Inflation and Professional Forecast Dynamics: An Evaluation of Stickiness, Persistence, and Volatility*

Elmar Mertens[†] AND **James M. Nason**[‡]

Current draft: June 16, 2019
First Draft: February 15, 2015

Abstract

This paper studies the joint dynamics of U.S. inflation and a term structure of average inflation predictions taken from the Survey of Professional Forecasters (SPF). We combine an unobserved components (UC) model of inflation and a sticky information forecast mechanism to study these dynamics. The UC model decomposes inflation into a trend and a gap component and measurement error. We innovate by endowing inflation gap persistence and the frequency of sticky information inflation forecast updating with drift. Stochastic volatility is imposed on the innovations to trend and gap inflation. The result is a non-linear state space model. The model is estimated on a sample from $1968Q4$ to $2018Q3$ using sequential Monte Carlo methods that include a particle learning filter and a Rao-Blackwellized particle smoother. Our estimates show that (*i*) longer horizon average SPF inflation predictions inform estimates of trend inflation; (*ii*) inflation gap persistence is procyclical before the Volcker disinflation and acyclical afterwards; (*iii*) by 1990 sticky information inflation forecast updating is less frequent than it was earlier in the sample; and (*iv*) the drop in the frequency of the sticky information forecast updating occurs at the same time persistent shocks become less important for explaining fluctuations in inflation. All told, the data calls for drift in inflation gap persistence and in the frequency of updating sticky information forecasts.

JEL Classification Numbers: $E31$, $C11$, $C32$.

Key Words: inflation; sticky information; professional forecasts; unobserved components; stochastic volatility; time-varying parameters; Bayesian; particle filter.

[†]*email*: elmar.mertens@bundesbank.de, *address*: Deutsche Bundesbank, Central Office–Research Centre, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany.

[‡]*email*: jmnason@ncsu.edu, *address*: Department of Economics, Campus Box 8110, NC State University, Raleigh, NC 27695 and Centre for Applied Macroeconomic Analysis.

*We thank Gregor Smith for several conversations that motivated this paper. We also received valuable comments from Frank Schorfheide (the co-editor), three anonymous referees, Todd Clark, Patrick Conway, Drew Creal, Bill Dupor, Andrew Filardo, Monica Jain, Alejandro Justiniano, and Wolfgang Lemke and suggestions from colleagues and participants at numerous seminars and conferences. Jim Nason thanks the Jenkins Family Economics Fund at North Carolina State University for financial support. The views herein are those of the authors and do not represent the views of the Deutsche Bundesbank or the Eurosystem. Supplementary material and results can be found in an online appendix available at `www.elmarmertens.com`.

1 Introduction

Central banks pay particular attention to inflation expectations. For example, Bernanke (2007) argues that well anchored inflation expectations are necessary for a central bank to stabilize inflation. A reason for the concern is that inflation expectations express private agents' beliefs about the underlying factors driving observed inflation dynamics. A problem is central bank policy makers lack direct knowledge of these latent factors. Instead, they have to infer the causes of inflation dynamics from other sources.

Surveys of professional forecasts are valuable sources of information about the path of future inflation. Among others, Faust and Wright (2013) and Ang, Bekaert, and Wei (2007) recognize surveys of professional forecasts yield predictions of inflation that often dominate model-based out of sample forecasts. The inflation forecasting literature also documents, as in Stock and Watson (2010) and Faust and Wright (2013), that there is time-variation in the long-run mean of inflation around which "good" forecasts should be centered. For example, Faust and Wright stress the value of survey expectations in tracking low-frequency variation in inflation, also known as movements in "trend inflation." In contrast, measures of real activity (*i.e.*, output and unemployment rate gaps) have been found to give only weak signals for inflation forecasting over above and beyond the information contained in survey forecasts.[1]

Notwithstanding the merits of surveys of professional forecasts to predict inflation, these forecasts are known to be inefficient. Coibion and Gorodnichenko (2012, 2015) argue persistence in survey forecast errors is consistent with imperfect information stories such as the sticky information framework of Mankiw and Reis (2002) (in partial equilibrium) or the noisy information (or rational inattention) models of Sims (2003) and Mackowiak and Wiederholt (2009) (in general equilibrium). Conveniently, sticky and noisy information can yield equivalent partial adjustment equations that relate the current survey forecast to a weighted average of the previous period's survey forecast and a rational expectations forecast.[2] Henceforth, we call the partial adjustment equation the "sticky information-law of motion of inflation forecasts" and its partial adjustment coefficient, which is a measure of the stickiness of inflation forecast updating, the "sticky information weight."

[1] Stock and Watson (1999) find a diminished role for activity-based forecasts of inflation, at least, since the mid 1980s. Their results are confirmed by Atkeson and Ohanian (2001), Stock and Watson (2009), Hansen, Lunde, and Nason (2011), and Faust and Wright (2013).

[2] As a result, we do not distinguish between the underlying sticky or noisy information models.

This paper studies the joint dynamics of realized inflation and inflation predictions of a survey of professional forecasters. Our approach to the inefficiency and persistence in surveys of professional forecasts of inflation is to link a term structure of average inflation predictions from the Survey of Professional Forecasters (SPF) to sticky information inflation forecasts to which we add measurement error. We innovate by letting the sticky information weight drift. Drift in the sticky information weight captures changes in the beliefs the sticky information forecaster has about the underlying dynamics of inflation. Our goal is to use estimates of drift in the sticky information weight to reexamine evidence Coibion and Gorodnichenko (2015) present that the frequency of sticky information inflation forecast updating approximates rational expectations in a recession. With estimates of drift in the sticky information, we explore the nature of the state dependency of sticky information inflation forecasts.

The sticky information-law of motion of inflation forecasts needs a rational expectations inflation forecast that we compute using a version of the Stock and Watson (2010) unobserved components (UC) model of inflation.[3] Our Stock-Watson (SW-)UC model of inflation includes the canonical non-linearities created by having innovations to trend and gap inflation that are subject to stochastic volatility. We introduce first-order autoregressive, AR(1), dynamics to the inflation gap, which is realized inflation net of trend inflation (up to a classical measurement error), in this model. We label the AR1 coefficient the "inflation gap persistence parameter", which can be static or drift.[4] We give the inflation gap inflation gap persistence to understand whether it contributes to smaller estimates of the inflation gap stochastic volatility compared with estimates found by Grassi and Proietti (2010), Stock and Watson (2010), Creal (2012), and Shephard (2013), among others. Meltzer (2014) provides another reason to include inflation gap persistence in our SW-UC model. He argues that the behavior of inflation over the business cycle changed during the 1980s. We use estimates of inflation gap persistence parameter to re-evaluate this argument of Meltzer.

We combine the SW-UC model, the sticky information-law of motion of inflation forecasts, and the term structure of average SPF inflation predictions to build non-linear state space models. Since we employ professional forecasts of inflation to study the un-

[3] Inflation's dependence on real activity is not modeled. We leave this task to the average SPF participant.

[4] Stickiness describes the smoothness of forecast updating. This paper defines persistence as the time needed to revert to trend (or steady state) in response to a shock. We take a different approach to measuring the persistence of inflation when reporting evidence on the share of the variation in inflation explained by persistent shocks to inflation in section 5 of the paper.

derlying dynamics of inflation, our modeling approach is similar to Kozicki and Tinsley (2012), Mertens (2016), Jain (2017), and Nason and Smith (2019). However, our state space models have non-linearities that always involve trend and gap inflation stochastic volatilities and in three of four specifications some combination of a drifting inflation gap persistence parameter and/or a drifting sticky information weight. We estimate the four non-linear state space models using the particle learning filter proposed by Carvalho, Johannes, Lopes, and Polson (2010) and a Rao-Blackwellized particle smoother developed Lindsten, Bunch, Särkkä, Schön, and Godsill (2016).[5] The particle learning filter and Rao-Blackwellized particle smoother are state of the art sequential Monte Carlo (SMC) methods that are new to empirical macroeconomics.[6]

We also discuss the identification of the inflation gap persistence parameter and sticky information weight. These parameters are identified in our state space models, in part, by the cross-section of the rational expectations and sticky information inflation forecasts. The forecasts are restricted by a common trend and a common cycle that aid in identifying the inflation gap persistence and sticky information weight.

The state space models are estimated on GNP/GDP deflator inflation and the average SPF nowcast and 1-, 2-, 3-, and 4-quarter ahead inflation predictions. The sample runs from $1969Q1$ to $2018Q3$. Our priors, state space models, and data provide strong evidence in favor of the state space model in which the inflation gap persistence parameter and sticky information weight drift. Estimates of this state space model show (*i*) the 4-quarter ahead SPF inflation prediction improves the efficiency of estimates of trend inflation, (*ii*) inflation gap persistence shifts from procyclical business cycle comovement before the Volcker disinflation to no comovement but substantial persistence after 1984, (*iii*) the frequency of sticky information inflation forecast updating falls from less than two quarters on average pre-1990 to about three to five quarters on average post-1990, except during the 2007–2009 recession, and (*iv*) the change in the frequency of sticky inflation forecast updating lines up with a fall in the share of the variation of realized inflation explained by persistent shocks to trend and gap inflation.

The structure of the paper follows. Section 2 builds a state space model in the observables of realized inflation and h-step ahead average SPF inflation predictions. We

[5] Our state space models can be Rao-Blackwellized because a subset of states are linear and Gaussian, given the non-linear states. A good introduction to Rao-Blackwellization of particle filters is Creal (2012). Lopes and Tsay (2011) discuss the role of Rao-Blackwellization in particle learning filters.

[6] A recent example is the work of Bonomolo, Ascari, and Lopes (2019) who estimate a small new Keynesian model with a particle learning filter, though without further utilization of a particle smoother.

sketch the SMC methods used to estimate the state space models in section 3. Results appear in sections 4 and 5. Section 6 offers our conclusions.

2 Statistical and Econometric Models

This section builds our baseline state space model using statistical and economic models. The statistical model is our SW-UC model. The economic model is the Coibion and Gorodnichenko (2015) version of the Mankiw and Reis (2002) sticky information law of motion. The baseline state space model has a static inflation gap persistence parameter, but the sticky information weight drifts. We label the baseline state space model $\mathcal{M}_0$.

2.1 A Stock and Watson UC Model of Inflation

Stock and Watson (2010), Grassi and Proietti (2010), Creal (2012), Shephard (2013), Cogley and Sargent (2015), and Mertens (2016) estimate versions of the SW-UC model that decompose realized inflation, π_t, into trend inflation, τ_t, and gap inflation, ε_t. These SW-UC models are non-linear because stochastic volatility affects innovations to τ_t and ε_t. We collect these features into our SW-UC model that also includes measurement error in π_t and AR(1) dynamics in ε_t

$$\pi_t = \tau_t + \varepsilon_t + \sigma_{\zeta,\pi}\zeta_{\pi,t}, \quad \zeta_{\pi,t} \sim \mathcal{N}(0,1), \tag{1.1}$$

$$\tau_{t+1} = \tau_t + \varsigma_{\eta,t+1}\eta_t, \quad \eta_t \sim \mathcal{N}(0,1), \tag{1.2}$$

$$\varepsilon_{t+1} = \theta\varepsilon_t + \varsigma_{\upsilon,t+1}\upsilon_t, \quad \upsilon_t \sim \mathcal{N}(0,1), \tag{1.3}$$

$$\ln \varsigma^2_{\ell,t+1} = \ln \varsigma^2_{\ell,t} + \sigma_\ell \xi_{\ell,t+1}, \quad \xi_{\ell,t+1} \sim \mathcal{N}(0,1), \ \ell = \eta, \upsilon, \tag{1.4}$$

where $\zeta_{\pi,t}$, $\varsigma_{\eta,t}$, and $\varsigma_{\upsilon,t}$ refer to measurement error in π_t, stochastic volatility in the innovation η_t of τ_t, and stochastic volatility in the innovation υ_t of ε_t. Equation (1.1) decomposes π_t into τ_t and ε_t plus $\zeta_{\pi,t}$. The random walk (1.2) describes the dynamics of τ_t. Its conditional expectation has the properties of the Beveridge and Nelson (1981) trend because, as Watson (1986) and Morley, Nelson, and Zivot (2003) note, τ_t is an I(1) state of a UC model. Equation (1.3) creates persistence in ε_t with stationary AR(1) dynamics by restricting the static inflation gap persistence parameter θ to $(-1, 1)$. Our SW-UC model includes inflation gap persistence to explore its impact on estimates of $\varsigma_{\upsilon,t}$ and behavior over the business cycle. Equation (1.4) is the random walk that generates

stochastic volatility as $\ln \varsigma_{\eta,t+1}^2$ $\left(\ln \varsigma_{\upsilon,t+1}^2\right)$ in trend (gap) inflation. We assume $\zeta_{\pi,t}$, η_t υ_t, $\xi_{\eta,t}$, and $\xi_{\upsilon,t}$ are uncorrelated at all leads and lags.

As a special case, consider a SW-UC model without persistence in the inflation gap. In addition to $\theta = 0$, shut down the stochastic volatilities, $\sigma_\eta = \varsigma_{\eta,t}$ and $\sigma_\upsilon = \varsigma_{\upsilon,t}$. The result is a fixed coefficient SW-UC model that has an IMA$(1,1)$ reduced form, $(1-\mathbf{L})\pi_t = (1-\varpi\mathbf{L})e_t$, where the MA1 coefficient $\varpi \in (-1,1)$, $\mathbf{L}$ is the lag operator, $\pi_{t-1} = \mathbf{L}\pi_t$, and the one-step ahead forecast error $e_t \equiv \pi_t - \mathbf{E}\left\{\pi_{t+1} \,|\, \pi^t\right\} = \varepsilon_t + \tau_t - \tau_{t-1|t-1}$.[7] The IMA$(1,1)$ implies a rational expectations inflation forecast updating equation, $\mathbf{E}\left\{\pi_{t+1} \,|\, \pi^t\right\} = (1-\varpi)\pi_t + \varpi\mathbf{E}\left\{\pi_t \,|\, \pi^{t-1}\right\}$, where π^t is the date t history of inflation, $\pi_1, \ldots, \pi_t$. In this case, ϖ corresponds to the Kalman gain of trend inflation, $\tau_{t|t} = \tau_{t-1|t-1} + \varpi e_t$.

Stock and Watson (2007), Grassi and Prioietti (2010), and Shephard (2013) note that stochastic volatility in the SW-UC model gives the MA1 coefficient a local time-varying parameter interpretation, ϖ_t, in the reduced form IMA$(1,1)$. By iterating the rational epectations updating equation $\mathbf{E}\left\{\pi_{t+1} \,|\, \pi^t,\, \varsigma_\eta^t,\, \varsigma_\upsilon^t\right\} = (1-\varpi_t)\pi_t + \varpi_t\mathbf{E}\left\{\pi_t \,|\, \pi^{t-1},\, \varsigma_\eta^{t-1},\, \varsigma_\upsilon^{t-1}\right\}$ backwards, we have an exponentially weighted moving average updating recursion or smoother

$$\mathbf{E}\left\{\pi_{t+1} \,|\, \pi^t,\, \varsigma_\eta^t,\, \varsigma_\upsilon^t\right\} = \sum_{j=0}^{\infty} \mu_{t,t-j}\pi_{t-j}, \tag{2}$$

that is similar to a forecasting tool traced to Muth (1960), where $\mu_{t,t} = 1 - \varpi_t$ at $j = 0$ and $\mu_{t,t-j} = (1-\varpi_t)\prod_{\ell=0}^{j-1}\varpi_{t-\ell}$ for $j \geq 1$. The exponentially weighted moving average smoother (2) generates a term structure of rational expectations inflation forecasts in which the discount, $\mu_{t,t-j}$, on past inflation, π_{t-j}, adjusts to changes in π^t, ς_η^t, and ς_υ^t. Extending these rational expectations inflation forecasts to the case with a non-zero inflation gap persistence parameter, we have an input into the sticky information-law of motion that computes sticky information inflation forecasts.

[7] Harvey (1991), Stock and Watson (2007), Grassi and Proietti (2010), and Shephard (2013) tie ϖ to the autocovariance functions (ACFs) of the IMA$(1,1)$ and fixed coefficient SW-UC model. At lags zero and one, the ACFs set $(1+\varpi^2)\,\sigma_e^2 = \sigma_\eta^2 + 2\sigma_\varepsilon^2$ and $-\varpi\sigma_e^2 = -\sigma_\varepsilon^2$. Substitute for σ_e^2 to find the quadratic equation $\varpi^2 - \left(2 + \sigma_\eta^2/\sigma_\varepsilon^2\right)\varpi + 1 = 0$. Its solution is $\varpi = \left[1 + 0.5\sigma_\eta^2/\sigma_\varepsilon^2\right] - \dfrac{\sigma_\eta}{\sigma_\varepsilon}\sqrt{1 + 0.25\sigma_\eta^2/\sigma_\varepsilon^2}$, given $\varpi \in (-1,1)$ and $\sigma_\eta, \sigma_\varepsilon > 0$.

2.2 The Sticky Information Prediction Mechanism of Inflation

Coibion and Gorodnichenko (2015) adapt the partial equilibrium model of sticky information by Mankiw and Reis (2002) to a setup in which agents' current forecast is their lagged sticky information forecasts at static probability λ and with probability $1-\lambda$ their rational expectations forecasts.[8] Averaging across forecasters gives the h-step ahead sticky information inflation prediction at time t, $F_t\pi_{t+h}$, and the static probability λ is the sticky information weight. Hence, $F_t\pi_{t+h} = \lambda F_{t-1}\pi_{t+h} + (1-\lambda)\mathbf{E}_t\pi_{t+h}$, which is a weighted average of the own lag, $F_{t-1}\pi_{t+h}$, and a rational expectations inflation forecast, $\mathbf{E}_t\pi_{t+h}$, $\lambda \in (0, 1)$, and $h = 1, \ldots, \mathcal{H}$. In this environment, the sticky information weight governs the average frequency, $1/(1-\lambda)$, at which $F_t\pi_{t+h}$ is updated.

We innovate on the static weight sticky information-law of motion by investing λ with drift. Drift in the sticky information weight is driven by an exogenous and bounded random walk, where the latter restriction ensures $\lambda_t \in (0, 1)$ for all dates t. We interpret λ_t as summarizing the beliefs the average sticky information forecaster holds about the underlying dynamics of inflation.

The sticky information-SPF block is built around the sticky information-law of motion with λ_t, its random walk, and a term structure of average SPF inflation predictions. The term structure links the average SPF participant's h-step ahead inflation predictions, $\pi^{SPF}_{t,t+h}$, to $F_t\pi_{t+h}$ plus a classical measurement error, $\zeta_{h,t}$. These elements form the system of equations

$$\pi^{SPF}_{t,t+h} = F_t\pi_{t+h} + \sigma_{\zeta,h}\zeta_{h,t}, \qquad \zeta_{h,t} \sim \mathcal{N}(0, 1), \tag{3.1}$$

$$F_t\pi_{t+h} = \lambda_t F_{t-1}\pi_{t+h} + (1-\lambda_t)\mathbf{E}_t\pi_{t+h}, \quad h = 1, \ldots, \mathcal{H}, \tag{3.2}$$

$$\lambda_{t+1} = \lambda_t + \sigma_\kappa\kappa_t, \qquad \kappa_t \sim \mathcal{TN}\left(0, 1; \lambda_{t+1} \in (0,1) \,\middle|\, \lambda_t, \sigma^2_\kappa\right), \tag{3.3}$$

where h belongs to the set of positive integers, $h \in \mathbb{Z}^+$, and κ_t is drawn from a truncated standard normal distribution ($\mathcal{TN}$) designed to guarantee $\lambda_{t+1} \in (0, 1)$.

Equations (3.1)–(3.3) define the sticky information prediction mechanism. Changes in λ_t and other state variables produce movements $F_t\pi_{t+h}$ that create fluctuations in

[8] Mankiw and Reis (2002) use the partial equilibrium sticky information-law of motion of inflation forecasts to construct a sticky information-Phillips curve. The sticky information-Phillips curve yields the dynamic response of inflation to a nominal shock that Mankiw and Reis compare with the predictions of the sticky price new Keynesian Phillips curve. Sims (2003) develops a dynamic general equilibrium model with a primitive information processing technology in which agents optimally react to shifts in the true model of the economy by smoothing their forecasts.

the observed term structure of SPF inflation predictions, $\pi^{SPF}_{t,t+h}$. The SPF term structure (3.1) includes measurement errors, $\zeta_{h,t}$, to capture deviations between $\pi^{SPF}_{t,t+h}$ and the sticky information term structure of inflation, $F_t\pi_{t+h}$, for $h = 1, \ldots, \mathcal{H}$. The sticky information-law of motion (3.2) generates updates of $F_t\pi_{t+h}$, for all h subject to drift in λ_t that is described by the random walk (3.3).

Updates of $F_t\pi_{t+h}$ rely, in part, on $\mathbf{E}_t\pi_{t+h}$. We assume the average SPF participant computes $\mathbf{E}_t\pi_{t+h}$ using the SW-UC model of equations (1.1)–(1.4). Absent any of the frictions captured by sticky information ($\lambda_t = 0$), the average SPF respondent knows the state variables of the SW-UC, τ_t, ε_t, $\varsigma_{\eta,t}$, and $\varsigma_{\upsilon,t}$. Nonetheless, we treat our SW-UC model as characterizing *reduced-form* inflation dynamics. Our SW-UC model is silent about the economic forces that may determine inflation, which, for example, might include an activity gap as in a Phillips curve, but it remains a good approximation of the inflation process as documented by Faust and Wright (2013), Stock and Watson (2009), and Grassi and Proietti (2010), among others.

