TI 2021-099/III
Tinbergen Institute Discussion Paper

An Empirical Assessment of the U.S. Phillips Curve over Time

Marente Vlekke[1]
Siem Jan Koopman[2]
Martin Mellens[3]

[1] CPB Netherlands Bureau for Economic Policy Analysis, Vrije Universiteit Amsterdam
[2] Vrije Universiteit Amsterdam, Tinbergen Institute, CREATES Aarhus University
[3] CPB Netherlands Bureau for Economic Policy Analysis

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at https://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

An empirical assessment of the U.S. Phillips curve over time*

Marente Vlekke† Siem Jan Koopman‡ Martin Mellens§

November 2021

Abstract

We assess the stability of the unemployment gap parameter using linear dynamic Phillips curve models for the United States. In this study, we allow the unemployment gap parameter to be time-varying such that we can monitor the importance of the Phillips curve over time. We consider different specifications and different measures for inflation. Furthermore, we include stochastic volatility for the observation errors. Our estimation results are based on practical Bayesian state space methods which include feasible testing and diagnostic checking procedures. A key finding is that the Phillips curve for U.S. headline inflation has remained empirically relevant over the years.

Keywords: Phillips curve, Inflation, Inflation Expectations, State space methods, Bayesian Gibbs sampling.

JEL classification: C18, C32, C52, E24, E31

*The authors would like to thank Leon Bettendorf, Dennis Bonam, Adam Elbourne, Rob Euwals, Lennart Hoogerheide, Gerhard Ruenstler and Loes Verstegen for helpful comments and suggestions.

†CPB Netherlands Bureau for Economic Policy Analysis; School of Business and Economics, Vrije Universiteit Amsterdam. E-mail: `m.vlekke@cpb.nl`

‡Corresponding author: Department of Econometrics, School of Business and Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands; CREATES Aarhus University, Denmark; Tinbergen Institute Amsterdam, The Nethderlands. E-mail: `s.j.koopman@vu.nl` Telephone: +31205986019.

§CPB Netherlands Bureau for Economic Policy Analysis.

1 Introduction

The relevance and stability of the Phillips curve in explaining and predicting inflation dynamics have been debated for decades. More than once, the statistical significance of this equation, which links inflation dynamics to real activity or 'slack', has been dismissed, only to be revived later on; see, for example, the discussions in Gordon (2013) and Hall (2013)). This debate has been rekindled in the aftermath of the Great Recession, when inflation in the United States and in euro area countries fell by less than what traditional Phillips curves would have predicted, giving rise to the "missing disinflation" puzzle. In more recent years, as the economies of these countries recovered, a twin-puzzle has emerged: inflation has increased by less than expected given the economic conditions. This gave rise to the notion that the relationship between real activity and inflation is unstable, and has faded over time. In the past two decades, the literature on the Phillips curve has been enriched by research that show this, but a more or less equal number of papers have appeared that rebut this claim.

To empirically assess the stability of the Phillips curve, we adopt Bayesian state space methods to estimate the time-varying coefficients in a linear dynamic model with stochastic volatility. Empirical studies employing a similar approach have found that the coefficient has declined over time; see, for example, Ball and Mazumder (2011) and Matheson and Stavrev (2013) whose approach was adopted by Blanchard, Cerutti, and Summers (2015), Blanchard (2016) and the IMF (2013) in their analysis of the Phillips curve in the World Economic Outlook of April 2013. While these studies differ in their choice of data, variables, model specification and estimation method, they all conclude that the slope of the Phillips curve has been declining over the past few decades in the United States. Blanchard, Cerutti, and Summers (2015) find similar evidence for other OECD countries. The literature has put forward a number of explanations for this. One is Bernanke's (2007) "anchored expectations" hypothesis, which states that the slope of the Phillips curve has become less strong because inflation expectations have become less informed by transitory shocks and recent past inflation, and more anchored to a fixed inflation rate or "target" set by a country's central bank. Ball and Mazumder (2011, 2019) and Blanchard (2016) provide some evidence for this.

