Federal Reserve Bank of Cleveland Working Paper Series

A Unified Framework to Estimate Macroeconomic Stars

Saeed Zaman

Working Paper No. 21-23R

August 2022

Suggested citation: Zaman, Saeed. 2022. "A Unified Framework to Estimate Macroeconomic Stars." Working Paper No. 21-23R. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202123r.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

A Unified Framework to Estimate Macroeconomic Stars

Saeed Zaman*
Federal Reserve Bank of Cleveland, USA

First version: June 30, 2021
This version: July 31, 2022

Abstract

We develop a flexible semi-structural time-series model to estimate jointly several macroeconomic *"stars"* — i.e., unobserved long-run equilibrium levels of output (and growth rate of output), the unemployment rate, the real rate of interest, productivity growth, price inflation, and wage inflation. The ingredients of the model are in part motivated by economic theory and in part by the empirical features necessitated by the changing economic environment. Following the recent literature on inflation and interest rate modeling, we explicitly model the links between long-run survey expectations and stars to improve the stars' econometric estimation. Our approach permits time variation in the relationships between various components, including time variation in error variances. To tractably estimate the large multivariate model, we use a recently developed precision sampler that relies on Bayesian methods. The by-products of this approach are the time-varying estimates of the wage and price Phillips curves, and the pass-through between prices and wages, both of which provide new insights into these empirical relationships' instability in US data. Generally, the contours of the stars echo those documented elsewhere in the literature – estimated using smaller models – but at times the estimates of stars are different, and these differences can matter for policy. Furthermore, our estimates of the stars are among the most precise. Last, we document the competitive real-time forecasting properties of the model and, separately, the usefulness of stars' estimates as steady-state values in external models.

JEL classification: C5, E4, E31, E24, O4

Keywords: state-space model, Bayesian analysis, time-varying parameters, natural rates, survey expectations, COVID-19 pandemic

---

I am extremely grateful to Joshua Chan, Todd Clark, Julia Darby, Ana Galvão, Gary Koop (my PhD advisor), Edward Knotek, Kurt Lunsford, Elmar Mertens, James Mitchell, Aubrey Poon, Ellis Tallman, Willem Van Zandweghe, Randal Verbrugge, and Ben Wong for valuable comments on the first draft. This research has benefited from discussions with Fabio Canova, Siddhartha Chib, Olivier Coibion, Sharada Davidson, Jesus Fernandez-Villaverde, Garo Garabedian, Paolo Gelain, Yuriy Gorodnichenko, Stuart McIntyre, Eric Sims, and Ping Wu. I also thank participants at the Society for Computational Economics 27th Conference 2021, NBER-NSF SBIES 2021, European Seminar on Bayesian Econometrics 2021, Computational and Financial Econometrics Conference 2021, 2022 North American Summer Meeting of the Econometric Society, 5th Vienna Workshop on High-Dimensional Time Series in Macroeconomics and Finance, Midwest Macroeconomics Spring 2022, Society for Nonlinear Dynamics and Econometrics 2022, ESCoE Conference on Economic Measurement 2022, and the Cleveland Fed Brown Bag. The views expressed herein are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

*Research Department, Federal Reserve Bank of Cleveland, Ohio, USA; email: saeed.zaman@clev.frb.org

1. Introduction

The estimates of long-run equilibrium levels of macroeconomic variables (often denoted with the "star" symbol) are of central importance in macroeconomics. These long-run levels are thought to reflect the fundamental structure of the economy in the absence of shocks. Hence, they are used as reference points, and deviations from these long-run levels reflect idiosyncratic and cyclical fluctuations, which typically serve as the source data for macroeconomic models (e.g., Smets and Wouters, 2003; Canova and Ferroni, 2011).

In this paper, we estimate jointly seven macroeconomic stars of broader interest to macroeconomists and policymakers: the level of potential output (gdp-star), the growth rate of potential output (g-star), the long-run equilibrium levels of the unemployment rate (u-star), the real short-term interest rate (r-star), labor productivity growth (p-star), price inflation (pi-star), and nominal wage inflation (w-star).[1] The assumption that a long-run equilibrium exists implies that in the long run, the economy is growing at potential, price inflation is growing at its trend rate, the unemployment rate has no cyclical pressure and only reflects structural factors, nominal wages grow at a rate equal to the sum of labor productivity growth and price inflation, and the real interest rate reflects the rate consistent with output growing at potential and stable inflation.[2]

In practice, determining the values of these stars is difficult because stars and some of their determinants are unobserved. To infer estimates of the stars, economists apply a range of econometric methods to observable historical data.[3] The multivariate unobserved components (UC) models, which are statistical models that use economic theory to frame the empirical specification, have been shown to provide reasonable estimates of the stars (e.g., Kuttner, 1994; Laubach and Williams, 2003; Basistha and Nelson, 2007; Chan, Koop, and Potter, 2016). Hence, they are the dominant methods for obtaining estimates of the stars. However, with few exceptions, the popular multivariate UC models that estimate time-varying stars focus on a small number of observables, often just two or three, and minimal structure (e.g., Laubach and Williams, 2003 [henceforth LW]). Studies that entertain more variables have abstracted from important empirical features such as time-varying parameters and stochastic volatility (e.g., Hasenzagl, Pellegrino, Reichlin, and Ricco, 2020; Del Negro et al., 2017; Fleischman and Roberts, 2011).

[1]The subset of these stars, p-star, g-star, u-star, and r-star, reflects the fundamental structural features of the economy, whereas others, pi-star and w-star, are thought to be influenced by central banks and monetary policy.

[2]The literature has referred to the concept of long-run equilibrium using different terminologies, such as "natural," "neutral," "trend," "steady-state," and "long-run." There are subtle differences among them, but they can be interpreted as the same for the purpose of this paper. In some studies, especially those using dynamic stochastic general equilibrium (DSGE) models, the concept of the natural rate refers to a medium-horizon equilibrium, and in these same models, the concept of the steady state is used to refer to the long-run equilibrium.

[3]The methods range from statistical univariate filters (e.g., Hodrick and Prescott, 1997; Ashley and Verbrugge, 2008) to multivariate models, including semi-structural time-series models (e.g., Pescatori and Turunen, 2016; Morley and Wong, 2020), and fully structural DSGE models (e.g., Del Negro, Giannone, Giannoni, and Tambalotti, 2017).

A priori, one would expect a framework based on greater information that explicitly permits (contemporaneous) interactions between stars and between cyclical components, and a richer structure to provide more reliable estimates of the objects of interest (e.g., stars) than frameworks that ignore them.

Accordingly, in this paper, we take on the challenge of jointly estimating several macroeconomic stars simultaneously, including g-star (and gdp-star), u-star, r-star, p-star, pi-star, and w-star, using a semi-structural time-series model. For each star, we formulate a rich structure whose elements are guided by past research and informed by economic theory. For example, econometric estimation of r-star is informed by various sources: the investment-savings (IS) equation, the Taylor-type rule, the equation linking g-star and r-star, and the equation relating r-star to survey expectations. We allow for time variation in important macroeconomic relationships and error variances. Fernández-Villaverde and Rubio-Ramírez (2010), and Carriero, Clark, and Marcellino (2019), among many others, highlight the importance of allowing for stochastic volatility in macroeconomic models. Incorporating these empirical features should better distinguish between cyclical fluctuations and lower-frequency movements in the macroeconomic aggregates considered in this paper.

We extend the Chan, Clark, and Koop (2018) – henceforth CCK – approach of using long-run survey expectations to improve pi-star precision to other macroeconomic stars. Specifically, for each macroeconomic variable of interest, we explicitly model the link between the unobserved "star" and the expectations about the star contained in the Blue Chip survey of economic forecasters (or as reported by the Congressional Budget Office [CBO] when the survey estimate is not available).[4] In a high-dimensional model like ours, the use of long-run survey expectations, which are direct measures of stars, could help anchor model-based estimates of stars to reasonable values (especially in times of heightened uncertainty) and potentially improve the precision of the estimates. We estimate the feature-rich UC model using Bayesian techniques, specifically the efficient sampling techniques developed in Chan, Koop, and Potter (2013) and the precision sampler proposed in Chan and Jeliazkov (2009).

All in all, the combination of time-varying parameters, SV, joint modeling of multiple stars, implementation of an expanded structure, and allowing for a direct connection between stars and long-term survey expectations is what differentiates our UC model from those in the existing literature. Many popular UC models could be viewed as special cases of our larger UC model, which facilitates model comparison. We note that among the stars considered, w-star has received less attention in the literature. Our UC model's ability to provide real-time estimates of w-star and its model-based decomposition into its determinants p-star and pi-star, as implied by economic theory, is a novel contribution. This specific decomposition is useful to monetary policymakers, who often refer to developments in nominal wages to support their forecasts and

[4] The long-run survey expectations can be thought of as a hybrid forecast because the survey combines judgment based on a range of information and forecasts derived from a range of modeling approaches. Our use of such a hybrid forecast implicitly serves as an additional channel through which the issue of omitted variable bias is mitigated.

related discussions on price inflation and employment.

Our results indicate that there are payoffs to modeling stars jointly using a larger multivariate UC model. The metric of Bayesian model comparison generally favors our larger UC model over smaller-scale UC models. The model yields credible estimates of stars and the output gap. For example, the output gap estimate is similar to the CBO estimate based on a production function approach. A deeper examination reveals that joint modeling of the unemployment rate and real GDP and allowing for SV in their cyclical components are the two key ingredients to obtaining credible estimates of the output gap.

Generally, the contours of stars echo those documented elsewhere in the literature, but at times they are different, and these differences can matter for policy. For example, let's consider pi-star. From 2000 to 2010, our UC model has an estimate of pi-star stable at or close to 2%, whereas pi-star from a popular univariate model (of Stock and Watson, 2007) displays notable fluctuations around 2%, and the bivariate model of CCK indicates a stable pi-star about a few tenths below 2%. These differences matter for central banks tasked with inflation targeting. Compared to some of the popular UC models and the smaller-scale restricted variants of our larger UC model, the precision estimates of the stars and the output gap from our UC model are among the most precise, where precision is measured as the width of 90% credible intervals. The model's reliance on long-term survey expectations data is the key reason for this improved precision.[5] Survey expectations have played a crucial role in guiding the model-based assessment of stars during the COVID-19 pandemic, a period of heightened uncertainty. The accuracy of our UC model's real-time point and density forecasts rivals and, in some cases, outperforms hard-to-beat benchmarks, including small-scale UC models.

We also demonstrate the usefulness of our estimated stars as terminal values for external models. Previous research shows that forecasting models, such as steady-state vector autoregressive (VAR) models, often improve their forecast accuracy by using external information about steady states informed by long-run survey expectations (e.g., Wright, 2013). Using a real-time forecasting comparison, we show that if we were to inform steady states in a VAR with the stars from our UC model, gains in forecast accuracy for some of the variables would be achieved compared to the standard approach relying on survey expectations. Hence, our framework provides a potential source for obtaining the stars' estimates in real time. An advantage of our framework compared to surveys is that it provides estimates of stars (steady states) for variables not covered by surveys (e.g., w-star) and offers both point and uncertainty estimates.

We summarize three additional findings. First, we find that the empirical evidence on the link between r-star and g-star, as implied by theory, is weak (consistent with Hamilton, Harris, Hatzius, and West, 2016; Lunsford and West, 2019), but by bringing survey expectations into the model, the link becomes stronger (providing support to Laubach and Williams, 2016). Second, our results indicate economically and statistically significant evidence of time variation in the

[5] Our precision estimates are on a par with recent studies highlighting the improved precision of stars derived from their approaches (e.g., r-star by Del Negro et al., 2017; u-star by Crump, Eusepi, Giannoni, and Şahin, 2019; pi-star by CCK).

model parameters capturing macroeconomic relationships and strong support for SV's inclusion in the model equations. It lends support to the popular narratives of: "The price Phillips curve has weakened over time," "The wage Phillips curve is alive," and "There is weakening in the procyclicality of labor productivity." Third, a comparison between final and real-time estimates of the stars indicates that their broad movements have generally tracked each other closely. We view this as a valuable finding because it suggests that we have made some progress in mitigating the well-known difficulties associated with the real-time estimation of the stars.

In recent years, advances in computational power and numerical methods have enabled researchers to estimate stars using UC models with more indicators and/or an expanded structure. For example, Johannsen and Mertens (2021) [henceforth JM], Pescatori and Turunen (2016), Del Negro et al. (2017), Brand and Mazelis (2019), González-Astudillo and Laforte (2020), among others, have examined the roles of additional determinants in explaining r-star.[6] None of these studies feature time-varying parameters, and only JM allows for SV, but their model size is significantly smaller than ours. Chan, Koop, and Potter (2016) [henceforth CKP] illustrate the value of modeling u-star and pi-star as bounded random walk processes in a bivariate Phillips curve. More recently, using fixed-parameter UC models, Crump et al. (2019) estimate u-star by combining a range of labor market indicators across demographic groups and survey expectations of inflation, and Hasenzagl et al. (2020) jointly estimate pi-star, u-star, gdp-star (and output gap).[7] Feunou and Fontaine (2021) develop a UC model with SV to examine the secular decline in bond yields by jointly modeling r-star, g-star, and pi-star.

The paper is organized as follows. The next section describes in detail the econometric model and its variants. Section 3 describes the data and estimation. Section 4 presents and discusses in detail the estimates of stars and other model parameters. Section 5 reports the real-time forecasting results and a discussion comparing real-time and final estimates of stars. Section 6 illustrates the ability of the model to handle the COVID-19 pandemic data. Section 7 concludes. This paper has a supplementary online appendix that lists detailed Bayesian estimation steps and additional results.

2. Empirical Macro Model and Variants

The ingredients of our macroeconomic econometric model are guided both by economic theory and by empirical considerations – namely, features that previous research has demonstrated to be empirically relevant. These features include stochastic volatility and time-varying pa-

[6] Pescatori and Turunen (2016) enrich the underlying structure to estimate r-star. In particular, to extract a reliable estimate of the output gap, they bring additional information from the CBO's estimate of the output gap by treating it as a noisy measure of the "true" output gap.

[7] Morley and Wong (2020) and Chan (2019) propose an alternative modeling framework based on VARs to estimate the long-run equilibrium values. The advantage of the VAR-based framework is the ability to handle larger amounts of information conveniently and flexibly compared to UC models. On the other hand, the advantage of UC modeling, as emphasized by CKP, is the availability of the direct estimates of stars, which in the case presented here proves quite convenient in allowing for direct modeling of the relationships between various stars.

rameters, which in turn imply time-varying predictability. Collectively, these empirical features permit modeling changing macroeconomic relationships in a flexible way.

We represent our empirical model using six sets of equations, which we denote blocks. These six blocks, which allow for *contemporaneous interactions* between them, characterize the joint dynamics of the unemployment rate, output growth, labor productivity growth, price inflation, nominal wage inflation, nominal interest rate, and corresponding stars. To be sure, the model assumes that all innovations are uncorrelated both serially and across equations. However, we emphasize that any assumed current-period correlations between the cyclical components and/or between stars are directly modeled via the model equations that define the contemporaneous relationships between the components (e.g., the cyclical output gap at time t with the cyclical unemployment gap at t; r-star and g-star).

Before we describe the model, we provide some necessary background information, including the econometric definition of the star and the usefulness of long-run survey expectations in the estimation of stars.

2.1. ***The econometric notion of a long-run equilibrium***

Following CCK, Mertens (2016), and Lee and Nelson (2007), among many others, this paper defines the long-run equilibrium (or star) of a particular macroeconomic series as its infinite-horizon forecast conditional on the current information set. This definition of a star is consistent with the notion of Beveridge-Nelson trend decomposition, and an extensive literature has adopted this approach to estimate stars. Equivalently, as commonly defined in the trend estimation literature, the infinite-horizon forecast could be viewed as an estimate of trend conditional on the current information set (e.g., CCK, Mertens, 2016, Lee and Nelson, 2007). As discussed in Mertens (2016), among others, different information sets would likely yield different estimates of the infinite-horizon forecast (or trend). Mertens showed that including survey projections of long-term inflation in the information set led to more precise and forward-looking estimates of trend inflation.

The link between the infinite-horizon forecast and the underlying trend is described well by the unobserved components (UC) model (see LW; Lee and Nelson, 2007; Mertens, 2016; CCK). In a UC model, a series (Y_t) is typically represented as the sum of a nonstationary trend component Y_t^*, which is assumed to evolve slowly, and a stationary cycle Y_t^c, whose infinite-horizon conditional expectation is assumed to be zero. Accordingly,

$$Y_t = Y_t^* + Y_t^c. \tag{1}$$

The trend component Y_t^* is interpreted as the limiting forecast of the series (conditional on the information set I_t) as the forecast horizon tends to infinity,

$$\lim_{j \to \infty} \mathbb{E}[Y_{t+j} | I_t] = Y_t^*. \tag{2}$$

Differencing the above equation yields,

$$Y_t^* = Y_{t-1}^* + \lim_{j\to\infty} \mathbb{E}[Y_{t+j}|I_t] - \lim_{j\to\infty} \mathbb{E}[Y_{t+j}|I_{t-1}] = Y_{t-1}^* + e_t, \;\; e_t \sim N(0, \sigma_e^2). \tag{3}$$

which suggests a random walk process for the trend Y_t^*.[8] It also suggests a stationary ergodic mean-zero process for Y_t^c.

Intuitively, the above set of assumptions implies that once the effects of the shocks have fully played out, the macroeconomic series of interest, Y_t, gravitates to its underlying trend level, Y_t^*.

As discussed in CCK, various statistical and econometric models could fit within the above-specified decomposition. This paper formulates a specific unobserved components time-series model and its variants.

2.2. ***The role of survey expectations***

As discussed in the introduction, an important contribution of this paper is to provide a direct role for long-run survey expectations in refining the stars' estimates. Specifically, we follow the approach of CCK (and Pescatori and Turunen, 2016). Those papers explicitly estimate an equation linking the observed measure of a long-run forecast obtained from external sources (survey in the case of CCK and the CBO's projection of the output gap in Pescatori and Turunen) to an unobserved object of interest. We extend their approach to the macroeconomic variables considered in this paper.

Several papers have documented an essential role of long-run survey (and institutional) forecasts in helping refine the econometric estimation of model parameters, including the latent components (e.g., pi-star: Kozicki and Tinsley, 2012; Mertens, 2016; Mertens and Nason, 2020; CCK; gdp-star: Pescatori and Turunen, 2016; r-star: Del Negro et al., 2017). Specifically, Mertens and Nason (2020), CCK, Mertens (2016), and Kozicki and Tinsley (2012), in using different methodologies (in combining survey data with model forecasts) to estimate the trend in US inflation, show that long-run survey forecasts of inflation deliver crucial additional information (beyond the recent inflation history) in refining trend estimates and improving model fit.[9] In a similar vein, Pescatori and Turunen (2016) document the usefulness of the CBO's projection of the potential output gap in improving their model's output gap precision. It is this particular literature that motivates us to consider long-run survey forecasts in our large-scale econometric model.

The advantage of survey (and institutional) forecasts stems from the fact that they could be

[8] The commonly adopted assumption of modeling Y_t^* as a random walk is partly due to consensus among macroeconomists that the factors driving the long-run equilibrium levels are perceived to be quite persistent (e.g., Lee and Nelson, 2007). In practice, the assumption of a (driftless) random walk has generally worked quite well, in that it has provided reasonable estimates of the stars (e.g., Clark, 1987; Kuttner, 1994; Mertens, 2016).

[9] An important empirical finding of CCK is that long-run survey expectations of inflation are a biased measure of the underlying trend inflation, at least at some times. Hence, simply equating pi-star with a long-run survey expectations (as is commonly done in macroeconomic models) may not be a reasonable strategy.

viewed as hybrid forecasts, i.e., a combination of judgment and forecasts derived from various modeling approaches. Furthermore, in high-dimensional models, such as the one developed in this paper, the use of long-run survey projections, which are targeted and direct measures of stars, could help anchor model-based estimates of stars to reasonable values and have the potential to improve precision of the estimates.

Accordingly, in this paper, with the exceptions of nominal wage inflation and labor productivity, for each of the remaining four variables, we model a direct link between long-run survey projections (or the long-run CBO projections in the years for which survey projections are unavailable) and the corresponding star using the following econometric equations:[10]

$$Z_t^j = C_t^j + \beta^j j_t^* + \varepsilon_t^{zj}, \;\; \varepsilon_t^{zj} \sim N(0, \sigma_{zj}^2), \;\; j = \pi, u, g, r \tag{4}$$

$$C_t^j = C_{t-1}^j + \varepsilon_t^{cj}, \;\; \varepsilon_t^{cj} \sim N(0, \sigma_{cj}^2), \;\; j = \pi, u, g, r \tag{5}$$

where π refers to price inflation, u refers to the unemployment rate, g refers to real GDP growth, r refers to the real short-term interest rate, Z_t^j refers to the long-run survey forecast corresponding to the variable j, and j_t^* is the unobserved j star.[11]

C_t^j is a time-varying intercept assumed to evolve as an RW process to possibly capture the permanent wedge between the survey estimate and the model-based star. This wedge can arise for several reasons, including the fact that the model-based star is assumed to be the infinite-horizon forecast, whereas the survey forecast refers to the average forecast for the five-year period starting seven years into the future in the case of BC and the ten-year-ahead forecast in the case of the SPF (for price inflation).