The sticky information-law of motion (3.2) implies also a exponentially weighted moving average smoother. Iterate (3.2) backwards and substitute the result into (3.2) repeatedly to produce the sticky information-exponentially weighted moving average smoother

$$F_t\pi_{t+h} = \sum_{j=0}^{\infty} \Lambda_{t,t-j}\mathbf{E}_{t-j}\pi_{t+h}, \tag{4}$$

where the time-varying discount $\Lambda_{t,t-j}$ is $1-\lambda_t$ for $j = 0$ and $\Lambda_{t,t-j} = \Lambda_{t,t}\prod_{\ell=0}^{j-1}\lambda_{t-\ell}$ otherwise. The sticky information-exponentially weighted moving average smoother (4) nests the rational expectations inflation forecast, $\lim_{\lambda_t \longrightarrow 0} F_t\pi_{t+h} = \mathbf{E}_t\pi_{t+h}$, and the pure sticky information inflation forecast, $\lim_{\lambda_t \longrightarrow 1} F_t\pi_{t+h} = \sum_{j=1}^{\infty}\Lambda_{t,t-j}\mathbf{E}_{t-j}\pi_{t+h} = F_{t-1}\pi_{t+h}$. The former limit equates $F_t\pi_{t+h}$ to $\mathbf{E}_t\pi_{t+h}$, as λ_t falls to zero. As λ_t moves to its upper bound, $F_{t-1}\pi_{t+h}$ becomes the sticky information inflation forecast because the weight on $\mathbf{E}_t\pi_{t+h}$ decreases while increasing on its lags.

Between these polar cases, shocks to λ_t alter the discount applied to the history of $\mathbf{E}_t\pi_{t+h}$ in the sticky information-exponentially weighted moving average smoother (4). This information aids in identifying the responses of $F_t\pi_{t+h}$ to changes in λ_t and variation in $\mathbf{E}_t\pi_{t+h}$. A similar relationship exists between $\mathbf{E}_t\pi_{t+h}$ and the time-varying discount generated by $\varsigma_{\eta,t}$ and $\varsigma_{\upsilon,t}$ in the exponentially weighted moving average smoother (2). This connects changes in $\varsigma_{\eta,t}$ and $\varsigma_{\upsilon,t}$ to movements in $F_t\pi_{t+h}$ through $\mathbf{E}_t\pi_{t+h}$ and

the sticky information-exponentially weighted moving average smoother (4). The upshot is the exponentially weighted moving average smoothers (2) and (4) place non-linear restrictions on $F_t\pi_{t+h}$ with respect to $\varsigma_{\eta,t}$, $\varsigma_{\upsilon,t}$, and λ_t. The term structure equation (3.1) shows $\pi^{SPF}_{t,t+h}$ responds to these non-linearities in the hidden factors that are responsible for changes in inflation dynamics.

2.3 The Benchmark State Space Model

This section builds our baseline state space model, $\mathcal{M}_0$, that combines our SW-UC model of inflation, the sticky information-law of motion of inflation forecasts, and the term structure of average SPF inflation predictions. Our SW-UC model is employed to construct a term structure of rational expectations inflation forecasts. This term structure has a two factor representation driven by τ_t and ε_t. We conjecture and verify state equations for sticky information trend and gap inflation, $F_t\tau_t$ and $F_t\varepsilon_t$, that are consistent with the sticky information-law of motion of inflation forecasts. An implication of the conjecture is $F_t\tau_t$ and $F_t\varepsilon_t$ are the factors of the term structure of sticky information inflation forecasts. In the term structure equation (3.1) of $\pi^{SPF}_{t,t+h}$, the sticky information states $F_t\tau_t$ and $F_t\varepsilon_t$ eliminate $F_t\pi_{t+h}$. The upshot is the term structure of rational expectations inflation forecast places cross-equation restrictions on the rational expectations and sticky information state equations and the term structure of average SPF inflation predictions. The cross-equation restrictions show $\mathcal{M}_0$ is built on internally consistent rational expectations, sticky information, and average SPF inflation forecasts. Moreover, we show the laws of motion of $F_t\tau_t$ and $F_t\varepsilon_t$ depend crucially on the linearity of $\mathbf{E}_t\pi_{t+h}$ with respect to τ_t and ε_t and the linearity of $F_t\pi_{t+h}$ with respect to $F_t\tau_t$ and $F_t\varepsilon_t$. As a result, $\mathcal{M}_0$ has seven state variables that we group into $\mathcal{X}_t = [\tau_t \;\; \varepsilon_t]'$, $F_t\mathcal{X}_t = [F_t\tau_t \;\; F_t\varepsilon_t]'$, and $\mathcal{V}_t = [\varsigma_{\eta,t} \;\; \varsigma_{\upsilon,t} \;\; \lambda_t]'$.

As already discussed, constructing $\mathcal{M}_0$ is a multi-step process. Start by rewriting the observation equation (1.1) of our SW-UC model

$$\pi_t \;=\; \delta_{\mathcal{X}}\mathcal{X}_t \;+\; \sigma_{\zeta,\pi}\zeta_{\pi,t}, \tag{5.1}$$

where $\delta_{\mathcal{X}} = [1 \;\; 1]$. Stack the random walk (1.2) of τ_{t+1} on top of equation (1.3), which is the AR(1) of ε_{t+1}, to create the state equations of our SW-UC model

$$\mathcal{X}_{t+1} \;=\; \boldsymbol{\Theta}\mathcal{X}_t \;+\; \boldsymbol{\Upsilon}_{t+1}\mathcal{W}_t, \tag{5.2}$$

where $\boldsymbol{\Theta} = \begin{bmatrix} 1 & 0 \\ 0 & \theta \end{bmatrix}$, $\boldsymbol{\Upsilon}_{t+1} = \begin{bmatrix} \varsigma_{\eta,t+1} & 0 \\ 0 & \varsigma_{\upsilon,t+1} \end{bmatrix}$, $\mathcal{W}_t = \begin{bmatrix} \eta_t \\ \upsilon_t \end{bmatrix}$, and the stochastic volatilities, $\xi_{\eta,t}$ and $\xi_{\upsilon,t}$, are random walks described by equation (1.4).

The term structure of rational expectations inflation forecasts is built using the observation equation (5.1) and state equations (5.2). Iterate the state equations (5.2) h-steps ahead, pass $\mathbf{E}_t\{\cdot\}$ through, and substitute to find $\mathbf{E}_t\mathcal{X}_{t+h} = \boldsymbol{\Theta}^h\mathcal{X}_t$, where $\mathbf{E}_t\{\cdot\}$ conditions on $\boldsymbol{\pi}^t$, $\varsigma_{\eta,t}$, and $\varsigma_{\upsilon,t}$. Push the observation equation (5.1) h-steps ahead, apply $\mathbf{E}_t\{\cdot\}$, and substitute for $\mathbf{E}_t\mathcal{X}_{t+h}$, to produce the rational expectations term structure of inflation forecasts $\mathbf{E}_t\pi_{t+h} = \delta_{\mathcal{X}}\boldsymbol{\Theta}^h\,\mathcal{X}_t$.

Next, the sticky information-law of motion (3.2) suggests the law of motion of sticky information-states is $F_t\mathcal{X}_{t+1} = \lambda_t F_{t-1}\mathcal{X}_{t+1} + (1-\lambda_t)\mathbf{E}_t\mathcal{X}_{t+1}$. The sticky information-exponentially weighted moving average smoother (4) is consistent with the law of motion of $F_t\mathcal{X}_{t+1}$. Iterate the latter law of motion backwards and substitute $\boldsymbol{\Theta}^{h+j}\mathcal{X}_{t-j}$ for $\mathbf{E}_{t-j}\mathcal{X}_{t+h}$ to obtain the exponentially weighted moving average smoother of the sticky information-states, $F_t\mathcal{X}_{t+h} = \boldsymbol{\Theta}^h\sum_{j=0}^{\infty}\Lambda_{t,t-j}\boldsymbol{\Theta}^j\,\mathcal{X}_{t-j}$. When $h = 0$, the exponentially weighted moving average smoother of the sticky information-states is

$$F_t\mathcal{X}_t \;=\; \sum_{j=0}^{\infty}\Lambda_{t,t-j}\,\boldsymbol{\Theta}^j\,\mathcal{X}_{t-j}, \tag{6}$$

Pull $\mathcal{X}_t$ out of the infinite sum of the sticky information-state equation exponentially weighted moving average smoother (6) to find $F_t\mathcal{X}_t = (1-\lambda_t)\mathcal{X}_t + \sum_{j=1}^{\infty}\Lambda_{t,t-j}\boldsymbol{\Theta}^j\,\mathcal{X}_{t-j}$. Changing the indexes $j = i+1$ and $\ell = s+1$ converts the infinite sum of the previous expression to $\lambda_t\boldsymbol{\Theta}\sum_{i=0}^{\infty}\Lambda_{t-1,t-i-1}\boldsymbol{\Theta}^i\,\mathcal{X}_{t-i-1}$. Since this infinite sum equals $F_{t-1}\mathcal{X}_{t-1}$, substitute $\lambda_t\boldsymbol{\Theta}F_{t-1}\mathcal{X}_{t-1}$ in the previous expression for $F_t\mathcal{X}_t$ to see that its law of motion is $F_t\mathcal{X}_t = (1-\lambda_t)\mathcal{X}_t + \lambda_t\boldsymbol{\Theta}F_{t-1}\mathcal{X}_{t-1}$. Finally, lead the law of motion of $F_t\mathcal{X}_t$ forward one period and substitute for $\mathcal{X}_{t+1}$ using the SW-UC model's state equations (5.2) to find

$$F_{t+1}\mathcal{X}_{t+1} \;=\; \lambda_{t+1}\boldsymbol{\Theta}F_t\mathcal{X}_t \;+\; (1-\lambda_{t+1})\boldsymbol{\Theta}\mathcal{X}_t \;+\; (1-\lambda_{t+1})\boldsymbol{\Upsilon}_{t+1}\mathcal{W}_t, \tag{7}$$

which is the system of sticky information-state equations. The sticky information state equations (7) display the cross-equation restrictions imposed by the term structure of rational expectations inflation forecast through the transition and impulse matrices, $\boldsymbol{\Theta}$ and $\boldsymbol{\Upsilon}_{t+1}$, of the state equations (5.2) of our SW-UC model.

The state equations of $\mathcal{M}_0$ are formed by stacking the state equations (5.2) of $\mathcal{X}_{t+1}$

on top of the sticky information state equations (7)

$$\mathcal{S}_{t+1} = \boldsymbol{\mathcal{A}}_{t+1}\mathcal{S}_t + \boldsymbol{\mathcal{B}}_{t+1}\mathcal{W}_t, \tag{8.1}$$

where $\mathcal{S}_t = \begin{bmatrix} \mathcal{X}_t \\ F_t\mathcal{X}_t \end{bmatrix}$, $\boldsymbol{\mathcal{A}}_{t+1} = \begin{bmatrix} \boldsymbol{\Theta} & \mathbf{0}_{2\times 2} \\ \left(1-\lambda_{t+1}\right)\boldsymbol{\Theta} & \lambda_{t+1}\boldsymbol{\Theta} \end{bmatrix}$, $\boldsymbol{\mathcal{B}}_{t+1} = \begin{bmatrix} \boldsymbol{\Upsilon}_{t+1} \\ \left(1-\lambda_{t+1}\right)\boldsymbol{\Upsilon}_{t+1} \end{bmatrix}$, and $\xi_{\eta,t+1}$, $\xi_{\upsilon,t+1}$, and λ_{t+1} evolve as the random walks (1.4) and (3.3). Drift in the sticky information weight and the stochastic volatilities create non-linearities in the state equations (8.1). However, $\mathcal{S}_{t+1}$ has linear dynamics conditional on a realization of $\mathcal{V}_{t+1}$.

We construct the observation equations of $\mathcal{M}_0$ using the observation equation (5.1) of our SW-UC model, SPF measurement equation (3.1), and rational expectations and sticky information term structures of inflation forecasts. The former term structure replaces $\mathbf{E}_{t-j}\pi_{t+h}$ with $\delta_{\mathcal{X}}\boldsymbol{\Theta}^h\,\mathcal{X}_{t-j}$ in the sticky inflation forecast-exponentially weighted moving average smoother (4), which yields $F_t\pi_{t+h} = \delta_{\mathcal{X}}\,\boldsymbol{\Theta}^h\sum_{j=0}^{\infty}\Lambda_{t,t-j}\,\boldsymbol{\Theta}^j\,\mathcal{X}_{t-j}$. Up to $\delta_{\mathcal{X}}\,\boldsymbol{\Theta}^h$, $F_t\pi_{t+h}$ equals the exponentially weighted moving average smoother (6) of the sticky information-states. The result is the term structure of sticky information inflation forecasts, $F_t\pi_{t+h} = \delta_{\mathcal{X}}\boldsymbol{\Theta}^h F_t\mathcal{X}_t$. We use it to eliminate $F_t\pi_{t+h}$ from the SPF term structure equation (3.1) to create $\pi^{SPF}_{t,t+h} = \delta_{\mathcal{X}}\boldsymbol{\Theta}^h F_t\mathcal{X}_t + \sigma_{\zeta,h}\zeta_{h,t}$, which shows the term structure of average SPF inflation prediction is driven by the sticky information-states and measurement errors. Put these SPF term structure equations beneath the observation equation (5.1) of our SW-UC model to produce the observation equations of $\mathcal{M}_0$

$$\mathcal{Y}_t = \boldsymbol{\mathcal{C}}\mathcal{S}_t + \boldsymbol{\mathcal{D}}\mathcal{U}_t, \tag{8.2}$$

where $\mathcal{Y}_t = \begin{bmatrix} \pi_t \\ \pi^{SPF}_{t,t+1} \\ \vdots \\ \pi^{SPF}_{t,t+\mathcal{H}} \end{bmatrix}$, $\boldsymbol{\mathcal{C}} = \begin{bmatrix} \delta_{\mathcal{X}} & \mathbf{0}_{1\times 2} \\ \mathbf{0}_{1\times 2} & \delta_{\mathcal{X}}\boldsymbol{\Theta} \\ \vdots & \vdots \\ \mathbf{0}_{1\times 2} & \delta_{\mathcal{X}}\boldsymbol{\Theta}^{\mathcal{H}} \end{bmatrix}$, $\boldsymbol{\mathcal{D}} = \begin{bmatrix} \sigma_{\zeta,\pi} & 0 & \dots & 0 \\ 0 & \sigma_{\zeta,1} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \sigma_{\zeta,\mathcal{H}} \end{bmatrix}$,

$\mathcal{U}_t = \begin{bmatrix} \zeta_{\pi,t} & \zeta_{1,t} & \dots & \zeta_{\mathcal{H},t} \end{bmatrix}'$, and $\boldsymbol{\Omega}_{\mathcal{U}} = \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{D}}'$. The observation equations (8.2) show the data, $\mathcal{Y}_t$, are linear in $\mathcal{S}_t$. Also, note the transition matrix $\boldsymbol{\Theta}$ of the rational expectations states $\mathcal{X}_t$ imposes cross-equation restrictions on the observation equations (8.2) and state equations (8.1).

2.4 Identification of the State Space Models

Identification of our baseline state space model, $\mathcal{M}_0$, depends on term structures of rational expectations and sticky information inflation forecasts. The linear states, τ_t, ε_t, $F_t\tau_t$, and $F_t\varepsilon_t$, are identified on (*i*) a two-factor term structure of rational expectations inflation forecasts produced by our SW-UC model, $\mathbf{E}_t\pi_{t+h} = \delta_{\mathcal{X}}\boldsymbol{\Theta}^h\,\mathcal{X}_t$, for $h = 1, \ldots, \mathcal{H}$ and (*ii*) across the same forecast horizons, the term structure of average SPF inflation predictions, $\pi^{SPF}_{t,t+h} = \delta_{\mathcal{X}}\boldsymbol{\Theta}^h F_t\mathcal{X}_t + \sigma_{\zeta,h}\zeta_{h,t}$, that also has a two-factor representation. These factors, $F_t\tau_t$ and $F_t\varepsilon_t$, are driven by a MA(∞) of τ_t or ε_t, where the MA coefficients are functions of θ and the history of λ_t.[9] Hence, information in the cross-section of average SPF inflation predictions aids in identifying the level and slope of the term structures of rational expectations and sticky information inflation forecasts.

We use τ_t, $F_t\tau_t$, ε_t, $F_t\varepsilon_t$, and restrictions embedded in equations (8.1)–(8.2) of $\mathcal{M}_0$ to identify λ_t and θ. Intuition for the identification is developed by fixing the sticky information weight, $\lambda_t = \lambda$. This assumption recovers the error correction mechanism, $F_t\tau_t - \tau_t = \lambda\left(F_{t-1}\tau_{t-1} - \tau_{t-1}\right) - \lambda\eta_t$, and common feature regression, $F_t\varepsilon_t - \varepsilon_t = \lambda\,\theta\left(F_{t-1}\varepsilon_{t-1} - \varepsilon_{t-1}\right) - \lambda\upsilon_t$, from the associated state space model.[10] The error correction mechanism identifies λ as its slope coefficient, which rests on the common trend restriction that cointegration places on $F_t\tau_t$ and τ_t. Once λ is known, θ is identified in the common feature regression because $F_t\varepsilon_t$ and ε_t share a common cycle.[11] Further analysis of this issue, which leans on results in Harvey (1991) and especially Komunjer and Ng (2011), is left to the online appendix.

[9] Own shocks and stochastic volatilities are propagated into movements in τ_t and ε_t as described by the first two rows of the state equations (8.1). The last two rows of these state equations map τ_t and ε_t into the associated sticky information-states, $F_t\tau_t$ and $F_t\varepsilon_t$. This mapping depends on time-variation in the sticky information weight, λ_t.

[10] Grant and Thomas (1999) argue weak forms of forecast efficiency demand that survey errors are stationary and independent of whether surveys provide optimal and efficient predictions of inflation. An equivalent restriction is surveys of inflation forecasts and realized inflation cointegrate. Kozicki and Tinsely (2012), Mertens (2016), and Nason and Smith (2019) apply this restriction to generate estimates of trend inflation from samples of realized inflation and surveys of inflation forecasts.

[11] Jain (2017) uses forecast revisions to identify predictability of individual SPF inflation forecasts. Applying Jain's approach to our state space model sets $\theta = (F_t\pi_{t+h+2} - F_t\pi_{t+h})/(F_t\pi_{t+h+1} - F_t\pi_{t+h})$, which assumes away measurement error. Given θ, the level and slope of the term structure of sticky information inflation forecasts are identified by $F_t\pi_{t+h}$ and two adjacent forecasts. Krane (2011) uses a similar approach, but identifies permanent and transitory components in revisions to the term structure of Blue Chip forecasts.

3 Econometric Methods

This section summarizes the SMC filtering and smoothing algorithms used to estimate our state space models. The interested reader should see the online appendix for a complete exegesis of our SMC methods.

We estimate a state space model's filtered states and static parameters with SMC methods proposed by Carvalho, Johannes, Lopes, and Polson (2010) and refined by Lopes and Tsay (2011).[12] This class of SMC methods is known as particle learning. Particle learning jointly filters the states, $\mathcal{S}_t$ and $\mathcal{V}_t$, and static parameters, Ψ_0, of $\mathcal{M}_0$ at every date $t = 1, \ldots, T$, where $\Psi_0 = \left[\sigma_\eta^2 \;\; \sigma_\upsilon^2 \;\; \sigma_{\zeta,\pi}^2 \;\; \sigma_{\zeta,1}^2 \;\; \sigma_{\zeta,2}^2 \;\; \sigma_{\zeta,4}^2 \;\; \sigma_{\zeta,5}^2 \;\; \sigma_\kappa^2 \;\; \theta\right]'$. [13] The result is a sequence of estimates of $\mathcal{S}_t$, $\mathcal{V}_t$, and Ψ_0 that condition on the same information at each date t, which is consistent with the real-time character of the SPF data.

Particle learning relies on sufficient statistics to construct the posterior of a state space model. Sufficient statistics are tied to prior beliefs about the state space model and its dynamic structure. We place priors on Ψ_0 that are conjugate. Conjugate priors yield analytic solutions that serve as laws of motion to update sufficient statistics of Ψ_0, which are coefficients of its priors. Filtered estimates of Ψ_0 are drawn from particle streams of the sufficient statistics. The result is an online process that learns about Ψ_0 by moving through the sample date by date from its start to its end.

Carvalho, et al advise, if feasible, to Rao-Blackwellize a state space model. Rao-Blackwellization exploits the conditional linear and Gaussian dynamics of $\mathcal{S}_t$, given $\mathcal{V}_t$ and Ψ_0. Hence, we refer to $\mathcal{S}_t$ and $\mathcal{V}_t$ as the "linear" and "non-linear" states. For a Rao-Blackwellized state space model, the Kalman filter analytically marginalizes out $\mathcal{S}_t$ and tracks its sufficient statistics. This improves the efficiency of the particle learning filter by lowering its sampling error. The non-linear states are estimated by simulating the random walks (1.4) and (3.3).