The measurement of inflation expectations has implications for the Phillips curve. Economic expectations play a central role in macroeconomic theory, as the perception about future economic conditions affects current economic decision-making. The Phillips curve theoretically models the decisions and hence expectations of firms. The measurement of expectations could therefore affect empirical outcomes in macroeconometric analyses of the Phillips curve. For most countries, there is no or very little data available on firm inflation expectations. Therefore, the literature resorts to proxies. While the earlier studies listed above simply use the lags of inflation (so-called 'backward-looking' expectations), surveys such as Survey of Professional Forecasters obtain *long-term professional* expectations. Coibion and Gorodnichenko (2015) show that the surveys measuring *short-term household* expectations may be a better proxy for the actual inflation expectations of firms. Their measure, obtained from the Michigan Survey of Consumers, accounts for the lack of disinflation during the Great Recession: household inflation expectations actually increased substantially in the crisis period, whereas professional expectations hovered around 2%. The authors find that including survey expectations yield a strong and stable negative relationship between inflation and unemployment.

Our aim is to make three contributions. First, we analyse the slope of the Phillips curve in a linear dynamic model that incorporates survey information on both professional and household inflation expectations for different expectation horizons. This systematic comparison of empirical results for these different expectation measures using unobserved components models with time-varying parameters is a novel contribution. Second, we propose a practical alternative to regular Bayesian model diagnostics, which is practical and easy to compute. In this respect, we address a less appealing feature of the Bayesian approach. It is not necessarily straightforward to obtain regular model diagnostics such as the R^2 goodness-of-fit statistic. The Bayesian equivalent would be a Bayes factor, computed by obtaining the marginal likelihoods of different specifications, but our sampling method makes it tedious and time-consuming to obtain the appropriate Bayes factor. Instead, we propose model diagnostics which rely on the Kalman filter and smoother and are *conditional* on the different moments in the posterior distributions of the parameters. Third, we follow a well-developed literature on the advantages of stochastic volatility by allowing the variance of the innovations to vary over time. While this has become an increasingly common addition to unobserved component models for inflation dynamics in general (see e.g. Stock and Watson (2007), Chan, Koop, and Potter (2013, 2016), Chan, Clark, and Koop (2018)), it is often absent in more traditional analyses of Phillips curve dynamics, even in papers that allow for time variation in the slope of the Phillips curve. As Cogley and Sargent (2005) and Primiceri (2005) point out, leaving out the stochastic volatility may lead to biased "fictitious" dynamics in the time-varying coefficients. Since including stochastic volatility renders the model to become nonlinear, we resort to Bayesian Gibbs sampler for estimation. The Gibbs sampler provides an efficient and practical algorithm to identify time-varying parameters from shocks in a nonlinear model with multiple unobserved components.

Our main findings are as follows. First, estimates of simple accelerationist Phillips curves for the U.S support the notion that the slope on the unemployment gap has declined since 1965. Our conditional Bayesian model diagnostics point to the importance of including expectations, particularly household expectations, and supply shock variables. The inclusion of professional expectations does not materially alter our conclusions based on the accelerationist specification. By contrast, specifications that also include household expectations point to a negative but volatile Phillips curve for headline inflation. Based on estimates for core inflation, we would conclude that the Phillips curve has weakened over time. This is because household expectations correlate less with core inflation than headline inflation. Second,we do not find convincing evidence supporting the anchored expectations hypothesis. While professional expectations have anchored, household expectations have not. Nevertheless, we find evidence for the notion, espoused by Coibion and Gorodnichenko (2015) that precisely because household expectations are not anchored that there was no disinflation after the Financial crisis. Third, allowing for stochastic volatility reduces the volatility and the width of the posterior distribution of the time-varying coefficient on the unemployment gap.

The remainder of the paper is organized as follows. In Section 2 we present our modelling framework. We discuss the data sources and the choices of variables for inflation, expectations and unemployment in Section 3. We present the empirical results in Section 4. We provide directions for further research in Section 5.

2 Methodology

We describe various specifications of our main model in section 2.1, and section 2.2 discusses the Gibbs sampler and our prior strategy. In section 2.4 we introduce our 'conditional model diagnostics', which we we will use to evaluate the fit of the various nested specifications.