The above set of equations defines a simplistic and flexible relationship between the long-run survey expectations and the star.

2.3. ***Unemployment block***

The long-run equilibrium level of unemployment (u-star) is the (nonzero) unemployment rate that prevails when output is growing at potential, and the economy adds jobs so as to maintain the full-employment level. As discussed in Crump et al. (2019), two approaches are commonly used to estimate u-star. The first approach applies UC modeling to detailed labor market data (such as job vacancies, firms' recruiting intensity, demographic changes, flows into and

[10]For the long-run inflation forecast, we use the Survey of Professional Forecasters (SPF) (and the PTR series that is available for download from the website of the Federal Reserve Board), and for the long-run forecasts of the other three variables, we use the Blue Chip (BC) survey.

[11]Because survey expectations of g-star, i.e., Z^g, are reported as annualized rates, the precise formulation of the equation linking survey expectations to g-star is $Z_t^g = C_t^g + \beta^g * 4 * g_t^* + \varepsilon_t^{zg}$. We note that r-star survey projections are not direct estimates; instead, they are inferred from the Blue Chip survey's long-run projections of the GDP deflator and short-term interest rates using the long-run Fisher equation. The inferred estimates of survey expectations for r-star go back to 1983.Q1. Please refer to the online appendix A9 for details on the procedure to back-cast estimates all the way back to 1959.

out of unemployment) to extract respective trends. These trends are used to construct implied estimates of u-star (e.g., Davis, Faberman, and Haltiwanger, 2013). The second approach uses a combination of information from prices (and or nominal wages, survey expectations) and the estimated Phillips curve relationship between price inflation and the aggregate unemployment rate to back out the estimate of u-star (e.g., Staiger, Stock, and Watson, 1996, Stella and Stock, 2015, CKP).[12] We adopt the latter approach.

Specifically, we posit that the observed unemployment rate is decomposed into a (bounded) RW trend component (u-star) and a stationary cyclical component.

$$U_t = U_t^* + U_t^c \tag{6}$$

The cyclical component is modeled as an AR(2) process.[13] Because we are also modeling the output gap (i.e., the level of real GDP minus the level of potential real GDP), we depart from the previous literature by augmenting the AR2 unemployment gap with the output gap (denoted ogap) as an additional explanatory variable.[14]

$$U_t - U_t^* = \rho_1^u(U_{t-1} - U_{t-1}^*) + \rho_2^u(U_{t-2} - U_{t-2}^*) + \phi^u ogap_t + \varepsilon_t^u, \;\; \varepsilon_t^u \sim N(0, e^{h_t^u}) \tag{7}$$

where, $\rho_1^u + \rho_2^u < 1$, $\rho_2^u - \rho_1^u < 1$, and $|\rho_2^u| < 1$.

The variance of the error term ε_t^u is allowed to change over time.[15]

Similarly, as shown later, we add information from the unemployment gap when modeling the output gap. The joint modeling of both the output gap and the unemployment gap allows us to estimate the strength of the relationship between the two cyclical components, popularly known as Okun's law. In equation (7), the coefficient ϕ_u captures the contemporaneous relationship between the output gap and the cyclical unemployment rate gap. The estimate, $\frac{1-\rho_1^u-\rho_2^u}{\phi^u}$, could be interpreted as the Okun's law coefficient.[16]

[12] As mentioned in Crump et al. (2019), one of the criticisms of this approach is that it will be affected by the breakdown of the Phillips curve relationship post-2007. However, by allowing time variation in the coefficients capturing the price and wage Phillips curve relationships, as we do, our approach should face less of a problem. In addition, as illustrated in Del Negro, Giannoni, and Schorfheide (2015), including information from long-run survey expectations of inflation (as we do) should further help capture the inflation's behavior in the post-2007 period.

[13] The use of a parsimonious (time-invariant) AR2 process to identify the cyclical component of the unemployment rate is a commonly used assumption, in our case motivated by a recent string of empirical studies, e.g., Lee and Nelson (2007), CKP, and Galí and Gambetti (2019). CKP explore the empirical importance of allowing for time variation in the parameters of an AR2 process, and find that the data prefer the time-invariant AR2 process, hence validating the widely used assumption of a simple AR2 process.

[14] Sinclair (2009), Grant and Chan (2017a), and Berger, Everaert, and Vierke (2016), among several others, document the empirical importance of jointly modeling the unemployment rate gap and the output gap.

[15] Mertens (2014), Stella and Stock (2015), and Berger, Everaert, and Vierke (2016) provide evidence in support of SV in the cyclical component of the unemployment rate.

[16] We note that when jointly modeling output and the unemployment rate, most researchers assume a common cyclical component between the two. However, in light of the empirical evidence that cyclical unemployment displays more persistence than the output gap (e.g., Berger, Everaert, and Vierke (2016)), we model two separate

U-star is modeled as a bounded RW, where the bounds' values are fixed at 3.5% (lower bound) and 7.5% (upper bound).[17]

$$U_t^* = U_{t-1}^* + \varepsilon_t^{u*}, \;\; \varepsilon_t^{u*} \sim TN(a_u - U_{t-1}^*, b_u - U_{t-1}^*; 0, \sigma_{u*}^2) \tag{8}$$

where the notation $TN(a, b; \mu, \sigma^2)$ refers to normal distribution with mean μ and variance σ^2 but truncated in the interval (a, b).[18]

2.4. ***Output block***

We are interested in both the potential output (i.e., gdp^*) and the growth rate in potential output (i.e., g^*). To feasibly estimate both of these latent processes, we follow the commonly adopted approach, which decomposes the level of aggregate output into the level of potential output and a cyclical component (output gap), where the cyclical component is defined as the deviation of the observed aggregate output level from potential output. This simple decomposition has a long tradition going back to Clark (1987).

$$gdp_t = gdp_t^* + ogap_t \tag{9}$$

where $gdp \equiv log(GDP)$ and gdp^* refers to potential output, which is unobserved.

Following Grant and Chan (2017b), gdp^* is assumed to follow a second-order Markov process.[19]

cycles linked to each other via the Okun's law relationship (similar to Sinclair (2009) and Grant and Chan (2017b)). As shown in Berger, Everaert, and Vierke (2016), in a specification that entertains two separate cycles (cyclical unemployment and the output gap), the data support a time-invariant parameter describing the Okun's law relationship. In contrast, a specification with a common cyclical component favored a time-varying Okun's law relationship (adding support to Knotek II, 2007). In a nutshell, the inference of Berger, Everaert, and Vierke (2016) suggests that time variation in the parameter capturing Okun's law reflects the sluggish response of the cyclical unemployment rate to movements in the output gap. Once they allowed for a sluggish response of the cyclical unemployment rate by adding persistence, via a one-period lag of the cyclical unemployment rate, evidence of a time-varying Okun coefficient disappears. We found similar evidence. Furthermore, the Bayesian model comparison assessment slightly preferred the approach of two separate cycles with a time-invariant Okun's law compared to a common cycle with a time-varying Okun's law parameter.

[17] These values are informed by estimating the CKP model over our estimation sample, and are close to values reported in CKP based on their estimation sample. As a further check, most estimates of the u-star reported in the commonly cited literature fall within the bounds we use in this paper.

[18] With the exception of CKP, most of the literature models u-star as a driftless RW. The use of an unrestricted RW process has empirically been shown to work well, but CKP show that modeling u-star as a bounded RW process is even better. They use bounds because, by construction, the unemployment rate is a bounded variable, which implies that the long-run equilibrium in the labor market would restrict the movements in u-star within a bounded interval. CKP argue that economic forces that govern the movements in u-star are slow-moving and those forces would not cause the unemployment rate to fall to levels close to zero or to levels that are higher than the previous peaks caused by recessions.

[19] This modeling assumption implies that all permanent shocks to output are shocks to g^*. Results are similar had we instead modeled gdp^* as an RW with a time-varying drift term, where the time-varying drift term (interpreted as g^*) is assumed to follow an RW process (to allow for a stochastic g^*). However, the metric of Bayesian model comparison slightly favors the assumption of a second-order Markov process for gdp^*, which is consistent with the findings of Grant and Chan (2017b). An advantage of modeling g^* as a second-order Markov process compared to an RW with time-varying drift is that it requires estimating a single shock parameter (σ_{gdp*}^2), as opposed to two for the latter (one for the shock to gdp^* and the other for the shock to the time-varying drift,

$$gdp_t^* = 2gdp_{t-1}^* - gdp_{t-2}^* + \varepsilon_t^{gdp*}, \;\; \varepsilon_t^{gdp*} \sim N(0, \sigma_{gdp*}^2) \tag{10}$$

which can be re-written as

$$\triangle gdp_t^* = \triangle gdp_{t-1}^* + \varepsilon_t^{gdp*}$$

If we define $g_t^* \equiv \triangle gdp_t^*$, where $\triangle$ is the first difference operator, then,

$$g_t^* = g_{t-1}^* + \varepsilon_t^{gdp*} \tag{11}$$

The cyclical component, ogap, is assumed to be a stationary AR(2) process augmented with additional explanatory variables: the real interest rate gap and the unemployment gap,

$$ogap_t = \rho_1^g(ogap_{t-1}) + \rho_2^g(ogap_{t-2}) + a^r(r_t^L - r_t^* - tp_t^*) + \lambda^g(U_t - U_t^*) + \varepsilon_t^{ogap} \tag{12}$$

where, $\varepsilon_t^{ogap} \sim N(0, e^{h_t^o})$, $\rho_1^g + \rho_2^g < 1$, $\rho_2^g - \rho_1^g < 1$, and $|\rho_2^g| < 1$

Equation (12) could be interpreted as defining an IS curve (as in LW and subsequent papers modeling r-star) that allows feedback (via parameter a^r) from the real interest rate gap to the output gap (i.e., the real interest rate gap responds to economic slack). The IS equation is inspired by LW but with two modifications. First, instead of using the interest rate gap based on the short-term real rate of interest, we use the long-term real interest rate (as in González-Astudillo and Laforte, 2020).[20] Specifically, the long-term real interest rate, r^L, is constructed as the difference between the nominal yield on a 10-year Treasury bond and 10-year inflation expectations (i.e., the PTR series for PCE inflation).[21] The long-run value of the term premium, tp^*, is treated as an exogenous variable and is constructed as the average of the differential between the long-term interest rate (i.e., 10-year Treasury bond) and the federal funds rate, similar to Johannsen and Mertens (2021).

Second, to improve the econometric estimation of the output gap, we enrich the IS equation by bringing in information from the unemployment gap (from the unemployment block) as an explanatory variable.[22] This latter addition is motivated by the approach taken in a long list of papers (e.g., Morley and Wong, 2020; Grant and Chan, 2017a; Fleischman and Roberts, 2011; Sinclair, 2009) that demonstrate the usefulness of the unemployment rate in improving

aka g^*). It is worth noting that the assumption of gdp^* following a second-order Markov process is consistent with the Beveridge-Nelson trend described in Section 2.1 (see Proietti, 1995).

[20] In theoretical models, the long-term interest rate influences household consumption decisions and business investment decisions.

[21] We also experimented with an alternative specification, in which the interest rate gap is constructed as the difference between the short-term federal funds rate and the first lag of four-quarter trailing PCE inflation, similar in spirit to LW. Based on model fit, this specification was slightly inferior. It is worth noting that had a longer history of long-term inflation expectations data been available at the time of writing, LW would have constructed the interest rate gap using the long-term interest rate (see page 1064 in LW).

[22] Model fit, the precision metric for u-star and the output gap, and the plausibility of the estimates of the output gap strongly support the joint modeling of the output gap and the unemployment gap.

the econometric estimation of the output gap. The coefficient λ^g captures the contemporaneous relationship between the output gap and the unemployment gap.[23]

2.5. ***Productivity block***

The estimates of the long-run level of (labor) productivity growth (p-star) are of considerable interest to policymakers.[24] This is because standard macroeconomic models tightly connect p-star to the long-run level of the real interest rate (i.e., r-star). In these models, a lower level of p-star implies a lower level of r-star, and a higher level of p-star implies a higher r-star (see Lunsford, 2017). However, based on post-1960 data, Lunsford found no statistical evidence supporting the link between p-star and r-star.

Several papers have endeavored to estimate the long-run level of productivity growth using various statistical and econometric models. To extract more precise and timely estimates of p-star, various authors (e.g., Kahn and Rich, 2007; Roberts, 2001) have proposed using additional variables alongside labor productivity (e.g., real compensation, real consumption, and average hours worked). On the other hand, Edge, Laubach, and Williams (2007) show that estimates of long-run productivity growth obtained from a simple trend-cycle univariate model solved with the Kalman filter do an adequate job of mirroring the long-run projections of productivity growth reported in the SPF and institutional forecasts (e.g., CBO).

On closer inspection, the ability of the Kalman filter to echo the predictions of the professional forecasters is not surprising. Productivity growth is a notoriously volatile series and is subject to extreme revisions from one vintage to another. So, distinguishing highly persistent fluctuations from truly permanent changes is a difficult job for professionals and models alike. Jacobs and van Norden (2016) discuss in detail some of these challenges when working with productivity data.

The findings in Lunsford (2017), Edge, Laubach, and Williams (2007), and Jacobs and van Norden (2016) motivate the formulation of a parsimonious structure for the productivity block relative to other blocks of the model. In particular, we abstract from explicit modeling of direct links between p-star and r-star and between p-star and g-star. However, in an alternative specification we allow for the latter, i.e., a direct link between p-star and g-star.[25] Our formulation is richer than that used in the cited literature, as we allow for time-varying parameters, including stochastic volatility.

[23]We note that innovations ε^2_{gdp*} and ε^2_{ogap} are uncorrelated. In an important contribution, Morley, Nelson, and Zivot (2003), who assume a deterministic g-star, show that this assumption matters for estimating potential output. However, Grant and Chan (2017a) show that in their specification, once a stochastic g-star is allowed for, the correlation between ε^2_{gdp*} and ε^2_{ogap} goes to zero. They also show that the model without correlation performs comparably to the model with correlated innovations based on Bayesian model comparison. Accordingly, to keep estimation tractable, we assume uncorrelated innovations.

[24]Labor productivity is defined here as output per hour worked.

[25]We abstract from the direct link between p-star and g-star in the baseline specification because doing so reduces the model fit and notably lowers the precision of the stars' estimates and the other model parameter estimates.

The productivity gap, which is defined as (nonfarm) labor productivity growth[26] (quarterly annualized) less p-star, is modeled as a function of a one-quarter lag in the productivity gap and the contemporaneous cyclical unemployment gap.

$$P_t - P_t^* = \rho^p (P_{t-1} - P_{t-1}^*) + \lambda_t^p (U_t - U_t^*) + \varepsilon_t^p, \;\; \varepsilon_t^p \sim N(0, e^{h_t^p}) \tag{13}$$

where, $|\rho^p| < 1$

The inclusion of the cyclical unemployment gap helps tease out movements in productivity associated with the business cycle.[27]

Galí and van Rens (2021) find a weakening in the correlation between labor productivity and the cyclical indicator, which motivates time variation in the coefficient λ^p.

$$\lambda_t^p = \lambda_{t-1}^p + \varepsilon_t^{\lambda p}, \;\; \varepsilon_t^{\lambda p} \sim N(0, \sigma_{\lambda p}^2) \tag{14}$$

The variance of the error term ε_t^p is allowed to change over time. Allowing for the time variation in the cyclical relationship and the error term lets the model better discriminate the cyclical movements and idiosyncratic movements in productivity from those associated with shifts in p-star.

P-star is modeled as a driftless random walk component, and the variance of the shocks to this component is assumed to be constant.

$$P_t^* = P_{t-1}^* + \varepsilon_t^{p*}, \;\; \varepsilon_t^{p*} \sim N(0, \sigma_{p*}^2) \tag{15}$$

Modeling p-star this way allows it to capture both unobserved and observed factors that are thought to be persistent but hard to measure. In particular, one factor is developments in fiscal policy; for example, high levels of government debt in the longer term tend to crowd out private investment, thereby reducing longer-term productivity growth.

Economic theory posits that the long-run nominal wage inflation equals the sum of long-run productivity growth and long-run price inflation. As discussed later in the wage inflation block, this theoretical restriction defines the law of motion for w-star and constitutes an additional

[26] As discussed in Kahn and Rich (2007), the focus outside of the farm sector is primarily on avoiding short-term transitory volatility in the farm sector that is heavily driven by weather and other nontechnological factors.

[27] The growth in labor productivity (and more generally aggregate productivity) has been shown to be procyclical to some degree (e.g., Roberts, 2001); it typically increases sharply at the onset of recoveries and falls during recessions. However, empirical evidence on the strength and the direction of the cyclical relationship is mixed. This mixed evidence stems from the use of different estimation samples and/or cyclical indicators (employment-based or output-based). For instance, Galí and van Rens (2021), using split sample estimation, illustrate empirically the significant weakening in the correlation between labor productivity and employment, especially post-1984. They find that the relationship has become countercyclical in the past three decades when using employment as the cyclical indicator. But it is slightly procyclical when using output as the cyclical indicator. In an alternative specification we replace cyclical unemployment with the output gap and obtain similar results. Galí and van Rens (2021), using a structural macro model, attribute the weakening procyclicality of labor productivity in part to the increased flexibility of the US labor market post-1984, which has enabled firms to make adjustments at the extensive margin quickly and easily in response to shocks.

channel influencing the dynamics of p-star.

2.6. *Price inflation block*

We use price inflation as measured by the personal consumption expenditures (PCE) price index, the inflation measure that the Federal Reserve targets. Our formulation for the price inflation block closely follows CKP and CCK, combining elements from both of these papers. Specifically, as in CKP, the stationary component, the inflation gap (defined as the deviation of inflation from pi-star),[28] is modeled as a function of the one-quarter lagged inflation gap, unemployment gap, and an error term, whose variance is allowed to vary over time.

The coefficient, ρ^π on the lagged inflation gap, which captures persistence in inflation dynamics, is allowed to vary over time.[29]

$$\pi_t - \pi_t^* = \rho_t^\pi(\pi_{t-1} - \pi_{t-1}^*) + \lambda_t^\pi(U_t - U_t^*) + \varepsilon_t^\pi, \;\; \varepsilon_t^\pi \sim N(0, e^{h_t^\pi}) \tag{16}$$

$$\rho_t^\pi = \rho_{t-1}^\pi + \varepsilon_t^{\rho\pi}, \;\; \varepsilon_t^{\rho\pi} \sim TN(0 - \rho_{t-1}^\pi, 1 - \rho_{t-1}^\pi; 0, \sigma_{\rho\pi}^2) \tag{17}$$

The innovations to the AR(1) coefficient, ρ^π, are truncated so that $0 < \rho_t^\pi < 1$, ensuring that the inflation gap (in equation 16) is stationary at each point in time t.

$$\lambda_t^\pi = \lambda_{t-1}^\pi + \varepsilon_t^{\lambda\pi}, \;\; \varepsilon_t^{\lambda\pi} \sim TN(-1 - \lambda_{t-1}^\pi, 0 - \lambda_{t-1}^\pi; 0, \sigma_{\lambda\pi}^2) \tag{18}$$

λ^π is the slope of the price Phillips curve and is constrained in the interval (-1,0).

The parameter λ estimates the price Phillips curve relationship. There is ample empirical evidence in support of a time-varying price Phillips curve (e.g., Stella and Stock, 2015; CKP), hence our choice of allowing for time variation in the parameter λ^π.

Pi-star is modeled as a driftless random walk component, and the variance of the shocks to this component is assumed to be constant (as in CKP). This latter assumption of homoscedastic errors is in contrast to Stock and Watson (2007), Mertens (2016), and several others. Our choice not to incorporate SV into shocks to pi-star is made to keep the estimation manageable and maintain consistency with our modeling assumptions for the stars.[30]

[28] Modeling inflation in gap form, where the gap is defined as the difference between inflation and a slowly moving trend, was popularized by Cogley and Sbordone (2008) (and Cogley, Primiceri, and Sargent, 2010), and since then has been a widely used approach to modeling inflation in macroeconomic models for policy and forecasting (e.g., Faust and Wright, 2013).

[29] Chan, Koop, and Potter (2013), CKP, and CCK have found strong empirical support for the time variation in the coefficient of the inflation gap. Our results reinforce the empirical importance of allowing for time variation in this coefficient.

[30] Allowing SV in the inflation gap component and not in the trend component is not without precedent. Besides CKP, Chan (2013) is a recent paper modeling SV only in the measurement equation (i.e., cyclical/transitory component). Berger, Everaert, and Vierke (2016) find support for SV in the inflation gap component but weak evidence for SV in the trend component. Our preliminary results indicate similar findings: that adding SV to the pi-star equation neither helps nor hurts the model fit.

$$\pi_t^* = \pi_{t-1}^* + \varepsilon_t^{\pi*}, \;\; \varepsilon_t^{\pi*} \sim N(0, \sigma_{\pi*}^2) \tag{19}$$

Last, as we show next (see equation 20), pi-star is restricted to satisfy the long-run restriction informed by theory.