We estimate the smoothed states of our Rao-Blackwellized state space models using an algorithm created by Lindsten, Bunch, Särkkä, Schön, and Godsill (2016). Their smoothing algorithm accounts for Rao-Blackwellization of a state space model by first forward filtering (*i.e.*, using a particle learning filter) of all the states, backward smooth-

[12] A state space model can be estimated by wrapping a Markov chain Monte Carlo (MCMC) sampler around a particle filter. Andrieu, Doucet, and Holenstein (2010) give particle MCMC (PMCMC) theoretical foundations. Schorfheide, Song, and Yaron (2018) put a PMCMC sampler into practice. A PMCMC algorithm can engender large computational costs because it runs a simulation inside a simulation.

[13] Parameter vectors associated with $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ are denoted Ψ_1, Ψ_2, and Ψ_3, respectively.

ing of the non-linear states, and forward smoothing of the linear states. We revise the Rao-Blackwellized particle smoother to marginalize out the sample uncertainty induced by estimating the static parameters with the particle learning filter.

3.1 Particle Filters and Particle Learning

When a state space model is non-linear and/or non-Gaussian, direct sampling from the posterior distribution of the states is often impossible. A particle filter proposes a density to sample streams of particles that represent the states to solve the problem. Carvalho, et al extend this approach to a particle learning filter that jointly samples $\mathcal{S}_t$, $\mathcal{V}_t$, and Ψ_0 while marginalizing Ψ_0 out of the posterior of $\mathcal{S}_t$ and $\mathcal{V}_t$.

The Bootstrap Particle Filter

We discuss the problem of sampling with a particle filter in the context of estimating the states of $\mathcal{M}_0$. Its state and observation equations (8.1) and (8.2) and random walks (1.4) and (3.3) are cast as the non-linear state space model

$$\begin{aligned} \mathcal{Y}_t \left| \mathcal{S}_t, \mathcal{V}_t \right. &\sim p\left(\mathcal{Y}_t \left| \mathcal{S}_t, \mathcal{V}_t; \Psi_0\right.\right), \\ \mathcal{S}_{t+1}, \mathcal{V}_{t+1} \left| \mathcal{S}_t, \mathcal{V}_t \right. &\sim p\left(\mathcal{S}_{t+1}, \mathcal{V}_{t+1} \left| \mathcal{S}_t, \mathcal{V}_t; \Psi_0\right.\right), \end{aligned}$$

for $t = 1, 2, 3, \ldots$, given Ψ_0 is known. A particle filter approximates the posterior density, $p\left(\mathcal{S}_{t+1}, \mathcal{V}_{t+1} \left| \mathcal{Y}^{t+1}; \Psi_0\right.\right)$, with a discrete and finite set of values of $\mathcal{S}_{t+1}$ and $\mathcal{V}_{t+1}$, where each particle is weighted by $w_{t+1}^{(i)} \in (0, 1)$, $i = 1, \ldots, M$, and $\mathcal{Y}^t = [\mathcal{Y}_1 \ldots \mathcal{Y}_t]'$.

A simple, albeit brute force particle filter bootstraps the posterior of the states.[14] A bootstrap particle filter first propagates and then resamples the states. Conditional on drawing M particles from the prior $p\left(\mathcal{S}_t, \mathcal{V}_t \left| \mathcal{Y}^t; \Psi_0\right.\right)$, propagate the states one-step ahead using the transition density $p\left(\mathcal{S}_{t+1}, \mathcal{V}_{t+1} \left| \mathcal{S}_t, \mathcal{V}_t; \Psi_0\right.\right)$. Next, given $\left\{\mathcal{S}_{t+1}^{(i)}, \mathcal{V}_{t+1}^{(i)}\right\}_{i=1}^{M}$, compute the likelihood, $p\left(\mathcal{Y}_{t+1} \left| \mathcal{S}_{t+1}^{(i)}, \mathcal{V}_{t+1}^{(i)}; \Psi_0\right.\right)$, and the importance sampling weight $w_{t+1}^{(i)} \propto p\left(\mathcal{Y}_{t+1} \left| \mathcal{S}_{t+1}^{(i)}, \mathcal{V}_{t+1}^{(i)}; \Psi_0\right.\right)$ M times. Draws from the posterior $p\left(\mathcal{S}_{t+1}, \mathcal{V}_{t+1} \left| \mathcal{Y}^{t+1}; \Psi_0\right.\right)$ involve resampling $\left\{\mathcal{S}_{t+1}^{(i)}, \mathcal{V}_{t+1}^{(i)}\right\}_{i=1}^{M}$ using $\left\{w_{t+1}^{(i)}\right\}_{i=1}^{M}$.

[14] Gordon, Salmond and Smith (1993) propose the bootstrap particle filter by combining sequential importance sampling with resampling with replacement.

Rao-Blackwellization of $\mathcal{S}_t$

A bootstrap particle filter, although a consistent estimator of the posterior of the states of a state space model, is not necessarily the most statistically efficient SMC method. Chen and Liu (2000) propose a mixture of Kalman filters that Rao-Blackwellizes a state space model to produce more efficient estimates of the states.

Our particle learning filter has a Rao-Blackwellization step that rests on the conditional linear and Gaussian structure of our state space models. Rao-Blackwellization of the posterior of the linear states, $\mathcal{S}_t$, decomposes the joint transition density of the states, $p\left(\mathcal{S}_t, \mathcal{V}_t \left| \mathcal{S}_{t-1}, \mathcal{V}_{t-1}; \Psi_0\right.\right)$, into $p\left(\mathcal{S}_t \left| \mathcal{S}_{t-1}, \mathcal{V}_t; \Psi_0\right.\right) p\left(\mathcal{V}_t \left| \mathcal{V}_{t-1}; \Psi_0\right.\right)$. The state equations (8.1) imply the Rao-Blackwellized transition density $p\left(\mathcal{S}_t \left| \mathcal{S}_{t-1}, \mathcal{V}_t; \Psi_0\right.\right)$ is multivariate normal with mean and variance linear in $\mathcal{S}_{t-1}$. Hence, draws from the posterior of $\mathcal{S}_t$ are multivariate normal and characterized by two sufficient statistics, given the date t particle histories of the non-linear states, $\{\mathcal{V}^{t,(i)}\}_{i=1}^{M}$.[15] Kalman filter recursions yield sufficient statistics that are the conditional posterior mean of $\mathcal{S}_t$, $\mathcal{S}_{t|t} = \mathbf{E}\left\{\mathcal{S}_t \left| \mathcal{Y}^t, \mathcal{V}^{t,(i)}; \Psi_0\right.\right\}$ $= \mathbf{E}\left\{\mathcal{S}_t \left| \mathcal{Y}^t, \mathcal{S}_{t-1|t-1}, \mathbf{\Sigma}_{t-1|t-1}; \Psi_0\right.\right\}$ and mean square error, $\mathbf{\Sigma}_{t|t} = \mathrm{Var}\left(\mathcal{S}_t \left| \mathcal{Y}^t, \mathcal{V}^{t,(i)}; \Psi_0\right.\right) =$ $\mathrm{Var}\left(\mathcal{S}_t \left| \mathcal{Y}^t, \mathcal{S}_{t-1|t-1}, \mathbf{\Sigma}_{t-1|t-1}; \Psi_0\right.\right)$. Since $\mathcal{S}_t$ is marginalized out analytically by the Kalman filter, Rao-Blackwellization lowers sampling error in the particle learning filter. The likelihood, $p\left(\mathcal{Y}_t \left| \mathcal{S}_t, \mathcal{V}_t; \Psi_0\right.\right)$, of the observation equations (8.2) is also multivariate normal with mean linear in $\mathcal{S}_t$. The posterior of $\mathcal{V}_t$ is not characterized analytically, but simulating the random walks (1.4) and (3.3) generates draws from the proposal $p\left(\mathcal{V}_t \left| \mathcal{V}_{t-1}; \Psi_0\right.\right)$.

Resampling, Propagation, and Proposal Densities

SMC methods need a proposal to create a particle stream of the states that is not "too far" from the true posterior. A bootstrap particle filter draws from $p\left(\mathcal{S}_{t+1}, \mathcal{V}_{t+1} \left| \mathcal{Y}^{t+1}; \Psi_0\right.\right)$ after propagating and resampling the states. An alternative is the auxiliary particle filter (APF) of Pitt and Shephard (1999, 2001). The APF is a resampling-propagation algorithm that can yield greater statistical efficiency compared with a bootstrap particle filter, given M.

Carvalho, et al build a particle learning filter using ideas adapted from the APF. Before resampling, a generic APF employs the Kalman filter predictive step to calculate

[15] Conditioning the sufficient statistics of $\mathcal{S}_t$ on $\mathcal{V}^{t,(i)}$ refers to a hypothetical history of the ith particle that ignores the effects of resampling. In practice, the actual sequence of the ith particle requires tracking $\mathcal{S}_{t|t}^{(i)}$ and $\mathbf{\Sigma}_{t|t}^{(i)}$ in the particle swarm.

the predictive likelihood, $\ell_t^{(i)} = p\left(\mathcal{Y}_{t+1} \middle| \mathcal{S}_t^{(i)}, \mathcal{V}_t^{(i)}; \Psi_0\right)$, that is proportional to the importance weight $w_{t|t+1}^{(i)} = \left(w_t^{(i)} \Big/ \sum_{i=1}^{M} w_t^{(i)}\right) \ell_t^{(i)}$, where $w_t^{(i)}$ is defined below. The linear and non-linear states are resampled using the normalized importance weights, $W_{t|t+1}^{(i)} = w_{t|t+1}^{(i)} \Big/ \sum_{i=1}^{M} w_{t|t+1}^{(i)}$. The Kalman filter is run a second time to produce the one-step ahead predictive likelihood $p\left(\mathcal{Y}_{t+1} \middle| \mathcal{S}_t^{(i)}, \mathcal{V}_t^{(i)}; \Psi_0\right)$ conditional on the resampled states and propagates these states into $\left\{\mathcal{S}_{t+1}^{(i)}, \mathcal{V}_{t+1}^{(i)}\right\}_{i=1}^{M}$. The weight $w_{t+1}^{(i)}$ is the ratio of the second-stage likelihood to resampled particles of $\ell_t^{(i)}$.

As already noted, state transition densities are exact under Rao-Blackwellization yielding more efficient estimates of the states. The Rao-Blackwellized transition density propagating the linear states, $p\left(\mathcal{S}_{t+1} \middle| \mathcal{S}_t, \mathcal{V}_t, \mathcal{Y}_{t+1}; \Psi_0\right)$, and drawing the non-linear states from the transition density $p\left(\mathcal{V}_{t+1} \middle| \mathcal{V}_t, \mathcal{Y}_{t+1}; \Psi_0\right)$ are also conditional on $\mathcal{Y}_{t+1}$ in the APF. Conditioning on $\mathcal{Y}_{t+1}$ propagates particles that carry greater weight in the likelihood. This can yield more efficient estimates of the states.[16]

Parameter Inference with Particle Learning

Our particle learning filter is grounded in algorithm 7 of Lopes and Tsay (2011). Their algorithm embeds Rao-Blackwellization and APF steps in a particle learning filter.[17]

Particle learning solves the problem that including proposals for Ψ_0 in the particle swarm of $\mathcal{S}_t$ and $\mathcal{V}_t$ requires propagating the static parameters. Without propagation, successive resampling of the particles of Ψ_0 would collapse the stream onto only a few particles. The result is a poor approximation of the posterior of Ψ_0. Particle learning avoids particle degeneracy by tracking a swarm of posterior distributions for Ψ_0, given $\left\{\mathcal{S}_{t|t}^{(i)}, \boldsymbol{\Sigma}_{t|t}^{(i)}, \mathcal{V}_{t+1}^{(i)}\right\}_{i=1}^{M}$. Similar to Rao-Blackwellization of $\mathcal{S}_t$, the posterior distribution of Ψ_0 is drawn from a vector of sufficient statistics that we label $\Gamma_t^{(i)}$, for $i = 1, \ldots, M$.

We place conjugate priors on the static parameters. The scale volatility parameters have inverse-gamma ($\mathcal{IG}$) priors, $\sigma_\ell^2 \sim \mathcal{IG}\left(\frac{\alpha_\ell}{2}, \frac{\beta_\ell}{2}\right)$, where α and β are scale and shape

[16] Resampling before propagation of the states mitigates particle degeneracy by conditioning $w_{t|t+1}^{(i)}$ on $\mathcal{Y}_{t+1}$ and $w_t^{(i)}$. Particle degeneracy occurs when a few particles carry most of the weight at future filtering steps. The result is unevenly distributed weights implying inadequate coverage of regions of high likelihood, which the APF aims to fix. However, Johansen and Doucet (2008) and Herbst and Schorfheide (2016) note the efficacy of the APF rests on $p\left(\mathcal{Y}_{t+1} \middle| \mathcal{S}_{t+1}^{(i)}, \mathcal{V}_{t+1}^{(i)}; \Psi_0\right)$ having thinner tails compared with $p\left(\mathcal{Y}_{t+1} \middle| \mathcal{S}_t^{(i)}, \mathcal{V}_t^{(i)}; \Psi_0\right)$.

[17] Early examples of particle learning are Liu and West (2001), Djuric and Miguez (2002), Fernhead (2002), and Storvik (2002). Särkkä (2013) has a useful summary of particle learning.

parameters and $\ell = \eta, \upsilon, \zeta_\pi, \zeta_1, \ldots, \zeta_5$, and κ. Our prior on θ is $\mathcal{TN}\left(\underline{\theta}, \underline{\sigma}^2_\theta; \theta \in (-1, 1)\right)$, where $\underline{\theta}$ and $\underline{\sigma}^2_\theta$ are its prior mean and variance. Since the priors of the static parameters are conjugate, the posterior distributions are analytic. This suggests laws of motion for parameters of the posterior distributions. For example, the posterior distribution of σ^2_η is $\sigma^{2(i)}_\eta \sim \mathcal{IG}\left(\frac{\alpha_t}{2}, \frac{\beta^{(i)}_{\eta,t}}{2}\right)$, where the laws of motion of the scale and shape parameters are $\alpha_t = \alpha_{t-1} + t - 1$ and $\beta^{(i)}_{\eta,t} = \sum_{\ell=1}^{t}\left[\sigma^{2(i)}_{\eta,\ell} - \sigma^{2(i)}_{\eta,\ell-1}\right]^2$. Hence, the particle $\beta^{(i)}_{\eta,t}$ becomes the sufficient statistic and its law of motion the transition equation to update the posterior of σ^2_η, $\sigma^{2(i)}_\eta \left| \mathcal{V}^{(i)}_{0,t}, \mathcal{V}^{(i)}_{t-1} \sim \mathcal{IG}\left(\frac{\alpha_t}{2}, \frac{\beta^{(i)}_{\eta,t}}{2}\right)\right.$.[18]

The particle learning filter employs the laws of motion of the sufficient statistics of Ψ_0 as proposals to update from $\Gamma^{(i)}_{t-1}$ to $\Gamma^{(i)}_t$, $i = 1, \ldots, M$. We denote the proposals as the system of transition equations $\Gamma^{(i)}_t = \oint\left(\Gamma^{(i)}_{t-1}, \mathcal{V}^{(i)}_t, \mathcal{V}^{(i)}_{t-1}, \mathcal{S}^{(i)}_t, \mathcal{Y}^t\right)$. The system of transition equations is appended to the process that draws $\mathcal{V}^{(i)}_{0,t}$ in the particle learning filter. The conditional posterior of the sufficient statistics obtained at date $t-1$, $\Gamma^{(i)}_{t-1}$, becomes the prior for updating to $\Gamma^{(i)}_t$. Subsequent to propagating the sufficient statistics, draw M particles of the static parameters from $\Psi^{(i)}_0 \sim p\left(\Psi_0 \left| \Gamma^{(i)}_t\right.\right)$. In essence, this equates $p\left(\Psi_0 \left| \mathcal{Y}^t, \mathcal{V}^{(i)}_t\right.\right)$ to $p\left(\Psi_0 \left| \Gamma^{(i)}_t\right.\right)$, which shows particle learning marginalizes the static parameters out of the posterior of the states. Hence, the particle learning filter yields posteriors of the filtered and smoothed states that fully reflect parameter uncertainty. Full sample estimates of Ψ_0 are obtained at date T and denoted $\hat{\Psi}_0$.

3.2 Estimating the Marginal Data Density

The likelihood of our baseline state space model is computed in the particle learning filter. The marginal data density (MDD), $p\left(\mathcal{Y}_t \left| \mathcal{Y}^{t-1}\right.\right)$, is estimated using advice in appendix A.2 of Pitt, dos Santos Silva, Giordani, and Kohn (2012). They compute the MDD of a state space model using the cloud of first- and second-stage weights, $\left\{w^{(i)}_{t-1|t}, w^{(i)}_t\right\}_{i=1}^M$. The weights yield an estimate of the MDD of a state space model

$$p\left(\mathcal{Y}_t \left| \mathcal{Y}^{t-1}\right.\right) = \left[\frac{1}{M}\sum_{i=1}^{M} w^{(i)}_t\right]\sum_{i=1}^{M} w^{(i)}_{t-1|t}. \tag{9}$$

We employ equation (9) to assess our baseline state space model against three alterna-

[18] The shape parameter is the numerator of the standard deviation of a random variable distributed $\mathcal{IG}$.

tive state space models that are described below.

3.3 A Rao-Blackwellized Particle Smoother

Rao-Blackwellization is not costless. Conditional linear states create problems for particle smoothing algorithms that merge forward filtering with backward simulation of particles drawn from $p\left(\mathcal{S}_t, \mathcal{V}_t \,|\, \mathcal{Y}^T, \Psi_0\right)$ as, for example, in Godsill, Doucet, and West (2004). This class of particle smoothers is known as forward-filtering with backward-simulation (FFBS).[19] A FFBS algorithm is a recursive decomposition of the posterior of the smoothed states that rely on Markovian state dynamics. However, this decomposition is unsuited for a conditionally linear and Gaussian state space model. The problem is marginalizing out $\mathcal{S}_t$ gives a likelihood conditional on $\mathcal{V}^t$ instead of conditioning only on $\mathcal{V}_t$.

Lindsten, Bunch, Särkkä, Schön, and Godsill (2016) propose a Rao-Blackwellized particle smoother to solve the problem. Their particle smoother extends the FFBS procedures of Godsill, Doucet, and West (2004) to a Rao-Blackwellized state space model by forward filtering of the linear and non-linear states, backward smoothing of the non-linear states, and forward smoothing of the linear states conditional on the smoothed non-linear states. Hence, Lindsten, et al call their algorithm a forward-backward-forward smoother.

We assign the initial forward filtering step of the Lindsten, et al particle smoother to our Rao-Blackwellized particle learning filter. This produces the joint posterior of $\mathcal{S}_t$ and $\mathcal{V}_t$, $p\left(\mathcal{S}_t, \mathcal{V}_t \,|\, \mathcal{Y}^t, \Psi_0\right)$, for a given realization of the parameter vector Ψ_0. After summarizing the approach of Lindsten, et al for generating smoothed draws of the states for a given parameter vector, we describe how we integrate out uncertainty over the parameter vector.

The backward smoothing step rests on the decomposition of the target density of the non-linear states, $p\left(\mathcal{V}^T \,\middle|\, \mathcal{Y}^T, \Psi_0\right)$, into $p\left(\mathcal{V}^t \,\middle|\, \mathcal{V}^{t+1}, \mathcal{Y}^T, \Psi_0\right) p\left(\mathcal{V}^{t+1:T} \,\middle|\, \mathcal{Y}^T, \Psi_0\right)$. However, Lindsten, et al initialize their Rao-Blackwellized particle smoother at date T by sampling from the filtered non-linear states, $\left\{\mathcal{V}_T^{(i)}\right\}_{i=1}^M$, to obtain smoothed non-linear states, $\left\{\tilde{\mathcal{V}}_T^{(i)}\right\}_{i=1}^M$. The factorization of $p\left(\mathcal{V}^T \,\middle|\, \mathcal{Y}^T, \Psi_0\right)$ is useful in smoothing backwards because $p\left(\mathcal{V}^t \,\middle|\, \mathcal{V}^{t+1}, \mathcal{Y}^T, \Psi_0\right)$ has information about the probabilities needed to draw

[19] Godsill, Doucet, and West (2004) is a classic example of a particle smoother built on FFBS methods. Carvalho, et al and Lopes and Tsay (2011) apply FFBS methods to particle learning filters.

$\left\{\tilde{\mathcal{V}}_t^{(i)}\right\}_{i=1}^{M}$ from $\left\{\mathcal{V}_t^{(i)}\right\}_{i=1}^{M}$. Since $p\left(\mathcal{V}^t \middle| \mathcal{V}^{t+1}, \mathcal{Y}^T, \Psi_0\right)$ is expensive to compute, Lindsten et al propose a simulator to backward filter the non-linear states.[20] The simulator relies on the decomposition $p\left(\mathcal{V}^t \middle| \mathcal{V}^{t+1}, \mathcal{Y}^T, \Psi_0\right) \propto p\left(\mathcal{Y}^{t+1:T}, \mathcal{V}^{t+1:T} \middle| \mathcal{V}^t, \mathcal{Y}^t, \Psi_0\right) p\left(\mathcal{V}^t \middle| \mathcal{Y}^t, \Psi_0\right)$, where $p\left(\mathcal{Y}^{t+1:T}, \mathcal{V}^{t+1:T} \middle| \mathcal{V}^t, \mathcal{Y}^t, \Psi_0\right) = \int p\left(\mathcal{Y}^{t+1:T}, \mathcal{V}^{t+1:T} \middle| \mathcal{S}_t, \mathcal{V}_t, \Psi_0\right) p\left(\mathcal{S}_t \middle| \mathcal{Y}^t, \mathcal{V}^t, \Psi_0\right) d\mathcal{S}_t$ is the predictive density that when normalized gives the probabilities of drawing a path for $\tilde{\mathcal{V}}^T$. Nonetheless, drawing $\tilde{\mathcal{V}}^{t+1:T}$ from $p\left(\mathcal{V}^{t+1:T} \middle| \mathcal{Y}^T, \Psi_0\right)$ is an approximation of the true density of the smoothed non-linear states. The Rao-Blackwellized particle smoother is repeated for dates $t = T-1, \ldots, 1$.

