation.

Equation (6) says that trend inflation follows a random walk with drift prior to 1998, and then follows a stationary process.¹⁵ Inflation's cyclical component depends on its own lagged values and also on lagged labor market slack, which itself follows an AR(2) process. As previously noted, long-run inflation expectations differ from trend inflation by a noise term, η_t^{LR} , but also by a term, c_t^{LR} , that follows an AR(1) process. Finally, revisions to headline inflation are allowed to be serially correlated.

That there was a break in the dynamics of trend inflation in the second half of the 1990s is acknowledged by recent Federal Reserve Chairs (Bernanke 2003; Yellen 2016). Econometric support comes from Koenig and Atkinson (2012) and Ball and Mazumder (2019), who perform Quandt-Andrews stability tests on a regression of long-horizon expected inflation on its own lag and measures of lagged realized core inflation.¹⁶ A break is also broadly consistent with recent empirical work by Nalewaik (2015), who finds that U.S. inflation

¹⁵An alternative approach—used by Stock and Watson (2007), for example—is to assume that trend inflation follows a random walk over the entire sample, but that the variance of trend-inflation innovations changes over time. Under that approach, current trend inflation is also (up to a constant) the far-forward expectation of PCE inflation, ex-food-and-energy CPI inflation, and professional forecasters' survey responses. Our estimate of θ_2 in Equation (6) is just 0.080, however, with standard error 0.008. In our model, then, trend inflation is a common factor, identified by covariance properties, but no longer a measure of far-forward expectations as in the random-walk case.

¹⁶Similarly, a Bai and Perron (1998) break test on the first difference of the HP trend of SPF long-term inflation expectations shows a break in 1999:Q1. The break date assumed in our exercise (1997:Q4) falls within the 95 percent confidence band of the break date estimated using the Bai and Perron (1998) approach. Our results don't change if we allow the break to take place one year earlier or one year later.

entered a stable-mean, low-variance regime in the 1990s, and the findings of Ball and Mazumder (2011), who report that near-term inflation expectations became better anchored over the course of the 1990s.

We assume that $u_{1,t}^\tau \sim iidN(0, \sigma_{u_1^\tau}^2)$ and $u_{2,t}^\tau \sim iidN(0, \sigma_{u_2^\tau}^2)$. The other transition-equation error terms are joint-normally distributed with mean zero and with no restrictions on their contemporaneous covariances. However, we require that $cov(u_t^i, u_s^j) = 0$ for all i, j when $s \neq t$.

The above set of observation and transition equations constitute our “baseline” or “BL” model. The full model can be put into state-space form and estimated using maximum likelihood via the Kalman filter.¹⁷

2.1 Four Variants

We think that our baseline specification is reasonable. As we will shortly see, it “works” in the sense that it produces reasonable estimates of trend inflation, core inflation, and slack, and produces more-accurate forecasts of headline inflation than do benchmark alternatives. Those results certainly don’t rule out the possibility that there are equally appealing specifications that work equally well—or better. To identify the contributions to real-time forecast performance from some of our modeling choices, and to offer some assurance that our findings concerning the role of slack in inflation forecasting are robust, we consider several variants of the baseline model. For example, to investigate the importance of modeling inflation revisions, we estimate a state-space model that drops Equations (5) and (10), and replaces Equation (1) with

$$\pi_t^{PCE} = \mu^{PCE} + \tau_t + c_t + \eta_t^{PCE}, \quad (12)$$

where π_t^{PCE} is, successively, the vintage- T , vintage- $(T+1)$, vintage- $(T+2)$, ... official estimate of quarter- t headline PCE inflation, and $T, T+1, \dots$ are the end dates of the different sample periods over which the model is estimated. Thus, we mimic how most analysts

¹⁷For the details on the estimation procedure, see chapter 2 of Kim and Nelson (2000).

do real-time estimation, using at each date the latest-vintage data then available. In tables, below, we label this variant the “BL-KK” model, short for “baseline model without the Kishor-Koenig model of data revisions.”

Similarly, to investigate whether our results are sensitive to the way that we model labor market slack, we consider a modified baseline model in which the unemployment gap, $U - U^*$, on the left-hand side of Equation (4) is replaced by the Stock and Watson (2010) “unemployment-recession gap,” which is the difference between the current unemployment rate and the minimum unemployment rate over the current and previous 11 quarters. We label this variant of the baseline model the “URG” model—it is identical to the BL model, except it uses a different observable slack indicator.

The baseline specification assumes a late-1990s break in the dynamics of trend inflation that is recognized by the time that our forecasting exercise begins. We argue that that is a reasonable approach, but as a robustness check we estimate a version of the baseline model in which trend inflation is obtained by applying a one-sided Hodrick-Prescott (HP) filter to long-forward inflation expectations. When forecasting, future trend inflation is set equal to the HP trend’s most-recent value. This variant of the baseline model is labeled “HP.” It is the baseline model with a flexible HP trend and no a priori break in trend dynamics.

Finally, to investigate the importance of labor market slack for successful inflation forecasting, we estimate a special case of each model that drops Equations (4) and (8) and sets $\delta_1 = \delta_2 = \delta_3 = \delta_4 = 0$ in Equation (7). We label these the “BL-S,” “BL-KK-S,” “URG-S,” and “HP-S” models, respectively. (Note, however, that the BL-S and URG-S models are identical.)

2.2 *The Data*

Our data run from 1984:Q1 through 2019:Q2. The start date is determined by the availability of survey measures of long-run inflation expectations. The sample is dominated by the Great Moderation period that has been found to be particularly challenging for inflation forecasters (Stock and Watson 2010).

Inflation data are expressed as annualized percent changes. Unemployment data are quarterly averages. Revisions to CPI inflation and unemployment data are relatively small and due solely to reestimation of seasonal factors. We ignore them. Real-time PCE inflation data are taken from the Federal Reserve Bank of Philadelphia's website (Croushore and Stark 2001). Our baseline model uses first-release headline PCE inflation and the third-release data available one quarter later.¹⁸ The BL-KK variant of the baseline model is estimated using end-of-sample-vintage PCE inflation—i.e., the latest data that would have been available to a forecaster in real time.

To help identify inflation's longer-run trend, we take the average of two measures of long-term inflation expectations. The most straightforward of these comes from a Blue Chip survey of long-range forecasts that is published twice each year, in early March and early October: It is the CPI inflation rate that respondents expect 7 to 11 years out. We use March survey results for the first and second quarters of each year, and October survey results for the third and fourth quarters.¹⁹ Our second measure of long-term inflation expectations is calculated from the Survey of Professional Forecasters (SPF) 10-year and 1-year CPI inflation expectations. Specifically, it is defined as $(10^*cpi10 - cpi1)/9$, where $cpi10$ and $cpi1$ are 10-year and 1-year median expected inflation rates, respectively—a formula that captures the expectation for inflation 2 to 10 years out implicit in SPF respondents' 10-year and 1-year inflation forecasts.²⁰ Importantly, both of our measures of long-term inflation expectations are *forward* rates—rates that exclude forecasters' expectations for the coming year. We don't want our own inflation forecasts, which extend out as far as four quarters, to “piggyback” on the near-term forecasts of professionals.