2.7. ***Wage inflation block***

The long-run equilibrium level of nominal wage inflation (w-star) is the nominal wage growth rate consistent with its fundamentals – p-star and pi-star. As noted earlier, in the long run, economic theory posits that nominal wage inflation equals the sum of the long-run growth rate of labor productivity and the long-run level of price inflation. In other words, in the long run, labor productivity growth is the only fundamental driver of real wages; therefore, price inflation and nominal wage inflation have to adjust relative to each other to maintain the fundamental relationship. In our setup, we impose this relationship to define w-star.

$$W_t^* = \pi_t^* + P_t^* + Wedge_t + \varepsilon_t^{w*}, \;\; \varepsilon_t^{w*} \sim N(0, \sigma_{w*}^2) \tag{20}$$

$$Wedge_t = Wedge_{t-1} + \varepsilon_t^{wlr}, \;\; \varepsilon_t^{wlr} \sim N(0, \sigma_{wlr}^2) \tag{21}$$

Because data for all three – nominal wage inflation, price inflation, and labor productivity growth – come from different sources and so differ in scope and coverage, a time-varying wedge, which is assumed to evolve as an RW process, is added to the above equation. The above equation implies that W^* adjusted for the wedge is approximately equal to the sum of $\pi_t^* + P_t^*$.

Equation (22) relates the nominal wage inflation gap – defined as the difference between nominal wage inflation and w-star – to its one-quarter lagged gap, the cyclical unemployment gap, and the price inflation gap. The variance of the error term, ε_t^w, is allowed to vary over time.

$$W_t - W_t^* = \rho_t^w(W_{t-1} - W_{t-1}^*) + \lambda_t^w(U_t - U_t^*) + \kappa_t^w(\pi_t - \pi_t^*) + \varepsilon_t^w, \;\; \varepsilon_t^w \sim N(0, e^{h_t^w}) \tag{22}$$

The findings in Knotek II and Zaman (2014) motivate the inclusion of a one-quarter lagged nominal wage inflation gap, with time variation in the parameter ρ^w; the latter quantifies the persistence in wage inflation dynamics.

$$\rho_t^w = \rho_{t-1}^w + \varepsilon_t^{\rho w}, \;\; \varepsilon_t^{\rho w} \sim TN(0 - \rho_{t-1}^w, 1 - \rho_{t-1}^w; 0, \sigma_{\rho w}^2) \tag{23}$$

The innovations to the AR(1) coefficient, ρ^w, are truncated so that $0 < \rho_t^w < 1$, to ensure that the wage gap (in equation 22) is stationary at each point in time t.

The parameter λ^w in equation (22) measures the strength of the cyclical relationship between the nominal wage gap and labor market slack (aka the wage Phillips curve). Many studies, both theoretical (e.g., Galí, 2011) and empirical (e.g., Knotek II and Zaman, 2014; Peneva and Rudd, 2017; Galí and Gambetti, 2019), have documented strong support for the existence of a wage Phillips curve in the US data. These studies have also demonstrated the instability of the wage Phillips curve, motivating the need for time variation in the parameter λ^w.[31]

$$\lambda_t^w = \lambda_{t-1}^w + \varepsilon_t^{\lambda w}, \;\; \varepsilon_t^{\lambda w} \sim TN(-1 - \lambda_{t-1}^w, 0 - \lambda_{t-1}^w; 0, \sigma_{\lambda w}^2) \tag{24}$$

where λ^w is the slope of the wage Phillips curve and is constrained in the interval (-1,0).

As discussed earlier in the price inflation block, both theory and empirical evidence point to the connection between price inflation and nominal wage inflation. The standard fully structural models describing the New Keynesian Phillips curve posit a tight relationship between price and wage inflation via the channel of current and expected future marginal costs. In these models, price inflation today is a function of expected price inflation and expected future marginal costs, where marginal costs are generally linked to wages. Knotek II and Zaman (2014) provide empirical evidence of the connection between nominal wage and price inflation. In particular, they show no clear evidence of one Granger-causing the other; instead, both wage and price inflation generally tend to move together. This reasoning would suggest the importance of modeling the direct relationship between wage inflation and price inflation. Hence, this motivates the inclusion of the price inflation gap in the measurement equation (22).

Several studies document a significant weakening in the empirical link between price inflation and nominal wage inflation since the 1980s (e.g., Peneva and Rudd, 2017; Knotek II and Zaman, 2014), motivating time variation in the parameter κ^w. The expression $\frac{\kappa_t^w}{1-\rho_t^w}$ could be interpreted as an estimate of the short-run pass-through from price inflation to wage inflation.

$$\kappa_t^w = \kappa_{t-1}^w + \varepsilon_t^{\kappa w}, \;\; \varepsilon_t^{\kappa w} \sim N(0, \sigma_{\kappa w}^2) \tag{25}$$

2.8. ***Interest rate block***

We close the model with the interest rate block characterizing the interest rate dynamics and the law of motion for r-star (the long-run equilibrium real short-term interest rate).

Our first equation of the block brings information from the nominal short-term interest rate via a Taylor-type rule (TR) to aid in identifying r-star. Specifically, this equation characterizes the dynamics of the short-term nominal interest rate gap, where the gap is the difference between the nominal short-term interest rate i, and the long-run level of the nominal neutral rate of interest, i-star. (i-star = pi-star + r-star). When modeling the nominal short-term interest

[31] The literature has posited various explanations for the instability of the wage Phillips curve, including downward nominal wage rigidities, where the degree of rigidity varies with the phase of the business cycle (see Daly and Hobijn, 2014).

rate, especially in a framework like ours, one must account for the effective lower bound (ELB) period.

The recent literature provides at least two options for handling the ELB. The first is to explicitly but separately model the observed short-term nominal rate, which cannot go below zero, and the "shadow interest rate," which is a hypothetical unobserved and unbounded counterpart. Wu and Xia (2016) popularized the concept of the shadow interest rate, and JM and González-Astudillo and Laforte (2020) are two recent approaches well suited for inclusion in UC models. The second approach is to treat the estimate of the "shadow rate" obtained from Wu and Xia (2016) as the measure of the short-term nominal interest rate in measurement equations such as the TR (e.g., Pescatori and Turunen, 2016).[32]

Given our model's size and complexity, we adopt the latter approach, which is simpler, though not perfect. Using a direct measure of the nominal shadow rate allows us to capture both conventional and unconventional monetary policy effects when the (observed) nominal federal funds rate is constrained at the ELB.[33]

Equation (26) relates the nominal interest rate gap (based on the shadow federal funds rate) to its one-period lagged interest rate gap, the current quarter inflation gap (i.e., the deviation of inflation from pi-star), and the unemployment rate gap (i.e., the deviation of the unemployment rate from u-star). This equation roughly characterizes the monetary policy reaction function as defined by Taylor (2001).[34] There is a broad consensus that policy adjustments outside of cyclical turning points are made very gradually. Hence, this motivates the inclusion of the lagged interest rate gap term.

Chan and Eisenstat (2018a,b) and JM document strong empirical support for constant parameters in the Taylor rule equation while allowing for stochastic volatility in the errors. Accordingly, we allow for SV in the interest rate equation. JM, González-Astudillo and Laforte (2020), and Brand and Mazelis (2019) document the usefulness of adding the TR equation to identify r-star. The latter two do not entertain SV, which JM has found to be empirically important. As discussed later, we also found that adding the TR equation improves the precision of the r-star estimates significantly, and the data strongly favor allowing for SV in the error process.

$$i_t - \pi_t^* - r_t^* = \rho^i(i_{t-1} - \pi_{t-1}^* - r_{t-1}^*) + \lambda^i(U_t - U_t^*) + \kappa^i(\pi_t - \pi_t^*) + \varepsilon_t^i, \;\; \varepsilon_t^i \sim N(0, e^{h_t^i}) \tag{26}$$

where ρ^i is truncated so that $0 < \rho^i < 1$.

[32]The estimates from Wu and Xia (2016) are publicly available and regularly updated. Treating the shadow rate as the measure of the short-term nominal rate in place of the federal funds rate is commonly done, and often academic papers report results indicating robustness to the use of Wu and Xia's shadow rate (e.g., Beyer and Wieland, 2019; Lewis and Vazquez-Grande, 2019)

[33]The nominal *shadow* federal funds rate is identical to the nominal federal funds rate when the effective lower bound is not binding.

[34]It is worth emphasizing that we denote this equation as a "Taylor-type rule" and not an exact Taylor rule because in our equation, pi-star refers to the estimate of trend inflation, which may or may not be equal to the central bank's long-run inflation goal.

Our second equation motivated by LW heeds the economic theory suggesting the role of various real factors in influencing movements in r-star. These factors include long-run output growth (and long-run productivity growth), trend labor force growth (reflecting shifts in demographics and net migration), taxation structure, government expenditure shifts, and shifts in liquidity preferences (e.g., Del Negro et al., 2017). Accordingly, equation (27) expresses r-star as a linear function of g-star and a "catch-all" component D. In our baseline specification, both g-star and D follow random walk processes similar to LW (and many other papers).[35]

$$r_t^* = \zeta g_t^* + D_t. \tag{27}$$

$$D_t = D_{t-1} + \varepsilon_t^d,\ \ \varepsilon_t^d \sim N(0, \sigma_d^2) \tag{28}$$

All in all, information from six sources and/or elements informs the econometric identification of r-star. These sources include: an IS equation (12); a TR equation (26), which allows for SV; an equation linking r-star to survey expectations; the shadow rate; and an equation relating r-star to g-star. As we will show shortly, all of these sources play a role in improving r-star's precision. To reiterate, in our framework, we use information from both short-term interest rates (via a TR equation) and long-term interest rates (via an IS equation) to inform the estimation of r-star.[36]

2.9. *Stochastic volatility in "gap" equations*

As discussed earlier, the variance of the error terms, ε_t^u, ε_t^{ogap}, ε_t^p, ε_t^π, ε_t^w, and ε_t^i is allowed to vary over time. This implies that the model permits changing size of the shocks to processes defining the cyclical unemployment gap, cyclical output gap, productivity gap, price inflation gap, nominal wage inflation gap, and nominal interest rate gap.

Following a long list of papers, we define the SV process as a driftless random walk in the log-variance.

$$h_t^{id} = h_{t-1}^{id} + \varepsilon_t^j,\ \ \varepsilon_t^j \sim N(0, \sigma_j^2) \tag{29}$$

where $id = \{u, ogap, p, \pi, w, i\}$, and $j = \{hu, ho, hp, h\pi, hw, hi\}$

[35] The RW assumption for D is an appropriate one, given that our focus is the long-run r-star that should, in principle, be influenced over time by permanent shifts in aggregate supply and demand (Laubach and Williams, 2016). Researchers have also explored an AR process for component D, which would be consistent if the interest is in medium-term r-star (see Lewis and Vazquez-Grande, 2019), as this would allow r-star to be influenced by the transitory shocks to aggregate demand (via the AR process) and permanent shocks to aggregate supply (via the RW process for g-star). In studies focused on the long-run notion of r-star, such as LW, Laubach and Williams (2016), Clark and Kozicki (2005), and Kiley (2020), specification based on the RW assumption has been shown to be empirically favored by the data compared to the AR assumption.

[36] Del Negro et al. (2017), JM, Bauer and Rudebusch (2020), and González-Astudillo and Laforte (2020) are recent studies that have highlighted the usefulness of exploiting information from both short-term and long-term interest rates in the identification of r-star.

2.10. ***Base model and its variants***

The equations (4), (5)...(29) define our baseline model formulation (denoted *Base*). Figure 1 provides a visual representation of our Base model. And Section A1.a. of the online appendix lists all of the equations for the Base model for easy reference. To assess the usefulness of survey information in the econometric estimation of our multivariate UC model, we also estimate a variant of the baseline model that excludes the equations linking long-run survey expectations to stars (i.e., excluding equations 4 and 5). We denote the latter specification as *Base-NoSurv*. The model specifications Base and Base-NoSurv constitute our two main model specifications. To assess the empirical support for numerous additional features (informed by theory and past empirical research) embedded in our modeling framework, we formulate several additional model specifications, each of which is a restricted variant of the Base. To keep the length of the paper manageable, we report selected results from the auxiliary model specifications in the main part of the paper with additional results included in the online appendix. Below, we summarize the description of the additional model variants.[37]

Base-NoSurv. To assess the usefulness of survey expectations, we estimate a variant of the baseline model that excludes all of the equations linking surveys to stars.

Base-NoSV. To assess the empirical support of stochastic volatility in shock variances, we estimate a variant of the baseline model with no SV in any of the measurement (gap) equations.

Base-W*RW. To assess the empirical support of the theoretical restriction defined by equation 20 (which defines w-star as the sum of pi-star and p-star), we estimate a variant of the baseline model that replaces equation 20 with a random walk assumption for w-star as defined by equation 20b.

$$W_t^* = W_{t-1}^* + \varepsilon_t^{w*}, \;\; \varepsilon_t^{w*} \sim N(0, \sigma_{w*}^2) \tag{20b}$$

Base-R*RW. To assess the empirical support for the theoretical restriction defined by equation 27 (the link between g-star and r-star), we estimate a model specification that replaces equation 27 with a random walk assumption for r-star as defined by equation 27b.

$$r_t^* = r_{t-1}^* + \varepsilon_t^{r*}, \;\; \varepsilon_t^{r*} \sim N(0, \sigma_{r*}^2) \tag{27b}$$

Base-NoLinkStars. To assess the empirical support of both theoretical restrictions defined by equations 20 and 27, we estimate a variant of the baseline model that combines *Base-W*RW* and *Base-R*RW*.

Base-G*LinkP*. To assess the empirical support for the theoretical link between g-star

[37] In the working paper version, we list several more specifications, but to keep the length of this paper manageable, here we focus on a smaller number of specs.

and p-star, we estimate a model specification that replaces equations 10 and 11 with equations 10b, 11b, and 11c.

$$gdp_t^* = gdp_{t-1}^* + g_t^* + \varepsilon_t^{gdp*}, \ \ \varepsilon_t^{gdp*} \sim N(0, \sigma_{gdp*}^2) \tag{10b}$$

$$g_t^* = \psi p_t^* + g_t^{other*} \tag{11b}$$

$$g_t^{other*} = g_{t-1}^{other*} + \varepsilon_t^{gother*}, \ \ \varepsilon_t^{gother*} \sim N(0, \sigma_{gother*}^2) \tag{11c}$$

Equation 11b expresses g-star as a linear function of p-star and a "catch-all" component g^{other*}, which captures the influence on g-star of all factors other than p-star. The parameter ψ captures the strength of the relationship between trend growth and trend productivity.

Base-NoBoundU*. To assess the empirical support for imposing bounds on the U*, we estimate a model specification without the bounds on U* process defined in eq. 8.

Base-PT-Wages-to-Prices. To assess the empirical support of allowing for pass-through from wages to prices, we estimate a model specification that replaces eq. 16 with eq. 16b, which adds the nominal wage inflation gap as an explanatory variable in the equation describing the price inflation gap. The parameter γ^π captures the strength of the relationship between the two cyclical inflation measures. The expression $\frac{\gamma^\pi}{1-\rho^\pi}$ can be interpreted as the pass-through from cyclical wage inflation to cyclical price inflation.[38]

$$\pi_t - \pi_t^* = \rho_t^\pi(\pi_{t-1} - \pi_{t-1}^*) + \lambda_t^\pi(U_t - U_t^*) + \gamma^\pi(W_t - W_t^*) + \varepsilon_t^\pi, \ \ \varepsilon_t^\pi \sim N(0, e^{h_t^\pi}) \tag{16b}$$

Base-NoPT. The Base model allows for pass-through from prices to wages. We assess the empirical support of this restriction by estimating a model specification that replaces eq. 22 with eq. 22b, which removes the price inflation gap in the equation describing the nominal wage inflation gap.

$$W_t - W_t^* = \rho_t^w(W_{t-1} - W_{t-1}^*) + \lambda_t^w(U_t - U_t^*) + \varepsilon_t^w, \ \ \varepsilon_t^w \sim N(0, e^{h_t^w}) \tag{22b}$$

3. Data and Bayesian Estimation

3.1. ***Data***

We estimate the empirical model using the following quarterly data: (1) the unemployment rate; (2) real GDP growth; (3) nonfarm labor productivity growth; (4) the inflation rate in personal consumption expenditures (PCE) price index; (5) average hourly earnings (AHE) of production and nonsupervisory workers (total private industries);[39] (6) the federal funds rate;

[38] We explored the possibility of allowing for time variation in γ^π but the estimation ran into difficulties; hence, we resort to time-invariant γ^π.

[39] Average hourly earnings (AHE) of production and nonsupervisory workers in total private industries goes back to 1964Q1. From 1959Q4 through 1963Q4, we use the AHE of production and nonsupervisory workers in

(7) the nominal yield on the 10-year Treasury bond; (8) the shadow federal funds rate from Wu and Xia (2016); (9) Blue Chip[40] (real-time) long-run projections of the three-month Treasury bill, real output growth, the unemployment rate, and GDP deflator inflation; (10) long-run inflation expectations of PCE inflation (PTR series). We also collect the real-time long-run CBO projections of real output growth, the level of real potential output, and the natural rate of unemployment. For forecast evaluation exercises, the real-time data vintages of real GDP growth, PCE inflation, the unemployment rate, AHE, and nonfarm labor productivity spanning 1998Q1 through 2019Q4 are downloaded from the ALFRED database maintained by the St. Louis Fed and the real-time database maintained by the Federal Reserve Bank of Philadelphia. For the data series labeled (1) through (7), which comprises our core data set, we collect two vintages of revised data: 2020Q2 and 2020Q4 vintages, respectively. We use data starting in 1959Q4 through 2019Q4 from the 2020Q2 vintage (which includes the third estimate of 2019Q4) as a featured sample for this paper. To show the implications of the COVID-19 data for our model estimates, we estimate our model(s) using the 2020Q4 vintage, which has data spanning 1959Q4 through 2020Q3. The vintages corresponding to the revised data are downloaded from Haver Analytics.

3.2. ***Bayesian estimation***

We use Bayesian estimation methods to fit our Base model and its variants. The use of inequality restrictions on latent parameters in our model(s) setup leads to a nonlinear state-space model, which renders estimation using standard Kalman filter methods infeasible. Accordingly, we implement our Markov chain Monte Carlo (MCMC) posterior sampler based on computational methods developed in Chan, Koop, and Potter (2013) and CKP, who use the band and sparse matrix algorithms detailed in Chan and Jeliazkov (2009). The CKP posterior sample developed for a relatively smaller-scale nonlinear state-space model is carefully extended to accommodate the additional structure and numerous features of our model(s). Since the computational methods used in this paper are based on CKP, we relegate the specific details of the sampler to online appendix A1.

In a methodological sense, this paper's novelty is in assembling the existing sampling algorithms based on the fast band and sparse matrix routines to solve a large nonlinear and a high-dimensional UC model. We found that the use of inequality restrictions, such as bounds on the u-star and other parameters, is crucial to estimating the model, especially in the Base-NoSurv model. Intuitively, features such as truncated distributions that we implement for some of the time-varying parameters, e.g., the Phillips curve (price and wage), persistence, and bounds on u-star, facilitate estimation by guiding the estimation procedure to the credible regions of the parameter space.

For each model, we simulate 1 million posterior draws from the MCMC posterior sampler.

goods-producing industries. We splice them together.

[40]Blue Chip Economic Indicators data published by Wolters Kluwer Legal and Regulatory Solutions U.S.

We then discard the first 500,000 draws, and of the remaining, we keep every 100th draw. Accordingly, all of the reported results for the Base model and its variants are based on 5000 retained draws.[41]

Bayesian model comparison is based on the marginal likelihood metric. In computing the marginal likelihood for various models, we use the approach proposed by CCK, which decomposes the marginal density of the data (e.g., inflation) into the product of predictive likelihoods; see appendix A1.d for details. An important advantage of the CCK approach is that it allows us to separately compute marginal data density for each variable of interest: inflation, nominal wages, interest rate, real GDP, the unemployment rate, and labor productivity. The variable-specific marginal densities prove useful because they allows for deeper insights into the source of the deficiencies, which helps differentiate models at a more granular level.

We note that our prior settings are similar to those used in CKP, CCK, and González-Astudillo and Laforte (2020). As discussed in CCK, UC models with several unobserved variables, such as the one developed in this paper, require informative priors. That said, our prior settings for most variables are only slightly informative. The use of inequality restrictions on some parameters such as the Phillips curve, persistence, and bounds on u-star could be viewed as additional sources of information that eliminate the need for tight priors, something also noted by CKP. For the parameters on which there is strong agreement in the empirical literature on their values, such as the Taylor-rule equation parameters, we use relatively tight priors, such that prior distributions are centered on prior means with small variance. In model comparison exercises, the priors are kept the same for the common parameters across models. We also perform prior sensitivity analysis, which is reported in appendix A2.