The third step of the Rao-Blackwellized particle smoother runs the Kalman filter forward to generate smoothed estimates of $\mathcal{S}_t$ and $\mathbf{\Sigma}_t$ by drawing from $p\left(\mathcal{S}_t \middle| \tilde{\mathcal{V}}^t, \mathcal{Y}^t, \Psi_0\right)$. These are the sufficient statistics of $\tilde{\mathcal{S}}_t$ employed in simulations to approximate the predictive density $p\left(\mathcal{Y}^{t+1:T}, \mathcal{V}^{t+1:T} \middle| \mathcal{V}^t, \mathcal{Y}^t, \Psi_0\right)$. Although $p\left(\mathcal{V}^t \middle| \tilde{\mathcal{V}}^{t+1}, \mathcal{Y}^T, \Psi_0\right)$ does not condition on $\tilde{\mathcal{S}}_t$, its estimates are needed to compute the probability of sampling $\tilde{\mathcal{V}}^{t:T}$.

Finally, given the ability to draw smoothed trajectories $\mathcal{V}^T$ and $\mathcal{S}^T$ from the joint conditional density of the states $p\left(\mathcal{V}^T, \mathcal{S}^T | \mathcal{Y}^T, \Psi_0\right)$ we integrate out uncertainty over the parameter vector using

$$p\left(\mathcal{V}^T, \mathcal{S}^T | \mathcal{Y}^T\right) = \int_{\Psi_0} p\left(\mathcal{V}^T, \mathcal{S}^T | \mathcal{Y}^T, \Psi_0\right) p(\Psi_0 | \mathcal{Y}^Y) d\Psi_0$$

via simulations using draws of Ψ_0 obtained from the particle learning filter. Further details are described in the online appendix.

4 The Model Space, Priors, Data, and Estimates

This section presents our priors on the baseline and the alternative state space models, the sample data, and estimates of the state space models. The posterior distributions of the filtered and smoothed states and static parameters are used to evaluate the state space models. We also report estimates of these posterior distributions for the state space model favored by the data.

[20] The Kalman filter creates an exact predictive density (up to a normalizing constant). However, computing the density involves running the Kalman filter across M particle streams while iterating it forward from the start of the sample to date T. These calculations are computationally costly, which motivate Lindsten, et al to approximate the predictive density with simulated sufficient statistics.

4.1 The Model Space

We evaluate our baseline state space model, $\mathcal{M}_0$, against three alternative state space models. Table 1 displays the restrictions on the inflation gap persistence parameter and sticky information weight that govern the transition and impulse dynamics of $\mathcal{S}_t$ in the state space models.[21] The first alternative state space model, $\mathcal{M}_1$, fixes the inflation gap persistence parameter and sticky information weight, $\theta_t = \theta$ and $\lambda_t = \lambda$. Letting both parameters drift defines the second alternative state space model, $\mathcal{M}_2$. Drift in the inflation gap persistence parameter evolves as the bounded random walk $\theta_{t+1} = \theta_t + \sigma_\phi \phi_{t+1}$, where $\phi_{t+1} \sim \mathcal{TN}\left(0,\, 1;\, \theta_{t+1} \in (-1,\, 1) \,\middle|\, \theta_t,\, \sigma_\phi^2\right)$.[22] The third alternative state space model, $\mathcal{M}_3$, holds λ fixed while the inflation gap persistence parameter drifts, θ_t.

Table 1: List of Baseline and Alternative state space models

	λ	
θ	Time-varying	Constant
Constant	$\mathcal{M}_0$	$\mathcal{M}_1$
Time-varying	$\mathcal{M}_2$	$\mathcal{M}_3$

Restrictions on the inflation gap persistence parameter and sticky information weight alter the non-linear states of a state space model. The baseline state space model has three non-linear states in $\mathcal{V}_t \equiv \mathcal{V}_{0,t} = \left[\varsigma_{\eta,t} \;\; \varsigma_{\upsilon,t} \;\; \lambda_t\right]'$. For $\mathcal{M}_1$, only the stochastic volatilities are non-linear states, $\mathcal{V}_{1,t} = \left[\varsigma_{\eta,t} \;\; \varsigma_{\upsilon,t}\right]'$. The non-linear state vectors are $\mathcal{V}_{2,t} = \left[\varsigma_{\eta,t} \;\; \varsigma_{\upsilon,t} \;\; \theta_t \;\; \lambda_t\right]'$ and $\mathcal{V}_{3,t} = \left[\varsigma_{\eta,t} \;\; \varsigma_{\upsilon,t} \;\; \theta_t\right]'$ for $\mathcal{M}_2$ and $\mathcal{M}_3$, respectively.

4.2 Priors and Initial Conditions

This section describes our priors on the static parameters of $\mathcal{M}_0$, $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$. Restrictions on the state space models imply the static parameter vectors Ψ_0, Ψ_1, Ψ_2, and Ψ_3 differ, but the state space models share many static parameters in common. For example, replace σ_κ^2 with λ in $\mathcal{M}_1$ to produce $\Psi_1 = \left[\sigma_\eta^2 \;\; \sigma_\upsilon^2 \;\; \sigma_{\zeta,\pi}^2 \;\; \sigma_{\zeta,1}^2 \;\; \sigma_{\zeta,2}^2 \;\; \sigma_{\zeta,4}^2 \;\; \sigma_{\zeta,5}^2 \;\; \lambda \;\; \theta\right]'$ that otherwise is identical to $\Psi_0 = \left[\sigma_\eta^2 \;\; \sigma_\upsilon^2 \;\; \sigma_{\zeta,\pi}^2 \;\; \sigma_{\zeta,1}^2 \;\; \sigma_{\zeta,2}^2 \;\; \sigma_{\zeta,4}^2 \;\; \sigma_{\zeta,5}^2 \;\; \sigma_\kappa^2 \;\; \theta\right]'$. Drifting inflation gap persistence parameter and sticky information weight in $\mathcal{M}_2$ yield $\Psi_2 =$

[21] We thank the editor and referees for suggesting the model space be expanded as shown in table 1.
[22] The online appendix solves for the state and observation equations of $\mathcal{M}_2$.

$\left[\sigma_\eta^2\ \sigma_\upsilon^2\ \sigma_{\zeta,\pi}^2\ \sigma_{\zeta,1}^2\ \sigma_{\zeta,2}^2\ \sigma_{\zeta,4}^2\ \sigma_{\zeta,5}^2\ \sigma_\kappa^2\ \sigma_\phi^2\right]'$ that equals Ψ_0 up to substituting σ_ϕ^2 for θ. Finally, swap drift in the sticky information weight for drift in the inflation gap persistence parameter in $\mathcal{M}_3$ to set $\Psi_3 = \left[\sigma_\eta^2\ \sigma_\upsilon^2\ \sigma_{\zeta,\pi}^2\ \sigma_{\zeta,1}^2\ \sigma_{\zeta,2}^2\ \sigma_{\zeta,4}^2\ \sigma_{\zeta,5}^2\ \lambda\ \sigma_\phi^2\right]'$.

We posit priors for the static volatility parameters and initial conditions of θ_t, λ_t, $\mathcal{S}_t$, and $\mathcal{V}_{k,t}$, $k = 0$, 1, 2, and 3. The scale volatility parameters on the stochastic volatilities, random walks of θ_t and λ_t, and measurement errors are given inverse gamma ($\mathcal{IG}$) priors. Table 2 reports the scale and shape parameters, α_ℓ and β_ℓ, of the $\mathcal{IG}$ priors along with the implied prior means, 5% and 95% quantiles.

Table 2: Priors on the Scale Volatility Parameters

					Quantiles	
Scale Volatility on Innovation to		α_ℓ	β_ℓ	Mean	5%	95%
Trend Inflation SV, $\ln \varsigma_{\eta,t+1}$:	σ_η^2	3.0	0.04	0.04	[0.005,	0.114]
Gap Inflation SV, $\ln \varsigma_{\upsilon,t+1}$:	σ_υ^2	3.0	0.04	0.04	[0.005,	0.114]
TVP-AR1 Coefficient, θ_{t+1}:	σ_ϕ^2	3.0	0.01	0.01	[0.001,	0.028]
Sticky Information Weight, λ_{t+1}:	σ_κ^2	3.0	0.01	0.01	[0.001,	0.028]
Measurement Error on π_t:	$\sigma_{\zeta,\pi}^2$	20.0	2.88	0.16	[0.092,	0.265]
Measurement Error on $\pi_{t,t+h}^{SPF}$:	$\sigma_{\zeta,h}^2$	20.0	2.88	0.16	[0.092,	0.265]

Note: Priors on the static volatility parameters are $\sigma_\ell^2 \sim \mathcal{IG}\left(\frac{\alpha_\ell}{2}, \frac{\beta_\ell}{2}\right)$, where $\ell = \eta,\ \upsilon,\ \kappa,\ \zeta_\pi$, and ζ_h for $h = 1, \ldots, 5$. For $\alpha_\ell > 2$, the mean of the $\mathcal{IG}$ distribution is $\beta_\ell/(\alpha_\ell - 2)$. See the online appendix for more about restricting σ_ϕ^2 and σ_κ^2 to guarantee $\theta_t \in (-1, 1)$ and $\lambda_t \in (0, 1)$.

Several features of our priors are worth discussing. First, we center the priors of σ_η^2 and σ_υ^2 around the fixed coefficient values used by Stock and Watson (2007). Next, the prior mean of 0.01 assigned to σ_κ^2 is smaller reflecting the bounded support of λ_t. Nonetheless, this prior admits substantial variation in λ_t between the bounds of zero and one when estimating $\mathcal{M}_0$ and $\mathcal{M}_2$. For similar reasons, the prior for the volatility of shocks to inflation gap persistence, σ_ϕ^2, in $\mathcal{M}_2$ and $\mathcal{M}_3$, has been centered around a value of 0.01. Second, our priors on σ_η^2 and σ_υ^2 deliver quantiles that exhibit greater variation compared with σ_ϕ^2 and σ_κ^2. Third, the quantiles of $\sigma_{\zeta,\pi}^2$, $\sigma_{\zeta,1}^2$, …, $\sigma_{\zeta,5}^2$ depict our belief that the measurement errors of π_t and $\pi_{t,t+h}^{SPF}$ are volatile.

Table 3 lists our priors for the static inflation gap persistence parameter, θ, and

sticky information weight, λ. The priors obey $\theta \in (-1, 1)$ and $\lambda \in (0, 1)$. For $\mathcal{M}_0$, and $\mathcal{M}_1$, we endow θ with a truncated standard normal prior. This prior has 5% and 95% quantiles at -0.87 and 0.87. For $\mathcal{M}_1$, and $\mathcal{M}_3$, the beta prior for λ has shape parameters of unity and is thus equivalent to the uniform distribution on the unit interval. In essence, we have non-informative priors over θ and λ.

Table 3: Priors on the Static Inflation Gap Persistence Parameters and Sticky Information Weight

					Quantiles	
		Distribution	Mean	STD	5%	95%
Gap Inflation AR(1):	θ	$\mathcal{TN}(-1, 1)$	0.0	1.0	$[-0.87,$	$0.87]$
Sticky Information Weight:	λ	$\text{Beta}(1, 1)$	0.5	–	$[0.05,$	$0.95]$

Note: The priors $\mathcal{TN}(-1, 1)$ and $\text{Beta}(1, 1)$ represent a truncated normal prior on the open interval $(-1, 1)$ and the beta distribution with shape parameters of unity.

Priors on the initial conditions of the linear and non-linear states appear in Table 4. The left side of the table lists priors on the initial conditions of the linear states, τ_0, ε_0, $F_0\tau_0$, and $F_0\varepsilon_0$. Initial conditions on the stochastic volatilities, $\ln \varsigma_{\eta,0}^2$ and $\ln \varsigma_{\upsilon,0}^2$, drifting sticky information weight, λ_0, and drifting inflation gap parameter, θ_0, are found on the right side of the Table 4. We draw τ_0 and $F_0\tau_0$ from normal priors. The prior means are two percent, which is about the mean of GNP deflator inflation on a $1958Q1$ to $1967Q4$ training sample. A variance of 100^2 yields an approximately flat prior over the relevant range of values for τ_0 and $F_0\tau_0$ in post-war U.S. data. The joint prior of ε_0 and $F_0\varepsilon_0$ is drawn from the ergodic bivariate normal distribution $\mathcal{N}(\mathbf{0}_{2\times 1}, \mathbf{\Sigma}_0)$; see the notes to table 4. The ergodic mean of gap inflation is zero by definition. Prior variances are produced by the distribution using the initial particle draws of $\mathcal{V}_{0,0}$, $\mathcal{V}_{1,0}$, $\mathcal{V}_{2,0}$, or $\mathcal{V}_{3,0}$.

The last two columns of Table 4 display our priors on initial conditions of the non-linear states. We endow priors of $\ln \varsigma_{\upsilon,0}^2$ and $\ln \varsigma_{\eta,0}^2$ with normal distributions. Prior means are calibrated to pre-1968 inflation data similar to Stock and Watson (2007).[23] Large variances reflect prior uncertainty about $\ln \varsigma_{\upsilon,0}^2$ and $\ln \varsigma_{\eta,0}^2$. Table 4 also shows the priors of θ_0 and λ_0 are drawn from $\mathcal{TN}$ distributions with means of zero and 0.5, unit

[23] One-third of the training sample variance of the first difference of GNP deflator inflation is attributed to trend volatility and the remaining two thirds to gap volatility.

variances, and are truncated to $(-1, 1)$ and $(0, 1)$, respectively. These priors are uninformative about initial conditions on the drifting inflation gap persistence parameter and sticky information weight, given the relevant bounds.

TABLE 4: PRIORS ON INITIAL CONDITIONS OF THE LINEAR AND NON-LINEAR STATES

Initial State		Prior Distribution	Initial State		Prior Distribution
τ_0	$\sim$	$\mathcal{N}\left(2.0,\ 100.0^2\right)$	$\ln \varsigma^2_{\eta,0}$	$\sim$	$\ln\mathcal{N}\left(\ln 0.2 - 5.0,\ 10.0\right)$
$F_0\tau_0$	$\sim$	$\mathcal{N}\left(2.0,\ 100.0^2\right)$	$\ln \varsigma^2_{\upsilon,0}$	$\sim$	$\ln\mathcal{N}\left(\ln 0.4 - 5.0,\ 10.0\right)$
ε_0	$\sim$	$\mathcal{N}\left(0.0,\ \sigma^2_{\varepsilon_0}\right)$	θ_0	$\sim$	$\mathcal{TN}\left(0.0,\ 1.0,\ -1.0,\ 1.0\right)$
$F_0\varepsilon_0$	$\sim$	$\mathcal{N}\left(0.0,\ \sigma^2_{F_0\varepsilon_0}\right)$	λ_0	$\sim$	$\mathcal{TN}\left(0.5,\ 1.0,\ 0.0,\ 1.0\right)$

Note: The priors on ε_0 and $F_0\varepsilon_0$ are drawn jointly from $\mathcal{N}\left(\mathbf{0}_{2\times 1}, \mathbf{\Sigma}_0^{(i)}\right)$, where $\sigma^2_{\varepsilon_0}$ and $\sigma^2_{F_0\varepsilon_0}$ are the diagonal elements of

$$\mathbf{\Sigma}_0^{(i)} = \sum_{j=0}^{\infty} \begin{bmatrix} \theta^{(i)} & 0 \\ (1-\lambda_0^{(i)})\theta^{(i)} & \lambda_0^{(i)}\theta^{(i)} \end{bmatrix}^j \begin{bmatrix} \varsigma^{2,(i)}_{\upsilon,0} & \lambda_0^{(i)}\varsigma^{2,(i)}_{\upsilon,0} \\ \lambda_0^{(i)}\varsigma^{2,(i)}_{\upsilon,0} & \lambda_0^{2,(i)}\varsigma^{2,(i)}_{\upsilon,0} \end{bmatrix} \begin{bmatrix} \theta^{(i)} & (1-\lambda_0^{(i)})\theta^{(i)} \\ 0 & \lambda_0^{(i)}\theta^{(i)} \end{bmatrix}^j,$$

and $\lambda_0^{(i)}$, and $\varsigma^{2,(i)}_{\upsilon,0}$ are the ith particle draws from priors on the associated initial conditions. If $\theta^{(i)} = 0$, the formula used to compute $\mathbf{\Sigma}_0^{(i)}$ collapses to the middle term above.

4.3 The Data

The data are real-time realized inflation, π_t, and h-step ahead average SPF inflation prediction, $\pi^{SPF}_{t,t+h}$, where $h = 1, 2, 3, 4, 5$. We obtain the data from the Real-Time Data Set for Macroeconomists (RTDSM) that is made available by the Federal Reserve Bank (FRB) of Philadelphia. The sample runs from $1968Q4$ to $2018Q3$ for π_t and $\pi^{SPF}_{t,t+h}$.[24]

We measure realized inflation using second-release GNP/GDP deflator inflation provided by the RTDSM.[25] The RTDSM compiles the second-release data of the GNP/GDP de-

[24]The SPF measured the price level of output with the implicit GNP deflator before $1992Q1$. From $1992Q1$ to $1996Q4$, the implicit GDP deflator played this role. It was replaced by the chain weighted GDP deflator beginning in $1997Q1$.

[25]These data are found at `https://www.philadelphiafed.org/-/media/research-and-data/real-time-center/real-time-data/data-files/files/xlsx/p_first_second_third.xlsx`. Second-release data for $1995Q3$ are unavailable because of a federal government shutdown. We fill in the missing observations with the corresponding third-release data collected by the RTDSM.

flator in growth rates. We convert the annualized growth rate, $G_t = 100((P_t/P_{t-1})^4 - 1)$, into continuously compounded growth rates using $\pi_t = \ln(1 + G_t/100)$. The SPF solicits forecasts for the GDP/GNP deflator in levels. We convert these price levels into expected growth rates by differencing the logs of the term structure of level forecasts.[26] Average SPF inflation predictions include a nowcast of the level of the GNP or GDP deflator and forecasts of these price levels at 1-, 2-, 3-, and 4-quarters ahead. We denote average SPF nowcast and 1-, 2-, 3-, and 4-quarter ahead forecasts $\pi^{SPF}_{t,t+1}$, $\pi^{SPF}_{t,t+2}$, $\pi^{SPF}_{t,t+3}$, $\pi^{SPF}_{t,t+4}$, and $\pi^{SPF}_{t,t+5}$, respectively.[27]

The RTDSM collects the quarter t SPF inflation predictions without full knowledge of π_t. Vintages of π_t reflect data releases that were publicly available around the middle of quarter t and most often the publicly available information contains observations through quarter $t-1$. We comply with this timing protocol by assuming $\pi^{SPF}_{t,t+1}, \ldots,$ and $\pi^{SPF}_{t,t+5}$ are formed conditional on information available at the end of quarter $t-1$.