2.1 Phillips curve models with time-varying features

The reduced-form expectations-augmented Phillips curve, first posited by Friedman (1968), is given by the model specification

$$\pi_t = \beta(u_t - u_t^*) + \pi_t^e + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2), \qquad t = 1, \dots, T, \tag{1}$$

where π_t denotes an observed measure of inflation, β is the coefficient of interest, $u_t - u_t^*$ represents the unemployment gap with u_t denoting the unemployment rate and u_t^* denoting the natural rate of unemployment or NAIRU (i.e. the rate corresponding to an economy in a steady state), and π_t^e is a measure of expected inflation in period t of period $t+1$. The disturbance term ε_t is a normally independently distributed (NID) disturbance term with mean zero and variance σ_ε^2. The time series length is T.

We examine possible time variation in β. For this purpose we allow this coefficient to vary over time, that is

$$\pi_t = \beta_t(u_t - u_t^*) + \pi_t^e + \varepsilon_t, \tag{2}$$

where β_t is now treated as a stochastically time-varying process. We typically assume that the coefficient β_t evolves as a random walk process

$$\beta_t = \beta_{t-1} + \eta_t \quad \eta_t \sim NID(0, \sigma_\eta^2), \tag{3}$$

where the disturbance term η_t is normally independently distributed with a mean of zero and variance σ_η^2. We assume that the disturbances ε_t and η_t are mutually and serially uncorrelated.

To introduce further flexibility in our model, we adopt a stochastic volatility

(SV) specification for ε_t with the effect of a time-varying variance, see for example Cogley and Sargent (2005), Primiceri (2005), Stock and Watson (2007), Berger, Everaert, and Vierke (2016), Chan, Koop, and Potter (2013, 2016)), Chan, Clark, and Koop (2018)). A key motivation is the improvement of the estimation of the time-varying coefficient: allowing for SV reduces the possibility that elements that belong in the disturbance term are included in our estimate of β_t, and SV reduces the need for ad-hoc dummies capturing one-off events. Since we impose a linear relationship between the left-hand and right-hand side of our models, another advantage is that stochastic volatility reduces the potential problem of misspecification in general.

The heteroskedastic unobservable shocks ε_t have variances $\sigma^2_{\varepsilon,t}$. We let $h_t = \log(\sigma_{\varepsilon,t})$ evolve as a random walk process given by

$$h_t = h_{t-1} + v_t, \qquad v_t \sim NID(0, \sigma^2_v), \tag{4}$$

where the disturbance term v_t is normally independently distributed with a mean of zero and variance σ^2_v. Allowing for stochastic volatility in the innovations renders the model non-linear, and modelling time varying parameters as states dramatically increases the number of parameters, both of which make the model more difficult to estimate. Bayesian methods can partially mitigate these problems, which is why we opt to use a Gibbs sampler for model inference. We discuss the relative advantages of our approach in section 2.2.

A baseline specification of equation (2) is the so-called *accelerationist* Phillips curve, where inflation only depends on the unemployment gap and backward-looking expectations, which are set equal to one or more lags of inflation or to $\frac{1}{4}\sum_{i=1}^{4}\pi_{t-i}$ (as in Ball and Mazumder (2011), Blanchard, Cerutti, and Summers (2015) and Blanchard (2016)). This equation hence relates the unemployment gap to the *change* in inflation rather than the level of inflation, that is

$$\pi_t - \frac{1}{4}\sum_{i=1}^{4}\pi_{t-i} = \beta_t(u_t - u_t^*) + \varepsilon_t. \tag{5}$$

In our empirical section, we will compare the accelerationist specification (5) with our main model of interest, equation (6). This model nests equation (5) and a number of other specifications that have appeared in the literature, particularly Ball and Mazumder (2011, 2019), Matheson and Stavrev (2013), Blanchard, Cerutti, and Summers (2015), Blanchard (2016), Coibion and Gorodnichenko (2015) and Coibion, Gorodnichenko, and Kamdar (2018). The specification is given by

$$\pi_t = \beta_t(u_t - u_t^*) + \theta_t \pi_t^{e,P} + \phi_t \pi_t^{e,H} + (1 - \theta_t - \phi_t)\pi_t^{e,B} + \boldsymbol{\gamma_t}'\boldsymbol{C_t} + \varepsilon_t, \qquad (6)$$

where π_t denotes annualized quarterly inflation and $u_t - u_t^*$ denotes the unemployment gap.