Trend unemployment, U_t^* , is a long-forward expected measure from the semi-annual Blue Chip survey of long-range forecasts. Specifically, it is the average unemployment rate expected over the

¹⁸Croushore (2008) discusses the timing and source of revisions to PCE inflation.

¹⁹The timing is conservative. First-quarter PCE inflation releases are unavailable until late April, and third-quarter PCE inflation releases are unavailable until late October.

²⁰SPF 10-year CPI inflation expectations are first available in 1991:Q4. Before then, we substitute Blue Chip 10-year, 1-year-forward CPI expectations.

Table 1. Baseline Model Parameters

Parameter	Estimate	S.E.
μ^{PCE}	-0.705	0.133
μ^{CPI}	-0.212	0.068
μ_1^T	-0.046	0.017
μ_2^T	2.210	0.007
μ^S	-0.006	0.022
$\Sigma\phi_i$	0.557	0.212
$\Sigma\beta_i$	0.958	0.328
$\Sigma\delta_i$	-0.061	0.028
θ_2	0.080	0.008
γ	0.031	0.033
κ	0.125	0.168
$\sigma_{\eta PCE}$	1.375	0.069
$\sigma_{\eta CPI}$	0.188	0.429
$\sigma_{\eta LR}$	0.012	6.210
$\sigma_{\eta U}$	0.082	1.620
$\sigma_{\mu_1^T}$	0.123	0.243
$\sigma_{\mu_2^T}$	0.064	0.697
σ_{u^c}	0.416	0.112
$\sigma_{u^{LR}}$	0.077	0.474
σ_{u^S}	0.134	0.353
σ_{u^ϵ}	0.243	0.098

five years beginning six years from the survey date. We use March survey results for the first and second quarters of each year, and October survey results for the third and fourth quarters.

2.3 Coefficient Estimates

Table 1 displays full-sample estimates of the baseline model's coefficients, and Table 2 shows the correlation matrices of the observation-equation and transition-equation error terms. Several features of those results are notable.

Trend inflation is more nearly a constant plus noise in the late sample period than a random walk: The coefficient, θ_2 , attached to lagged trend inflation in Equation (6) is estimated at just 0.080 (S.E. = 0.008). With the constant term, μ_2^T , estimated to

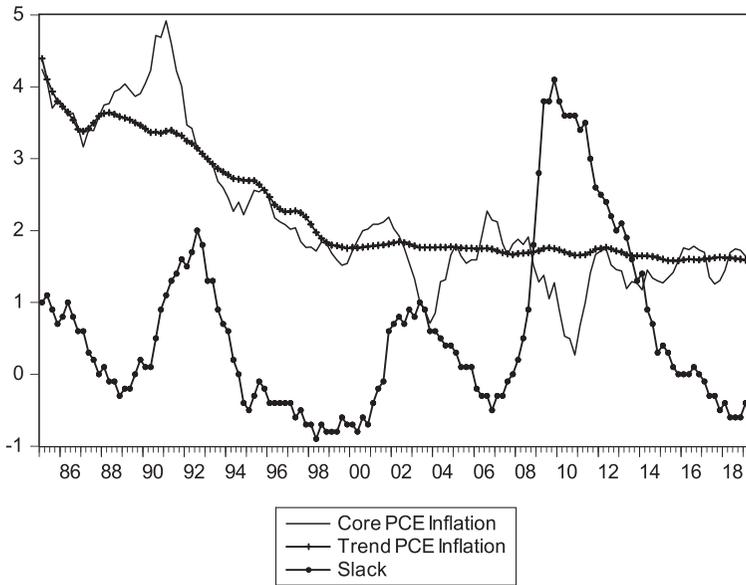
Table 2. Correlation Matrices

Observation-Equation Errors				
	η^{PCE}	η^{CPI}	η^{LR}	η^U
η^{PCE}	1			
η^{CPI}	0.231 (0.052)	1		
η^{LR}	0.719 (0.012)	0.818 (0.026)	1	
η^U	-0.182 (0.020)	-0.998 (0.021)	-0.790 (0.057)	1
Transition-Equation Errors				
	u^c	u^S	u^{LR}	u^ε
u^c	1			
u^S	0.397 (0.029)	1		
u^{LR}	-0.351 (0.019)	0.028 (0.061)	1	
u^ε	-0.189 (0.033)	0.178 (0.029)	0.044 (0.059)	1

equal 2.210 (S.E. = 0.007), trend inflation would converge to $2.210/(1 - 0.080) = 2.40$ percent in the absence of new shocks. The estimates of μ^{PCE} and μ^{CPI} in Equations (1) and (2) then imply that headline PCE inflation and ex-food-and-energy CPI inflation will converge to 1.70 percent and 2.19 percent, respectively. The former figure is 30 basis points below the Federal Open Market Committee's (FOMC's) announced 2.0 percent long-run PCE inflation target, suggesting that that target may not be fully credible. With the estimated value of γ very close to zero in Equation (9) (0.031, S.E. = 0.033), long-forward inflation expectations (Equation (3)) are essentially trend inflation plus noise. There is only weak evidence of serial correlation in revisions to headline PCE inflation ($\kappa = 0.125$ in Equation (10), S.E. = 0.168).

Inflation is sensitive to slack ($\sum_{i=1}^4 \delta_i = -0.061$ in Equation (7), S.E. = 0.028), but the long-run slope of the Phillips curve is modest, at $-0.061/(1 - 0.557) = -0.138$ (S.E. = 0.090). Slack is close to a random walk, with $\beta_1 + \beta_2$ in Equation (8) estimated at 0.958 (S.E. = 0.328).

With three exceptions ($\sigma_{\eta^{PCE}}, \sigma_{u^c}, \sigma_{u^\varepsilon}$), the standard deviations of the observation-equation and transition-equation error terms are

Figure 3. Core Inflation, Trend Inflation, and Slack

estimated imprecisely. The variation in headline PCE inflation unexplained by movements in inflation's trend and cycle is enormous ($\sigma_{\eta_{PCE}} = 1.375$, S.E. = 0.069): There is a great deal of transitory variation in quarterly headline inflation.

There are strong correlations between several of the error terms, especially in the observation equations. Shocks to headline PCE inflation and to ex-food-and-energy CPI inflation are strongly positively correlated with shocks to survey long-forward inflation expectations. Shocks to the unemployment gap are strongly negatively correlated with shocks to ex-food-and-energy CPI inflation and long-forward inflation expectations.²¹

Figure 3 shows full-sample estimates of four-quarter UC-filtered core inflation (π_t^{CORE}) plotted along with trend PCE inflation

²¹We estimated three restricted versions of the baseline model: one with orthogonal transition-equation shocks, one with orthogonal observation-equation shocks, and one with orthogonal shocks in both the transition and observation equations. The likelihoods of these restricted models were -157.34 , -169.70 , and -190.85 , respectively, as compared with -120.91 for the baseline model. Likelihood-ratio tests strongly reject each set of restrictions.

$(\mu^{PCE} + \tau_t)$ and slack (S_t). The difference between the first two series is cyclical inflation (c_t). The negative relationship between slack and cyclical inflation is evident: Core inflation tends to run below trend inflation following recessions, when slack is elevated, and to run above trend inflation late in expansions, when slack is low.