Figure 1 displays π_t and the average SPF predictions. The data has several interesting features. First, figure 1 shows π_t is more volatile than the average SPF inflation predictions. Second, the average SPF inflation predictions are smoother and more centered on π_t as h increases. For example, figure 1(a) shows the average SPF inflation nowcast, $\pi^{SPF}_{t,t+1}$, moving with π_t, during the inflation spikes of the 1973–1975 recession and the double dip recessions of the early 1980s. However, the spikes in the average SPF inflation predictions around these recessions are inversely related to h as depicted in figure 1(b), 1(c), and 1(d). Figure 1(d) also shows π_t fluctuating around $\pi^{SPF}_{t,t+5}$ subsequent to the Volcker disinflation.[28]

4.4 Posterior Estimates of the Static Parameters and Model Fit

Table 5 lists moments of the posterior distributions of the elements of $\hat{\Psi}_0$, $\hat{\Psi}_1$, $\hat{\Psi}_2$, and $\hat{\Psi}_3$. The posterior moments are medians and 90 percent uncertainty bands (*i.e.*, five and 95 percent quantiles) that are displayed in brackets. Log MDDs are reported at the bottom

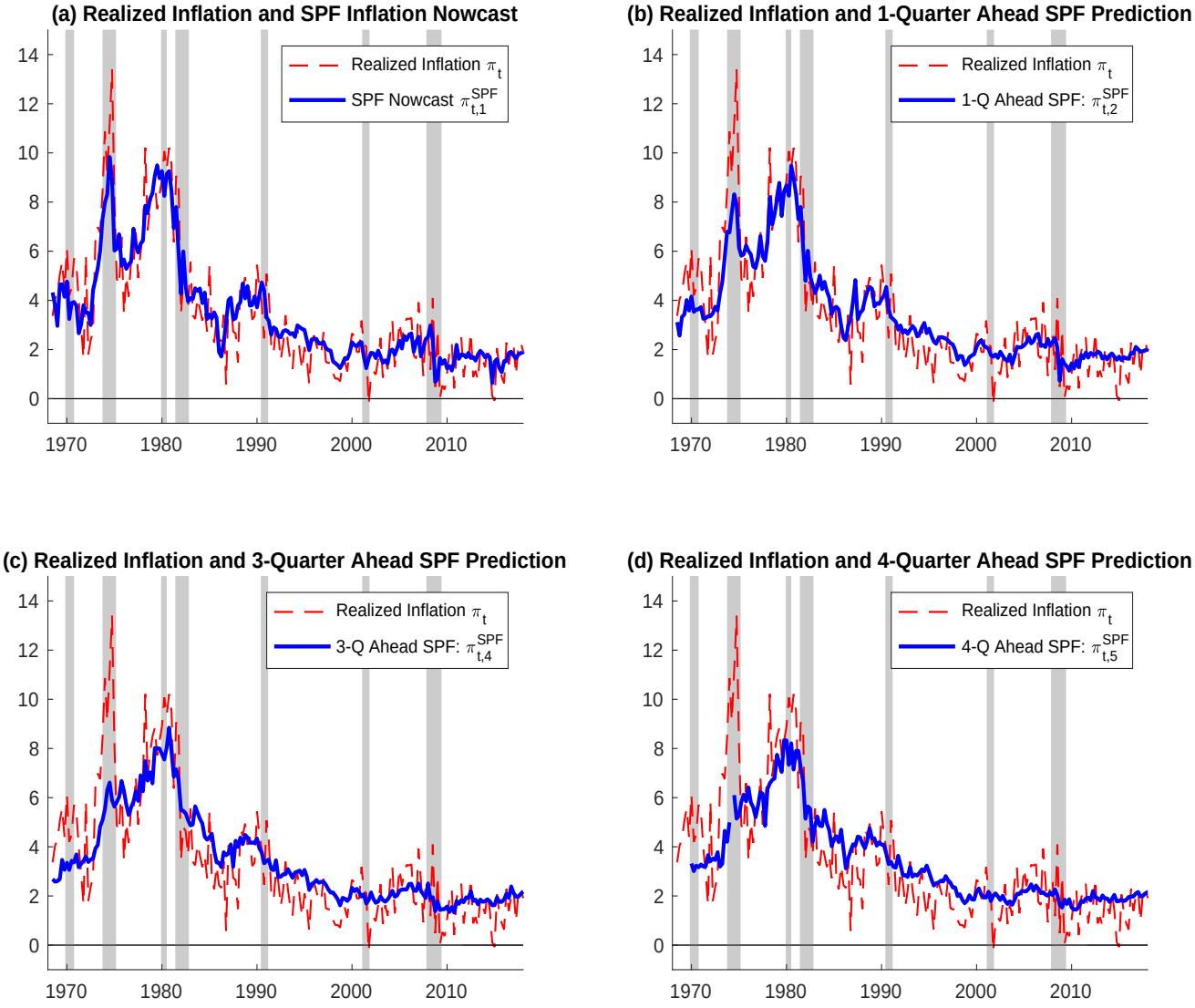
[26] Given the levels forecast of the GNP/GDP deflator for two adjacent quarters, $F_t P_{t+h}$ and $F_t P_{t+h-1}$, the sticky information inflation forecast is $F_t \pi_{t+h} = 400 \ln(F_t P_{t+h}/F_t P_{t+h-1})$. As is common in the literature that evaluates SPF forecasts, this procedure ignores Jensen-inequality effects on inflation forecasts arising from converting the level of forecasts into log differences; for example, see Aruoba (2019).

[27] Observations are missing in $\pi^{SPF}_{t,t+5}$ during 1969, 1970, and 1974. We modify the Kalman filter in the particle learning filter and Rao-Blackwellized particle smoother to accommodate these observations.

[28] Meltzer (2014, p. 1209) marks 1986 as the end of the Volcker disinflation.

of table 5. Each state space model has been estimated with $M = 100{,}000$ particles.

Figure 1: Realized Inflation and SPF Inflation Predictions, 1968$Q4$ to 2018$Q3$



Note: The plots contain vertical gray bands that denote NBER dated recessions.

Estimates of the static parameters differ in several ways. Table 5 indicates $\mathcal{M}_2$ is responsible for the smallest posterior median of σ_η^2 and the largest posterior median of σ_υ^2. This suggests τ_t and $\varsigma_{\eta,t}$ are smoother and ε_t and $\varsigma_{\upsilon,t}$ are more volatile in the state space model that has drift in the inflation gap persistence parameter and sticky information weight compared with $\mathcal{M}_0$, $\mathcal{M}_1$, or $\mathcal{M}_3$. Compared with the other static parameters, estimates of σ_η^2 and σ_υ^2 also display more variation across the four state space models. Estimates of the other static parameters are broadly similar, except under $\mathcal{M}_1$. This state space model, which fixes $\theta = \theta_t$ and $\lambda = \lambda_t$, produces the largest posterior median estimate of $\sigma_{\zeta,\pi}^2$ and smallest posterior median estimates of $\sigma_{\zeta,1}^2, \ldots, \sigma_{\zeta,5}^2$. The reason is a greater share of the variation in π_t is loaded onto $\zeta_{\pi,t}$ while the average SPF inflation predictions are more responsive to $\mathcal{S}_t$ and $\mathcal{V}_{0,t}$.

The log MDDs of $\mathcal{M}_0$, $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ are estimated using equation (9). The log MDDs are labeled $\ln \text{MDD}\left(\mathcal{M}_i \middle| \mathcal{Y}^T\right)$ for $i = 0$, 1, 2, and 3 at the bottom of table 5. We average over 250 repetitions of the particle learning filter to calculate the log MDDs. Uncertainty over the log MDDs resulting from the SMC approximation is measured using the numerical standard errors of the log MDD estimates computed on the 250 repetitions.[29] Table 5 shows that the difference between the log MDD associated with $\mathcal{M}_0$ and the other three alternatives is always greater than 6.5, which, in the language of Kass and Raftery (1995), consitutes very strong evidence in favor of $\mathcal{M}_2$. Moreover, the numerical standard errors of the log MDD estimates are thwarted by these differences in the log MDD. Hence, our discussion is centered on estimates produced using $\mathcal{M}_2$.[30]

Figure 2 plots particle learning filter estimates of the scale volatility parameters σ_η^2, σ_υ^2, σ_ϕ^2, and σ_κ^2 of $\mathcal{M}_2$. Paths of these estimates appear in figures 2(a), 2(b), 2(c) and 2(d) along with 68% and 90% uncertainty bands in dark and light shades. The plots show estimates of σ_η^2, σ_ϕ^2, and σ_κ^2 that cease to change much, if at all, after 1983. The exception is σ_υ^2. In figure 2(b), the static scale volatility of $\ln \varsigma_{\upsilon,t}^2$ in $\mathcal{M}_2$ drifts up during the sample and especially after 1983. Between 1983 and the end of the sample, σ_υ^2 almost doubles in size from about 0.025 in $1983Q1$ to 0.045 at $2018Q3$ while its 90% uncertainty bands are widest running from 0.02 to 0.15 after the 2007–2009 recession.

[29] The use of numerical standard errors for gauging the uncertainty of simulation-based estimates is grounded in the work of Geweke (1989), see also Fuentes-Albero and Melosi (2013), and Herbst and Schorfheide (2014) for applications in the context of log MDD estimates.

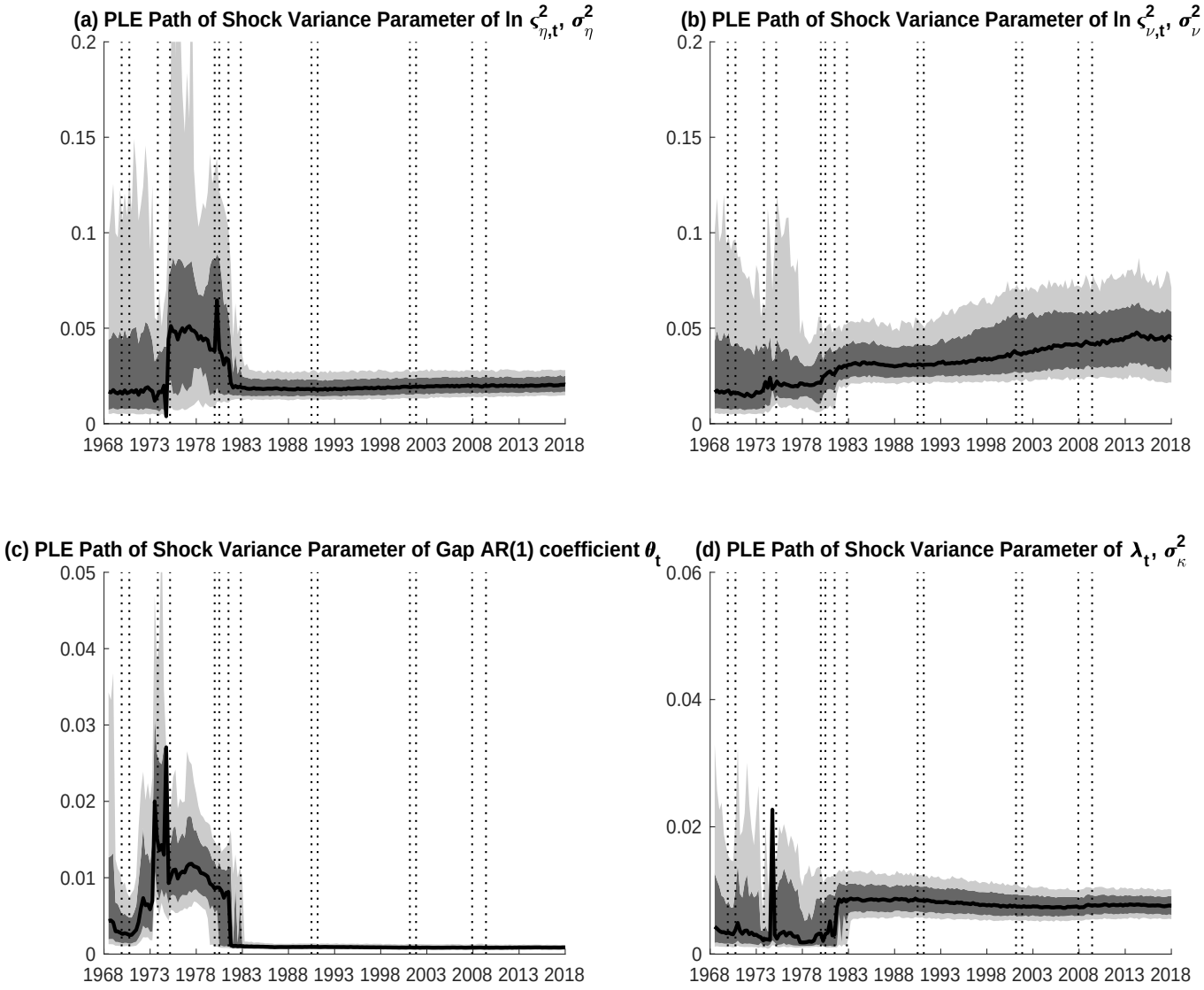
[30] Estimates of $\mathcal{M}_0$, $\mathcal{M}_1$, and $\mathcal{M}_3$ are in the online appendix. The online appendix also shows broadly similar estimates from state space models that omit measurement error, $\zeta_{\pi,t}$, from the observation equation (1.1) of π_t. These state space models are not favored by the data in terms of log MDDs.

Table 5: Parameter Estimates and Log MDDs of the state space models

Parameter	Models			
	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$
Variances of shocks to SV processes				
σ_η^2 (Trend SV)	0.037	0.051	0.020	0.018
	[0.026, 0.048]	[0.036, 0.150]	[0.015, 0.028]	[0.013, 0.027]
σ_v^2(Gap SV)	0.006	0.232	0.045	0.036
	[0.005, 0.009]	[0.026, 0.329]	[0.021, 0.074]	[0.026, 0.052]
Persistence of inflation gap				
θ	0.758	0.707	–	–
	[0.663, 0.854]	[0.635, 0.780]		
σ_ϕ^2	–	–	0.001	0.003
			[0.001, 0.001]	[0.002, 0.004]
Forecast stickiness				
λ	–	0.307	–	0.324
		[0.175, 0.373]		[0.268, 0.383]
σ_κ^2	0.005	–	0.008	–
	[0.004, 0.007]		[0.005, 0.010]	
Measurement error variances				
$\sigma_{\zeta,\pi}^2$	0.517	0.778	0.612	0.486
	[0.429, 0.621]	[0.644, 0.939]	[0.519, 0.723]	[0.403, 0.592]
$\sigma_{\zeta,1}^2$	0.113	0.008	0.101	0.136
	[0.096, 0.135]	[0.007, 0.010]	[0.085, 0.122]	[0.116, 0.161]
$\sigma_{\zeta,2}^2$	0.042	0.008	0.042	0.042
	[0.035, 0.051]	[0.007, 0.010]	[0.036, 0.051]	[0.035, 0.051]
$\sigma_{\zeta,3}^2$	0.044	0.008	0.044	0.043
	[0.037, 0.052]	[0.007, 0.010]	[0.037, 0.053]	[0.036, 0.051]
$\sigma_{\zeta,4}^2$	0.047	0.008	0.044	0.050
	[0.040, 0.055]	[0.007, 0.010]	[0.037, 0.054]	[0.041, 0.060]
$\sigma_{\zeta,5}^2$	0.063	0.008	0.066	0.059
	[0.054, 0.075]	[0.007, 0.010]	[0.054, 0.079]	[0.049, 0.074]
$\ln \text{MDD}\left(\mathcal{M}_i \middle\| y_{1:T}\right)$	−528.964	−535.401	−520.613	−527.144
	(0.421)	(0.486)	(0.349)	(0.394)

Note: The table contains posterior moments and log MDDs for the state space models $\mathcal{M}_0$, $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$. Estimates of the static scale volatility parameters are produced using M = 100,000 particles. The main entry for every static parameter reports its posterior median estimated on the full sample that begins in $1968Q3$ and ends at $T = 2018Q3$. Five and 95 percent quantiles appear in the brackets below the posterior medians. Log MDDs for $\mathcal{M}_i$, $i = 0$, 1, 2, and 3 are denoted $\ln \text{MDD}\left(\mathcal{M}_i \middle\| y^T\right)$ and computed using equation (9). The reported values are the average estimates obtained from 250 repetitions of the particle learning filter, and the associated numerical standard errors appear in parentheses below each estimate.

Figure 2: Particle Learning Filter Estimates of Static Parameters, 1968$Q4$ to 2018$Q3$



Note: Posterior quantiles of several static parameters of the state space model $\mathcal{M}_2$. In each panel, the solid line depicts the posterior median while the dark and light shaded areas correspond to 68% and 90% uncertainty bands, respectively. Dotted vertical lines denote NBER recession peaks and troughs. Estimates of the particle learning filter paths of the static parameters are denoted PLE.

4.5 Trend and Gap Inflation

We present filtered estimates of rational expectations and sticky information trend and gap inflation produced by $\mathcal{M}_2$ in figure 3. This figure plots π_t, the average SPF inflation nowcast, $\pi^{SPF}_{t,t+1}$, 4-quarter ahead average SPF inflation prediction, $\pi^{SPF}_{t,t+5}$, filtered rational expectations and sticky information trend inflation, $\tau_{t|t}$ and $F_{t|t}\tau_t$, and filtered rational expectations and sticky information gap inflation, $\varepsilon_{t|t}$ and $F_{t|t}\varepsilon_t$, on the $1968Q4$ to $2018Q3$ sample. These plots depict $\pi^{SPF}_{t,t+1}$ and $\pi^{SPF}_{t,t+5}$ with solid (red) lines, $F_{t|t}\tau_t$ and $F_t\varepsilon_{t|t}$ with dot-dashed (black) lines, and τ_t and $\varepsilon_{t|t}$ with dotted (red) lines.

Figures 3(a) and 3(b) give evidence of the role $\pi^{SPF}_{t,t+5}$ has in estimating $F_{t|t}\tau_t$. Before the Volcker disinflation, $\pi^{SPF}_{t,t+1}$ often deviates substantially from $F_{t|t}\tau_t$. This relationship is reversed mostly from the end of the Volcker disinflation to $2018Q3$. The spread between $\pi^{SPF}_{t,t+5}$ and $F_{t|t}\tau_t$ is smaller during the 1973–1975 recession, the double-dip recessions of the early 1980s, and the Volcker disinflation. From 1990 to the end of the sample, $\pi^{SPF}_{t,t+5}$ fluctuates around $F_{t|t}\tau_t$. This suggests $F_{t|t}\tau_t$ depends on the relative smoothness of the 4-quarter ahead average SPF inflation prediction.

Compared with $\tau_{t|t}$ and $F_{t|t}\tau_t$, π_t is relatively volatile in figure 3(c). During the first oil price shock, π_t is a third or more greater than $\tau_{t|t}$ and $F_{t|t}\tau_t$. However, $F_{t|t}\tau_t$, π_t explain much of the increases in π_t in the late 1970s and early 1980s. Subsequently, π_t is often less than $\tau_{t|t}$ and $F_{t|t}\tau_t$ during the rest of the sample.

Figure 3(d) depict $\varepsilon_{t|t}$ and $F_{t|t}\varepsilon_t$ as rising from about -2% in $1968Q4$ to almost 4% in 1970. The 1973–1975 recession sees the largest spikes in $\varepsilon_{t|t}$ and $F_{t|t}\varepsilon_t$ reaching 9% or more before falling to about -2.5 percent by 1976. Between the Volcker disinflation and the end of the 1980s, $\varepsilon_{t|t}$ and $F_{t|t}\varepsilon_t$ are negative with a trough below -3.0% in 1986.

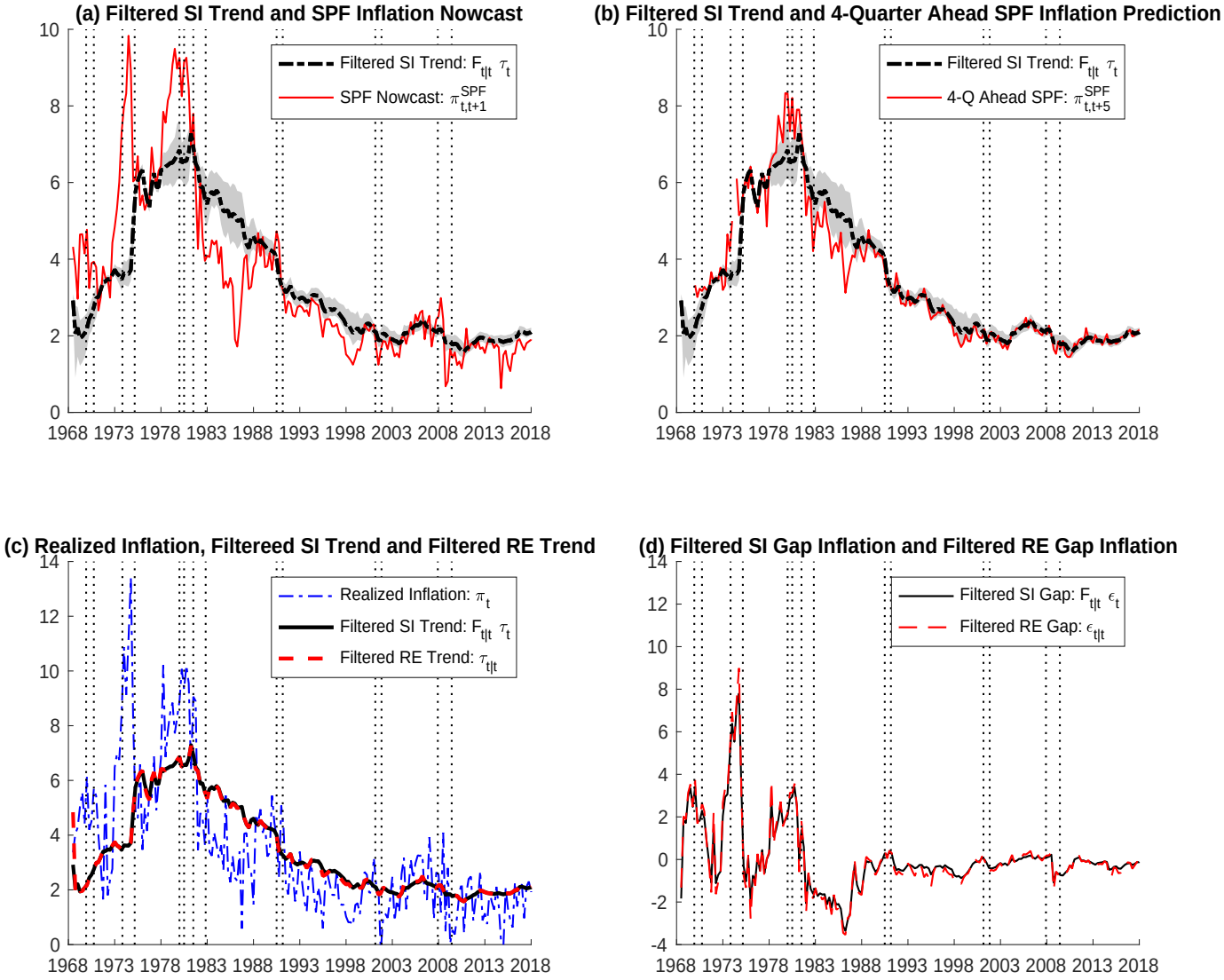
Applying our SMC methods to $\mathcal{M}_2$ also produce estimates of $\tau_{t|t}$ and $F_{t|t}\tau_t$ in figure 3(c) and of $\varepsilon_{t|t}$ and $F_{t|t}\varepsilon_t$ in figure 3(d) that are nearly identical on the $1968Q4$–$2018Q3$ sample. The reason is that the rational expectations and sticky information states are driven by the same set of shocks up to the innovations to λ_t, as the state equations (8.1) show. When λ_t is small, the transition and impulse dynamics of the rational expectations and sticky information states are similar. Otherwise, $F_{t|t}\tau_t$ $(F_{t|t}\varepsilon_t)$ are close to $\tau_{t|t}$ $(\varepsilon_{t|t})$ because the sticky information states adjust slowly to almost the identical set of sets driving the rational expectations states. As we discuss below, drift in the sticky information weight is important for capturing low frequency fluctuations in the term structure of average SPF inflation predictions.