4. Full Sample Estimation Results

This section discusses the results obtained by estimating the models using the sample from 1959 through the end of 2019.

The importance of SV in model equations

Our results provide strong evidence of the importance of allowing for SV in all six equations describing the gap components. Figure 2 plots the posterior mean estimates of the time-varying standard deviation of the innovations to the unemployment rate gap, output gap, labor productivity gap, price inflation gap, nominal wage inflation gap, and the nominal interest rate gap. Also, plotted are the corresponding 90% credible intervals. The plots indicate statistically

[41] In the appendix, we report the efficiency diagnostics of our MCMC algorithm. Those diagnostics, which include inefficiency factors and convergence metrics, indicate good convergence properties (and low autocorrelation) of our sampler for both the Base and the Base-NoSurv models. Regarding computational time, given the high dimensionality of our model(s) and the number of posterior simulations we require, the speed is quite fast (in our assessment). When applied to the Base model, the MCMC algorithm, which is implemented in Matlab, takes about 350 seconds to generate 10,000 posterior draws using a laptop computer with an Intel(R) Xeon(R) E-2176M CPU @ 2.70 GHz processor. To generate 1 million posterior draws, it takes less than 10 hours.

significant evidence of time variation in the volatility of the idiosyncratic components in all of the gap equations. For example, in the case of the price inflation gap, estimates imply high volatility during the period of the Great Inflation that fell subsequently. Inflation volatility increased sharply again during the Great Recession but has trended lower since then. By 2019, inflation volatility had returned to the low levels of the early 2000s but it remains shy of the historic lows of the mid-1960s and mid-1990s. The model comparison further provides evidence supporting SV inclusion, as is evident from the Base-NoSV model's significantly inferior fit to the data compared to the Base and model variants that allow for SV (see Table 1).

Time variation in parameters defining macro relationships

Figure 3 presents the time-varying posterior mean estimates (and the 90% credible intervals) of the parameters from the Base model describing the persistence in the price and nominal wage inflation gaps, the price and wage Phillips curves, the short-run pass-through from prices to wages, and the cyclical dynamics of labor productivity. A quick visual inspection indicates both statistically and economically significant evidence of time variation in these parameters that capture important empirical relationships.

Price inflation gap persistence (Panel a). There is strong evidence of time variation in parameter ρ^{π}, inflation gap persistence. For example, gap persistence was low (0.25) in early 1960, but from there on began to increase steadily, reaching close to 0.83 by early 1970, and remained at that level through the early 1980s. Subsequently, the persistence had declined steadily to 0.3 by the early 2000s, and it remained close to that level through the end of 2019.

Nominal wage gap persistence (Panel b). The posterior mean estimate of the parameter ρ^{w}, capturing the persistence in the nominal wage inflation gap, indicates increasing persistence in the wage inflation gap beginning in early 1960 and peaking in mid-1980. From the mid-1980s to the early 1990s, the persistence steadily declines but after that increases through the mid-2000s; from there on through the early 2010s, the estimated persistence in the nominal wage gap falls to levels seen in the mid-1970s. Since then, it has been slowly increasing. It is worth noting that the credible intervals around the posterior mean are wide, suggesting high uncertainty in the inference about the estimated persistence.

Price Phillips curve (Panel c). The plot indicates strong evidence of time variation in the slope of the Phillips curve. For example, the model estimates a steeper Phillips curve in the 1960s that subsequently weakens (becomes less negative) over time through 2010. From there on, it slowly begins to become steeper (more negative), ending 2019 at -0.23, which is still weak, historically speaking, and is surrounded by wide intervals spanning -0.05 to -0.52.

Wage Phillips curve (Panel d). The plot provides strong evidence supporting the existence of the wage Phillips curve in the post-war data. Notably, the plot also offers convincing evidence of time variation in this relationship. According to the posterior mean estimate, from the early 1960s through the early 1980s, the model implies that the strength of the wage Phillips curve is at a moderate level, but from there on through the mid-1980s, the wage Phillips curve steepened sharply. By the mid-1980s, the posterior mean of the wage Phillips curve parameter is estimated to be -0.5, with 90% credible intervals ranging from -0.1 to -0.7. The estimated strength in the relationship remained at that level through the mid-2000s. From there on, it began to flatten rapidly until early 2010. In 2010, the model estimates the posterior mean of the Phillips curve parameter at -0.25. The prevalence of the downward wage rigidities during the Great Recession is among the primary explanations for the flattening of the wage Phillips curve (see Daly and Hobijn, 2014). From 2010 onward, with an improving economy, the estimated wage Phillips curve has steadily steepened. Our empirical evidence on the wage Phillips curve is consistent with the findings of Peneva and Rudd (2017) and Galí and Gambetti (2019).

Short-run pass-through from prices to wages (Panel e). The posterior estimate of pass-through (defined by $\frac{\kappa_t^w}{1-\rho_t^w}$) indicates a weakening relationship between cyclical nominal wage inflation and cyclical price inflation over the estimation sample, confirming the evidence presented in Peneva and Rudd (2017). The relationship between the two was strong in the 1970s to the mid-1980s, but since then, it has gradually weakened such that it has been nonexistent (i.e., the pass-through is estimated to be zero) for the past decade. This period of the breakdown in the relationship between the two cyclical components has coincided with a period of low and stable price inflation. The literature has offered various explanations for this breakdown in the relationship, including an improved anchoring of inflation expectations (Peneva and Rudd, 2017) and an amplification of downward wage rigidities during low levels of price inflation (Daly and Hobijn, 2014).

Cyclical dynamics of labor productivity (Panel f). The plot of the estimate of the parameter λ^p, which relates cyclical unemployment to the productivity gap, indicates a high level of uncertainty around the estimate of λ^p. The 90% credible intervals are wide, such that they include both positive and negative values, complicating reliable inference. Going just by the posterior mean estimate, the evidence suggests a countercyclical behavior of labor productivity, with this relationship weakening over time.[42] We also explored the possibility that the SV may be soaking up the variation in productivity, which otherwise would have been attributed to the cyclical component of productivity. The model specification Base-NoSV, which shuts down SV in the idiosyncratic component of productivity (and other model equations), yields estimates of λ^p similar to Base, suggesting that SV is not contributing to the ambiguous result on the cyclicality of labor productivity.

Links between survey expectations and stars

Our model estimates indicate that for the four stars, where the model structure permits information from survey expectations, the estimated relationships between survey expectations and stars are found to be strong, suggesting an influential role of survey expectations in

[42]We explored an alternative model specification, Base-P*CycOutputgap, which relates productivity to the output gap instead of the unemployment gap and the results indicate generally similar inference about the cyclical nature of labor productivity (corroborating the evidence presented in Galí and van Rens (2021)). Please refer to appendix A13.

econometric estimation of the model-based stars. For example, the model estimation yields a posterior mean estimate of 0.99 for β^{π}, with 90% credible intervals spanning 0.92 to 1.07 (also reported in Table 2). The estimates of the time-varying intercepts (e.g., C_t^{π}) indicate evidence of considerable time variation in the estimated relationship between the survey expectations and model-based stars. Figure A6 in the online appendix plots the posterior estimates of the coefficients capturing the estimated relationship between survey expectations and model-based stars.

Bayesian model comparison

As discussed in the next section, bringing in additional information from surveys leads to more reasonable and precise estimates of stars; however, the Bayesian model comparison indicates comparable support in the data for both Base and Base-NoSurv. And both models have marginally higher support over Base-NoLinkStars, as shown in Table 1. The breakdown of the marginal data density by variables suggests that the Base model has a better fit to the nominal interest rate data compared to Base-NoSurv, but that improved fit is offset by its inferior fit to the unemployment rate, resulting in comparable overall fit across the two models. As noted above, the model specification without SV has a significantly inferior fit to the data compared to the other three specifications.

Next, we sequentially discuss (full sample) estimates for each of the six stars.

4.1. ***Estimation results for u-star***

Figure 4 plots the evolution of u-star (and its uncertainty) covering the period 1960 through 2019. Panel (a) plots the posterior estimates from the Base model and panel (b) from the Base-NoSurv model. Also plotted are the corresponding 90% credible intervals. In the past six decades, the (posterior mean of) u-star has fluctuated between 4.0% and 7.0%, peaking in the early 1980s and troughs at the end of our sample period. The contours of u-star from both models are generally similar; however, the level of u-star can differ notably in some periods. From 1960 through the late 1970s, u-star has gradually increased (from 5.4% to 6.5% in Base and 5.3% to 7.1% in Base-NoSurv). But since the mid 1980s through the late 1990s, u-star has steadily drifted lower (to 5.0%). This downward trend in the later period is also documented in the u-star literature based on job-flows data, which attributes the decline in u-star to declining trends in job-separation and job-finding rates (e.g., Crump et al., 2019).

From early 2000 through early 2010, u-star has slowly trended higher, with a sharp pickup during the Great Recession period. Since 2010, u-star has steadily drifted lower. By the end of 2019, both Base and Base-NoSurv have u-star at 4.2% (with a 90% interval covering 3.7% to 4.9%). As shown in Figure 5, at the end of 2019, the unemployment gap implied by both models is negative, i.e., the unemployment rate is below the estimated u-star.

The use of survey information in the Base model mainly contributes to the difference in the levels of u-star across the two models. To facilitate comparison, panel (a) also plots u-star from the survey. As is evident from the plot, u-star from the survey displays more pronounced shifts in u-star than the model-based estimates. However, due to a strong estimated relationship between the survey u-star and Base u-star (i.e., posterior mean of $\beta^u = 0.95$), the Base estimate of u-star reflects the contours in survey u-star. As can be seen in comparing the width of the credible bands between panels (a) and (b), taking survey information on board improves the precision of u-star notably.

Surprisingly, based on the Bayesian model comparison, the Base model has an inferior fit to the unemployment data compared to Base-NoSurv. As shown later, this result contrasts with the results for pi-star, r-star, and g-star, for which survey information helps improve the model fit or at least does not worsen the fit.

Sensitivity of u-star to modeling assumptions including information set

We explore several variants of the Base model to examine the sensitivity of u-star to modeling assumptions and the informational aspect of joint modeling. We also compare our model estimates to the u-star estimate from the CBO. For the sake of brevity we have included the results and discussion in online appendix A13.a, but here we highlight one noteworthy finding related to the implementation of bounds on u-star. We find that using bounds on u-star is extremely important in the Base-NoSurv model, as it helps keep the estimation tractable. At the time of writing this paper, the world was hit with a COVID-19 shock, an extreme and unprecedented global health shock along various dimensions, leading several analysts to call it a "once-in-a-lifetime upheaval." As we will show in Section 6, the implementation of bounds on u-star is part of the story in preventing our models from blowing up in response to COVID-19 data.

Cyclical unemployment

Figure 5 presents the posterior mean estimate of the unemployment gap and the corresponding 90% credible intervals. The top panel plots the estimates from the Base model, and the bottom panel from the Base-NoSurv model. A visual inspection indicates that the movements in cyclical unemployment correspond quite well with the NBER's business cycle dating. For instance, cyclical unemployment falls in economic expansions and rises during recessions. Both models show a significant spike in the cyclical unemployment rate in the 1982-83 and 2007-09 recessions, and a sharper recovery following the 1982-83 recession but a more gradual recovery following the Great Recession. The figure also highlights that both models produce similar estimates of cyclical unemployment.

4.2. ***Estimation results for g-star and the output gap***

Panel (a) in Figure 6 plots the g-star estimates from Base, Base-NoSurv, and the univariate model (of Grant and Chan, 2017b), and panel (b) plots the corresponding precision. As is evident, g-star estimates from our two main models indicate a steady decline throughout the sample, except a temporary rise in the late 1990s, which the literature has attributed to the technology boom. According to the posterior mean estimates of g-star from our Base and Base-NoSurv models, the growth rate of potential output has continuously drifted lower from an annualized rate of close to 4.5% in early 1960 to 1.2% by 2012. Thereafter, it gradually moves up reaching 1.9% (1.7% in Base-NoSurv) by the end of 2019.[43] From 1960 through early 1980, the story is generally similar based on the inference from the univariate model, which could be viewed as a nested specification of the Base model. However, from 1980 onward, the inference from the univariate model is notably different.

Not surprisingly, the precision of the g-star estimates displays patterns that align well with intuition. For instance, the Base model that incorporates survey expectations yields more precise estimates than specifications that ignore survey data. Base and Base-NoSurv models that relies on multivariate information generate significantly more precise estimates than the univariate model. Table 3 reports the assessment of model fit to the GDP data for the various model specifications. As is evident, Base model has the best fit to GDP data. The model specification denoted Bivariate, which builds on the univariate GDP model by adding the unemployment rate, has a better fit than the univariate model. And adding SV to univariate and bivariate models significantly improves their respective fits to the GDP data.

Output gap estimates: Base vs. CBO and LW

Next, we examine the estimates of the output gap. Panel (c) in Figure 6 presents the output gap estimates from the Base model, the CBO, and the LW model. A few observations immediately stand out. First, the estimate from the Base model accords well with the NBER recession dates and lends support to the business cycle asymmetry, in that recessions are shorter in duration but deeper than expansions in the US (Morley and Piger, 2012). It is instructive to highlight that estimates imply a more negative output gap (of -8.8%; posterior mean) during the 1981-82 recession compared to the Great Recession period (-7%) when output fell more dramatically. At a first pass, this may seem odd. But a closer inspection reveals that in comparison to the 1981-82 recession, during the Great Recession, g-star fell significantly (as can be seen in panel a), resulting in a smaller negative output gap; in contrast, during the 1981-82 recession, g-star is estimated to have remained stable. Second, the output gap estimate implied from the LW model is notably different over the second half of the sample period. In particular, during the Great Recession period, the output gap from LW turned slightly negative, while other estimates implied larger negative gaps. The slight negative gap in the LW model is the result of the LW model estimating a dramatic fall in potential output, in line with the collapse in actual output.

[43]This continuous reduction in the growth rate of potential output has been extensively documented elsewhere (e.g., Berger, Everaert, and Vierke, 2016; Grant and Chan, 2017b; Coibion, Gorodnichenko, and Ulate, 2018) and in particular the decline since 2009 has been of great concern among policymakers. Several other researchers, including Pescatori and Turunen (2016); Laubach and Williams (2016); Antolin-Diaz, Drechsel, and Petrella (2017); Holston, Laubach, and Williams (2017), have also documented the secular decline in g-star over the past two decades.

Third, with few exceptions, the estimate from the Base model is generally similar to the CBO's estimate. This close similarity is notable because the CBO's approach is based on an entirely different methodology, a production function approach (see Shackleton, 2018). However, similar to our framework, the CBO also relies on multivariate information to infer the output gap. Our supplementary analysis suggests that joint modeling of real GDP and the unemployment rate and allowing for SV in their respective cyclical components are the key ingredients to obtaining a credible output gap estimate (i.e., resembling the CBO's estimate).[44]

Our paper's result indicating a close resemblance of our models' output gap estimates to the CBO's output gap provides evidence supporting the common practice of using output gap estimates from the CBO as an exogenous variable in empirical macroeconomic models (e.g., JM; Stock and Watson, 2020). We view this result as a useful contribution to the applied macroeconomics literature.

Posterior parameter estimates for the output block

Next, we discuss the Base model's parameter estimates of the output block that drive the dynamics of g-star and the output gap. The posterior mean estimates of parameters ρ_1^g and ρ_2^g indicate a high degree of persistence ($\rho_1^g + \rho_2^g = 0.73$) and suggest a hump-shaped response of the output gap to shocks (as $\rho_1^g > 1$). These parameters are precisely estimated as evidenced by tight posterior credible intervals. The posterior estimate of parameter λ^g (the coefficient on the unemployment gap in the output gap equation) is negative and highly significant statistically. The estimated posterior mean of λ^g is -0.48 (with 90% interval -0.61 to -0.35). Similarly, the parameter ϕ^u (the coefficient on the output gap in the unemployment equation), discussed earlier, is also negative and highly significant statistically. Together, these estimates indicate a strong Okun's law relationship in the data. The implied posterior mean estimate of the Okun's law coefficient, $\frac{(1-\rho_1^u-\rho_2^u)}{\phi_u}$ is -2.2, with 90% credible intervals spanning -2.4 to -1.9. This estimated coefficient is strikingly identical to the conventional estimate often discussed in macroeconomic textbooks. Therefore, not surprisingly, both the estimated output gap and the unemployment gap (shown earlier) reveal similar cyclical dynamics. For instance, according to both cyclical measures, the 1981-82 recession is estimated to have been deeper than the Great Recession.

The parameter a^r, which relates the output gap to the real rate gap (characterizing the IS relation), is negative and much smaller than the prior mean. The estimated posterior mean of a^r is -0.04 (with 90% interval -0.09 to -0.00). The posterior mean estimate of $E(\sigma^2_{gdp*})$, the variance parameter of the innovations to the process governing the evolution of g-star (and gdp-star), is nearly identical across the Base and Base-NoSurv models: 0.02^2. For comparison, the

[44] Morley and Wong (2020), who estimate the output gap using a large BVAR, also found that the unemployment rate is the most crucial indicator for the output gap. (In online appendix A8, we compare our model estimates with additional estimates, including Morley and Wong's). Recently, Barbarino, Berge, Chen, and Stella (2020) use a range of small-scale UC models (without SV) to estimate the output gap and similarly find that the unemployment rate is the most valuable indicator.

prior mean $E(\sigma^2_{gdp*})$ is 0.01^2. The estimation results suggest that the data are quite informative in influencing the dynamics of both the output gap and g-star, confirming Kiley (2020).

4.3. *Estimation results for p-star*

Figure 7 presents posterior estimates of p-star and other parameters of the productivity block. Panel (a) presents p-star estimate from the Base model. Both the posterior mean and the 90% credible intervals are shown. Also plotted is the actual labor productivity series. A visual inspection of the actual series indicates the unusually high volatility of the quarterly productivity data. Not surprisingly, researchers have emphasized that these difficulties of extreme volatility, extensive revisions, and real-time measurement issues with productivity data complicate its trend-cycle decomposition (e.g., Edge, Laubach, and Williams, 2007; Kahn and Rich, 2007).

Our model-based estimates reflect these challenges. For instance, the estimate of the parameter ρ^p, reported in Table 2, indicates close to zero persistence in the labor productivity data, defined as the difference between the growth rate in labor productivity and p-star. Similarly, our estimation indicates that the labor productivity data have very little influence on the estimate of p-star. Put differently, the data are too volatile to allow for a meaningful identification of trend in the productivity data. The posterior mean of $E(\sigma^2_{p*})$, the variance of the shock process for p-star, is essentially the same as the prior mean.[45] As a result, the degree of time variation in p-star is primarily influenced by the prior setting. So conditional on our prior belief, which allows p-star to evolve slowly from one quarter to the next, we find considerable evidence of gradual time variation in p-star over the post-war sample. The evidence of time variation is economically significant and is consistent with the findings of Roberts (2001), Benati (2007), Edge, Laubach, and Williams (2007), and Fernald (2007).

Panel (b) presents posterior mean estimates of p-star from the Base-NoSurv and and Base-W*RW models. The latter model removes the restriction that the long-run w-star grows at a rate equal to the sum of pi-star and p-star. So removing this restriction eliminates the direct influence on p-star from wages and prices. As evident from the Bayesian model comparison reported in Table 4, the elimination of this restriction marginally improves the fit of the model to the productivity data but significantly reduces the overall model fit to other data, particularly interest rates – via the changes in pi-star and u-star in the Taylor-type rule equation. That said, all three models generally indicate similar broad patterns in p-star.

After averaging between 2% and 3% in the 1960s, the models indicate that p-star experienced a sharp deceleration in the 1970s through the mid-1980s, mirroring the dramatic fall in productivity growth. Both Base and Base-NoSurv estimates show p-star trending lower from 2.4% (2.3%) in early 1970 to 1.2% by the mid-1980s, whereas Base-W*RW has it falling close to 1.5%, with wide 90% credible intervals that range from 0.6% to 2.3%. From there on through to the late 1990s, p-star increased sharply, at a pace roughly equivalent to its deceleration in

[45] We tried different values for the prior mean on this parameter and found that the posterior moves with the prior.

prior periods, to reach a level of 2.4% by 1999. The literature attributes part of this acceleration in the latter half of the 1990s to the information technology boom. Roberts (2001), Edge, Laubach, and Williams (2007), and Benati (2007) document estimates of trend productivity generally similar to the p-star implied from the Base-W*RW model.[46]

In the 2000s, the models have p-star gradually declining to a level close to 1.2% by 2012. It remained close to that level through most of the past decade, but since 2018, it has steadily increased. At the end of our sample, all three models estimate the posterior mean of p-star at, or close to, 1.5%. As we show in appendix A12, these estimates of p-star are consistent with the narrative implied by the two-regime Markov-switching model of Kahn and Rich (2007), an influential contribution to the trend productivity literature.