3. Real-Time Estimation and Forecasting

By estimating the baseline model using data running from $t = 1$ to $t = T$, we obtain vintage- T estimates of all of the coefficients appearing in the observation and transition equations, along with vintage- T estimates of the state variables τ_t , c_t , S_t , and c_t^{LR} , for $t = 1, 2, \dots, T$, and of $\varepsilon_{T,T+1}$. With those coefficient and state-variable estimates in hand, the transition equations are used to project each state variable into the future. Those forecasts are, in turn, substituted into Equations (1) and (5), producing forecasts of how headline PCE inflation will be reported after revision. (The observation equations also produce forecasts of first-release PCE inflation, core CPI inflation, long-forward inflation expectations, and the unemployment gap.) The sample period is then extended by one quarter, to $t = 1, 2, \dots, T + 1$, and the whole process is repeated. In our forecast-comparison exercises we start with $T = 2002:Q1$, forecasting PCE inflation in the five quarters from 2002:Q1 through 2003:Q1. We conclude with $T = 2018:Q2$, forecasting inflation from 2018:Q2 through 2019:Q2.

By construction, the baseline model forecasts headline PCE inflation as it will appear one quarter after initial release (i.e., after two official rounds of revision). Because subsequent revisions are difficult to predict, a forecast of PCE inflation as it appears then also performs well as a forecast of later-vintage PCE inflation.²² Indeed, by modeling the revision process only up to the point where the

²²We test the predictability of revisions to early PCE inflation releases by regressing the revisions on a constant using the generalized method of moments. Instruments include inflation and slack data that would have been available when the inflation data were published. “No predictability” is rejected if the regression’s J statistic shows statistically significant correlation between the regression errors and the instruments. Using that methodology, revisions to first-release inflation estimates are predictable, while revisions to inflation estimates published one quarter after the first releases are not.

real-time official data become efficient estimates of “truth,” we strip unforecastable noise from the system, improving the accuracy with which the model’s coefficients are estimated (Koenig, Dolmas, and Piger 2003). When evaluating the forecast performance of different models, we compare the models’ forecasts with the latest headline PCE inflation data not to have undergone a comprehensive revision.²³

An issue when estimating our model is how to handle the 1997:Q4 break in the dynamics of inflation’s trend component. Our baseline assumption is that by the time our real-time forecasting exercise begins, in 2002, an analyst would have recognized the existence of a break. That assumption is broadly consistent with the policy discussion of the time. In January 2003, for example, Federal Reserve Governor Kohn cited a late-1990s shift in the behavior of long-term interest rates as evidence that “the U.S. economy has enjoyed most of the benefits ascribed to inflation targeting in terms of anchoring inflation expectations as well as inflation itself” (Kohn 2003). The following month, then-Federal Reserve Governor Bernanke credited the FOMC’s “constrained discretion” policy strategy with “providing an anchor for inflation expectations,” as evidenced by a reduced sensitivity of core inflation to commodity-price and exchange rate fluctuations, and of inflation, inflation expectations, and long-term interest rates to large changes in short-term interest rates (Bernanke 2003). In September 2003, a European Central Bank analysis of long-term inflation expectations and their relationship with short-term expectations concluded that “with regard to the two countries in our review where no numerical value for the inflation objective was announced, the United States and Japan, inflation expectations appear to be well anchored in the former but not in the latter” (Castelnuovo, Nicoletti-Altamari, and Rodriguez-Palenzuela 2003). As a check, we repeated the Koenig-Atkinson (2012) analysis of the relationship between long-term inflation expectations, lagged expectations, and recent inflation realizations over a sample

²³We take this approach rather than compare with current-vintage headline inflation because comprehensive revisions to the national income and product accounts often include definitional changes.

ending in 2001:Q4.²⁴ Consistent with the Koenig-Atkinson results, the Quandt-Andrews test rejects stability at the 1 percent significance level and identifies a mid-1990s break in the dynamics of inflation expectations, documenting a shift from unanchored to anchored expectations.²⁵ These results are not definitive—they apply to long-term inflation expectations rather than to trend inflation as defined here—but suggest that our baseline assumption is reasonable. To be on the safe side, one of our robustness exercises takes an agnostic approach to trend inflation.

Figure 4 compares our real-time UC estimates of core PCE inflation (π_t^{CORE}) with first-release ex-food-and-energy PCE inflation (both measured over four-quarter intervals), and our real-time estimates of labor market slack with real-time Congressional Budget Office (CBO) estimates of the unemployment gap. The two inflation series generally move similarly, but ex-food-and-energy inflation shows more volatility than UC core during the expansion that followed the Great Recession, and runs above UC core inflation during and in the immediate lead-up to that recession. The two slack series generally move together, too, but the CBO series is perhaps more cyclical: It tends to show more slack during and after recessions (especially the 2008–09 Great Recession), and somewhat less slack in the late-expansion years of 2005–06 and 2018–19.

Finally, Figure 5 documents the stability of the relationship between cyclical inflation and slack by showing real-time estimates of the sum of lagged slack coefficient estimates, $\sum_{i=1}^4 \delta_i$, in Equation (7). Note, in particular, that there was no marked weakening of the relationship during the expansion that followed the Great Recession.

²⁴For this exercise we follow Ball and Mazumder (2019) and use weighted-median CPI inflation as our measure of realized inflation, rather than trimmed-mean PCE inflation. The trimmed-mean data are subject to revision and available in real time only after 2004.

²⁵Specifically, the most likely break date is 1995:Q3 in this exercise, but 1997:Q4—the break date identified by Koenig and Atkinson (2012) and adopted here—lies within the 95 percent confidence interval. Before the break, the hypothesis that the coefficients on lagged expectations and lagged realizations sum to 1 cannot be rejected. After the break, the sum is significantly less than 1. Very similar results are obtained as the sample period is extended beyond 2001:Q4.