In summary, our estimates of the linear states are a counterpoint to studies that find trend or gap inflation dominate movements in realized inflation. Stock and Watson (2007, 2010) present estimates of trend inflation that track realized inflation while Cogley and Sbordone (2008) report estimates of trend inflation that are smooth and point to the dominance of gap inflation in explaining realized inflation. Figures 3(c) and 3(d) show that conditioning estimates of $\mathcal{M}_2$ on average SPF inflation predictions produces estimates of trend and gap inflation that fall somewhere between these polar cases. For example, most of the inflation spike around the 1973–1975 recession is attributed to $\varepsilon_{t|t}$ and $F_{t|t}\varepsilon_t$ by $\mathcal{M}_2$. However, by the late 1970s and 1980 recession, $\tau_{t|t}$ and $F_{t|t}\tau_t$ account for a larger share of the surge in realized inflation.

Our estimates of trend and gap inflation can also be used to infer the beliefs the average SPF respondent has about the future path of realized inflation. For example, $\varepsilon_{t|t}$ and $F_{t|t}\varepsilon_t$ are negative for most of the 1980s. These estimates suggest the average SPF participant anticipated the Volcker disinflation would only produce a transitory drop in realized inflation, which is consistent with Goodfriend and King (2005) and Meltzer (2014, p. 1131). They argue households, firms, and investors expected the Volcker disinflation would only produce a transitory decline in inflation after 1984. Similarly, $\tau_{t|t}$ and $F_{t|t}\tau_t$ are often greater than π_t after 2007 indicating the average member of the SPF expected realized inflation to rise back to trend in the short- to medium-run.[31]

[31] The Beveridge and Nelson (1981) decomposition is the source of these predictions; see Nelson (2008). There is a Beveridge-Nelson trend in our state space models because τ_t is a random walk (1.2).

Figure 3: Realized Inflation, SPF Inflation Predictions, and Estimates of Trend and Gap Inflation, 1968$Q4$ to 2018$Q3$



Note: The top row of charts contains light gray shaded areas that represent 68 percent uncertain bands around estimates of filtered sticky information trend inflation, $F_{t|t}\tau_t$. The vertical dotted bands denote NBER dated recessions in the four charts.

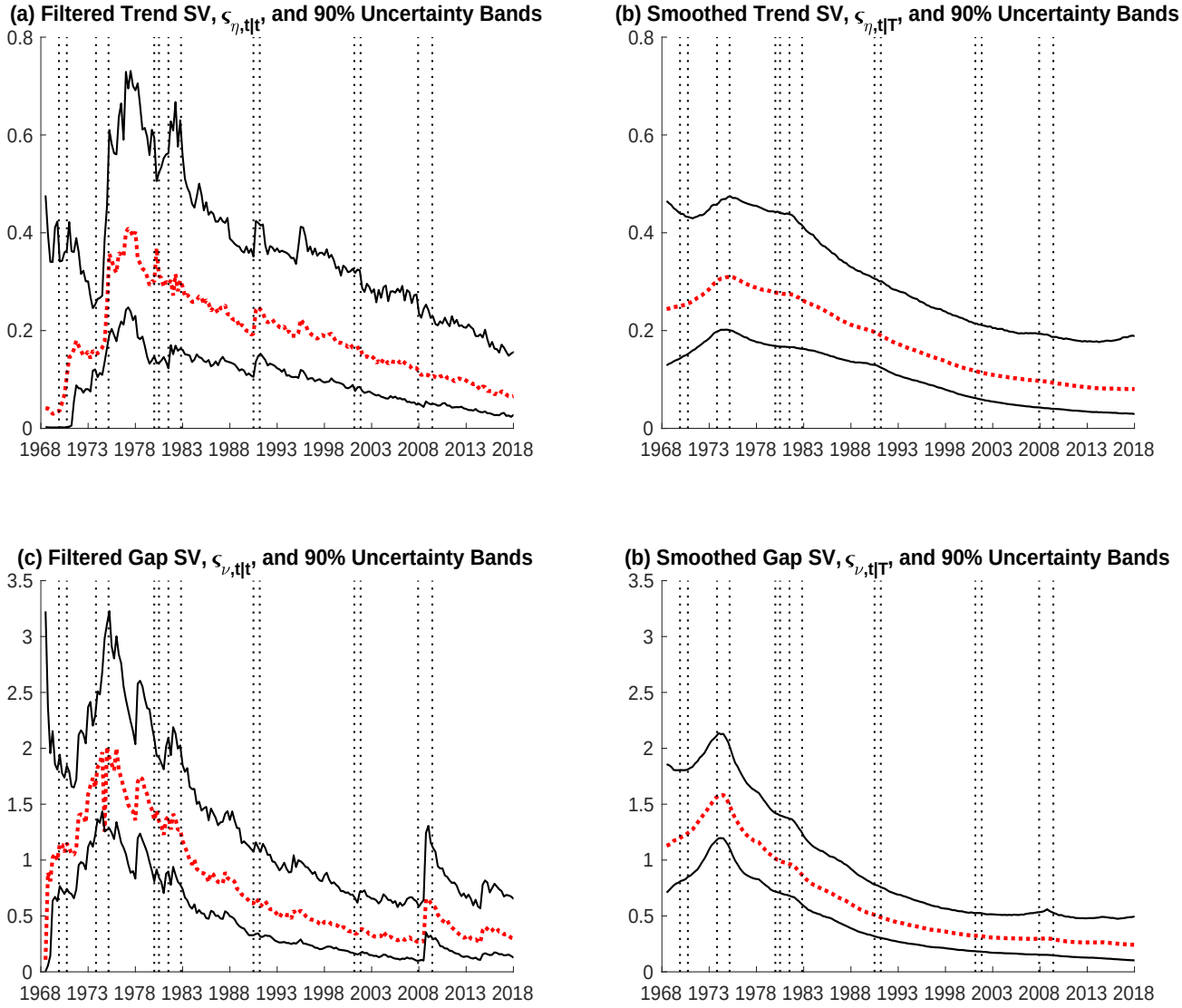
4.6 Trend and Gap Inflation Volatilities

The state space model $\mathcal{M}_2$ yields estimates of filtered and smoothed stochastic volatilities that appear in figure 4. Figures 4(a) and 4(c) has dotted lines that are filtered trend and gap inflation stochastic volatilities, $\varsigma_{\eta,t|t}$ (purple) and $\varsigma_{\upsilon,t|t}$ (teal). Smoothed stochastic volatilities, $\varsigma_{\eta,t|T}$ and $\varsigma_{\upsilon,t|T}$, are the dot-dashed lines (purple and teal) in figures 4(b) and 4(d). The thinner solid (black) lines in figure 4 are 90% uncertainty bands.

Figure 4 shows $\varsigma_{\eta,t|t}$ and $\varsigma_{\eta,t|T}$ are often smaller than $\varsigma_{\upsilon,t|t}$ and $\varsigma_{\upsilon,t|T}$ during the sample. The largest peaks in $\varsigma_{\eta,t|t}$ occur in 1978 and 1980 as displayed in figure 4(a) while $\varsigma_{\upsilon,t|t}$ is dominated by spikes in 1975–1976 and 1978 in figure 4(c). Another revealing feature of figures 4(a) and 4(c) is that $\varsigma_{\eta,t|t}$ and $\varsigma_{\upsilon,t|t}$ often rise during or after a NBER recessions. However, figures 4(b) and 4(d) display a single peak in $\varsigma_{\eta,t|T}$ and $\varsigma_{\upsilon,t|T}$ that occurs in the 1973–1975 recession. Subsequently, $\varsigma_{\eta,t|T}$ and $\varsigma_{\upsilon,t|T}$ decline steadily to the end of the sample.

Our estimates of trend and gap inflation stochastic volatilities differ from Grassi and Proietti (2010), Stock and Watson (2010), and Cogley, Primiceri and Sargent (2010), among others. These authors report trend stochastic volatility dominates inflation gap stochastic volatility from the 1970s well into the late 1990s. In contrast, figure 4 indicates that the jumps in inflation volatility during the 1970s, 1980s, and 1990 was transitory rather than reflecting volatility in trend inflation. This suggests that our approach to estimating the inflation gap persistence parameter has a substantial impact on estimates of the trend volatilities affecting the innovations to trend and gap inflation.

Figure 4: Estimates of the Stochastic Volatility of Trend and Gap Inflation, $1968Q4$ to $2018Q3$



Note: The solid thin (black) lines around estimates of filtered and smoothed trend and gap inflation stochastic volatility are lower and upper bounds on 90% uncertainty bands. The four plots contain vertical dotted bands that denote NBER dated recessions.

4.7 Drifting Inflation Gap Persistence and Sticky Information Updating

This section presents evidence to illustrate why the data prefers $\mathcal{M}_2$. By comparing estimates of θ_t and λ_t to estimates of θ and λ gleaned from $\mathcal{M}_1$ and $\mathcal{M}_0$, we argue procyclical movements of the drifting inflation gap persistence parameter pre-Volcker disinflation and upward drift in the sticky information weight post-1990 explain the data favoring $\mathcal{M}_2$ over the other state space models.

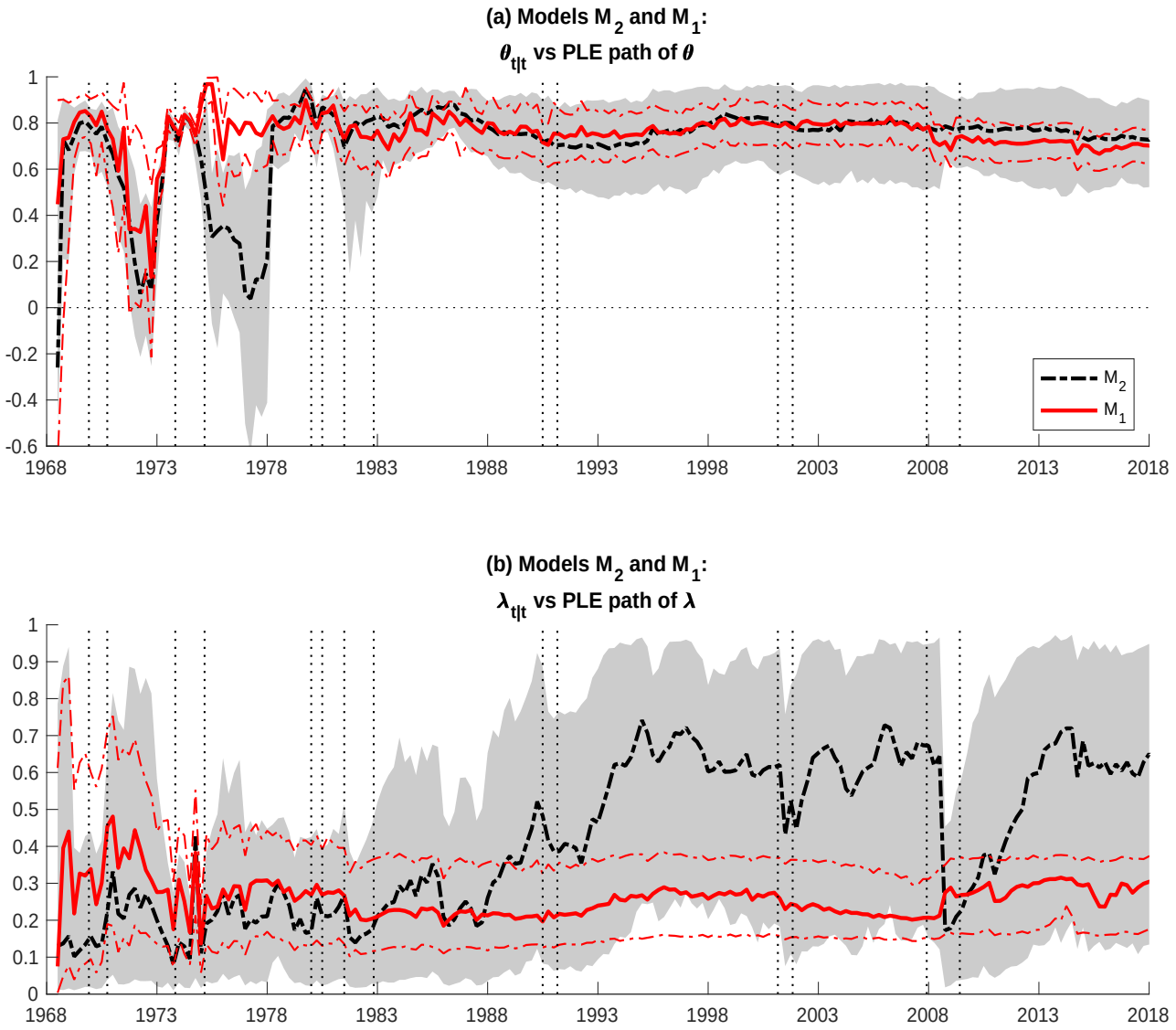
Figures 5(a) and 5(b) begins the discussion with two different kinds of estimates of the inflation gap persistence parameter and sticky information weight. These figures contain dot-dash (black) lines that are filtered estimates of the drifting inflation gap persistence parameter and sticky information weight, $\theta_{t|t}$ and $\lambda_{t|t}$, of $\mathcal{M}_2$. Applying our particle learning filter to $\mathcal{M}_1$ produces paths of θ and λ that are the solid (red) lines in figures 5(a) and 5(b). The paths of θ and λ are surrounded by 90% uncertainty bands denoted by thin dot-dash (red) lines while the light gray shading represent 90% uncertainty bands of $\theta_{t|t}$ and $\lambda_{t|t}$.

The estimates of $\mathcal{M}_1$ indicate the path of θ shifts from procyclical to acyclical comovement with the NBER dated recessions by the early 1980s. Figure 5(a) depicts the path of θ peaking at more than 0.8 during the 1969–1970 and 1973–1975 recessions while dropping almost to zero in 1972 and early 1973. From 1976 to the end of the sample, this path often wobbles, but is never less than 0.6 or greater than 0.8. The 90% uncertainty bands for the path of θ are narrow for most of the sample.

Figure 5(a) also shows that by 1980 $\theta_{t|t}$ experiences a similar change in its comovement with the business cycle. During the 1969–1970 and 1973–1975 recessions, $\theta_{t|t}$ peaks. In the expansions that follow these recessions, there are troughs in $\theta_{t|t}$ that are near zero in 1972 and 1977. There is a third peak in $\theta_{t|t}$ that almost reaches one (its upper bound) in 1979. Subsequently, $\theta_{t|t}$ falls to near 0.7 at the start of the 1981–1982 recession and fluctuates between 0.65 and 0.85 for the rest of the sample. Hence, $\theta_{t|t}$ is procyclical pre-1980 and acyclical post-1980. The sampling uncertainty surrounding $\theta_{t|t}$ is widest after the 1973–1975 recession and during the 1981–1982 recession.

The shift in the behavior of inflation gap persistence over the business cycle is consistent with Meltzer (2014, p. 1006 and p. 1207). He contends that Fed monetary policy produced procyclical inflation during the 1970s, but this changed with the Volcker disinflation. Our estimates of the inflation gap persistence parameter in $\mathcal{M}_1$ and $\mathcal{M}_1$ indicates the average member of the SPF agrees with Meltzer.

Figure 5: Estimates of Drifting and Static Inflation Gap Persistence Parameter and Sticky Information Weight, 1968$Q4$ to 2018$Q3$



Note: The dot-dash (black) lines are filtered estimates of the inflation gap persistence parameter and sticky information weight, $\theta_{t|t}$ and $\lambda_{t|t}$, of $\mathcal{M}_2$ in panels (a) and (b). The light gray shaded areas are 90% uncertainty bands of the non-linear states. Panels (a) and (b) include solid (red) lines that are particle learning filter estimates of θ and λ produced by $\mathcal{M}_1$. These estimates are surrounded by thin dot-dash (red) lines that cover 90% uncertainty bands. The vertical dotted bands denote NBER dated recessions.

The sticky information weights produced by estimating $\mathcal{M}_2$ and $\mathcal{M}_1$ are found in Figure 5(b). These estimates show the paths of $\lambda_{t|t}$ and λ deviate after 1988. During the first half of the sample, λ and $\lambda_{t|t}$ are close and always less than a half. The path of λ continues to show that sticky information forecast were updated quite frequently until the end of the sample. Strikingly, $\lambda_{t|t}$ begins to increase after 1988. By 1995, the frequency of sticky information inflation forecast updating is about three quarters on average. This holds for the rest of the sample with one exception. Between 2009 and 2014, $\lambda_{t|t}$ follows a V-shaped path. The frequency of sticky information inflation forecast updating approximates rational expectations in 2009, but bounces back to the pre-recession frequency of sticky information inflation forecast updating by 2014.

Filtered estimates of the drifting inflation gap persistence parameter and sticky information weight provide evidence for the preference of the data for $\mathcal{M}_2$. Figure 5 presents two kinds of evidence. One piece of evidence is the procyclical behavior of $\theta_{t|t}$ pre-Volcker disinflation and not after in figure 5(a). Next, the drop in the frequency of sticky information inflation forecast updating occurs about the same time estimates of rational expectations and sticky information trend inflation shown in figure 3(c) and of filtered and smoothed stochastic volatilities plotted in figure 4 also are falling.

Figure 5(b) displays 90% uncertainty bands of $\lambda_{t|t}$ that are almost always wider than the 90% uncertainty bands of the path of λ. The sampling uncertainty around $\lambda_{t|t}$ maps into an average frequency of sticky information inflation forecast updating that runs from one to 10 quarters on average after 1990. The exception is the 2007–2009 recession when the 90% uncertainty bands of $\lambda_{t|t}$ narrow to between near zero and 0.4.

We include figure 6 to understand better estimates of the sticky information weight extracted from $\mathcal{M}_2$. Figure 6(a) repeats figure 6(b), but replaces the particle learning filter estimates of λ with smoothed estimates of the drifting sticky information weight, $\lambda_{t|T}$ from $\mathcal{M}_2$. The dot-dashed (red) line represents $\lambda_{t|T}$ and the thin solid (red) lines are its 90% uncertainty bands. Figures 6(b) plots accumulated changes of the smoothed drifting sticky information weight, $\lambda_{t|T} - \lambda_{1|T}$ with dark and light gray areas representing 68% and 90% uncertainty bands. Figures 6(c) and 6(d) report similar filtered and smoothed estimates of the drifting sticky information weight, but these are produced using our baseline state space model, $\mathcal{M}_0$, in which the inflation gap persistence parameter is static, $\theta_t = \theta$. Figures 6(a) and 6(c) show estimates of $\lambda_{t|t}$ and $\lambda_{t|T}$ taken from $\mathcal{M}_0$ and $\mathcal{M}_2$ differ before the Volcker disinflation. Movements in $\lambda_{t|t}$ and $\lambda_{t|T}$ are procyclical in the 1970s as shown in figure 6(a) while this is less evident in figure 6(b).

We see this as evidence the sticky information weight becomes a sink for business cycle fluctuations conditional on $\theta_t = \theta$ in $\mathcal{M}_0$. After the 1981–1982 recession, estimates of the sticky information weight are qualitatively and quantitatively similar across figures 6(a) and 6(c). However, $\mathcal{M}_1$ produces different estimates of the path of the static sticky information weight as seen in figure 5(b).

Figures 6(a) and 6(c) display 90% uncertainty bands of $\lambda_{t|t}$ and $\lambda_{t|T}$ that are wide except for several NBER dated recessions. The width of the uncertainty bands is especially striking because during the latter half of the 1990s the Fed engaged in a policy of "opportunistic disinflation", according to Meyer (1996) and Orphanides and Wilcox (2002). Moreover, the early 2000s witnessed the "considerable" and "extended" policy episodes of the Greenspan and Bernanke Feds, and post-2007 the Bernanke and Yellen Feds employed a host of unconventional policies.[32] Rather than having greater certainty, our estimates indicate there was substantial sampling uncertainty around the frequency of sticky information inflation forecast updating during the last 30 years.