We include three types of expectations: survey expectations of professionals such as economists and professional forecasters, denoted by $\pi_t^{e,P}$, household inflation expectations $\pi_t^{e,H}$ obtained from household surveys, and purely backward-looking expectations $\pi_t^{e,B}$, which again we set equal to $\frac{1}{4}\sum_{i=1}^{4}\pi_{t-i}$. The literature actually distinguishes two types of expectations: 'backward-looking' or 'adaptive' expectations and purely 'forward-looking expectations'. A Phillips curve that only depends on backward-looking expectations $\pi_t^{e,B}$ takes the form of equation (5). In line with New-Keynesian theory, many empirical macroeconomic papers use a measure of forward-looking expectations instead. Forward-looking expectation measures are supposed to capture all information relevant to future inflation that in theory is not sensitive to recent shocks in past inflation or supply shocks. However, given the persistence of inflation, this specification is usually rejected by the data. Therefore, a now common approach is to assume that both forward- and backward-looking behavior play a role in the formation of expectations, resulting in a 'hybrid' or 'expectations-augmented' New-Keynesian Phillips curve as in equation (6).

Forward-looking expectations can be measured in several ways. In this study we follow the papers mentioned above and opt for measures derived from surveys of professionals ($\pi_t^{e,P}$) and households ($\pi_t^{e,H}$). A number of studies, summarized in Coibion, Gorodnichenko, and Kamdar (2018), find that including survey-based inflation expectations improves model fit, increases parameter precision and stability, reduces the need for ad-hoc lags and decreases forecast error.

Since the New-Keynesian Phillips curve is derived from the firm's optimization problem, a drawback of both household and professional expectation measures is that neither provides a direct measure of the expectations of firms. Data on the inflation expectations of firms is sparse, and there are currently no time series on firm expectations available for the countries in our sample[1]. Nevertheless, as Coibion and Gorodnichenko (2015) and Coibion, Gorodnichenko, and Kamdar (2018) concluded for the United States, survey expectations may serve as a proxy for the expectations of firms. We expect that, depending on the size and type of activities of firms, their expectations are in between those of professionals and households. We discuss the survey data in greater detail in section 3.3. We discuss the disadvantages of alternative measures, such as market-based measures or endogenized expectations in Appendix B.1.1.

The vector $\boldsymbol{C_t}$ contains supply shock control variables. Since Gordon (1982) pointed to the importance of supply shocks in the Phillips curve in his 'triangle model' of inflation, it is common to include these as controls. Following the literature (see e.g. Blanchard, Cerutti, and Summers (2015), Blanchard (2016), Gordon (2013), Coibion and Gorodnichenko (2015)), we include relative import price inflation[2] and oil price inflation[3]. Supply shocks can shift the level of inflation for a given unit of real activity, which leads to the spurious conclusion that the Phillips curve has flattened if these shocks are not included as control variables. A common alternative is to filter supply shocks out of the dependent variable by using measures for core inflation, usually by inflation excluding food and energy, or by computing the *median*[4] inflation rate. However, since a shock in for example oil prices could still affect, either directly or indirectly, the components that make up common measures of core inflation, neither of these measures may not be completely inoculated against the effect of supply shocks[5].

[1]The Federal Reserve Bank of Atlanta administers a survey of firm expectations, but only since 2011 and only for its own district. This survey is therefore of limited use to this paper.

[2]Relative import inflation is measured as the the annualized quarter-on-quarter change of the import price index divided by the GDP deflator index, as in Matheson and Stavrev (2013), Blanchard, Cerutti, and Summers (2015) and Gordon (2013).

[3]We use the West Texas Intermediate oil price.