Figure 4. Alternate Measures of Core Inflation and Slack

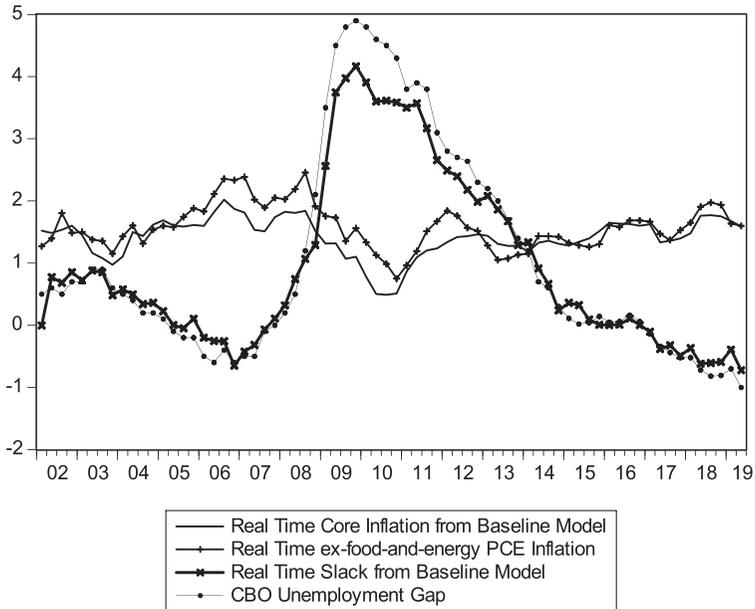
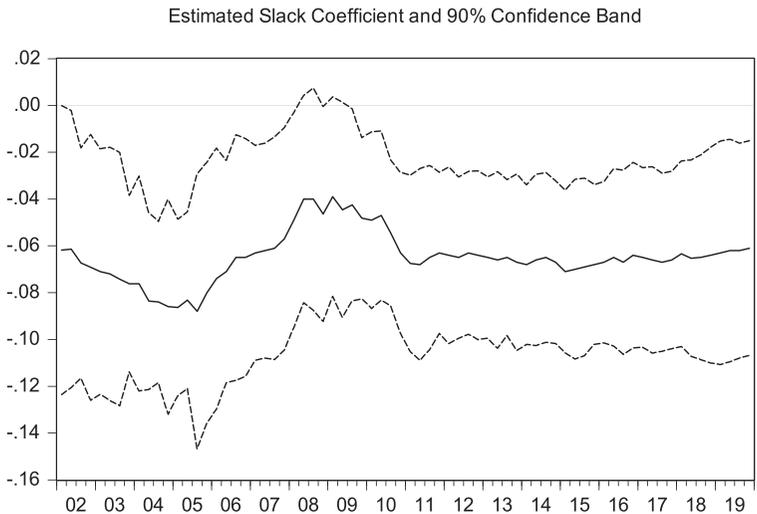


Figure 5. Recursive Estimates of the Slack Coefficient



3.1 Main Results

We compare the real-time forecasting performance of our baseline UC model with the forecasting performance of simple AR(1) models in headline PCE inflation and ex-food-and-energy PCE inflation, estimated using the latest data that would have been available in real time.²⁶ We also look at what happens to the performance of the AR(1) models when they are expanded to include labor market slack, and at what happens to the performance of the baseline model when slack is excluded from that model. Briefly, we find that the UC model performs *better* with slack than without, whereas the AR(1) models in headline and ex-food-and-energy PCE inflation perform *worse* with slack than without. The baseline UC model significantly outperforms the AR(1) models.

Table 3 shows mean-square forecast errors (MSEs) for the two AR(1) models (“PCE” and “XFE”), the AR(1) models supplemented with labor market slack as measured by the lagged gap between the unemployment rate and Blue Chip long-forward unemployment expectations (“PCE+S” and “XFE+S”), and the baseline UC model both with slack (“BL”) and without (“BL-S”). Forecast performance is compared at horizons ranging from zero to four quarters ($h = 0, 1, 2, 3, 4$) and over the current and subsequent four quarters combined ($h = 0-4$). The lowest MSE at each forecast horizon is shown in bold.

The table confirms the Atkeson and Ohanian (2001) and Ang, Bekaert, and Wei (2007) result that adding slack to simple autoregressive models of headline or ex-food-and-energy inflation does not improve forecast performance: Performance deteriorates at every

²⁶We also undertook comparisons with spliced Federal Reserve Greenbook and SPF inflation forecasts. However, the timing of the Greenbooks and SPF surveys is such that Federal Reserve staff and professional forecasters have a distinct information advantage in forecasting current-quarter inflation. Beyond the current quarter we found that the Greenbook/SPF forecasts were consistently and substantially *worse* than those generated by a simple AR(1) model in headline inflation. Our result contrasts sharply with Faust and Wright (2009), perhaps because that study examines GDP-deflator inflation rather than PCE inflation, or because the sample over which it evaluates relative forecasting performance ends in 2000, whereas ours begins in 2002. In the interests of brevity, we don't report the Greenbook/SPF forecast comparisons.

Table 3. Real-Time MSEs of Alternative PCE-Inflation Forecasting Models

Horizon	PCE	PCE+S	XFE	XFE+S	BL-S	BL
0	2.693<	2.863	2.622*	2.676>>	2.423**†	2.412**††
1	2.669<	3.082	2.571*	2.706>>	2.519*>	2.453**†
2	2.527<	2.987	2.597<	2.830>>	2.413**†>	2.378**††
3	2.558<<	3.089	2.629<	2.842>>	2.498†>	2.412**††
4	2.555<<	3.321	2.560<	2.816>>	2.529>	2.400**†
0–4	0.694<<	1.170	0.723<	0.899>>	0.608**††>>	0.545**††

Note: The first set of forecasts is for 2002:Q1–2003:Q1; the final set is for 2018:Q2–2019:Q2. “+S” indicates that the unemployment gap, $U - U^*$, is included in the model. “-S” indicates that labor market slack, S , is excluded from the model. *Outperforms the AR(1) headline inflation model (“PCE”) at the 10 percent level. **Outperforms the AR(1) headline inflation model (“PCE”) at the 5 percent level. †Outperforms the AR(1) ex-food-and-energy inflation model (“XFE”) at the 10 percent level. ††Outperforms the AR(1) ex-food-and-energy inflation model (“SFE”) at the 5 percent level. >The MSE difference between adjacent entries is significant at the 10 percent level. >>The MSE difference between adjacent entries is significant at the 5 percent level. The smallest MSE in each row is bolded.

forecast horizon when slack is added to the PCE and XFE models.²⁷ The deterioration is statistically significant at all horizons for the PCE model and at longer horizons for the XFE model. The table also confirms the result from Crone et al. (2013) that there is little payoff to forecasting headline inflation using a model of ex-food-and-energy inflation: Only at the very shortest horizons ($h = 0, 1$) does the XFE model outperform the PCE model. However, the picture changes radically when headline inflation is split into trend, cyclical, and noise components; trend and cyclical inflation are separately forecast, taking into account slack’s influence on the latter; and first-release headline inflation data aren’t taken at face value. Thus, our baseline (BL) model outperforms the PCE and XFE models at every forecast horizon. Moreover, those performance differences are

²⁷Stock and Watson (2019) report similar results for PCE inflation excluding food and energy. Theirs is a “pseudo out-of-sample forecasting exercise” that uses latest-vintage data rather than the data that would have been available in real time.

statistically significant.²⁸ Finally, although adding slack to the PCE and XFE models causes their performance to deteriorate, *cutting* slack from the baseline model worsens that model's forecast performance at every horizon—significantly so at all but the very shortest horizon ($h = 0$). (See the results in the columns headed “BL” and “BL-S”.)

In summary, filtering out noise and controlling for slack are highly useful when forecasting headline inflation, provided that one also controls for inflation's trend and gives careful attention to data revisions. We look, next, at the sensitivity of those results to modifications to the baseline model.

3.2 Robustness: Data Revisions

First, we drop the Kishor-Koenig treatment of data revisions from the baseline model and, instead, take at face value the latest inflation data that would have been available to an analyst in real time. The new forecasting results are displayed in the middle two columns of Table 4, under the headings “BL-KK-S” and “BL-KK.” To facilitate comparison, that table also includes results for the AR(1) headline inflation model (“PCE”) and for the baseline model (“BL”) that previously appeared in Table 3. We want to know whether taking latest-available data at face value harms forecasting performance, and whether it is still true that slack is helpful for forecasting inflation once inflation is decomposed into trend, cycle, and noise.