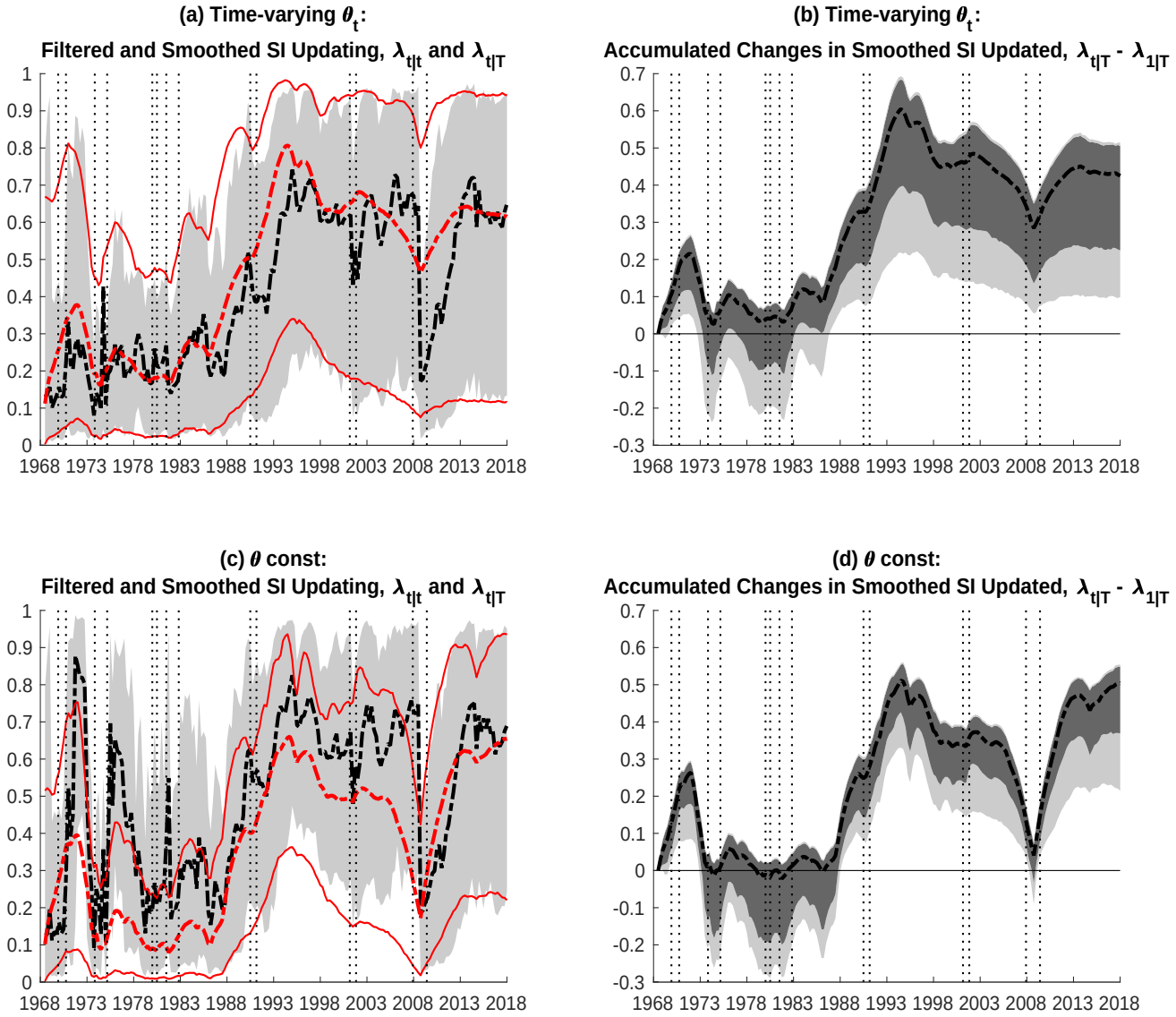
The same cannot be said about $\lambda_{t|T} - \lambda_{1|T}$ that are plotted in figures 6(b) and 6(d). These figures show 68% and 90% uncertainty bands of $\lambda_{t|T} - \lambda_{t|T}$ that are tight compared with the uncertainty bands of $\lambda_{t|T}$ in figures 6(a) and 6(b). For example, only from the 1973-1975 recession to about 1998 do these uncertainty bands cover zero except during the 2007–2009 recession in figures 6(b).

We contend the uncertainty bands of $\lambda_{t|T} - \lambda_{t|T}$ in figures 6(b) and 6(d) are additional evidence our estimates of λ_t represent the beliefs the average SPF respondent has about the underlying dynamics of inflation. Figures 6(b) and 6(d) tell a familiar story about movements in the frequency of sticky inflation forecast updating. The smoothed estimates of the accumulated changes in the frequency of sticky information inflation forecast updating are near zero from 1973 to 1988, begin to increase in the early 1990s reaching a plateau by 1995. This plateau is roughly maintained until $\lambda_{t|T} - \lambda_{1|T}$ takes a V-shaped plunge during the 2007–2009 recession that is followed by a snap back to pre-recession levels by 2014. Since this narrative is grounded in more precise estimates, $\lambda_{t|T} - \lambda_{t|T}$ is additional evidence the inflation predictions of the average member of the SPF became stickier at about the same time persistent shocks to inflation become noticeably smaller after the Volcker disinflation. We also argue this comovement of the drifting sticky information weight with variation in persistent shocks to inflation contributes to

[32] This line of research is beyond the scope of the paper. Interest rates are needed to evaluate the impact of monetary policy on inflation dynamics as studied, for example, by Leeper and Zha (2003).

the data favoring $\mathcal{M}_2$ over the three other state space models.

Figure 6: Filtered and Smoothed Estimates of the Drifting Sticky Information Weight, 1968$Q4$ to 2018$Q3$



Note: The dark (light) gray areas are 68(90)% uncertainty bands of $\lambda_{t|t}$ as dash (black) lines. The left column of charts also displays solid thin (red) lines around $\lambda_{t|T}$ as dot-dashed (red) lines that are lower and upper bounds on 90% uncertainty bands. The four plots contain vertical dotted bands that denote NBER dated recessions.

5 Trend Inflation Uncertainty and Forecast Stickiness

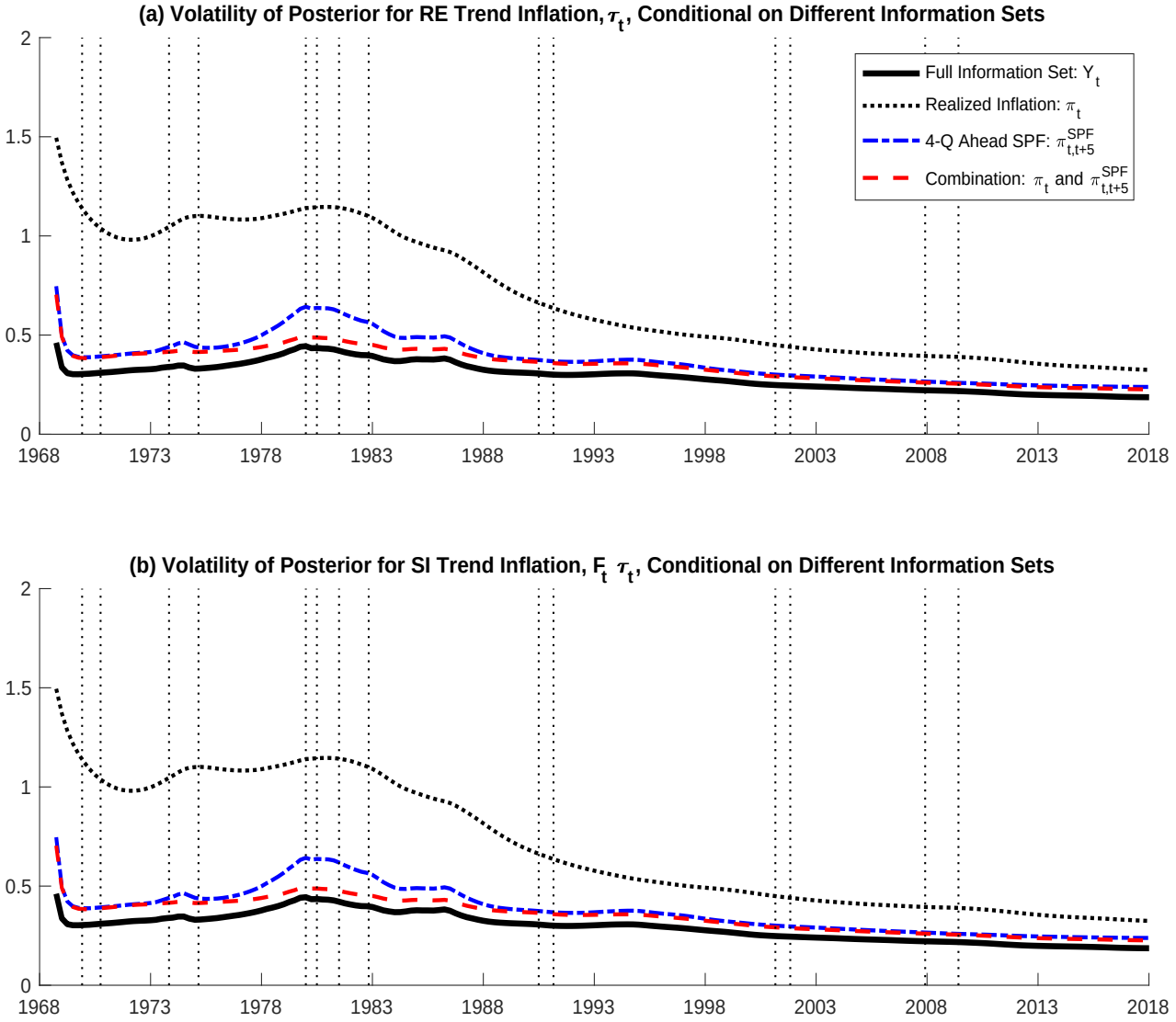
This section assesses the impact of uncertainty and sticky information on estimates of our state space models. The role of uncertainty is judged on the efficiency of trend inflation estimates conditional on various subsets of the data. We evaluate sticky information by measuring the response of realized inflation to the state of the economy. A small forecasting exercise also appears that compares average SPF inflation predictions with forecasts produced from our state space model.

5.1 SPF Inflation Predictions and Trend Inflation Uncertainty

Figure 7 displays conditional volatilities of rational expectations trend inflation, τ_t, and sticky information trend inflation, $F_t\tau_t$. The plots quantify uncertainty over time in τ_t and $F_t\tau_t$ conditional on the history of y_t, or histories of subsets of its elements, smoothed estimates of the non-linear states, and full sample estimates of the static parameters, $\hat{\Psi}$. We measure the volatility of τ_t with $\text{Var}(\tau_t \,|\, y^t, \tilde{\mathcal{V}}_{t|T}, \hat{\Psi})$, where the entire information set is from the first observation to quarter t and the smoothed non-linear states begin at quarter t and end at $T = 2018Q3$. Similar computations are used to produce the conditional volatility of $F_t\tau_t$. Thus, the paths of the non-linear states and parameter estimates are held fixed across subsets of the sample data that are fed into the Kalman filter to produce estimates of the conditional volatilities of τ_t and $F_t\tau_t$.

Figure 7(a) plots the conditional volatilities of τ_t. The conditional volatilities of $F_t\tau_t$ are found in figure 7(b). In these figures, the solid (black), dashed (gray), dotted (blue), and dot-dashed (red) lines are the volatility of trend inflation conditional on the full information set, $\text{Var}(x \,|\, y^t, \tilde{\mathcal{V}}_{t|T}, \hat{\Psi})$, only on realized inflation, $\text{Var}(x \,|\, \pi^t, \tilde{\mathcal{V}}_{t|T}, \hat{\Psi})$, only on the 4-quarter ahead average SPF inflation prediction, $\text{Var}(x \,|\, \pi^{SPF,t}, \tilde{\mathcal{V}}_{t|T}, \hat{\Psi})$, and on realized inflation and the 4-quarter ahead average SPF inflation prediction, $\text{Var}(x \,|\, \pi^t, \pi^{SPF,t}, \tilde{\mathcal{V}}_{t|T}, \hat{\Psi})$, respectively, where $x = \tau_t$ or $F_t\tau_t$.

Figure 7: Uncertainty Measure of Trend Inflation Conditional on Different Information Sets, 1968$Q4$ to 2018$Q3$



Note: The two plots contain vertical dotted bands that denote NBER dated recessions.

Figures 7(a) and 7(b) show $\pi^{SPF}_{t,t+5}$ contributes the bulk of the information pertinent for estimating τ_t and $F_t\tau_t$ efficiently. The dot-dashed (blue) lines, $\text{Var}(x \,|\, \pi^{SPF,t}_{t+5}, \widetilde{\mathcal{V}}_{t|T}, \widehat{\Psi})$, and the dashed (red) lines, $\text{Var}(x \,|\, \pi^t, \pi^{SPF,t}_{t+5}, \widetilde{\mathcal{V}}_{t|T}, \widehat{\Psi})$, are close to the solid black lines that are estimates conditioned on the entire information set, $\text{Var}(x \,|\, y^t, \widetilde{\mathcal{V}}_{t|T}, \widehat{\Psi})$. Conditioning only on $\pi^{SPF}_{t,t+5}$ produces more volatility in τ_t and $F_t\tau_t$ during the double-dip recessions of the early 1980s that is manifested as a hump in the dot-dashed (blue) lines of figures 7(a) and 7(b). In contrast, these figures show a large gap between the dotted (gray) lines, $\text{Var}(x \,|\, \pi^t, \widetilde{\mathcal{V}}_{t|T}, \widehat{\Psi})$, and the solid (black) lines, $\text{Var}(x \,|\, y^t, \widetilde{\mathcal{V}}_{t|T}, \widehat{\Psi})$. The distance indicates that prior to the Volcker disinflation there is insufficient information in π_t alone to estimate τ_t and $F_t\pi_t$ efficiently. We conclude $\pi^{SPF}_{t,t+5}$ has useful information for lowering uncertainty surrounding estimates of τ_t and $F_t\tau_t$.

5.2 Forecasting Inflation with Our State Space Models

This section reports on a small forecasting exercise. We compare the h-quarter ahead inflation forecasts produced by $\mathcal{M}_0$, $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ to the average SPF inflation predictions of the same horizon on the $1968Q4$–$2018Q3$ sample. The forecast comparisons appear in table 6. The table includes ratios of root mean square errors (RMSEs) of the h-quarter ahead inflation forecasts produced by our state space models to RMSEs of $\pi^{SPF}_{t,t+h}$, for $h = 1, \ldots, 5$. Differences in the squared losses of the forecasts appear in brackets. The parentheses contain Newey-West standard errors for the Diebold and Mariano (1995) test computed with $h + 1$ lags.

The forecasting exercise leads us to two conclusions. At low-order forecast horizons, table 6 shows the average SPF inflation predictions dominate the relative RMSEs and produce smaller squared losses. At $h \geq 3$, this changes. Our state space models yield better forecasts in every comparison save one and always have lower squared losses, according to table 6. We interpret these results as indicating the SPF is useful for predicting a nowcast of inflation and inflation one-quarter ahead while our state space models yield competitive inflation forecasts, especially at longer forecast horizons.

Table 6: Inflation Forecast Comparisons of the SPF and State Space Models, $1968Q4$ to $2018Q3$

	Models			
horizon	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$
1	**1.13**	**1.03**	**1.09**	**1.10**
	[0.319]	[0.080]	[0.211]	[0.255]
	(0.102)	(0.052)	(0.066)	(0.071)
2	**1.01**	**1.00**	**1.00**	**1.02**
	[0.035]	[0.002]	[0.011]	[0.065]
	(0.056)	(0.050)	(0.041)	(0.043)
3	**0.98**	**0.99**	**0.98**	**1.00**
	[-0.107]	[-0.065]	[-0.090]	[-0.025]
	(0.068)	(0.070)	(0.054)	(0.057)
4	**0.96**	**0.99**	**0.98**	**0.99**
	[-0.226]	[-0.076]	[-0.130]	[-0.051]
	(0.084)	(0.067)	(0.064)	(0.061)
5	**0.98**	**0.99**	**0.98**	**0.99**
	[-0.214]	[-0.129]	[-0.161]	[-0.151]
	(0.138)	(0.114)	(0.110)	(0.112)

Note: Relative root mean square error (RMSE) predictions generated by our state space models evaluated against the average SPF inflation predictions, $\pi_{t,t+h}^{SPF}$, for $h = 1, \ldots, 5$. Values less than 1.00 indicate a lower RMSE for the state space model forecasts. Below the relative RMSEs are square brackets that contain the difference in squared losses of both forecasts, with the associated standard errors in parentheses. The standard errors are computed using a Diebold-Mariano test with Newey-West standard errors using $h+1$ lags.

5.3 Forecast Stickiness and Persistent Shocks to Inflation

Up to this point, we have documented significant time-variation in forecast stickiness, as measured by the frequency of sticky information forecast updating, λ_t. However, we find a noticeable drop in forecast stickiness around the 2007–2009 recession. Coibion and Gorodnichenko (2015) report broadly similar evidence of time variation in λ_t and point to recessions and the decline in macroeconomic volatility since the mid 1980s, also known as the Great Moderation, as proximate causes of forecast stickiness.[33] This section complements their analysis by juxtaposing our estimates of λ_t with characteristics of the inflation process implied by our state space models.

We focus on two metrics characterizing the inflation process that should matter for forecasters. One metric is the mean squared error of the forecast (MSE). The second is the share of forecast error variance attributable to non-persistent shocks. We argue this forecast error variance is a suitable proxy for inflation persistence. Estimates of stickiness, the MSE, and the variance share of non-persistent shocks are jointly drawn from the posterior distribution of $\mathcal{M}_2$. However, the priors of $\mathcal{M}_2$ view stickiness as unrelated to the underlying dynamics of the inflation. Recall that for parsimony and not to impose a specific form of state dependence on inflation forecast stickiness, we assume λ_t is driven by an exogenous shock. Nevertheless, there is a striking resemblance between time-variation in estimates of inflation stickiness and the importance of persistent shocks in accounting for the MSE of inflation.

Coibion and Gorodnichenko (2015) use the MSE of the sticky information forecaster's h-step ahead prediction at a given date t as the dependent variable in their regression analysis to evaluate forecast stickiness.[34] As is standard, this MSE can be decomposed into the sum of squared bias and forecast error variance

$$\text{MSE}_{t,h} \equiv \mathbf{E}_t\left\{(\pi_{t+h} - F_t\pi_{t+h})^2\right\} = \left(\mathbf{E}_t\pi_{t+h} - F_t\pi_{t+h}\right)^2 + \text{Var}_t(\pi_{t+h}), \qquad (10)$$

where $\left(\mathbf{E}_t\pi_{t+h} - F_t\pi_{t+h}\right)^2$ and $\text{Var}_t(\pi_{t+h})$ are squared bias and forecast error variance.

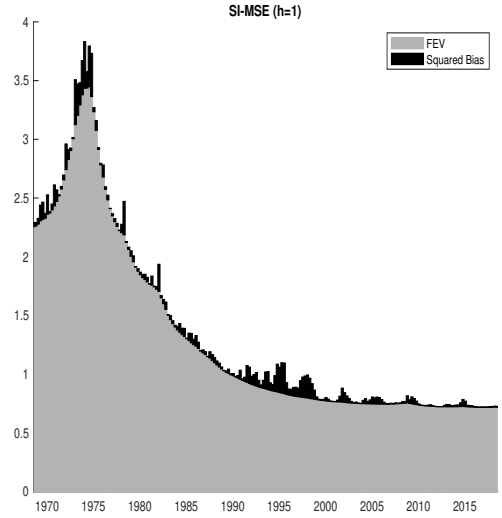
The decomposition of equation (10) is the source of the plots in figures 8(a) and

[33] Coibion and Gorodnichenko (2015) gauge time variation in the forecast stickiness with rolling window regressions projecting $\pi_{t+h} - F_t\pi_{t+h}$ onto $F_t\pi_{t+h} - F_{t-1}\pi_{t+h}$ with slope coefficient β. The sticky information law of motion 3.2 implies the regression slope coefficient $\beta = \lambda/(1-\lambda)$.

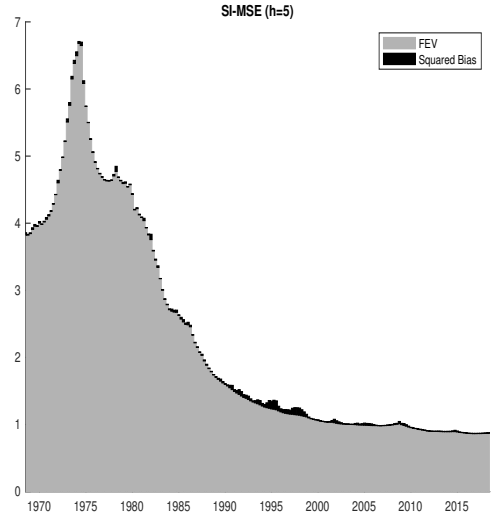
[34] Absent a specific model of the sticky information forecaster's loss function, squared errors are only a proxy for actual losses. Since the rational expectations benchmark minimizes squared errors, this makes measuring the consequences of deviations from rational expectations a natural starting point for assessing the forecast errors of a sticky information forecaster.