The uncertainty around the posterior mean estimates of p-star is large. Panel (c) quantifies this uncertainty by reporting the width of the 90% credible intervals corresponding to all three models. The plots provide evidence that the theoretical restriction (defined by eq. 20) contributes to improved precision of p-star (just as it does for pi-star and w-star), as is evident in the Base and Base-NoSurv plots lying below the Base-W*RW.

4.4. ***Estimation results for π-star***

Panels (a) and (b) of Figure 8 plot the posterior mean estimates of pi-star along with the 90% credible intervals from the Base and Base-NoSurv model specifications, respectively. Panel (c) plots the corresponding precision estimates, defined as the width of the 90% intervals. A quick visual inspection shows that the pi-star from the Base specification is significantly more precise than that from the Base-NoSurv, as evidenced by narrower credible bands and the precision plot corresponding to Base lying below the Base-NoSurv plot. Based on the marginal likelihood criteria, the fit of the inflation equation to the data in the case of Base is marginally better than that of Base-NoSurv (as reported in Table 1). Our finding that adding survey expectations improves both the model fit and the precision of pi-star is consistent with CCK.

The broad contours reflected in the posterior mean pi-stars from the two models are similar to those documented elsewhere in the literature (e.g., CCK). For instance, pi-star was low in the 1960s, high in the 1970s, fell sharply in the 1980s, continued a steady deceleration in the 1990s, fluctuated in a narrow range between 2.0% and 2.5% in the 2000s, and has been below 2% since 2012. This general pattern is consistent with the widely held view. Focusing on the specifics, unlike some papers (e.g., Stock and Watson, 2007; Mertens, 2016), which show two peaks in pi-star, one in the mid-1970s and another in the early 1980s, our model-based estimates (both with and without survey data) do not show the earlier peak (similar to CCK). Relatedly, in those same papers, pi-star is estimated to peak at a level of 10% or higher; in contrast, the mean estimate of pi-star in our model specifications peaks at a lower level (similar to CCK and

[46] Edge, Laubach, and Williams (2007), who collect real-time estimates of long-run productivity from various sources, including historical Economic Reports of the President, document a similar pattern in the trend productivity estimates.

Mertens, 2016 – in his model specification that is augmented with survey data).[47]

Comparing estimates from the Base and Base-NoSurv specifications, the level of pi-star is similar in the 1960s but starting in early 1970, pi-star from the Base specification sharply accelerates, to peak at 5.8% in early 1980, while the estimates of pi-star from the Base-NoSurv specification also accelerate but peak at a lower level of 3.7%. As shown, uncertainty about pi-star increases sharply in early 1980, with the Base-NoSurv estimates experiencing a much more dramatic rise. It is the case that uncertainty around pi-star (as measured by the width of the 90% credible intervals) inferred from the Base-NoSurv is higher compared to the Base throughout the estimation sample. But in early 1980, the differential in uncertainty is twice as large, as can be seen comparing the dotted and solid plots in panel (c).

A similar rise in model-based estimates of pi-star uncertainty in the late 1970s through early 1980 (known as the Great Inflation period) has been noted elsewhere (e.g., Mertens, 2016). A subset of the literature attributes the rise in pi-star uncertainty to the un-anchoring of inflation expectations during the Great Inflation period. Beginning in early 1980 and continuing through early 2000, both models have (posterior mean of) pi-star steadily declining to 2%. Between 2000 and 2012, whereas in the case of Base, pi-star is flat at 2%, in Base-NoSurv, it is stable at a slightly higher level of 2.4%. Since 2012, pi-star has slowly moved down to reach 1.6% (in Base) and 1.7% (in Base-NoSurv).

In the online appendix A11.a, we include the results and a discussion comparing pi-star estimates from the Base model to external models: CCK, CKP, and UCSV. The estimates indicate that the CCK model generates the most precise pi-star, followed by the Base model, CKP, and UCSV. Overall, comparing across Base and Base-NoSurv specifications, and comparing the Base specification with outside models, strongly suggest the usefulness of survey forecasts in improving the econometric estimation of pi-star (i.e., survey forecast information yields sensible estimates of pi-star and improved precision), hence, corroborating the evidence in CCK, Mertens, 2016, and Nason and Smith (2021).

In the online appendices A11.b and A11.c, we explore and discuss the sensitivity of pi-star to modeling assumptions such as the usefulness of the theoretical restriction imposed by (20). The Bayesian model comparison provides evidence supporting the long-run theoretical restriction defined by equation (20).

4.5. ***Estimation results for W-star***

In modeling w-star, a novel feature of our framework is the decomposition of w-star into its fundamental components, pi-star and p-star. Figure 9 presents posterior estimates of w-

[47] Since both Stock and Watson (2007) and Mertens (2016) endow the RW process governing pi-star with SV, whereas we do not, this difference in the modeling assumption may explain the difference in the pi-star estimates around the Great Inflation period. However, CCK, who also allow SV in the pi-star process, yield pi-star estimates generally similar to ours, suggesting that the SV assumption is likely not the answer.

star along with the decomposition. The first row in the figure plots estimates from the Base model, and the second row plots estimates from the Base-NoSurv. The third row plots p-star estimates from other model variants alongside the Base and Base-NoSurv models, and also presents precision estimates of w-star.

The estimates imply that w-star increased steadily in the 1970s and peaked in the early 1980s. This increasing w-star reflected an upward drift in pi-star that more than offset the downward drift in p-star, as evidenced by the widening in the shaded area representing pi-star and the slight narrowing of the shaded area representing p-star. The Base model implies that w-star peaked at 7.1% in the early 1980s, while Base-NoSurv has w-star peaking at 5.8% in the later half of 1970. Both models have w-star sharply drifting lower through much of the 1980s, to reach near 3.4% by early 1990. From there on, the path of w-star across the two models is very similar and indicates a gradual slowing to 2.7% by the end of 2017. W-star moved up to 3.0% by the early 2018.

Not surprisingly, w-star is more precisely estimated in the Base model than in the Base-NoSurv model, as shown in panel (f). The considerable uncertainty around w-star, implied by the Base-NoSurv model during the 1970s, is mostly driven by pi-star. As noted earlier, the pi-star estimate from the Base-NoSurv model was highly imprecise during the 1970s. Interestingly, despite the inferior precision of the Base-NoSurv compared to Base, the Bayesian model comparison suggests a marginally better fit of the Base-NoSurv model to the nominal wage data than Base (see Table 5).

Sensitivity of w-star to modeling assumptions

Panel (e) plots estimates of w-star from two additional Base model variants: Base-W*RW and Base-NoPT. The Base-W*RW model eliminates the long-run restriction that w-star is the sum of p-star and pi-star (on average) and instead models w-star as an RW process. Interestingly, the path of w-star implied by Base-W*RW is below other models. Since then, it is similar to Base and Base-NoSurv. Although in the first half of the estimation sample, w-star from Base-W*RW is less precisely estimated than Base, in the second half of the sample, it is more precise. According to the Bayesian model comparison, the fit of the Base-W*RW to the nominal wage data is substantially inferior to both Base-NoSurv and Base.

The Base-NoPT model eliminates the pass-through from prices, i.e., it removes the price inflation gap from the equation describing the wage inflation gap. In other words, the direct link between the cyclical components of prices and nominal wages is eliminated, but the connection between the permanent components pi-star and w-star remains. Doing so has notable implications for the estimate of w-star and the model's fit. As shown in panel (e), w-star implied from Base-NoPT is higher than that implied by the other models through the first half of the sample. While the Base model has w-star peaking at a little above 7% in the early 1980s, the Base-NoPT model implies a higher peak of 7.8%. The acceleration in w-star implied by the Base-NoPT model during the 1970s is much stronger than that implied by the Base model estimates. This stronger path of w-star is associated with more precise estimates of w-star in the 1970s compared to the Base and other models, as can be seen in panel (f). However, according to the Bayesian model comparison reported in Table 5, removing the connection between the cyclical components negatively impacts the model fit, as evidenced by the substantially inferior fit of the Base-NoPT model to data compared to the Base model.

4.6. ***Estimation results for r-star***

Figure 10 presents r-star and the "catch-all" component D estimates for our two main model specifications. The top row of the figure plots the estimates of r-star (panel a) and component D (panel b) from the Base model. Also included in panel (a) are the survey expectations of r-star (which enters our Base model). As can be seen, the contours of (posterior mean) r-star from the Base model track the survey estimate; this suggests that survey data play an influential role in guiding the model's assessment of r-star. The posterior mean estimate from the Base model shows r-star staying relatively flat at 3.4% in the 1960s, and then slowly trending down through the 1970s, to reach 2.5% by early 1980. Thereafter, it fluctuates in a range between 2.0% and 3.0% until the beginning of 2000. From there on, it steadily declines, reaching 1.0% at the end 2019.

Panel (b) plots the estimate of component D, whose dynamics are shaped by the survey expectations data and by information from the Taylor rule and IS equations. As can be seen in the figure, component D is imprecisely estimated. According to the posterior mean estimate, in the 1960s, component D exerts slight upward pressure on r-star that is mostly offset by downward pressure coming from g-star (via equation 27), helping to keep r-star relatively flat. After that, with D remaining flat through 2000, developments in g-star shape the trajectory of r-star. From 2000 onward, all forces (as captured through the model structure) work in the same direction to push r-star steadily downward. The estimated link between r-star and g-star is of moderate strength (posterior mean of parameter $m = \frac{\zeta}{4} = 0.66$); see Table 2); therefore, movements in g-star play an influential role in driving r-star.

Moving on to the Base-NoSurv model specification, in panel (c), the mean estimate shows r-star rising from 1.6% in early 1960 to 2.4% through early 1980 and then remaining stable through early 2000. From 2000 onward, r-star steadily declines, reaching 1.3% at the end of 2019. The trajectory of r-star from 2000 onward is similar to that from the Base model. It is worth noting that our models' indication of a secular decline in r-star beginning in 2000 is also documented elsewhere in the literature tackling r-star (the exception being JM).[48] However, the extent of decline varies considerably across studies. The literature offers various explanations for this secular decline in r-star, including a trend decline in g-star (e.g., Laubach and Williams (2016));

[48] JM document that r-star from their preferred specification (which allows for SV in the TR equation) is generally flat over their sample, spanning 1960 through 2018. However, in an alternative specification, which does not permit SV, the r-star estimate exhibits a decline in r-star similar to that documented elsewhere. In our examination, r-star's trajectory is little changed comparing between the Base specification and Base without SV (i.e., Base-NoSV).

rising premiums for convenience yield, i.e., increased demand for safe and liquid Treasury bonds (see Del Negro et al., 2017); and excess global savings (Pescatori and Turunen, 2016).

The uncertainty around the r-star estimate from the Base-NoSurv model is substantially higher than that in Base, as can be seen by comparing panel (a) and panel (c) of Figure 10, and also shown in panel (f). The increased uncertainty in r-star comes from component D, which is imprecisely estimated without the survey data. Furthermore, based on the marginal likelihood criteria, the Base model is favored over the Base-NoSurv model (see Table 1).

Without the survey information about r-star, the estimated link between g-star and r-star is significantly weaker (posterior mean of $m = \frac{\zeta}{4} = 0.35$; see Table 2), which is consistent with the evidence documented in Hamilton et al. (2016) and Lunsford and West (2019). Therefore, the movements in component D significantly dominate the contours of r-star in the Base-NoSurv model. Both the IS curve and Taylor-rule equations shape the evolution of component D. The hump-shaped patterns in both D and r-star reflect the trends in real long-term interest rates (informed from the IS equation) and short-term interest rates (from the Taylor-rule equation). It is interesting to note that the r-star estimates from JM (and González-Astudillo and Laforte, 2020) – who use the Taylor-rule equation and information from both short- and long-term interest rates – also exhibit hump-shaped behavior (though, in the case of JM, it is only slight). As we show shortly, the prior setting on the shock process for r-star (in our Base and Base-NoSurv cases, component D) plays an essential role in shaping the contours of r-star.

As discussed earlier, the estimates of g-star from both the Base and the Base-NoSurv models are quite similar. Therefore, the primary source of the differential in the r-star estimates between Base and Base-NoSurv is the quantitatively weaker relationship estimated between r-star and g-star in Base-NoSurv than in Base.

Random walk assumption for r: Base vs. Base-NoLinkStars*

As is commonly done when estimating stars, a random walk assumption for r-star is a popular choice (e.g., Kiley, 2020; JM) and our model specification Base-NoLinkStars embeds this assumption. By adopting an RW assumption for r-star, the model will be unable to uncover any specific causes of movements in estimates of r-star.

Figure 10, panel (e) plots the posterior mean r-star estimate from the Base-NoLinkStars model. To facilitate comparison, also shown are estimates from the Base and Base-NoSurv models. Panel (f) plots the corresponding r-star precision estimates. A few observations immediately stand out. First, the estimated r-star from Base-NoLinkStars, although exhibiting broadly similar contours, is higher than those obtained from the specifications that impose a relationship between r-star and g-star. Second, from 2000 and onward, r-star estimates from the specifications with the assumed link between r-star and g-star experience a more stark decline than those from the specification with the RW assumption. For instance, by late 2019, the estimate of r-star from Base-NoLinkStars settle at close to 1.5%, whereas those from the Base and Base-NoSurv models fall further to a range of 1.0 to 1.3%. This differential is mostly explained by the lack of direct downward pressure from g-star in the specification with the RW assumption. Third, bringing in information from surveys improves the precision of the r-star estimates substantially, irrespective of whether r-star is modeled simply as an RW process or as a combination of an RW component and a component linking r-star to g-star.

Based on the model comparison metric reported in Table 1, the Base model yielding the most precise r-star does not necessarily rank as the best fitting model to the interest rate data. It's fit to the interest rate data is inferior to that of the Base-NoLinkStars model. However, the overall fit of the Base model is slightly better than that of the Base-NoLinkStars model.

Taken together, the evidence described above suggests that the RW assumption for r-star is a viable option, and bringing in information from surveys helps improve the precision of the estimates of r-star substantially.[49]

We also explored the role of data versus priors in determining r-star, and in the interest of brevity, the results are relegated to the online appendix (see A10.a). In online appendix A10.b, we compare our model(s) estimates of r-star with those of external models. In the online appendix A10.c, using our two main models we provide an assessment of the stance of monetary policy, defined as the deviation of the short-term nominal interest rate from the long-run nominal neutral rate of interest (i.e., sum of r* and pi*).

Overall, we find that when it comes to r-star, specification choices matter a lot (something also highlighted by Clark and Kozicki (2005), Beyer and Wieland (2019), and Kiley (2020)). The model specifications that include survey expectations yield estimates that are both reasonable and the most precise. Since 2000, the best fitting model specifications indicate a steady decline in r-star.

5. Real-time Estimates and Forecasting

In this section, we perform two real-time, out-of-sample forecasting exercises. In the first exercise, we compare the real-time forecasting performance of our two main models, Base and Base-NoSurv. We evaluate the accuracy of both the point and the density forecasts for real GDP growth, PCE inflation, the unemployment rate, nominal wage inflation, labor productivity growth, and the shadow federal funds rate. We show that the Base model is generally more accurate on average compared to Base-NoSurv for all variables of interest except the shadow federal funds rate. We also document our Base model's superior forecasting properties relative to "hard to beat" benchmarks, including some of the recently proposed UC models for inflation forecasting. By-products of our real-time forecasting exercise are the real-time estimates of the stars from 1999 through 2019. We compare these real-time estimates to the final (smoothed) estimates – based on the entire sample spanning 1959Q4 through 2019Q4.

In the second forecasting exercise, we illustrate the efficacy of the stars' estimates produced

[49]In a supplementary set of exercises, we illustrate the usefulness of the Taylor-type rule equation and the equation linking r-star to survey expectations for identifying r-star. The addition of the Taylor-type rule equation turns out to be crucial to yielding plausible and precise estimates of r-star.

from our models by demonstrating their usefulness in forecasting with external models (e.g., steady-state VARs). We find that the quality of our estimates of the stars from the Base model is generally competitive with the survey estimates, which are commonly used as proxies for stars in VAR forecasting models. For purposes of brevity, we relegate the discussion of forecasting results corresponding to the first and second exercises to the online appendices A4 and A5, respectively.

Real-time versus final estimates

Up to this point, we only have examined the smoothed estimates of the stars inferred using all of the sample data, i.e., from 1959Q4 through 2019Q4, which we denote here as final estimates. As discussed in CKP and Clark and Kozicki (2005), the examination of final estimates is beneficial for "historical analysis," such as the evaluation of past policy. But for real-time analysis, such as forecasting and policymaking, real-time estimates at time t – estimates based on data and model estimation through time t (instead of through 1:T) – are the relevant measures. In estimating the stars, a voluminous number of papers have documented the typical pattern of notable differences between real-time and final estimates; e.g., see Clark and Kozicki (2005) and Beyer and Wieland (2019) for r-star.[50]

Relatedly, several researchers have attributed the inability to precisely know the location of these stars in real time to past policy mistakes; see Powell (2018) and references therein. The documented differences between the real-time and final estimates, which at times could be dramatic, and the recognition of these differences by policymakers have been the primary reason limiting the usefulness of real-time estimates of the stars in policy discussions in recent years (see Powell, 2018). Hence, there is a strong preference for methods that can provide more credible inferences about stars in real time.[51]

Comparing real-time and final estimates of the stars from our Base model suggests that we have made some progress in mitigating the difficulties in previous real-time estimation of the stars. However, in the Base-NoSurv model, there is less success in mitigating this issue. We believe a big reason for this lack of success in the latter case is that we estimate a very high-dimensional model with a lot less data (as will be the case when stopping estimation at earlier periods). An artifact of this is that it requires the imposition of very tight priors in earlier periods than when estimating with more recent periods, which, in turn, affects the posterior estimates of model parameters and the stars. This latter fact mechanically contributes to more considerable observed differences between real-time and final estimates in the first half of the

[50]Both revisions to past data and the accrual of additional data could contribute to the observed differences between the real-time and final estimates. Many have found the estimation with additional data to be the primary factor causing revisions to historical estimates of the stars and contributing to divergence between the real-time and final estimates (see Clark and Kozicki, 2005).

[51]The issue of imprecision in the estimation of stars is an important one. It has been long recognized that considerable uncertainty surrounds the estimated stars complicating reliable inference. Despite the stars' imprecision, they continue to be used as inputs into policymaking and for other purposes. After all, as discussed in Mester (2018), uncertainty about the stars is just one source of uncertainty among many that confront policymakers.

sample period analyzed. In the case of the Base model, the use of survey information helps anchor the estimates to more reasonable values even in the face of tight priors.

Figure 11 plots the real-time posterior mean estimates of r-star, u-star, and the output gap from 1999Q1 to 2019Q4. Also plotted are the corresponding final (posterior mean) smoothed estimates and the 68% and 90% credible bands, respectively. The real-time estimates are the end-sample posterior mean (of the smoothed) estimates at any given period. For example, the 1999Q1 estimate corresponds to estimating the model(s) from 1959Q4 to 1999Q1; similarly, the 1999Q2 estimate corresponds to estimating the models from 1959Q4 to 1999Q2. Figures A2 and A3 in the online appendix A6 plot the estimates for pi-star, p-star, w-star, and g-star. As can be seen, for the most part (with the exception of g-star), the real-time estimates of the stars generally remain within the credible intervals, especially the 68% credible sets implied based on full-sample information.

Beginning with r-star (panel a), the contours of the real-time and final estimates are remarkably similar. Between 2000 and 2005, and post-2014, the real-time estimate closely tracked the final estimate. From 2006 through 2013, the real-time estimate averaged 20 basis points higher than the final estimate. In our assessment, the magnitude of the gap between real-time and final estimates is relatively small compared to the uncertainty estimate around the posterior mean and estimates of uncertainty typically reported in papers estimating r-star, e.g., Clark and Kozicki (2005), Laubach and Williams (2016), and Lubik and Matthes (2015). Furthermore, the real-time estimate of r-star remained within the 68% credible intervals throughout the sample period considered. The width of the estimated 68% (and 90%) intervals from the Base model has been remarkably stable between 1.0% and 1.3% (1.7% and 2.2%) over the last 25 years. For reference, the typical estimates of 90% bands from popular models such as LW and Lubik and Matthes (2015) have a width averaging more than 4% and 3.5%, respectively.[52] Given that the 68% and 90% credible intervals are significantly narrower compared to typical estimates reported elsewhere in the literature, we view the evidence of our real-time r-star remaining inside the estimated credible intervals as encouraging.

In the case of u-star (panel b), in the first half of the forecast evaluation sample, the real-time estimate generally tracked the posterior mean of the final estimate closely. But post-Great Recession, although the movements between the two track each other, it is the case that the real-time estimate is notably higher and mostly fell outside the posterior credible bands.