The answer to the first of those questions is “yes”: At every forecast horizon but one ($h = 1$), the MSE from the BL model is lower than that from the BL-KK model. In particular, at $h = 0$ the BL model successfully second-guesses the first-release official estimate of PCE inflation that the BL-KK model takes at face value. The performance differences are statistically significant at $h = 0, 2, 3,$ and 4 . Still, the BL-KK model consistently outperforms the simple PCE model. Those performance differences are statistically significant at every horizon except $h = 4$.

²⁸For non-nested models, we use the Diebold and Mariano (1995) and West (1996) forecast comparison test. To perform comparison of nested models, we use the Clark and West (2007) test.

Table 4. Real-Time MSEs: Effect of Dropping KK Model of Data Revisions

Horizon (h)	PCE	BL-KK-S	BL-KK	BL
0	2.693	2.503* >	2.459* >	2.412**
1	2.669	2.522* >>	2.326**	2.453**
2	2.527	2.484	2.457* >	2.378**
3	2.558	2.501	2.485* >	2.412**
4	2.555	2.500	2.499 >	2.400**
0-4	0.694	0.629* >>	0.556**	0.545**

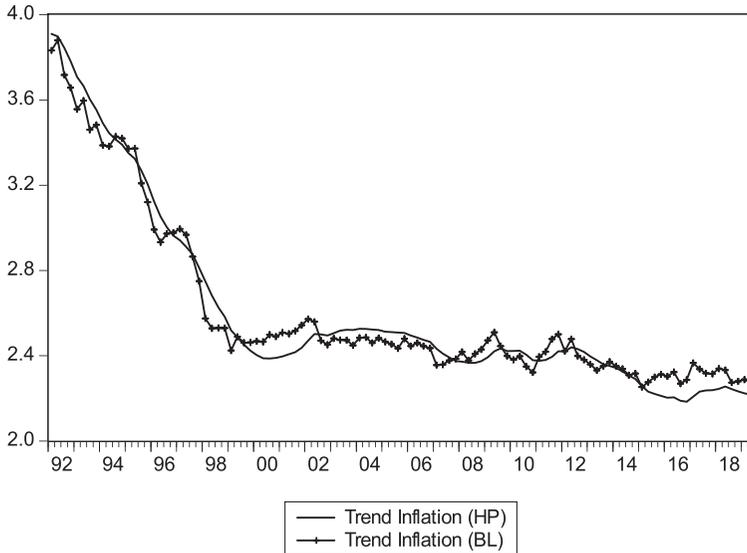
Note: The first set of forecasts is for 2002:Q1–2003:Q1; the final set is for 2018:Q2–2019:Q2. “-S” indicates that labor market slack, S, is excluded from the model. *Outperforms the AR(1) headline inflation model (“PCE”) at the 10 percent level. **Outperforms the AR(1) headline inflation model (“PCE”) at the 5 percent level. >The MSE difference between adjacent entries is significant at the 10 percent level. >>The MSE difference between adjacent entries is significant at the 5 percent level. The smallest MSE in each row is bolded.

What of slack? Eliminating slack from the BL-KK model causes forecast performance to deteriorate at every horizon, but especially at shorter horizons. (Compare the results in the columns headed “BL-KK” and “BL-KK-S.”) Performance differences are statistically significant for $h = 0, 1,$ and $0-4$.

In summary, filtering out noise and controlling for slack are highly useful when forecasting headline inflation, provided that one also controls for inflation’s trend. Explicit modeling of data revisions is helpful, too, but neglecting that step (as in the BL-KK model) still yields a real-time forecasting performance superior to that of the benchmark AR(1) model.

3.3 Robustness: Trend Inflation

Our baseline (BL) model allows for a shift in trend inflation’s dynamics at the close of 1997. By the time that we initiate real-time inflation forecasting, in 2002:Q1, the analyst is assumed to recognize that trend dynamics have changed. As a robustness exercise we consider an analyst who, instead, estimates trend inflation by applying the HP filter directly to long-forward inflation expectations, and who,

Figure 6. UC-Filtered and HP-Filtered Trend Inflation

when forecasting, assumes that in coming quarters trend inflation will hold steady at its latest observed value.

Real-time estimates of trend inflation from the BL and HP models are compared in Figure 6.²⁹ The HP-filtered trend is smoother than the UC-filtered trend, but the two series generally behave similarly. In particular, both show substantial declines during the mid-1990s and transition to comparative stability by the end of the decade. Over the past five years, however, the UC trend has run almost 10 basis points higher than the HP trend.

Mean-square forecast errors for the HP model are shown in Table 5, which has a format similar to that of Table 4: MSEs for the PCE and BL models bracket two new columns of results, one for the HP model and the other for the HP-S model. Of interest is how

²⁹To generate the real-time UC-filtered trend, we estimate the BL model over successively longer samples. In that exercise, we assume that the analyst wouldn't have recognized the 1998:Q1 shift in trend inflation's dynamics until 2002:Q1.

Table 5. Real-Time MSEs: Effect of Using HP Filter to Identify Trend Inflation

Horizon (<i>h</i>)	PCE	HP-S	HP	BL
0	2.693	2.464*>	2.438**	2.412**
1	2.669	2.409**	2.419**	2.453**
2	2.527	2.487>>	2.427*>	2.378**
3	2.558	2.582>>	2.452*>	2.412**
4	2.555	2.508*>	2.466>	2.400**
0–4	0.694	0.624*>>	0.567**>	0.545**

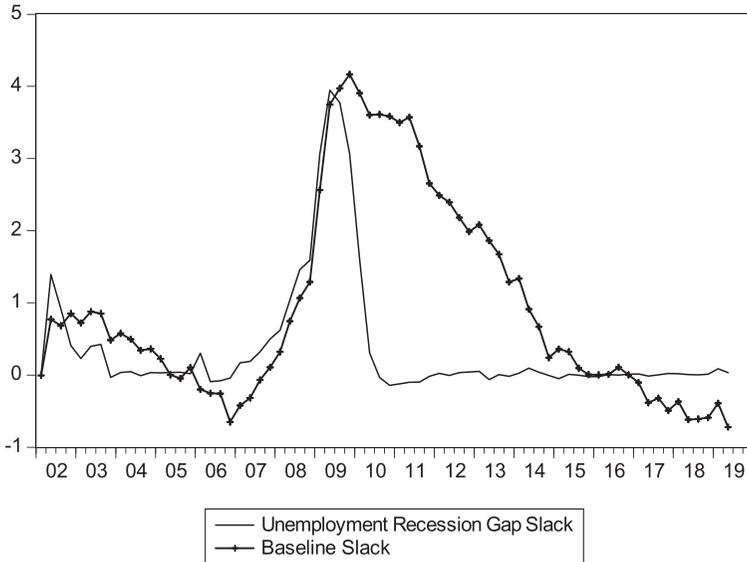
Note: The first set of forecasts is for 2002:Q1–2003:Q1; the final set is for 2018:Q2–2019:Q2. “-S” indicates that labor market slack, S, is excluded from the model. *Outperforms the AR(1) headline inflation model (“PCE”) at the 10 percent level. **Outperforms the AR(1) headline inflation model (“PCE”) at the 5 percent level. >The MSE difference between adjacent entries is significant at the 10 percent level. >>The MSE difference between adjacent entries is significant at the 5 percent level. The smallest MSE in each row is bolded.

the different treatment of trend inflation affects forecasting performance, and whether it changes our conclusions about the usefulness of slack in predicting inflation.