FIGURE 8: Forecast Stickiness and the Contribution of Persistent Shocks to Inflation

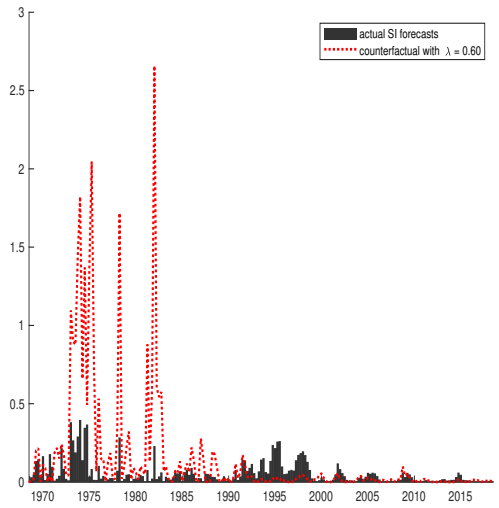
(A) Sticky Information-MSE for $h = 1$



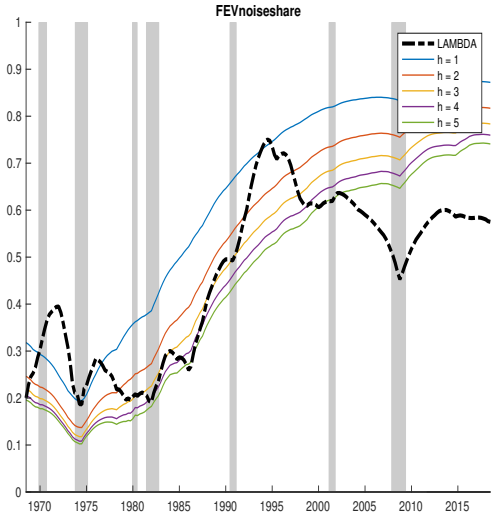
(B) SI-MSE for $h = 5$



(C) Actual and Counterfactual Square SI Bias



(D) FEV Share of Noise Shocks



Note: Figures 8(a) and 8(b) depict the sticky information-MSE and its two components defined in equation (10) for $h = 1$ and $h = 5$, respectively. Figure 8(c) displays squared bias for the actual sticky information forecasts with $h = 1$ as well as a sequence of counterfactual sticky information forecasts constructed using $\lambda = 0.8$. Figure 8(d) compares the shares of unpredictable measurement error disturbances, $\zeta_{\pi,t}$, to realized inflation in forecast error variances at different horizons against smoothed estimate of the sticky information weight, $\lambda_{t|T}$, of $\mathcal{M}_2$. The vertical dotted bands in figure 8(d) denote NBER dated recessions.

8(b). These figures depict variation in the MSE defined in equation (10) for $h = 1$ and $h = 5$. The contributions to the MSEs from squared bias and forecast error variance are also portrayed. We produce these figures using smoothed estimates of $\mathcal{M}_2$ conditioned on full sample information to compute rational expectations and sticky information forecasts as well the MSEs. This conditioning implies our estimates of MSE are $\mathbf{E}_t\left\{\left(\pi_{t+h} - F_t\pi_{t+h}\right)^2 \mid \mathcal{Y}^T\right\}$ rather than $\mathbf{E}\left\{\left(\pi_{t+h} - F_t\pi_{t+h}\right)^2 \mid \mathcal{Y}^T\right\}$. Hence, the conditional moments $\mathbf{E}_t\{\cdot\}$ and $\mathrm{Var}_t(\cdot)$ are evaluated from the perspective of a forecaster who knows the current level of inflation, its decomposition into trend, gap and noise, the stochastic volatilities, and static parameters of $\mathcal{M}_2$. This approach is intended to capture the potentially richer information set of the average SPF respondent, relative to the econometrician, that drives forecast updating.[35]

Figures 8(a) and 8(b) display sizable variation in the MSEs over the entire sample. In line with Stock and Watson (2007) and Clark, McCracken, and Mertens (2019), uncertainty around inflation forecasts is greatest during the 1970s. For $h = 1$ and $h = 5$, the MSE peaks around the 1973–1975 recession and remain elevated for the remainder of the 1970s. With the onset of the Volcker disinflation, forecast uncertainty, as measured by the MSE, declines steadily over the remainder of the sample with one exception. The steady decline in the MSE is interrupted by brief spikes tied to a short-lived increase in bias and a minimal rise around the onset of the 2007–2009 recession for $h = 1$.

Considering the relative contributions of bias and forecast error variance, the MSE is dominated by forecast error variance with squared bias playing only a minor role.[36] Strikingly, extended periods of large biases (in absolute value) are rare throughout the sample. Moreover, squared bias becomes negligible since the late 1990s, which is when our filtered and smoothed estimates of λ_t increase to 0.6 or more. This evidence is consistent with the notion that the frequency of sticky information inflation forecast updating may not have risen by accident, but rather reflects an evaluation of costs and benefits of forecast updating, as suggested, for example, by the theories and models of Woodford (2003), Sims (2003), and Mackowiak and Wiederholt (2009).

We confirm the hypothesis that variation in sticky information inflation forecasts may be related to changes in the inflation process by considering the counterfactual ex-

[35] Further details about the computations are provided in the online appendix.

[36] For $h > 1$, the estimated bias converges to the difference between sticky information and rational expectations *trends*, which is relatively small, while the forecast error variance is strictly increasing. Figure 8(b) shows the MSE at $h = 5$ is dominated by the forecast error variance to a greater extent than is shown in figure 8(a) for $h = 1$.

periment that takes as given our estimates of trend and gap inflation. The hypothetical sticky information inflation forecasts are based on the sticky information law of motion (3.2), but fixes λ to a moderately high value of 0.6. This sets the frequency of sticky information inflation forecast updating to 2.5 quarters on average. Given estimates of trend and gap inflation and $\lambda = 0.6$, the MSE of the h-step ahead rational expectations and sticky information inflation forecasts are generated to measure bias between these forecasts. Figure 8(c) compares the actual squared sticky information bias for $h = 1$ against the counterfactual measure that would have resulted if λ had been equal to 0.6.[37] The upshot is that if inflation forecast stickiness had been that high during the latter 1970s, forecast bias would have been substantially higher than it actually was. This bias would have been as high as the forecast-error variance associated with the optimal rational expectations inflation forecast displayed in figure 8(a). This stands in contrast to the actual and counterfactual bias, which are fairly small, during the second half of the sample, even though both reflect elevated levels of inflation forecast stickiness.[38] All told, the estimated increase in forecast stickiness occurs at the same time the constant frequency of sticky information inflation forecast updating produces little difference in rational expectations and sticky information inflation forecasts.

A key concern is the importance of persistent shocks in the inflation process. In our state space models, inflation is the sum of three components: trend, gap and serially uncorrelated measurement error. Hence, only shocks to trend and gap cause variations in rational expectations forecasts. Remember from equation (3.2) that $\pi_t = \tau_t + \varepsilon_t + \sigma_{\zeta,\pi}\zeta_{\pi,t}$, where $\zeta_{\pi,t} \sim \mathcal{N}(0, 1)$. Since all three inflation components are assumed to be uncorrelated, it is straightforward to decompose the forecast error variance into the contributions from each component,

$$\mathrm{Var}_t(\pi_{t+h}) \;=\; \mathrm{Var}_t(\tau_{t+h}) \;+\; \mathrm{Var}_t(\varepsilon_{t+h}) \;+\; \sigma^2_{\zeta,\pi}\,,$$

where stochastic volatility causes the share of forecast error variances explained by each shock to vary. As described in section 2, optimal rational expectations inflation forecasts are linear combinations of τ_t and ε_t. Hence, greater volatility in the shocks to

[37] The actual squared sticky information bias shown in figure 8(c) corresponds to the squared bias shown in the MSE decomposition of figure 8(a).

[38] The online appendix considers an alternative counterfactual generated by setting $\lambda = 0.2$. The resulting counterfactual sticky information bias is quantitatively similar to the actual bias generated by estimates of the sticky information weight that are 0.6 or greater by the second half of the sample.

these states increases the variability of optimal rational expectations inflation forecasts.

We contend that when shocks to τ_t and ε_t account for a larger share of variations in π_t a sticky information forecaster has a greater incentive to update more frequently. As a result, forecast stickiness should be positively related to the share of forecast error variance due to measurement error in realized inflation. This is measured by

$$\text{FEV-share-noise}_{t,h} = \frac{\sigma^2_{\zeta,\pi}}{\text{Var}_t(\pi_{t+h})}.$$

Indeed, as figure 8(d) shows, the share of the one-step ahead forecast error variance due to measurement error mirrors variations in our smoothed estimates of the sticky information weight, $\lambda_{t|T}$ in figure 6(a).[39] Although $\lambda_{t|T}$ suggests low levels of sticky information during the latter half of the 1970s and the Volcker disinflation, the stickiness of inflation forecast updating has risen since 1990 and remained high except for a brief dip during the 2007–2009 recession. Figure 8(d) also shows that the share of the sticky information forecast error variance due to unpredictable measurement error in inflation was relatively low during the first half of the sample. This share increases at the same time as the Volcker disinflation takes hold and remains high for the rest of the sample with only a brief drop during the 2007–2009 recession proving the exception. In contrast, figures 8(a) and 8(b) depict the MSE of total forecast uncertainty as declining steadily since the 1970s and has only a minimal increase during the 2007–2009 recession. In summary, we obtain substantial evidence of an increase in the stickiness of inflation that lines up with the Volcker disinflation.

6 Conclusions

This paper studies the joint dynamics of realized inflation and a term structure of average inflation predictions found in the Survey of Professional Forecasters (SPF). We build non-linear state space models using the Stock and Watson unobserved components model and the Coibion and Gorodnichenko (2015) version of the Mankiw and Reis (2002) sticky information model to estimate trend and gap inflation, the stochas-

[39] The online appendix documents similar patterns for the share of noise shocks in the forecast error variances when $h > 1$. We also report results based on an alternative decomposition of the forecast error variances in the online appendix that is akin to one used by Cogley, Primiceri, and Sargent (2010). It measures persistence with respect to changes in one-step ahead expectations. Our results are unchanged using this measure of persistence.

tic volatility affecting these states, inflation gap persistence parameter, and the sticky information weight. The state space models are estimated on a sample of real-time realized inflation and averages of SPF inflation predictions from $1968Q4$ to $2018Q3$ using sequential Monte Carlo methods. The sequential Monte Carlo methods consist of a particle learning filter proposed by Carvalho, Johannes, Lopes, and Polson (2010) and a Rao-Blackwellized particle smoother developed by Lindsten, Bunch, Särkkä, Schön, and Godsill (2016).

We draw five headline results from estimates of the state space models. First, the data prefer the state space model that includes drift in the inflation gap persistence parameter and sticky information weight. Second, the 4-quarter ahead average SPF inflation prediction has information that increases the efficiency of estimates of trend inflation. Third, inflation gap persistence flips from procyclical before the Volcker disinflation to lacking any business cycle comovement during the rest of the sample. There is also substantial persistence in the inflation gap after the Volcker disinflation. The years after the Volcker disinflation is when the stickiness of inflation forecasts increases. This remains the case to the end of the sample except for a transitory decline in the sticky information weight during the 2007–2009 recession. Fifth, shifts in the stickiness of inflation forecasts occur at the same time the importance of persistent shocks for explaining the variation in realized inflation begins to decline. We interpret this evidence as suggesting sticky information inflation forecasts are state dependent, but lack business cycle dependence.

Our results fit into a literature that finds permanent shocks matter more to professional forecasters compared with transitory shocks. An example is that the frequency of sticky inflation forecast updating is estimated to change at the same time persistent shocks become less important for inflation dynamics. In our view, this evidence should point future research toward endogenizing the sticky information weight, perhaps, in the tradition of rational inattention models of Sims (2003), Woodford (2003), and Mackowiak and Wiederholt (2009). In the same way that this research motivated us, we hope that our paper stimulates further work on the ways in which professional forecasters and other economic agents process information to form beliefs and predictions about future economic outcomes and events.

References

Ascari, G., Bonomolo, P., and H. F. Lopes, (2019), "Walk on the wild side: Temporarily unstable paths and multiplicative sunspots", *American Economic Review* 109(5), 1805–'842.

Andrieu, C., A. Doucet, and R. Holenstein (2010), "Particle Markov chain Monte Carlo methods." *Journal of the Royal Statistical Society, Series B*, 72(3), 269–342.

Ang, A., G. Bekaert, and M. Wei (2007), "Do macro variables, asset markets, or surveys forecast inflation better?" *Journal of Monetary Economics*, 54(4), 1163–1212.

Aruoba, S.B. (2019), "Term structures of inflation expectations and real interest rates," *Journal of Business & Economic Statistics*, forthcoming.

Atkeson, A. and L.E. Ohanian (2001), "Are Phillips curves useful for forecasting inflation?" *Quarterly Review*, Federal Reserve Bank of Minneapolis, 25, 2–11.

Bernanke, B.S. (2007), "Inflation Expectations and Inflation Forecasting." Remarks given at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts (July 10). Available at `http://www.federalreserve.gov/newsevents/speech/Bernanke20070710a.htm`.

Beveridge, S. and C.R. Nelson (1981), "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle." *Journal of Monetary Economics*, 7(2), 151–174.

Carvalho, C.M., M.S. Johannes, H.F. Lopes, and N.G. Polson (2010), "Particle learning and smoothing." *Statistical Science*, 25(1), 88–106.

Chen, R. and J.S. Liu (2000), "Mixture Kalman filters." *Journal of the Royal Statistical Society, Series B*, 62(3), 493–508.

Clark, T.E., M.W. McCracken, and E. Mertens (2019), "Modeling time-varying uncertainty of multiple-horizon forecast errors." *The Review of Economics and Statistics*, forthcoming.

Cogley, T., G. Primiceri, and T.J. Sargent (2010), "Inflation-gap persistence in the US." *American Economic Journal: Macroeconomics*, 2(1), 43–69.

Cogley, T. and T.J. Sargent (2015), "Measuring price-level uncertainty and instability in the U.S., 1850–2012." *Review of Economics and Statistics*, 97(4), 827–838.

Cogley, T. and A. Sbordone (2008), "Trend inflation, indexation, and inflation persistence in the new Keynesian Phillips curve inflation-gap persistence in the US." *American Economic Review*, 98(5), 2101–2126.

Coibion, O. and Y. Gorodnichenko (2015), "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review*, 105(8), 2644–2678.

Coibion, O. and Y. Gorodnichenko (2012), "What can survey forecasts tell us about informational rigidities?" *Journal of Political Economy*, 120(1), 116–159.

Creal, D. (2012), "A survey of sequential Monte Carlo methods for economics and finance." *Econometric Reviews*, 31(), 245–296.

Diebold, F. X. and R. S. Mariano, (1995), "Comparing predictive accuracy", *Journal of Business & Economic Statistics* 13(3), 253–63.

Djuric, P. and J. Miguez (2002), "Sequential particle filtering in the presence of additive Gaussian noise with unknown parameters." *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2, 1621–1624.

Fearnhead, P. (2002), "Markov chain Monte Carlo, sufficient statistics, and particle filters." *Journal of Computational and Graphical Statistics*, 11(4), 848–862.

Faust, J. and J.H. Wright (2013), "Forecasting inflation." In *Handbook of Economic Forecasting, vol. 2*, Ch. 1 (Elliot, G. and A. Timmermann, eds.), 2–56, Elsevier Science.

Fuentes-Albero, C. and L. Melosi, (2013), "Methods for computing marginal data densities from the Gibbs output", *Journal of Econometrics* 175(2), 132–141.

Geweke, J. (1989), "Bayesian inference in econometric models using Monte Carlo integration", *Econometrica* 57(6), 1317–1339.

Godsill, S.J., A. Doucet, and M. West (2004), "Monte Carlo smoothing for nonlinear time series." *Journal of the American Statistical Association*, 99, 156–168.

Goodfriend, M. and R.G. King (2005), "The incredible Volcker disinflation." *Journal of Monetary Economics*, 52(5), 981–1015.

Gordon N., D. Salmond, and A.F.M. Smith (1993), "Novel approach to nonlinear/non-Gaussian Bayesian state estimation." *IEE Proceedings F, Radar Signal Processing*, 140(2), 107–113.

Grant, A.P. and L.B. Thomas (1999), "Inflationary expectations and rationality revisited." *Economics Letters*, 62(3), 331–338.

Grassi, S. and T. Proietti (2010), "Has the volatility of U.S. inflation changed and how?" *Journal of Time Series Econometrics*, 2(1), article 6.

Hansen, P.R., A. Lunde, and J.M. Nason, (2011), "The model confidence set." *Econometrica*, 79(2), 453–497.

Harvey, A.C. (1991), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

Herbst, E. and F. Schorfheide (2014), 'Sequential Monte Carlo sampling for DSGE models', *Journal of Applied Econometrics* 29(7), 1073–1098.

Herbst, E. and F. Schorfheide (2016), *Bayesian inference for DSGE models*. Princeton University Press.

Jain, M. (2017), "Perceived inflation persistence." *Journal of Business & Economic Statistics*, forthcoming.

Johansen, A.M. and A. Doucet (2008), "A note on auxiliary particle filters." *Statistics and Probability Letters*, 78(12), 1498–1504.

Kass, R. E. and A. E. Raftery, (1995), 'Bayes factors', *Journal of the American Statistical Association* 90(430), 773–795.

Komunjer, I., and S. Ng (2011), "Dynamic identification of dynamic stochastic general equilibrium models." *Econometrica*, 79(6), 1995–2032.

Kozicki, S. 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), 145–169.

Krane, S.D. (2011), "Professional forecasters' views of permanent and transitory shocks to GDP." *American Economic Journal: Macroeconomics*, 3(1), 184–211.

Leeper, E.M. and T. Zha (2003), "Modest policy interventions." *Journal of Monetary Economics*, 50(8), 1673–1700.

Lindsten, F., P. Bunch, S. Särkkä, T.B. Schön, and S.J. Godsill (2016), "Rao-Blackwellized particle smoothers for conditionally linear Gaussian models." *IEEE Journal of Selected Topics in Signal Processing*, 10(2), 353–365.

Liu, J. and M. West (2001), "Combined parameters and state estimation in simulation based filtering." In *Sequential Monte Carlo Methods in Practice*, Chapter 10 (A. Doucet, N. de Freitas, and N. Gordon, eds.), 197–223, Springer-Verlag.

Lopes, H.F. and R.S. Tsay (2011), "Particle filters and Bayesian inference in financial econometrics." *Journal of Forecasting*, 30(1), 168–209.

Mackowiak, B. and M. Wiederholt (2009), "Optimal sticky prices under rational inattention." *American Economic Review*, 99(3), 769–803.

Mankiw, N.G. and R. Reis (2002), "Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve." *Quarterly Journal of Economics*, 117(4), 1295–1328.

Meltzer, A.H. (2014), *A History of the Federal Reserve: Volume 2, Book II, 1970–1986*. The University of Chicago Press.

Mertens, E. (2016), "Measuring the level and uncertainty of trend inflation." *The Review of Economics and Statistics*, 98(5), 950–967.

Meyer, L. (1996), "Monetary policy objectives and strategy." Remarks given at the National Association of Business Economists 38th Annual Meeting, Boston, MA (September 8). Available at `http://www.federalreserve.gov/boarddocs/speeches/1996/19960908.htm`.

Morley, J.C., C.R. Nelson, and E. Zivot (2003), "Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different? *Review of Economics and Statistics*, 85(2), 235–243.

Muth, J.F. (1960), "Optimal properties of exponentially weighted forecasts." *Journal of the American Statistical Association*, 55(290), 299–306.

Nason, J.M. and G.W. Smith (2019), "Measuring the slowly evolving trend in US inflation with professional forecasts." Manuscript, Department of Economics, Queen's University.

Nelson, C.R. (2008), "The Beveridge-Nelson decomposition in retrospect and prospect." *Journal of Econometrics*, 146(2), 202–206.

Orphanides, A. and D. Wilcox (2002), "The opportunistic approach to disinflation." *International Finance*, 5(1), 47–71.

Pitt, M.K., R. dos Santos Silva, P. Giordani, and R. Kohn (2012), "On some properties of Markov chain Monte Carlo simulation methods based on the particle filter." *Journal of Econometrics*, 171(2), 134–151.

Pitt, M.K. and N. Shephard (2001), "Auxiliary variable based particle filters." In *Sequential Monte Carlo Methods in Practice*, Chapter 13 (Doucet, A., N. de Freitas, and N. Gordon, eds.), 273–294, Springer-Verlag.

Pitt, M. K. and N. Shephard (1999), "Filtering via simulation: auxiliary particle filters." *Journal of the American Statistical Association*, 94(446), 590–599.

Särkkä, S. (2013), *Bayesian Filtering and Smoothing*. Cambridge University Press.

Schorfheide, F., D. Song, and A. Yaron (2018), "Identifying long-run risks: A Bayesian mixed-frequency approach." *Econometrica*, 86(2), 617–654.

Shephard, N. (2013), "Martingale unobserved component models." Economics Series Working Papers 644, Department of Economics, University of Oxford.

Sims, C.A. (2003), "Implications of rational inattention." *Journal of Monetary Economics*, 50(3), 665–690.

Stock, J.H. and M.W. Watson (1999), "Forecasting inflation." *Journal of Monetary Economics*, 44(2), 293–335.

Stock, J.H. and M.W. Watson (2007), "Why has US inflation become harder to forecast?" *Journal of Money, Credit and Banking*, 39(S1), 3–33.

Stock, J.H. and M.W. Watson (2009), "Phillips curve inflation forecasts." In *Understanding Inflation and the Implications for Monetary Policy – A Phillips Curve Retrospective*, Ch. 3 (J. Fuhrer, Y.K. Kodrzycki, J.S. Little, and G.P. Olivei, eds.), 99–184, MIT Press.

Stock, J.H. and M.W. Watson (2010), "Modeling inflation after the crisis." In Macroeconomic Challenges: The Decade Ahead, Chapter 3, (C.S. Hakkio and E.S. Knotek, II, eds.), 173–220, Federal Reserve Bank of Kansas City.

Storvik, G. (2002), "Particle filters for state-space models with the presence of unknown static parameters." *IEEE Transactions on Signal Processing*, 50(2), 281–289.

Watson, M.W. (1986), "Univariate detrending methods with stochastic trends." *Journal of Monetary Economics*, 18(1), 49–75.

Woodford, M. (2003), "Imperfect common knowledge and the effects of monetary policy." In Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps (P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds.), 25–58. Princeton University Press.