Panel (c) plots the estimates for the output gap, and to illustrate the usefulness of the SV, panel (d) plots the corresponding estimates for the output gap from the variant of the Base model without SV in output and the unemployment rate (but keeps SV in other model variables). As can be seen comparing the two panels, allowing for SV significantly improves the real-time reliability of the output gap estimates, especially during the Great Recession period and the subsequent recovery (i.e., both real-time and final estimates track each other closely in

[52] We note that recent approaches to model r-star such as JM and Del Negro et al. (2017) also generate precise intervals similar to ours, with JM marginally less precise and Del Negro and all measurably more precise than ours.

the Base model compared to the model variant that does not allow SV).

Overall, the real-time estimates of stars and forecast evaluation based on the past 20 years of data provide empirical evidence supporting the competitive forecasting and real-time properties of the Base model. Unfortunately, the high dimensionality of our models and the limited availability of real-time data on nominal wage inflation prevent the evaluation of the models' forecasting properties over a more extended historical sample.

6. The Implications of the COVID-19 Pandemic for Stars

At the time of writing this paper, the global economy is in the midst of an ongoing global pandemic crisis (GPC), which has continued to inflict significant disruption on economic activity both in the US and globally. The GPC, which started in early 2020, contributed to extreme movements in many US macroeconomic indicators, including those used in this paper. For instance, US real GDP growth in quarterly annualized terms declined from -5% in Q1 to -31% in Q2, the deepest contraction in the post-war data (COVID-19 recession). And in Q3, growth rebounded to $+33\%$, a record increase in the post-war data. These extreme movements, which were several standard deviations away from their historical averages, contributed to the breakdown of many conventional time-series models, especially the time-invariant VAR models estimated with monthly data; see Lenza and Primiceri (2020) and Carriero, Clark, Marcellino, and Mertens (2021).

Up to this point, our analysis has focused on the pre-GPC data. In light of recent work documenting the difficulties of the standard time-series models in handling pandemic data, we are naturally curious to see how our two main models respond to the COVID-19 GPC data. We find that the Base model handled the pandemic data quite well, whereas Base-NoSurv did less well. We had to tone down the Base-NoSurv model (by removing SV in output and the unemployment rate) to estimate it with pandemic data.

We believe that the rich features of our models helped position our models, especially the Base model, to handle the pandemic data well.[53] In the interests of brevity, we refer the reader to online appendix A8 for the results comparing the pre-pandemic and pandemic periods, and results comparing estimates from the Base model to outside sources, including the CBO.

7. Conclusion

This paper takes up the challenge of developing a large-scale UC model to jointly estimate the dynamics of inflation, nominal wages, labor productivity, the unemployment rate, real GDP, interest rates, and their respective survey expectations to back out estimates of the long-run

[53]The rich features include: (1) modeling the changing economic relationships via the implementation of time-varying parameters; (2) allowing for the changing variance of the innovations to various equations (i.e., SV); (3) imposing bounds on some of the random walk processes; (4) joint modeling of the output gap and unemployment gap in particular; and (5) using survey forecasts;

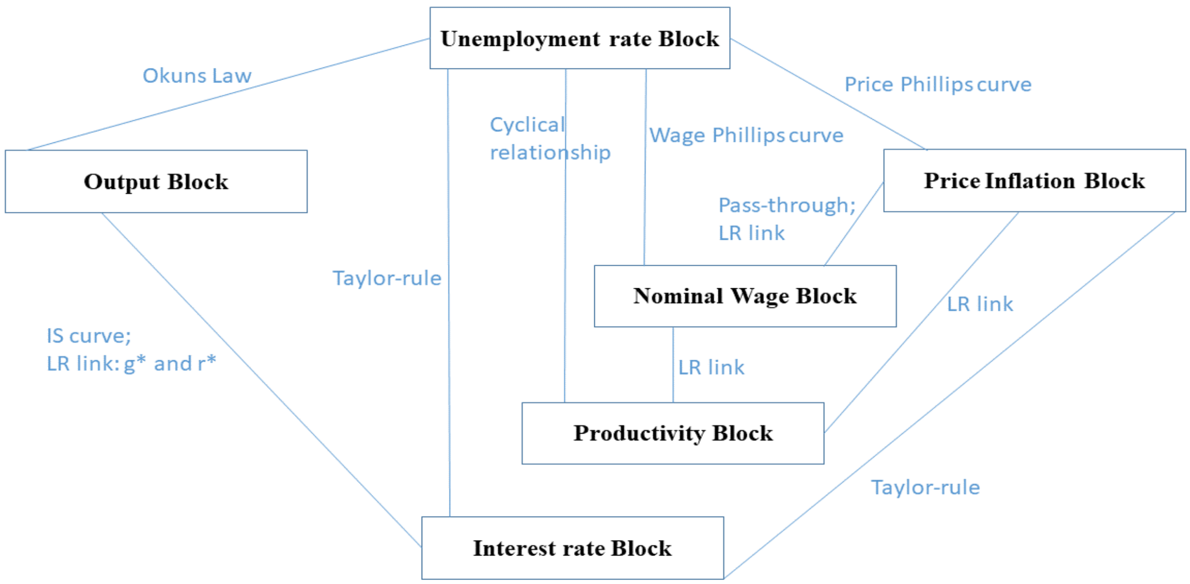
counterparts of these variables. These long-run counterparts include potential output (gdp-star), the growth rate of potential output (g-star), the equilibrium levels of the unemployment rate (u-star), the real short-term interest rate (r-star), price inflation (pi-star), labor productivity growth (p-star), and nominal wage inflation (w-star). The structure of our UC model is guided by economic theory and past empirical research, which has highlighted strong evidence of changing macroeconomic relationships and allows for stochastic volatility in the shocks to cyclical components of a range of macroeconomic indicators. Accordingly, our model structure permits time-varying parameters and stochastic volatility in the main model equations.

An essential feature of our model structure is the explicit role of long-run survey expectations in possibly informing the econometric estimation of the stars. We show significant improvements in the precision of the stars' estimates by bringing in additional information from survey expectations. Our estimates of the stars generally echo the contours of stars documented elsewhere in the literature – estimated using smaller-scale UC models – but at times the estimates of the stars are different, and these differences can matter for policy. We also show that our baseline model held up well when including the COVID-19 pandemic data. The rich set of features embedded in our UC model helped handle the pandemic data without any difficulties.

We explore the empirical relevance of various features incorporated in our baseline model by estimating several variants of the baseline model. The Bayesian model comparison results provide strong support to model features informed by past research and confirm findings documented elsewhere. For instance, we find that allowing for SV in the model equations is very important. Similarly, we find economically and statistically significant evidence of a time-varying price Phillips curve, wage Phillips curve, the evolving cyclicality of labor productivity, a changing pass-through relationship between wages and prices, and evolving persistence in price inflation and wage inflation gaps. Given the richness of our model, we document an expansive set of empirical results that we hope will prove helpful for both applied and theoretical macroeconomists alike.

Last, we document the competitive real-time forecasting properties of both our main model and, separately, the estimates of the stars, if they were to be used as steady-state values in external models.

Fig. 1. Visual Overview of Interactions Between Blocks



Notes: The solid lines represent a contemporaneous relationship between the element(s) of the blocks. LR link denotes long-run relationship, i.e., link between stars.

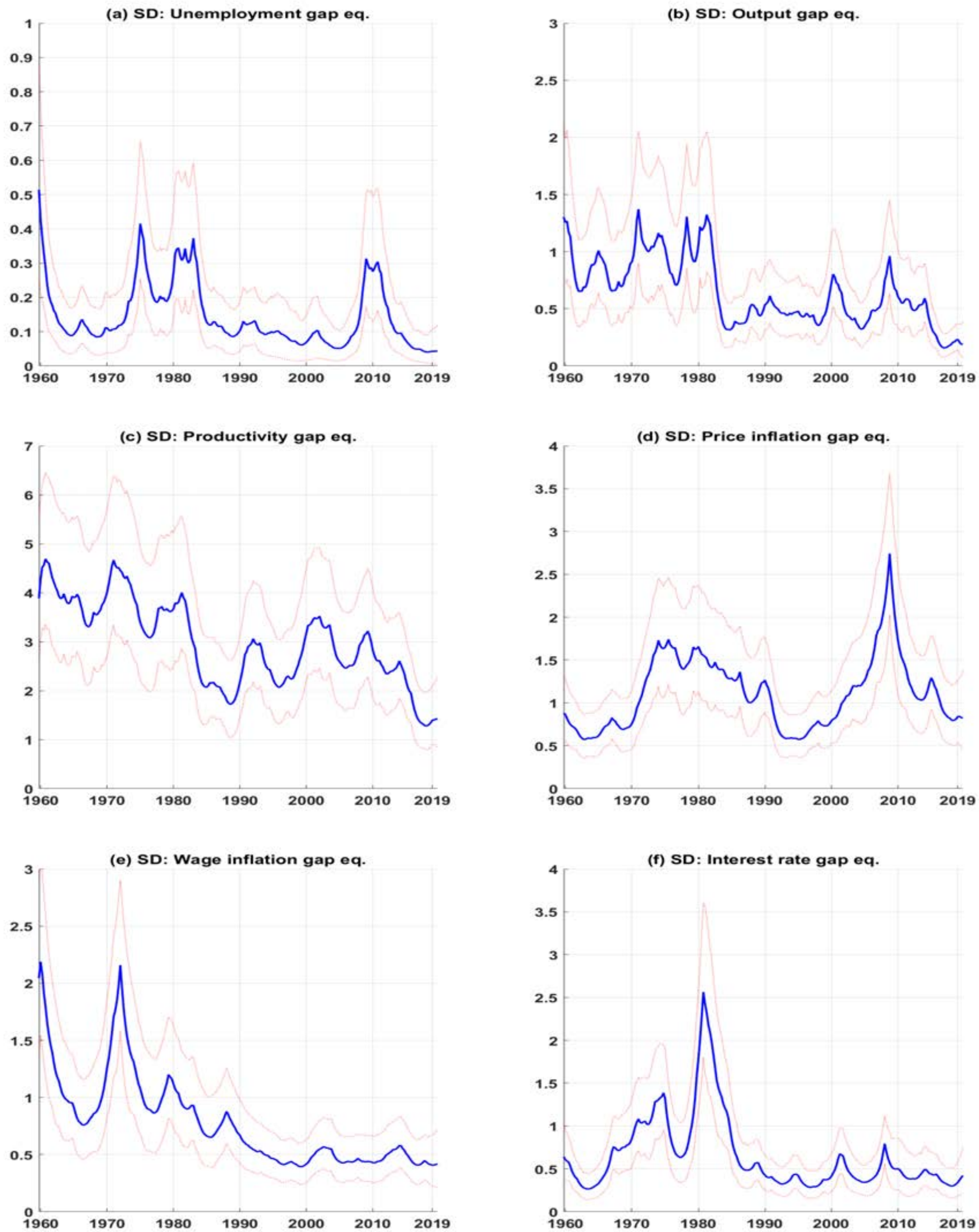
Table 1: Bayesian Model Comparison: Main Models and Selected Variants

	Base	**Base-NoSurv**	Base-NoLinkStars	Base-NoSV
MDD of Inflation	-365.2	-365.6	-365.6	-412.5
MDD of Productivity	-604.3	-603.7	-602.3	-630.1
MDD of Nominal Wage	-275.3	-274.1	-283.7	-341.6
MDD of Unemployment	89.4	93.1	89.3	40.9
MDD of Interest rate	-211.7	-215.5	-205.8	-323.3
MDD of GDP	-220.1	-221.4	-220.3	-273.0
MDD	-1587.3	-1587.1	-1588.4	-1939.7

Table 2: Parameter Estimates

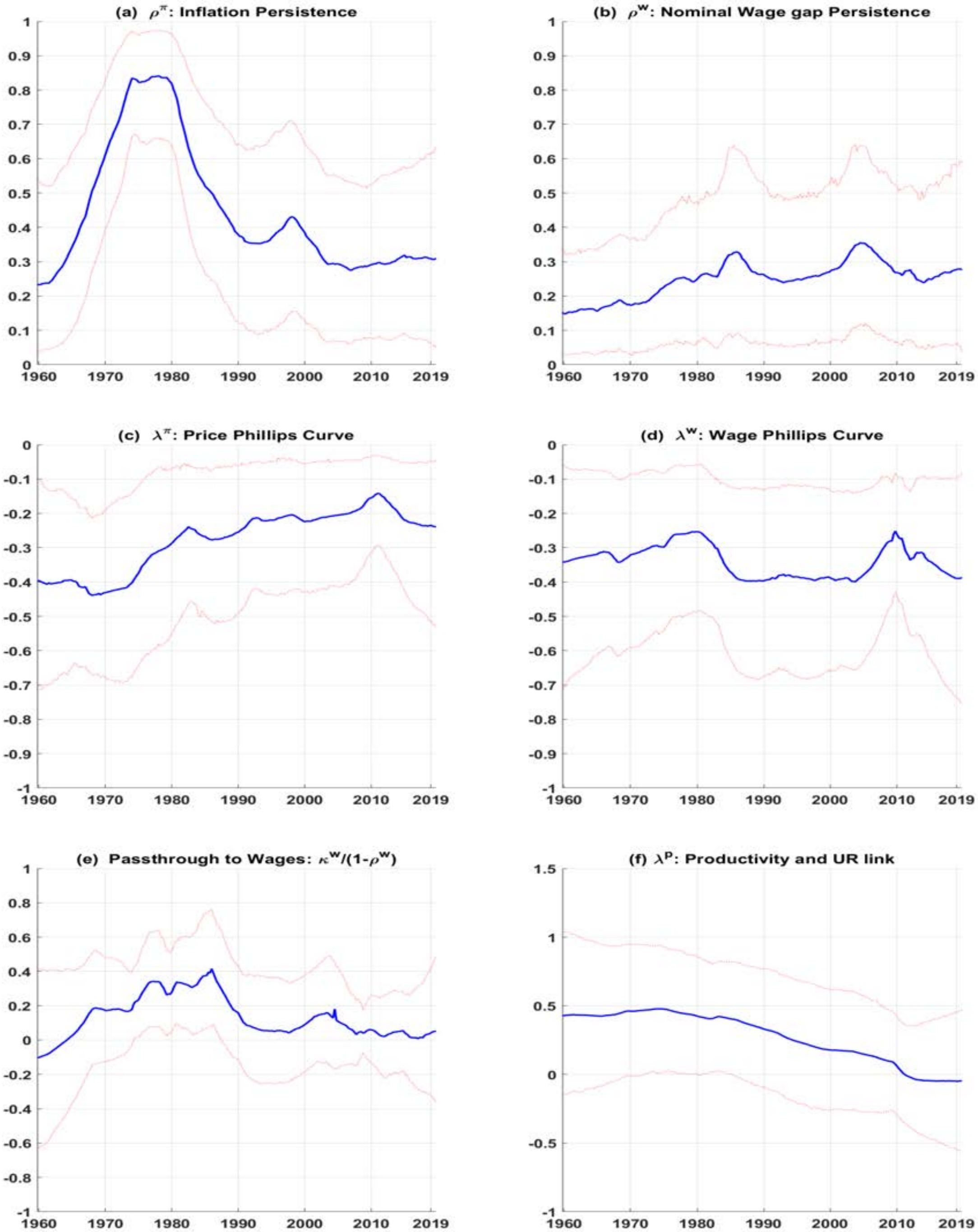
Parameter	Parameter description	Posterior estimates					
		Base			Base-NoSurv		
		Mean	5%	95%	Mean	5%	95%
a^r	Coefficient on interest rate gap	-0.041	-0.086	-0.001	-0.022	-0.066	0.019
$\rho_1^g + \rho_2^g$	Persistence in output gap	0.726	0.666	0.788	0.708	0.642	0.774
ρ_1^u	Lag 1 coefficient on UR gap	1.269	1.224	1.315	1.252	1.206	1.295
ρ_2^u	Lag 2 coefficient on UR gap	-0.502	-0.539	-0.466	-0.508	-0.543	-0.474
$\rho_1^u + \rho_2^u$	Persistence in UR gap	0.767	0.726	0.809	0.744	0.704	0.779
ρ^p	Persistence in productivity gap	-0.023	-0.139	0.092	-0.025	-0.140	0.093
$m = \frac{\zeta}{4}$	Relationship between r* and g*	0.658	0.576	0.740	0.346	0.227	0.464
ρ^i	Persistence in interest rate gap	0.876	0.836	0.913	0.866	0.828	0.903
λ^i	Interest rate sensitivity to UR gap	-0.254	-0.303	-0.206	-0.272	-0.320	-0.225
κ^i	Interest rate sensitivity to inflation	0.059	0.012	0.103	0.085	0.040	0.130
λ^g	Output gap response to UR gap	-0.479	-0.613	-0.349	-0.538	-0.689	-0.386
ϕ^u	UR gap response to output gap	-0.108	-0.127	-0.088	-0.120	-0.140	-0.103
$\frac{(1-\rho_1^u-\rho_2^u)}{\phi_u}$	Implied Okun's Law	-2.172	-2.421	-1.937	-2.141	-2.375	-1.923
β^g	Link between g* and survey	0.875	0.723	1.030	—	—	—
β^u	Link between u* and survey	0.950	0.880	1.021	—	—	—
β^r	Link between r* and survey	1.032	0.932	1.137	—	—	—
β^π	Link between π^* and survey	0.993	0.916	1.072	—	—	—
$\sigma^2_{\pi^*}$	Variance of the shocks to π^*	0.119^2	0.100^2	0.140^2	0.118^2	0.082^2	0.162^2
$\sigma^2_{p^*}$	Variance of the shocks to p^*	0.142^2	0.110^2	0.180^2	0.140^2	0.108^2	0.178^2
$\sigma^2_{u^*}$	Variance of the shocks to u^*	0.093^2	0.079^2	0.106^2	0.121^2	0.103^2	0.139^2
$\sigma^2_{gdp^*}$	Variance of the shocks to gdp^*	0.024^2	0.018^2	0.030^2	0.024^2	0.017^2	0.033^2
σ^2_d	Variance of the shocks to d	0.094^2	0.078^2	0.112^2	0.100^2	0.076^2	0.128^2
$\sigma^2_{w^*}$	Variance of the shocks to w^*	0.032^2	0.024^2	0.041^2	0.032^2	0.024^2	0.041^2
σ^2_{ho}	Var. of the Volatility – Output gap eq.	0.522^2	0.438^2	0.615^2	0.523^2	0.438^2	0.616^2
σ^2_{hu}	Var. of the Volatility – UR gap eq.	0.599^2	0.487^2	0.723^2	0.663^2	0.524^2	0.822^2
σ^2_{hp}	Var. of the Volatility – Productivity eq.	0.275^2	0.220^2	0.339^2	0.276^2	0.220^2	0.338^2
σ^2_h	Var. of the Volatility – Price Inf. eq.	0.297^2	0.238^2	0.363^2	0.298^2	0.236^2	0.366^2
σ^2_{hw}	Var. of the Volatility – Wage Inf. eq.	0.294^2	0.233^2	0.361^2	0.294^2	0.235^2	0.362^2
σ^2_{hi}	Var. of the Volatility – Interest rate eq.	0.380^2	0.297^2	0.472^2	0.383^2	0.297^2	0.485^2
$\sigma^2_{\lambda\pi}$	Var. of the shocks to TVP λ^π	0.042^2	0.032^2	0.053^2	0.042^2	0.033^2	0.054^2
$\sigma^2_{\lambda w}$	Var. of the shocks to TVP λ^w	0.041^2	0.032^2	0.052^2	0.044^2	0.033^2	0.059^2
$\sigma^2_{\lambda p}$	Var. of the shocks to TVP λ^p	0.045^2	0.034^2	0.058^2	0.045^2	0.034^2	0.059^2
$\sigma^2_{\kappa w}$	Var. of the shocks to TVP κ^w, PT	0.041^2	0.032^2	0.052^2	0.042^2	0.032^2	0.053^2
$\sigma^2_{\rho w}$	Var. of the shocks to TVP ρ^w	0.042^2	0.032^2	0.054^2	0.042^2	0.032^2	0.054^2
$\sigma^2_{\rho\pi}$	Var. of the shocks to TVP ρ^π	0.048^2	0.036^2	0.060^2	0.048^2	0.036^2	0.060^2

Fig. 2. Full Sample Estimates of Stochastic Volatility



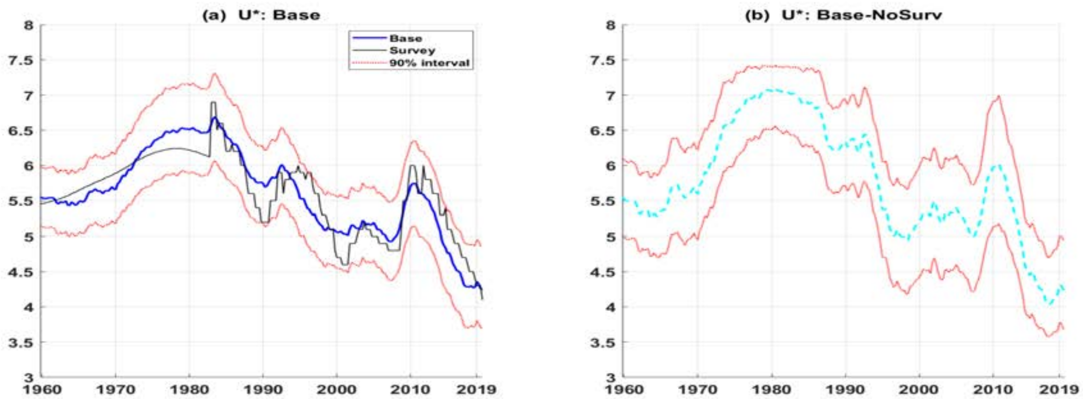
Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