The table shows that the real-time performance of the HP model is worse than that of the BL model at every horizon except $h = 1$. The differences in MSE between the two models are statistically significant at the $h = 2$, $h = 3$, $h = 4$, and $h = 0–4$ forecast horizons. So, there is a price to be paid for adopting a less-sophisticated model of trend inflation. Still, the HP model outperforms the PCE model across the board, and those performance differences are statistically significant except when $h = 4$.

The answer to the second question is an unambiguous “no”: The modified-trend model with slack (HP) outperforms the same model without slack (HP-S) at every forecast horizon except $h = 1$, where the MSE difference is small. In those cases where including slack improves forecast performance, the improvement is statistically significant.

In summary, our main result—that one can improve on the forecast performance of an AR(1) inflation model by splitting headline inflation into trend, cycle, and noise; separately forecasting trend

Figure 7. Different Measures of Slack

and cyclical inflation while taking into account slack's influence on the latter; and not taking first-release headline inflation data at face value—does not depend on strong a priori assumptions about the dynamics of trend inflation.

3.4 Robustness: The Unemployment Recession Gap

The URG model is identical to the baseline, or BL model, except it uses the Stock and Watson (2010) unemployment recession gap instead of $U - U^*$ as an observable indicator of labor market slack.³⁰ Figure 7 shows the differences between the real-time UC slack estimates produced by the two models. In the URG model, slack is

³⁰One of the attractive features of the unobserved-common-components approach is that this need not be an either/or choice: In principle, we could include both slack indicators in an expanded version of model. Similarly, we could include one or more additional inflation indicators, such as median CPI inflation. The computational intensity of estimation increases greatly, though, in the absence of a priori restrictions on the variance/covariance matrix of the disturbances.

Table 6. Real-Time MSEs: Effect of Using an Alternative Slack Indicator

Horizon (h)	PCE	URG-S	URG	BL
0	2.693	2.423*	2.473*>	2.412**
1	2.669	2.519*>>	2.406**	2.453**
2	2.527	2.413*>	2.396*	2.378**
3	2.558	2.498>	2.425*	2.412**
4	2.555	2.529>	2.475*>	2.400**
0–4	0.694	0.608**>>	0.564**>	0.545**

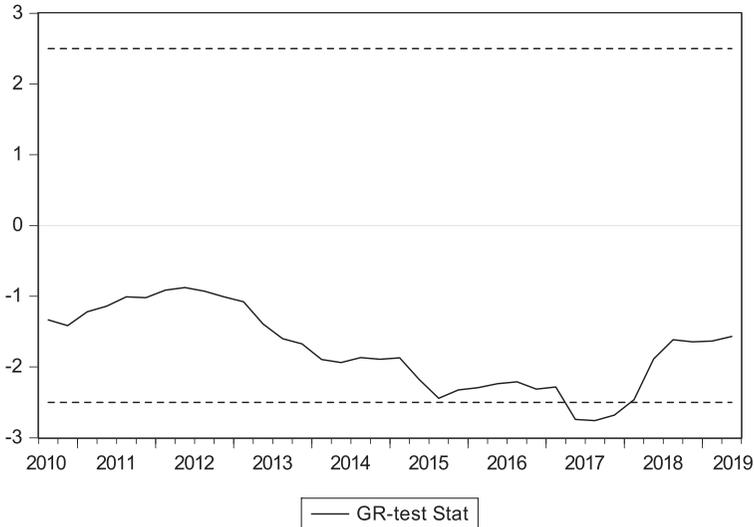
Note: The first set of forecasts is for 2002:Q1–2003:Q1; the final set is for 2018:Q2–2019:Q2. “-S” indicates that labor market slack, S, is excluded from the model. *Outperforms the AR(1) headline inflation model (“PCE”) at the 10 percent level. **Outperforms the AR(1) headline inflation model (“PCE”) at the 5 percent level. >The MSE difference between adjacent entries is significant at the 10 percent level. >>The MSE difference between adjacent entries is significant at the 5 percent level. The smallest MSE in each row is bolded.

more asymmetric: There is no such thing as a “tight” labor market. On the other hand, URG labor market slack disappears much more quickly after recessions. Neither difference is hugely surprising given the different observable slack indicators that feed into the two models.

Are the large differences evident in Figure 7 meaningfully reflected in real-time inflation-forecasting performance? Mean-square forecast errors are displayed in Table 6, which has a format similar to Tables 4 and 5. Comparing the columns labeled “BL” and “URG,” substituting the unemployment recession gap for $U - U^*$ causes forecast performance to deteriorate at every horizon except $h = 1$, and the deterioration is statistically significant for $h = 0$, $h = 4$, and $h = 0-4$. However, the URG model, like the BL model, significantly outperforms the benchmark AR(1) forecasting model, PCE, at every horizon.

Comparing the columns labeled “URG” and “URG-S,” excluding slack from the URG model leads to a statistically significant deterioration in real-time forecast performance at every horizon except $h = 0$.

In summary, our main results are not sensitive to using Stock and Watson’s unemployment recession gap in place of the unemployment gap as the observable, real-time indicator of labor

Figure 8. Fluctuation Test Statistic

market slack. Still, the model that uses the unemployment gap performs marginally better.

3.5 Robustness: Subsample Forecast Performance

Giacomini and Rossi (2010), Rossi and Sekhposyan (2010), and Rossi (2019) provide numerous examples where relative forecast performance is not robust to evaluation period: One model produces out-of-sample forecasts that are superior to another's, on average, over some interval, but the performance advantage disappears or is even reversed in one or more subintervals. Giacomini and Rossi (2010) propose a “fluctuations test” for detecting such performance instabilities. Figure 8 plots the Giacomini-Rossi test statistic, with critical values, for a comparison of our baseline UC model's real-time inflation forecasts against the real-time forecasts of the benchmark AR1 model.³¹ The figure shows that the baseline UC model

³¹The forecasts are of headline PCE inflation at the 0-to-4-quarter horizon, compared over rolling 32-quarter intervals. The interval length is such that the first and last intervals don't overlap. We use the variant of the fluctuations test that is appropriate for models that are estimated recursively. See Giacomini and Rossi (2010, pp. 600–601).

outperforms the AR1 model over every subinterval within our evaluation period, but the difference is typically not statistically significant. Notably, performance differences are most pronounced over subintervals that include the recovery from the 2008–09 Great Recession. Those results are encouraging but go only so far. We have strong evidence that the transition equation governing trend inflation shifted in the 1990s. A significant change in monetary policy strategy could cause it to shift again. Depending on the nature of that shift and the amount of time required to recognize it, the performance of our UC model might very well suffer for a time (though it is not obvious that it would suffer relative to benchmark models). These issues merit investigation in future research.