[4]The Federal Reserve Bank of Cleveland developed a 'median' CPI inflation index, a measure of which Ball and Mazumder (2011, 2019) suggest it is a better measure of core inflation, because it is less affected by volatile transitory components. We also estimated our models for this measure. The results are in the Appendix.

[5]For example, Coibion and Gorodnichenko (2015) find that most of the differences in inflation

2.2 The Gibbs sampler

Inference with the Bayesian Gibbs sampler, and Bayesian MCMC samplers in general, have a number of advantages compared to frequentist methods. First, for the non-stationary unobserved processes included in the model, such as equation (3), the Maximum Likelihood point estimate of the standard deviation of the innovation is biased towards 0, a phenomenon the literature refers to as the *pile-up problem*; see, for example, Sargan and Bhargava (1983), Stock and Watson (1998) and Shephard and Harvey (1990). Second, allowing for stochastic volatility renders the model non-linear, which makes the likelihood of the model more difficult to maximize. It is not unlikely that such a model has a likelihood with multiple peaks. If these peaks are very wide, the parameters become difficult to identify (the 'flat likelihood' phenomenon). Conversely, if these peaks are very narrow, the likelihood may attain a maximum value in an unreasonable region of the parameter space. This is especially a risk for models with multiple unobserved components. A Bayesian MCMC method such as the Gibbs sampler splits the original estimation problem into multiple steps of smaller and less complex estimation problems, by drawing from conditional posteriors with a lower dimension than the joint posterior of the whole parameter set. Therefore the Gibbs sampler can deal with issues related to non-linearity and high dimensionality in an efficient manner. Moreover, the prior in a Bayesian approach can be utilized to ensure that the posterior distribution of a parameter does not attain values from an unreasonable region of the parameter space.

The Gibbs sampler algorithm used to evaluate the posterior distributions of the states and hyperparameters is based on the approach of Primiceri (2005) and Del Negro and Primiceri (2015). The results reported below are based on 60,000 Gibbs sampler iterations, of which we discarded the first 10,000 draws and stored every 5^{th} of the remaining 50,000 iterations, resulting in 10,000 draws used for the evaluation of the posterior distributions of the states and hyperparameters. As shown in Table 3 in the Appendix, the convergence checks are satisfactory as the sample autocorrelations decay fast.

expectations between households and professionals can be explained by oil price changes. Hence, oil prices are not only important to include because of direct effects on inflation, but also because of its role in the formation of inflation expectations.

To calibrate the prior distributions of $B_0 = [\beta_0, \theta_0, \phi_0, \gamma_0]$, h_0, Σ_η and σ_v^2, we use Maximum Likelihood (ML) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) estimates of equation (7) below based on the first 10 years of data (40 observations, dating from 1955Q1 to 1964Q4). Our approach builds on the prior settings of Cogley and Sargent (2005) and Primiceri (2005). They use OLS estimates of a time *invariant* model of the first 40 observations of his sample, thereby making assumptions about the order of magnitude of the scale parameters of the prior Inverse Wishart distributions of the disturbance variances. They also assume a scaled identity matrix for the scale parameter of the Inverse Wishart prior distribution for σ_v^2. In this paper we attempt to avoid these assumptions, and instead rely on ML and subsequently GARCH estimates of a model with time-varying parameters to inform the priors. To our knowledge, ours is the first paper to employ this strategy.

We use the same prior for the different specifications. The main reason for this is that using different priors for different specifications adds one extra source of potential variation between estimation results, which muddles our analysis.

To obtain our prior information, we estimate equation (7), which is a special case of equation (6), where ϕ_t is set to 0 and $\boldsymbol{C_t}$ only contains import price inflation, that is

$$\pi_t = \beta_t(u_t - u_t^*) + \theta_t \pi_t^{e,P} + (1-\theta_t)\pi_t^{e,B} + \gamma_t \pi_t^M + \varepsilon_t. \tag{7}$$

where π_t denotes headline CPI inflation, $u_t - u_t^*$ is the CBO estimate of the unemployment gap, $\pi_t^{e,P}$ is a combination of the Livingston Survey and the Survey of Professional Forecasters of one-year-ahead inflation expectations, $\pi_t^{e,B}$ is defined as described below equation (6), and the control variable π_t^M is relative import inflation. For convenience it is assumed that $\varepsilon_t \sim NID(0, \sigma_\varepsilon^2)$, i.e. we assume that there is no stochastic volatility. After all, allowing for stochastic volatility results in a non-linear model, which is one of the reasons why we opted for a Bayesian approach. However, as we will elaborate on below and in Appendix A.2, we consider a GARCH model for the residuals of this model, as denoted by $\hat{\varepsilon}_t$, to inform the prior for σ_v^2.