Fig. 3. Full Sample Estimates of Time-Varying Parameters



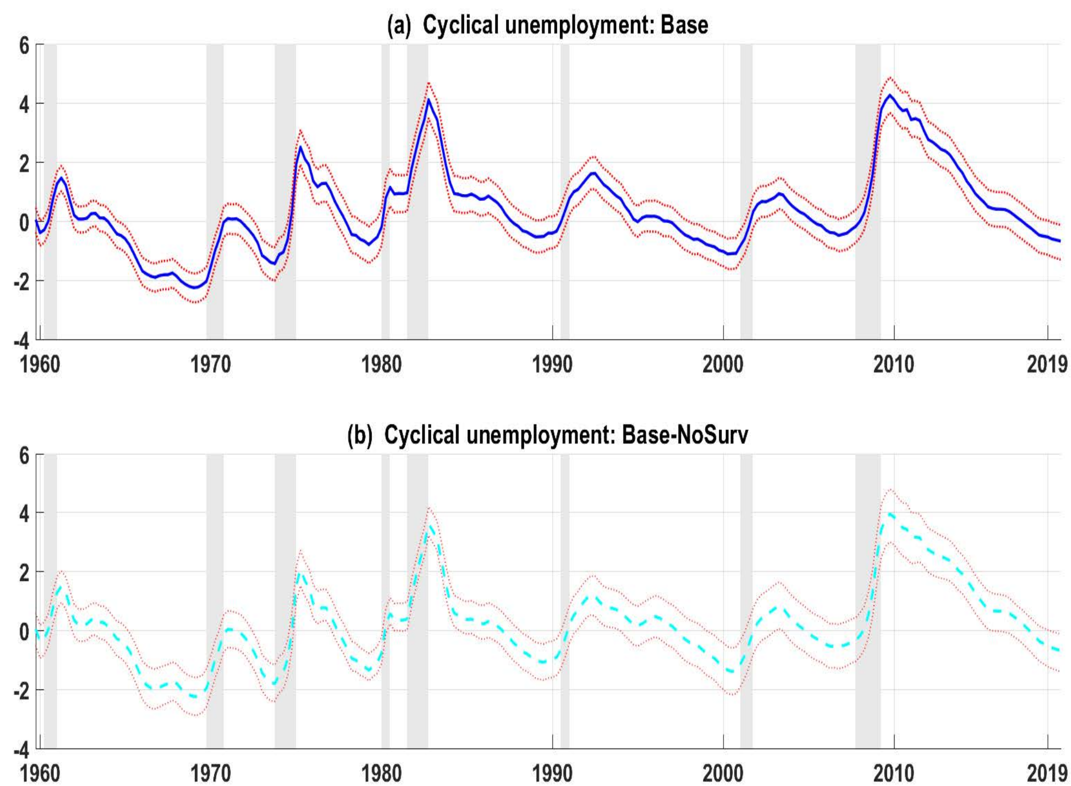
Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

Fig. 4. Full Sample Estimates for Unemployment Rate Block



Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

Fig. 5. Full Sample Estimates for Cyclical Unemployment



Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4). The shaded areas represent NBER recession dates.

Fig. 6. Full Sample Estimates for Output Block

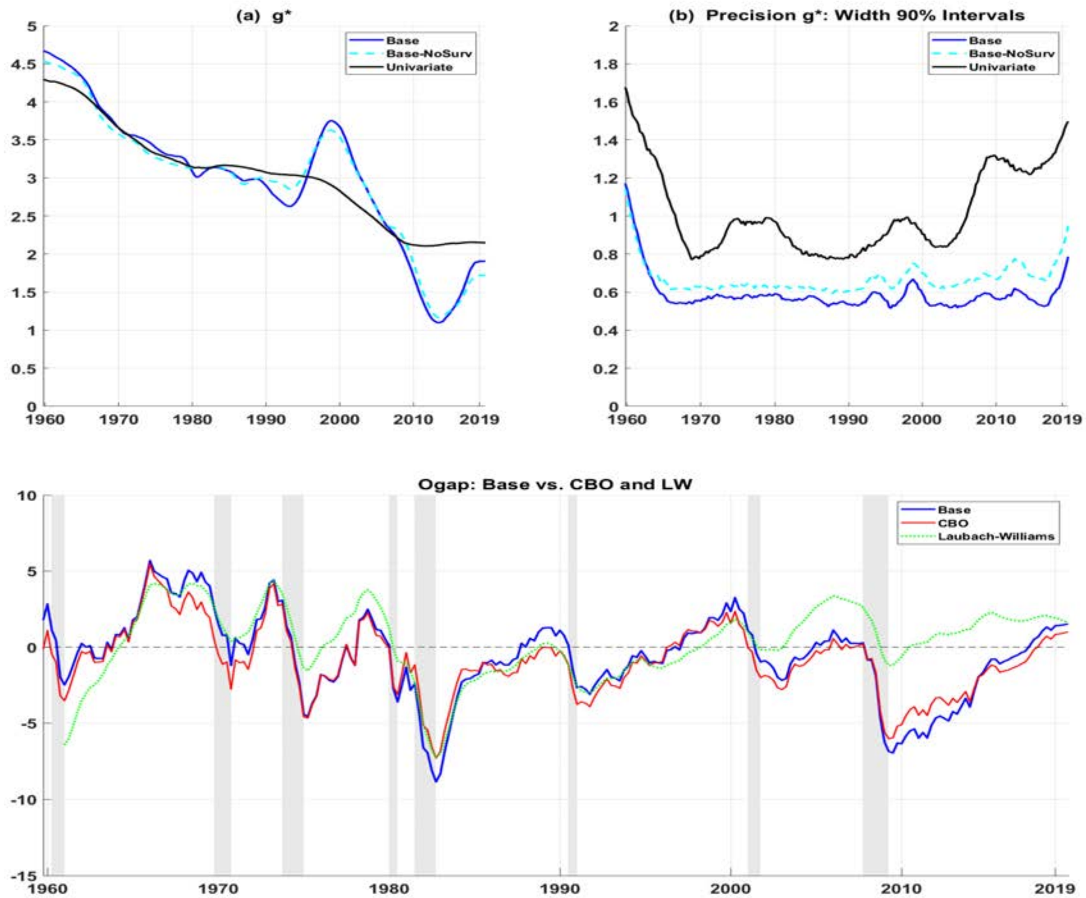
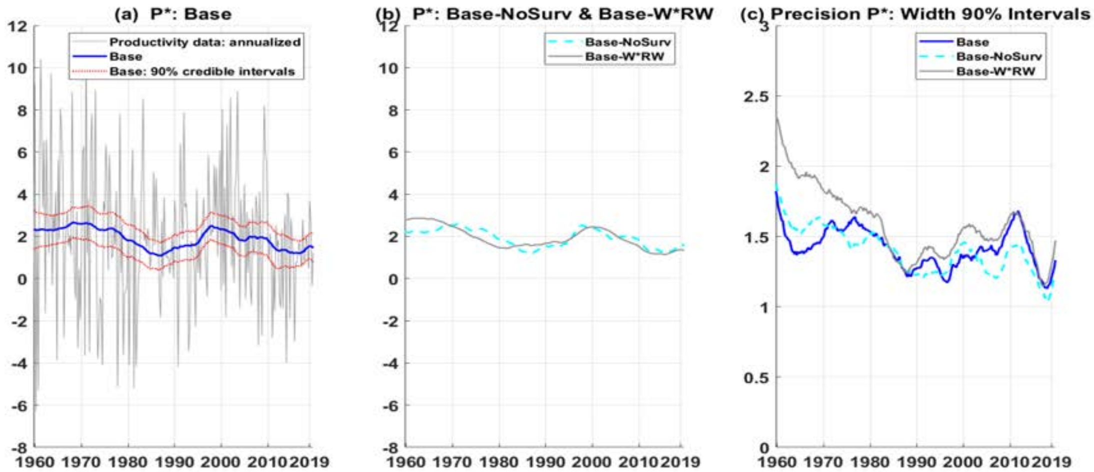


Table 3: Model Comparison: Variants Focused on GDP

	Base	Base-NoSurv	Bivariate-SV	Univariate-SV	Bivariate	Univariate
MDD of GDP	-220.1	-221.4	-223.2	-239.6	-280.4	-296.5

Fig. 7. Full Sample Estimates for Productivity Block

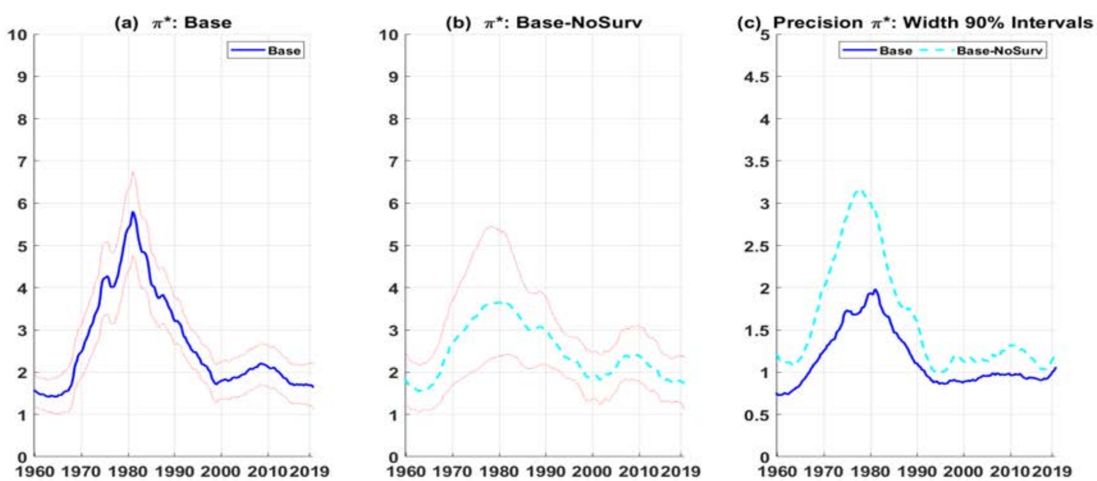


Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

Table 4: Model Comparison: Variants Focused on Labor Productivity

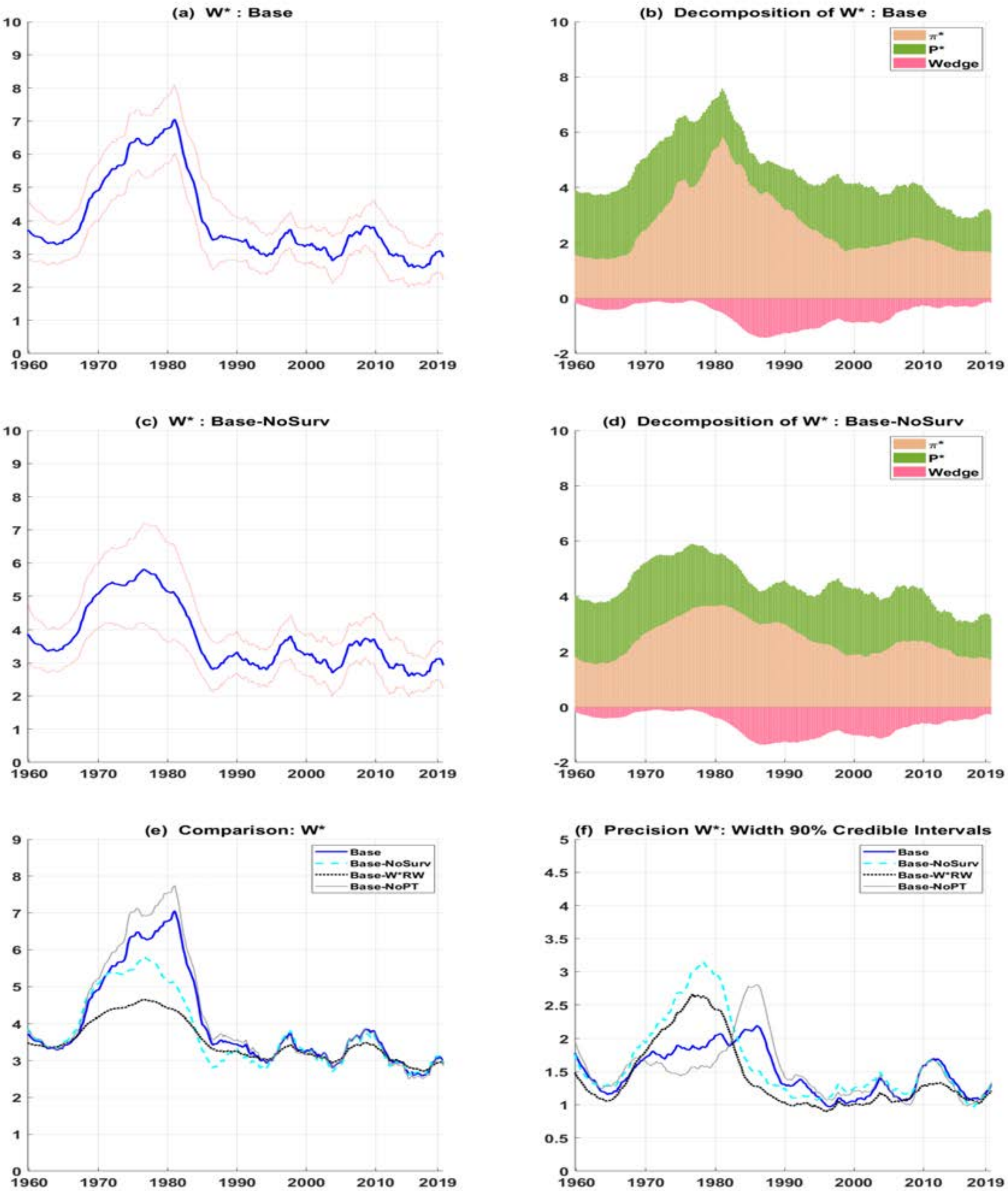
	Base	Base-NoSurv	Base-W*RW	Base-NoSV
MDD of productivity	-604.3	-603.7	-602.3	-630.1
MDD of model	-1587.3	-1587.1	-1594.3	-1939.7

Fig. 8. Full Sample Estimates for Price Inflation Block



Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

Fig. 9. Full Sample Estimates for Nominal Wage Block

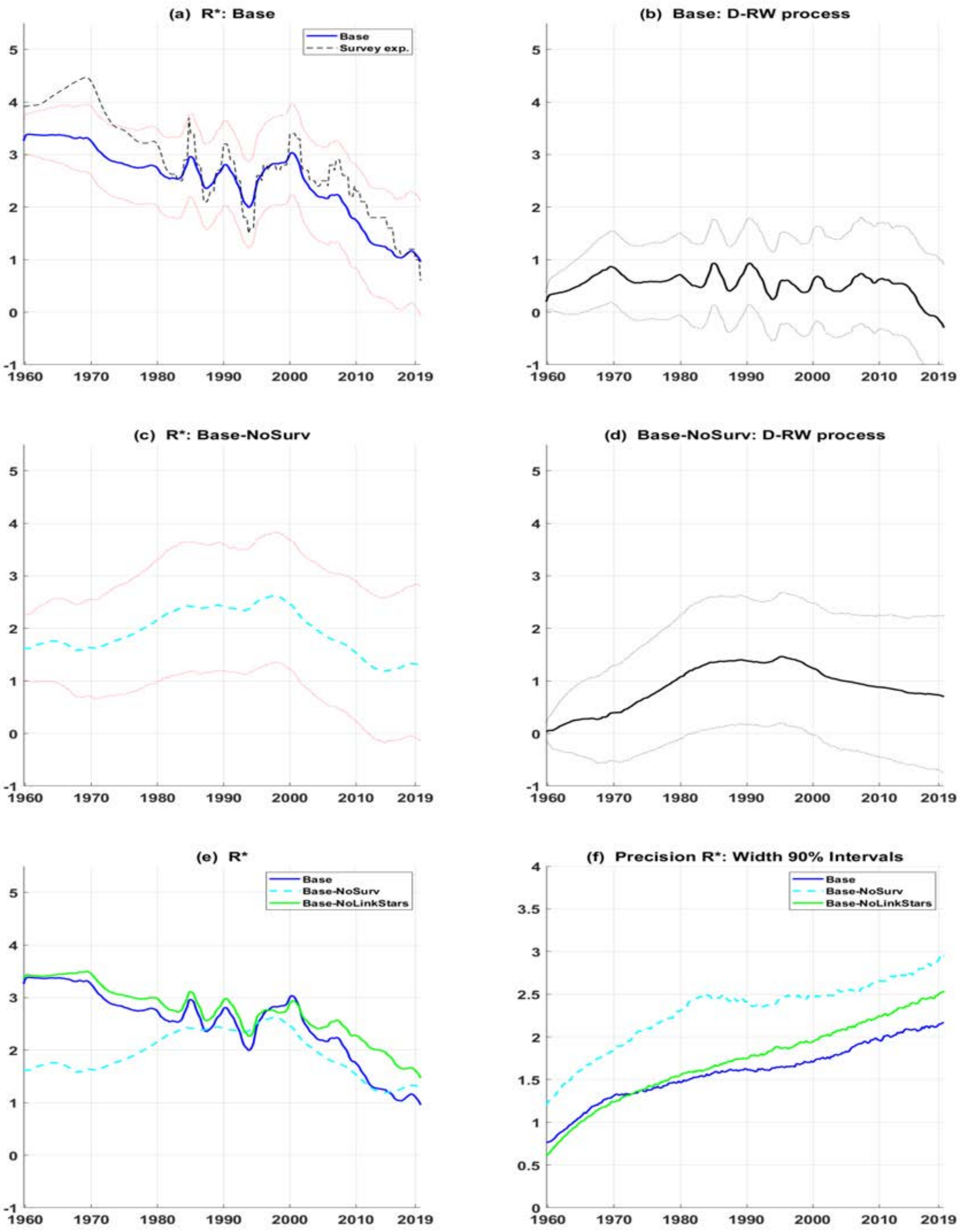


Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4).

Table 5: Model Comparison: Variants Focused on Nominal Wages

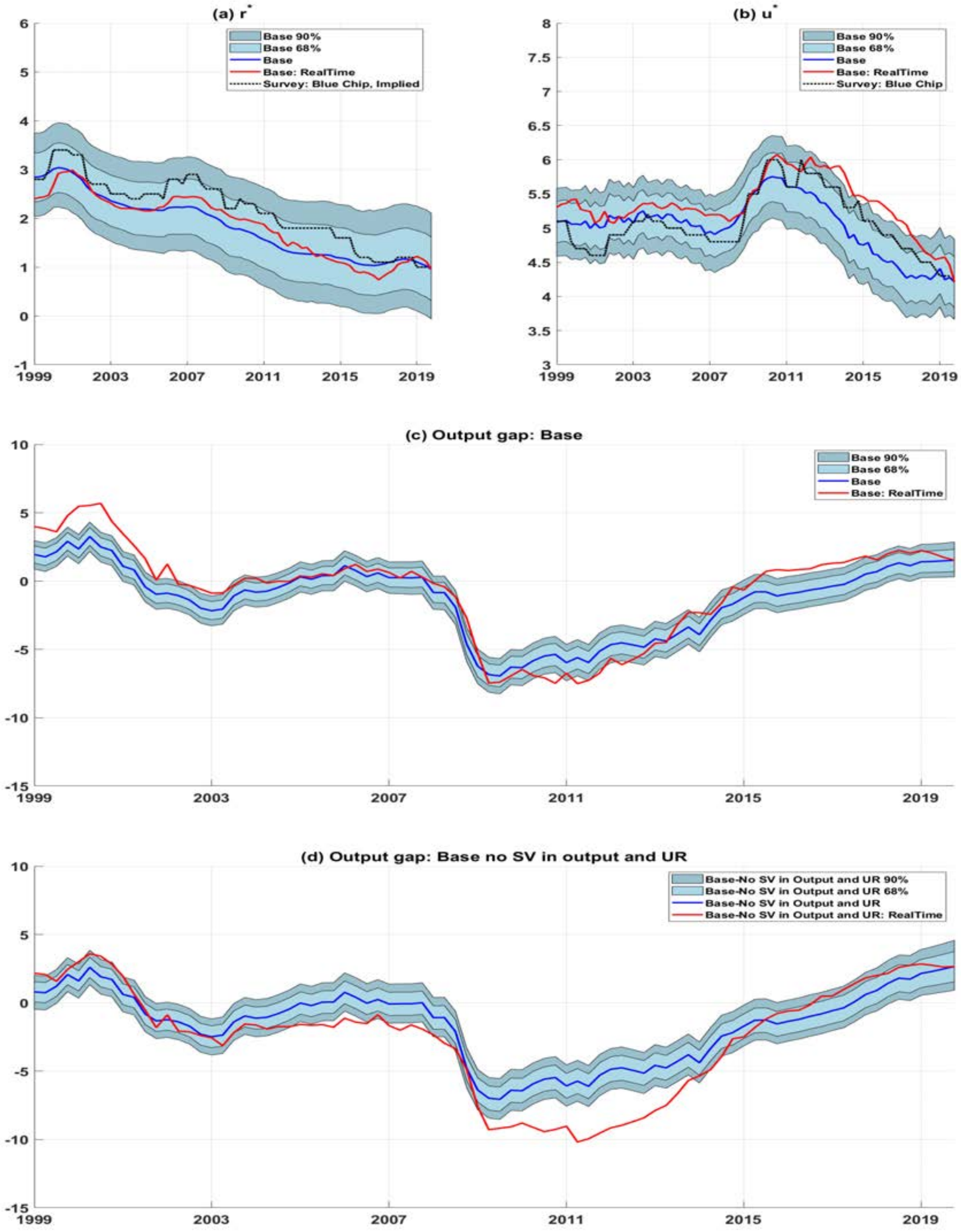
	Base	Base-NoSurv	Base-W*RW	Base-NoPT	Base-NoSV
MDD of nominal wages	-275.3	-274.1	-283.9	-279.3	-341.6
MDD of model	-1587.3	-1587.1	-1594.3	-1592.3	-1939.7

Fig. 10. Full Sample Estimates for Interest Rate Block



Notes: The posterior estimates are based on the full sample (from 1959Q4 through 2019Q4). In the top panel, the plot labeled ”survey exp.” is an implied estimate. It is inferred from the Blue Chip survey long-run estimates of the GDP deflator and short-term interest rates (3-month Treasury bill) using the long-run Fisher equation, specifically, the long-run forecast of the 3-month Treasury bill less the long-run forecast of the GDP deflator. To this differential, we add +0.3 to reflect the average differential between the federal funds rate and the 3-month Treasury bill.