4. *Why Does Slack Matter?*

The above robustness exercises demonstrate that labor market slack is helpful in real-time inflation forecasting even when latest-available inflation data are taken at face value, even when inflation's trend is approximated by applying the HP filter to long-term inflation expectations, and even when we use a different measure of slack. Each of those exercises makes a marginal change to our baseline UC model. For a more-complete understanding of our results concerning slack, in this section we examine the impact of two *non*-marginal changes to our inflation-forecasting methodology. The exercises demonstrate that neither detrending nor stripping out noise is, by itself, sufficient to overturn the conventional wisdom that slack is of no use in real-time inflation forecasting: Finding a role for slack requires isolating inflation's *cyclical* component, which is inflation stripped of both its trend *and* transitory noise.

Our first exercise considers whether detrending alone is sufficient to find a role for slack. We strip headline PCE inflation of the trend identified in our baseline UC analysis, separately forecast trend inflation and detrended headline inflation, and then add the two forecasts together to produce a forecast of headline inflation.³²

³²The exercise is artificial in that the analyst is given the real-time trend estimates generated by the BL model, but is denied (or ignores) that model's estimates of inflation's cyclical and noise components. We further assume that the

Table 7. Real-Time MSEs: Detrending and Noise Filtering, Separately Considered

Horizon (h)	Detrend PCE	Detrend PCE+S	Core PCE	Core PCE+S
0	2.572	2.630	2.556	2.565
1	2.614	2.674	2.507	2.549
2	2.517	2.529	2.517	2.579
3	2.489	2.515	2.525	2.557
4	2.452	2.468	2.482	2.467
0–4	0.649	0.698	0.638	0.673

Note: The first set of forecasts is for 2002:Q1–2003:Q1; the final set is for 2018:Q2–2019:Q2. “+S” indicates that labor market slack, S, is included in the model.

In one version of this exercise, detrended headline inflation is forecasted using only its own lagged values. In another version, lagged values of the slack variable identified by our baseline UC analysis are also included in the regression. The analysis is recursive, using real-time data. As shown in the columns labeled “Detrend PCE” and “Detrend PCE+S” in Table 7, the model without slack outperforms the model with slack at every horizon.³³ By itself, then, detrending headline inflation is not enough to make slack useful in real-time forecasting.

Our second exercise considers whether stripping noise from headline inflation is sufficient to find a role for slack. We estimate a forecasting equation for core inflation as identified by our baseline UC model (cf. Equation (11)), and see how well that equation predicts headline inflation in real time.³⁴ In one version of the exercise the forecasting equation includes lags of slack along with lags of core inflation. In another version, slack is excluded from the forecasting

analyst has access to the BL model real-time estimate of what current-quarter headline inflation will look like after revision.

³³Both models forecast better than the headline inflation models without detrending that we examined earlier: Compare the results in Table 7 with those for the PCE and PCE+S models in Table 3.

³⁴This exercise is also deliberately artificial. Now, the analyst is handed the BL model’s real-time estimates of core inflation, but is denied (or ignores) that model’s estimates of trend inflation.

equation. Comparing the columns labeled “Core PCE” and “Core PCE+S” in Table 7, the model without slack outperforms the model with slack at every forecast horizon but one ($h = 4$).³⁵ Simply stripping noise from headline inflation is not enough to make slack useful in real-time forecasting.

The baseline UC model, of course, both detrends inflation *and* removes noise, and in that setting slack proves to be consistently and almost always significantly helpful in forecasting. (Compare the columns labeled “BL” and “BL-S” in Table 3.) Moreover, the forecasting results produced by the baseline model are more accurate than those produced by detrending alone or by stripping out noise alone. (Compare the “BL” column in Table 3 with any of the columns in Table 7.)

5. Conclusion

Consistent with the existing literature, we find that adding labor market slack to a simple autoregressive model of headline PCE inflation is counterproductive in real-time inflation forecasting, and that substituting ex-food-and-energy inflation for headline inflation is unhelpful. We have greater success using unobserved-components analysis to separate inflation into trend, cycle, and noise, and then separately forecasting trend and cycle using slack to help forecast the latter. Slack *does* matter for inflation, and stripping out noise *is* of help in identifying the inflation–slack connection and in improving inflation forecasts, but only if one also controls for inflation’s trend. Results are especially favorable when a simple model of inflation revisions is incorporated into the unobserved-components model.

As a practical matter, the unobserved-components approach requires that the analyst make modeling choices that are open to question. In that connection it is encouraging that our conclusions are robust to an alternative treatment of data revisions, to an alternative, more flexible treatment of inflation’s trend, and to an alternative indicator of labor market slack, and that they hold up well during the recovery from the Great Recession. Our conclusions,

³⁵ Both models forecast headline inflation better than models estimated using ex-food-and-energy PCE inflation: Compare the results in Table 7 with those for the XFE and XFE+S models in Table 3.

moreover, are broadly consistent with results from a small number of studies that have looked at the influence of labor market slack on deviations of trimmed-mean or weighted-median inflation from survey measures of long-run inflation expectations.

Appendix A. Mean-Square Forecast Errors from Inflation Gap, IMA(1,1), and Random-Walk Models

Table A.1 presents real-time forecasting results from our benchmark AR(1) model of headline PCE inflation (the column headed “PCE”); from the approach advocated here, which uses UC filtering to split headline inflation into trend, cycle, and noise, and models the data-revision process (the column headed “BL-S”); and from three alternative approaches that have appeared in the inflation-forecasting literature. (Forecast periods are the same as those in Tables 3–6.) Inflation-gap models forecast the gap between inflation and trend inflation, and then typically add the most-recent estimate of trend inflation back in to generate a headline inflation forecast (Stock and Watson 2010; Faust and Wright 2013). Results reported in the column labeled “GAP” apply that approach to headline PCE inflation using an average of SPF and Blue Chip long-forward inflation expectations to measure trend inflation. The GAP model does better than the AR(1) model at all but one forecast horizon ($h = 1$), but does worse than our UC model at every horizon but one ($h = 4$). An IMA(1,1) model, similar to the stochastic volatility model favored by Stock and Watson (2007), does poorly relative to the PCE, BL-S,

Table A.1. Real-Time Forecasting Results from Benchmark AR(1) Model of Headline PCE Inflation

Horizon (h)	PCE	BL-S	GAP	IMA(1,1)	RW
0	2.693	2.423	2.681	2.881	3.719
1	2.669	2.519	2.693	3.008	5.372
2	2.527	2.413	2.491	2.796	5.056
3	2.558	2.498	2.517	2.717	4.998
4	2.555	2.529	2.521	2.619	4.264
0–4	0.694	0.608	0.677	0.974	2.804

and GAP models at all horizons over our forecast period. Finally, the column labeled “RW” reports results from a simple random-walk model of headline inflation. The random-walk forecast performs worse than every alternative model at every horizon.