Fig. 11. Real-time Recursive Estimates of Stars: Base Model



Notes: The plots labeled Base are posterior estimates reflecting information in the full sample (from 1959Q4 through 2019Q4). The plots labeled Base:RealTime are posterior estimates reflecting information available at a given point in time (i.e., truly real time).

References

Antolin-Diaz, Juan, Thomas Drechsel, and Ivan Petrella (2017). "Tracking the slowdown in long-run GDP growth." *The Review of Economics and Statistics*, 99(2), pp. 343–356. doi:10.1162/REST_a_00646.

Ashley, Richard and Randal J. Verbrugge (2008). "Frequency dependence in regression model coefficients: An alternative approach for modeling nonlinear dynamic relationships in time series." *Econometric Reviews*, 28(1-3), pp. 4–20. doi:10.1080/07474930802387753.

Barbarino, Alessandro, Travis J. Berge, Han Chen, and Andrea Stella (2020). "Which output gap estimates are stable in real time and why?" Technical Report 2020-101, Board of Governors of the Federal Reserve System. doi:10.17016/FEDS.2020.102.

Basistha, Arabinda and Charles R. Nelson (2007). "New measures of the output gap based on the forward-looking new Keynesian Phillips curve." *Journal of Monetary Economics*, 54(2), pp. 498–511. doi:10.1016/j.jmoneco.2005.07.021.

Bauer, Michael D. and Glenn D. Rudebusch (2020). "Interest rates under falling stars." *American Economic Review*, 110(5), pp. 1316–1354. doi:10.1257/aer.20171822.

Benati, Luca (2007). "Drift and breaks in labor productivity." *Journal of Economic Dynamics and Control*, 31(8), pp. 2847–2877. doi:10.1016/j.jedc.2006.11.004.

Berger, Tino, Gerdie Everaert, and Hauke Vierke (2016). "Testing for time variation in an unobserved components model for the U.S. economy." *Journal of Economic Dynamics and Control*, 69, pp. 179–208. doi:10.1016/j.jedc.2016.05.017.

Beyer, Robert C. M. and Volker Wieland (2019). "Instability, imprecision and inconsistent use of equilibrium real interest rate estimates." *Journal of International Money and Finance*, 94, pp. 1–14. doi:10.1016/j.jimonfin.2019.01.005.

Brand, Claus and Falk Mazelis (2019). "Taylor-rule consistent estimates of the natural rate of interest." Working Paper 2257, European Central Bank. doi:10.2866/44361.

Canova, Fabio and Filippo Ferroni (2011). "Multiple filtering devices for the estimation of cyclical DSGE models: Multiple filtering devices for DSGE estimation." *Quantitative Economics*, 2(1), pp. 73–98. doi:10.3982/QE36.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2019). "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors." *Journal of Econometrics*, 212(1), pp. 137–154. doi:10.1016/j.jeconom.2019.04.024.

Carriero, Andrea, Todd E. Clark, Massimiliano Marcellino, and Elmar Mertens (2021). "Addressing COVID-19 outliers in BVARs with stochastic volatility." Working Paper 21-02, Federal Reserve Bank of Cleveland. doi:10.26509/frbc-wp-202102.

Chan, Joshua C. C. (2013). "Moving average stochastic volatility models with application to inflation forecast." *Journal of Econometrics*, 176(2), pp. 162–172. doi:10.1016/j.jeconom.2013.05.003.

Chan, Joshua C. C. (2019). "Large hybrid time-varying parameter VARs." Technical Report 2019-77, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The

Australian National University.

Chan, Joshua C. C., Todd E. Clark, and Gary Koop (2018). "A new model of inflation, trend inflation, and long-run inflation expectations." *Journal of Money, Credit and Banking*, 50(1), pp. 5–53. doi:10.1111/jmcb.12452.

Chan, Joshua C. C. and Eric Eisenstat (2018a). "Bayesian model comparison for time-varying parameter VARs with stochastic volatility." *Journal of Applied Econometrics*, 33(4), pp. 509–532. doi:10.1002/jae.2617.

Chan, Joshua C. C. and Eric Eisenstat (2018b). "Comparing hybrid time-varying parameter VARs." *Economics Letters*, 171, pp. 1–5. doi:10.1016/j.econlet.2018.06.031.

Chan, Joshua C. C. and Ivan Jeliazkov (2009). "Efficient simulation and integrated likelihood estimation in state space models." *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1/2), p. 101. doi:10.1504/IJMMNO.2009.030090.

Chan, Joshua C. C., Gary Koop, and Simon M. Potter (2013). "A new model of trend inflation." *Journal of Business & Economic Statistics*, 31(1), pp. 94–106. doi:10.1080/07350015.2012.741549.

Chan, Joshua C. C., Gary Koop, and Simon M. Potter (2016). "A bounded model of time variation in trend inflation, Nairu and the Phillips curve." *Journal of Applied Econometrics*, 31(3), pp. 551–565. doi:10.1002/jae.2442.

Clark, Peter K. (1987). "The cyclical component of U.S. economic activity." *The Quarterly Journal of Economics*, 102(4), p. 797. doi:10.2307/1884282.

Clark, Todd E. and Sharon Kozicki (2005). "Estimating equilibrium real interest rates in real time." *The North American Journal of Economics and Finance*, 16(3), pp. 395–413. doi:10.1016/j.najef.2005.04.002.

Cogley, Timothy, Giorgio E. Primiceri, and Thomas J. Sargent (2010). "Inflation-gap persistence in the US." *American Economic Journal: Macroeconomics*, 2(1), pp. 43–69. doi:10.1257/mac.2.1.43.

Cogley, Timothy and Argia M. Sbordone (2008). "Trend inflation, indexation, and inflation persistence in the new keynesian phillips curve." *American Economic Review*, 98(5), pp. 2101–2126. doi:10.1257/aer.98.5.2101.

Coibion, Olivier, Yuriy Gorodnichenko, and Mauricio Ulate (2018). "The cyclical sensitivity in estimates of potential output." *Brookings Papers on Economic Activity*, 2018(2), pp. 343–441. doi:10.1353/eca.2018.0020.

Crump, Richard, Stefano Eusepi, Marc Giannoni, and Ayşegül Şahin (2019). "A unified approach to measuring U*." Working Paper 25930, National Bureau of Economic Research. doi:10.3386/w25930.

Daly, Mary C. and Bart Hobijn (2014). "Downward nominal wage rigidities bend the Phillips curve." *Journal of Money, Credit and Banking*, 46(S2), pp. 51–93. doi:10.1111/jmcb.12152.

Davis, Steven J., R. Jason Faberman, and John C. Haltiwanger (2013). "The establishment-level behavior of vacancies and hiring." *The Quarterly Journal of Economics*, 128(2), pp.

581–622. doi:10.1093/qje/qjt002.

Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti (2017). "Safety, liquidity, and the natural rate of interest." *Brookings Papers on Economic Activity*, 2017(1), pp. 235–316. doi:10.1353/eca.2017.0003.

Del Negro, Marco, Marc P. Giannoni, and Frank Schorfheide (2015). "Inflation in the Great Recession and new Keynesian models." *American Economic Journal: Macroeconomics*, 7(1), pp. 168–196. doi:10.1257/mac.20140097.

Edge, Rochelle M., Thomas Laubach, and John C. Williams (2007). "Learning and shifts in long-run productivity growth." *Journal of Monetary Economics*, 54(8), pp. 2421–2438. doi:10.1016/j.jmoneco.2007.01.003.

Faust, Jon and Jonathan H. Wright (2013). "Forecasting inflation." In *Handbook of Economic Forecasting*, volume 2, pp. 2–56. Elsevier. doi:10.1016/B978-0-444-53683-9.00001-3.

Fernald, John G. (2007). "Trend breaks, long-run restrictions, and contractionary technology improvements." *Journal of Monetary Economics*, 54(8), pp. 2467–2485. doi:10.1016/j.jmoneco.2007.06.031.

Fernández-Villaverde, Jesús and Juan Rubio-Ramírez (2010). "Macroeconomics and volatility: Data, models, and estimation." Working Paper 16618, National Bureau of Economic Research. doi:10.3386/w16618.

Feunou, Bruno and Jean-Sebastien Fontaine (2021). "Secular Economic Changes and Bond Yields." *Review of Economics and Statistics*. doi:10.1162/rest_a_01034.

Fleischman, Charles A. and John M. Roberts (2011). "From many series, one cycle: Improved estimates of the business cycle from a multivariate unobserved components model." Technical Report 2011-46, Board of Governors of the Federal Reserve System.

Galí, Jordi (2011). "The return of the wage Phillips curve." *Journal of the European Economic Association*, 9(3), pp. 436–461. doi:10.1111/j.1542-4774.2011.01023.x.

Galí, Jordi and Luca Gambetti (2019). "Has the U.S. wage Phillips curve flattened? A semi-structural exploration." Working Paper 25476, National Bureau of Economic Research. doi:10.3386/w25476.

Galí, Jordi and Thijs van Rens (2021). "The vanishing procyclicality of labour productivity." *The Economic Journal*, 131(633), pp. 302–326. doi:10.1093/ej/ueaa065.

González-Astudillo, Manuel and Jean-Philippe Laforte (2020). "Estimates of R* consistent with a supply-side structure and a monetary policy rule for the U.S. economy." Technical Report 2020-085, Board of Governors of the Federal Reserve System. doi:10.17016/FEDS.2020.085.

Grant, Angelia L. and Joshua C. C. Chan (2017a). "A Bayesian model comparison for trend-cycle decompositions of output." *Journal of Money, Credit and Banking*, 49(2-3), pp. 525–552. doi:10.1111/jmcb.12388.

Grant, Angelia L. and Joshua C. C. Chan (2017b). "Reconciling output gaps: Unobserved components model and Hodrick–Prescott filter." *Journal of Economic Dynamics and Control*, 75, pp. 114–121. doi:10.1016/j.jedc.2016.12.004.

Hamilton, James D., Ethan S. Harris, Jan Hatzius, and Kenneth D. West (2016). "The equilibrium real funds rate: Past, present, and future." *IMF Economic Review*, 64(4), pp. 660–707. doi:10.1057/s41308-016-0015-z.

Hasenzagl, Thomas, Filippo Pellegrino, Lucrezia Reichlin, and Giovanni Ricco (2020). "A model of the Fed's view on inflation." *The Review of Economics and Statistics*, pp. 1–45. doi:10.1162/rest_a_00974.

Hodrick, Robert J. and Edward C. Prescott (1997). "Postwar U.S. business cycles: An empirical investigation." *Journal of Money, Credit and Banking*, 29(1), p. 1. doi:10.2307/2953682.

Holston, Kathryn, Thomas Laubach, and John C. Williams (2017). "Measuring the natural rate of interest: International trends and determinants." *Journal of International Economics*, 108, pp. S59–S75. doi:10.1016/j.jinteco.2017.01.004.

Jacobs, Jan P.A.M. and Simon van Norden (2016). "Why are initial estimates of productivity growth so unreliable?" *Journal of Macroeconomics*, 47, pp. 200–213. doi:10.1016/j.jmacro.2015.11.004.

Johannsen, Benjamin K. and Elmar Mertens (2021). "A time-series model of interest rates with the effective lower bound." *Journal of Money, Credit and Banking*, 53(5), pp. 1005–1046. doi:10.1111/jmcb.12771.

Kahn, James A. and Robert W. Rich (2007). "Tracking the new economy: Using growth theory to detect changes in trend productivity." *Journal of Monetary Economics*, 54(6), pp. 1670–1701. doi:10.1016/j.jmoneco.2006.07.008.

Kiley, Michael T. (2020). "What can the data tell us about the equilibrium real interest rate?" *International Journal of Central Banking*, 16(3), pp. 181–209.

Knotek II, Edward S. (2007). "How useful is Okun's law?" *Economic Review*, 92(Q IV), pp. 73–103.

Knotek II, Edward S. and Saeed Zaman (2014). "On the relationships between wages, prices, and economic activity." *Economic Commentary (Federal Reserve Bank of Cleveland)*, pp. 1–6. doi:10.26509/frbc-ec-201414.

Kozicki, Sharon and P. A. Tinsley (2012). "Effective use of survey information in estimating the evolution of expected inflation." *Journal of Money, Credit and Banking*, 44(1), pp. 145–169. doi:10.1111/j.1538-4616.2011.00471.x.

Kuttner, Kenneth N. (1994). "Estimating potential output as a latent variable." *Journal of Business & Economic Statistics*, 12(3), pp. 361–368. doi:10.1080/07350015.1994.10524551.

Laubach, Thomas and John C. Williams (2003). "Measuring the natural rate of interest." *Review of Economics and Statistics*, 85(4), pp. 1063–1070. doi:10.1162/003465303772815934.

Laubach, Thomas and John C. Williams (2016). "Measuring the natural rate of interest redux." *Business Economics*, 51(2), pp. 57–67. doi:10.1057/be.2016.23.

Lee, Jaejoon and Charles R. Nelson (2007). "Expectation horizon and the Phillips curve: The solution to an empirical puzzle." *Journal of Applied Econometrics*, 22(1), pp. 161–178. doi:10.1002/jae.928.

Lenza, Michele and Giorgio Primiceri (2020). "How to estimate a VAR after march 2020." Working Paper 27771, National Bureau of Economic Research / National Bureau of Economic Research. doi:10.3386/w27771.

Lewis, Kurt F. and Francisco Vazquez-Grande (2019). "Measuring the natural rate of interest: A note on transitory shocks." *Journal of Applied Econometrics*, 34(3), pp. 425–436. doi:10.1002/jae.2671.

Lubik, Thomas A. and Christian Matthes (2015). "Calculating the natural rate of interest: A comparison of two alternative approaches." *Richmond Fed Economic Brief*, (Oct).

Lunsford, Kurt G. (2017). "Productivity growth and real interest rates in the long run." *Economic Commentary (Federal Reserve Bank of Cleveland)*, pp. 1–6. doi:10.26509/frbc-ec-201720.

Lunsford, Kurt G. and Kenneth D. West (2019). "Some evidence on secular drivers of US safe real rates." *American Economic Journal: Macroeconomics*, 11(4), pp. 113–139. doi:10.1257/mac.20180005.

Mertens, Elmar (2014). "On the reliability of output gap estimates in real time." *URL https://wwwelmarmertenscom/research/workingpapers#hicmo08w5r6zr*.

Mertens, Elmar (2016). "Measuring the level and uncertainty of trend inflation." *Review of Economics and Statistics*, 98(5), pp. 950–967. doi:10.1162/REST_a_00549.

Mertens, Elmar and James M. Nason (2020). "Inflation and professional forecast dynamics: An evaluation of stickiness, persistence, and volatility." *Quantitative Economics*, 11(4), pp. 1485–1520. doi:10.3982/QE980.

Mester, Loretta J. (2018). "Demographics and their implications for the economy and policy." *Cato Journal*, 38(2), pp. 399–413.

Morley, James C., Charles R. Nelson, and Eric Zivot (2003). "Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different?" *Review of Economics and Statistics*, 85(2), pp. 235–243. doi:10.1162/003465303765299765.

Morley, James C. and Jeremy Piger (2012). "The asymmetric business cycle." *Review of Economics and Statistics*, 94(1), pp. 208–221. doi:10.1162/REST_a_00169.

Morley, James C. and Benjamin Wong (2020). "Estimating and accounting for the output gap with large Bayesian vector autoregressions." *Journal of Applied Econometrics*, 35(1), pp. 1–18. doi:10.1002/jae.2733.

Nason, James M. and Gregor W. Smith (2021). "Measuring the slowly evolving trend in US inflation with professional forecasts." *Journal of Applied Econometrics*, 36(1), pp. 1–17. doi:10.1002/jae.2784.

Peneva, Ekaterina V. and Jeremy B. Rudd (2017). "The passthrough of labor costs to price inflation." *Journal of Money, Credit and Banking*, 49(8), pp. 1777–1802. doi:10.1111/jmcb.12449.

Pescatori, Andrea and Jarkko Turunen (2016). "Lower for longer: Neutral rate in the U.S." *IMF Economic Review*, 64(4), pp. 708–731. doi:10.1057/s41308-016-0017-x.

Powell, Jerome H. (2018). "Monetary policy in a changing economy." In *Changing Market*

Structure and Implications for Monetary Policy, volume 2018. Federal Reserve Bank of Kansas City, Jackson Hole, WY.

Proietti, Tommaso (1995). "The beveridge-nelson decomposition: Properties and extensions." *J It Statist Soc*, 4(1), pp. 101–124. doi:10.1007/BF02589061.

Roberts, John M. (2001). "Estimates of the productivity trend using time-varying parameter techniques." *The BE Journal of Macroeconomics*, 1(1), p. 153460051,014. doi:10.2202/1534-6005.1014.

Shackleton, Robert (2018). "Estimating and projecting potential output using CBO's forecasting growth model." Working Paper 2018-03, Congressional Budget Office.

Sinclair, Tara M. (2009). "The relationships between permanent and transitory movements in U.S. output and the unemployment rate." *Journal of Money, Credit and Banking*, 41(2-3), pp. 529–542. doi:10.1111/j.1538-4616.2009.00220.x.

Smets, Frank and Raf Wouters (2003). "An estimated dynamic stochastic general equilibrium model of the Euro area." *Journal of the European Economic Association*, 1(5), pp. 1123–1175. doi:10.1162/154247603770383415.

Staiger, Douglas, James H. Stock, and Mark Watson (1996). "How precise are estimates of the natural rate of unemployment?" Working Paper 5477, National Bureau of Economic Research. doi:10.3386/w5477.

Stella, Andrea and James H. Stock (2015). "A state-dependent model for inflation forecasting." In Siem Jan Koopman and Neil Shephard, editors, *Unobserved Components and Time Series Econometrics*, pp. 14–29. Oxford University Press. doi:10.1093/acprof:oso/9780199683666.003.0003.

Stock, James H. and Mark W. Watson (2007). "Why has U.S. inflation become harder to forecast?" *Journal of Money, Credit and Banking*, 39, pp. 3–33. doi:10.1111/j.1538-4616.2007.00014.x.

Stock, James H. and Mark W. Watson (2020). "Slack and cyclically sensitive inflation." *Journal of Money, Credit and Banking*, 52(S2), pp. 393–428. doi:10.1111/jmcb.12757.

Taylor, John B. (2001). "A historical analysis of monetary policy rules." In John B. Taylor, editor, *Monetary Policy Rules*, volume 31 of *Studies in Business Cycles*, pp. 319–347. University of Chicago Press.

Wright, Jonathan H. (2013). "Evaluating real-time VAR forecasts with an informative democratic prior." *Journal of Applied Econometrics*, 28(5), pp. 762–776. doi:10.1002/jae.2268.

Wu, Jing Cynthia and Fan Dora Xia (2016). "Measuring the macroeconomic impact of monetary policy at the zero lower bound." *Journal of Money, Credit and Banking*, 48(2-3), pp. 253–291. doi:10.1111/jmcb.12300.