References

- Ang, A., G. Bekaert, and M. Wei. 2007. “Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?” *Journal of Monetary Economics* 54 (4): 1163–1212.
- Atkeson, A., and L. E. Ohanian. 2001. “Are Phillips Curves Useful for Forecasting Inflation?” *Quarterly Review* (Federal Reserve Bank of Minneapolis) 25 (1): 2–11.
- Bai, J., and P. Perron. 1998. “Estimating and Testing Linear Models with Multiple Structural Changes.” *Econometrica* (1): 47–78.
- Ball, L., and S. Mazumder. 2011. “Inflation Dynamics and the Great Recession.” IMF Working Paper No. 11/121.
- . 2019. “A Phillips Curve with Anchored Expectations and Short-Term Unemployment.” *Journal of Money, Credit and Banking* 51 (1): 111–37.
- Basistha, A., and C. R. Nelson. 2007. “New Measures of the Output Gap Based on the Forward-looking New Keynesian Phillips Curve.” *Journal of Monetary Economics* 54 (2): 498–511.
- Basistha, A., and R. Startz. 2008. “Measuring the NAIRU with Reduced Uncertainty: A Multiple-Indicator Common-Cycle Approach.” *Review of Economics and Statistics* 90 (4): 805–11.
- Bernanke, B. 2003. “‘Constrained Discretion,’ and Monetary Policy.” Remarks before the Money Marketmakers of New York University, New York, New York, February 3.
- Bryan, M. F., and S. G. Cecchetti. 1993. “Measuring Core Inflation.” NBER Working Paper No. 4303.
- Castelnuovo, E., S. Nicoletti-Altimari, and D. Rodriguez-Palenzuela. 2003. “Definition of Price Stability, Range and Point Inflation Targets: The Anchoring of Long-term Inflation Expectations.” Working Paper No. 273, European Central Bank.
- Chan, J. C. C., T. E. Clark, and G. Koop. 2015. “A New Model of Inflation, Trend inflation, and Long-run Inflation Expectations.” Working Paper No. 15-20, Federal Reserve Bank of Cleveland.

- Clarida, R. 2020. "The Federal Reserve's New Monetary Policy Framework: A Robust Evolution." Remarks at the Peterson Institute for International Economics, Washington, D.C., August 31.
- Clark, P. K. 1989. "Trend Reversion in Real Output and Unemployment." *Journal of Econometrics* 40 (1): 15–32.
- Clark, T. E., and T. Doh. 2014. "Evaluating Alternative Models of Trend Inflation." *International Journal of Forecasting* 30 (3): 426–48.
- Clark, T. E., and K. D. West. 2007. "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models." *Journal of Econometrics* 138 (1): 291–311.
- Cogley, T., G. Primiceri, and T. J. Sargent. 2010. "Inflation-Gap Persistence in the U.S." *American Economic Journal: Macroeconomics* 2 (1): 43–69.
- Crone, T., N. Neil K. Khettry, L. J. Mester, and J. A. Novak. 2013. "Core Measures of Inflation as Predictors of Total Inflation." *Journal of Money, Credit and Banking* 45 (2–3): 505–19.
- Croushore, D. 2008. "Revisions to PCE Inflation Measures: Implications for Monetary Policy." Working Paper No. 08-8, Federal Reserve Bank of Philadelphia.
- Croushore, D., and T. Stark. 2001. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics* 105 (1): 111–30.
- Diebold, F. X., and R. S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics* 13 (3): 253–63.
- Dolmas, J. 2005. "Trimmed Mean PCE Inflation." Working Paper No. 506, Federal Reserve Bank of Dallas.
- Faust, J., and J. Wright. 2009. "Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset." *Journal of Business and Economic Statistics* 27 (4): 468–79.
- . 2013. "Forecasting Inflation." In *Handbook of Economic Forecasting*, Vol. 2, ed. G. Elliott and A. Timmermann. Elsevier.
- Giacomini, R., and B. Rossi. 2010. "Forecast Comparisons in Unstable Environments." *Journal of Applied Econometrics* 25 (4): 595–620.
- Hasenzagl, T., F. Pellegrino, L. Reichlin, and G. Ricco. 2018. "A Model of the Fed's View on Inflation." CEPR Discussion Paper No. 12564.

- Kiley, M. T. 2008. "Estimating the Common Trend Rate of Inflation for Consumer Prices and Consumer Prices Excluding Food and Energy Prices." Finance and Economics Discussion Series No. 2008-38, Board of Governors of the Federal Reserve System.
- Kim, C.-J., and C. R. Nelson. 2000. *State-Space Models with Regime Switching*. Cambridge, MA: MIT Press.
- Kishor, N. K., and E. F. Koenig. 2012. "VAR Estimation and Forecasting When Data Are Subject to Revision." *Journal of Business and Economic Statistics* 30 (2): 181–90.
- . 2014. "Credit Indicators as Predictors of Economic Activity: A Real-Time VAR Analysis." *Journal of Money, Credit and Banking* 46 (2–3): 545–64.
- Koenig, E. F., and T. Atkinson. 2012. "Inflation, Slack and Fed Credibility." Staff Paper No. 16, Federal Reserve Bank of Dallas.
- Koenig, E. F., S. Dolmas, and J. Piger. 2003. "The Use and Abuse of Real-Time Data in Economic Forecasting." *Review of Economics and Statistics* 85 (3): 618–28.
- Kohn, D. 2003. "Comments on Marvin Goodfriend's 'Inflation Targeting in the United States'." Remarks at the NBER Conference on Inflation Targeting, Bal Harbour, Florida, January 25.
- Kozicki, S., and P. A. Tinsley. 2001. "Shifting Endpoints in the Term Structure of Interest Rates." *Journal of Monetary Economics* 47 (3): 613–52.
- . 2005. "What Do You Expect? Imperfect Policy Credibility and Tests of the Expectations Hypothesis." *Journal of Monetary Economics* 52 (2): 421–47.
- Mertens, E. 2016. "Measuring the Level and Uncertainty of Trend Inflation." *Review of Economics and Statistics* 98 (5): 950–67.
- Morley, J. C., C. R. Nelson, and E. Zivot. 2003. "Why Are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP So Different?" *Review of Economics and Statistics* 85 (2): 235–43.
- Nalewaik, J. 2015. "Regime Switching Models for Estimating Inflation Uncertainty." Working Paper, Federal Reserve Board.
- Phillips, A. W. 1958. "The Relationship between Unemployment and the Rate of Change of Money Wages in the United Kingdom 1861–1957." *Economica* 25 (100): 283–99.

- Rossi, B. 2019. "Forecasting in the Presence of Instabilities: How We Know Whether Models Predict Well and How to Improve Them." Barcelona GSE Working Paper Series No. 1162.
- Rossi, B., and T. Sekhposyan. 2010. "Have Economic Models' Forecasting Performance for U.S. Output Growth and Inflation Changed Over Time, and When?" *International Journal of Forecasting* 26 (4): 808–35.
- Stock, J. H., and M. W. Watson. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39 (s1): 3–34.
- . 2010. "Modeling Inflation After the Crisis." In *Macroeconomic Challenges: The Decade Ahead*. Proceedings of the 2010 Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, held in Jackson Hole, Wyoming, August 26–28.
- . 2016. "Core Inflation and Trend Inflation." *Review of Economics and Statistics* 98 (4): 770–84.
- . 2019. "Slack and Cyclically Sensitive Inflation." NBER Working Paper No. 25987.
- West, K. D. 1996. "Asymptotic Inference about Predictive Ability." *Econometrica* 64 (5): 1067–84.
- Yellen, J. L. 2016. "Macroeconomic Research After the Crisis." Remarks at the 60th annual economic conference sponsored by the Federal Reserve Bank of Boston, Boston, Massachusetts, October 14.