Equation (7) is based on the models of Matheson and Stavrev (2013) and

Blanchard, Cerutti, and Summers (2015). We chose this specification, because it is relatively concise compared to the general equation (6), but still contains at least one variable from the different possibilities for π_t^e and $\boldsymbol{C_t}$. For models that include more than one expectation or control variable, we set the prior distribution of the parameters on these variables equal to the prior distributions we obtain for comparable variables in equation (7). For example, the prior for ϕ_0 is set equal to the prior for θ_0, and the prior of the time-varying parameter on oil prices is set equal to the prior for that parameter on relative import prices.

Define $B_{ML,t} = [\beta_t, \theta_t, \gamma_t]'$ and $\tilde{y}_t = [\pi_1, \pi_2, \ldots, \pi_t]'$. Then, the elements in B_0 have a Normal conjugate prior for which we set the mean and variance equal to $\mathbb{E}(B_{ML,40}|\tilde{y}_{39}) = b_{40}$ and $\text{Var}(B_{ML,40}|\tilde{y}_{39}) = P_{40}$, respectively. We derive the underlying scale parameters of $\Sigma_{\eta,ML} = \text{diag}(\sigma^2_{\eta,ML,1}, \sigma^2_{\eta,ML,2}, \sigma^2_{\eta,ML,3})$ and use these to inform the prior, instead of using the ML estimates directly as input for the priors of Σ_η (as is done in e.g. Primiceri (2005)). Σ_η has an inverse-Wishart conjugate prior distribution for which the degrees of freedom T_0 are set to 2 + the dimension of the matrix, and a diagonal scale matrix Θ derived from the ML point estimates $\hat{\sigma}^2_{\eta,ML,i}$. The derivation is described in detail in Appendix A.2.2. For the prior for h_0 we use the smoothed disturbances $\hat{\varepsilon}_{ML,t} = \sigma_{\hat{\varepsilon},t}\epsilon_t$, $\epsilon_t \sim NID(0,1)$ to estimate a GARCH(1,1) model. We use the estimated parameters to compute $\mathbb{E}(\sigma^2_{\hat{\varepsilon},t}) = \mathbb{E}(\hat{\varepsilon}^2_{ML,t})$ and $\text{Var}(\sigma^2_{\hat{\varepsilon},t})$ to compute a prior for h_0. Recall that $h_t = \ln(\sigma_{\varepsilon,t})$. Hence, using the Delta method we derive $h_0 \sim N(\ln(\sqrt{\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})}), \widehat{\text{Var}}(\hat{\sigma}^2_{\hat{\varepsilon},t}) \cdot \frac{1}{2\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})})$, see Appendix A.2.3.

Estimating the GARCH(1,1) model results in an estimated series of $\hat{\sigma}^2_{\hat{\varepsilon},t}$, see Appendix A.2.4. To compute a prior for σ^2_v, analogous to the specification for h_t, we assume the logarithm of this series is generated by a random walk which we estimate with maximum likelihood. From this exercise we obtain $\hat{\sigma}^2_v$ and compute the the prior according to the procedure used to obtain priors for Σ_η, assuming an inverse-Gamma distribution for σ^2_v with degrees of freedom T_0 and scale parameter ϑ^v. In sum, the priors are specified as follows:

- $B_0 \sim N(b_{40}, P_{40})$, $\Sigma_\eta \sim IW(T_0, \Theta)$, $\sigma^2_v \sim IG(\frac{T_0}{2}, \frac{\vartheta^v}{2})$
- $h_0 \sim N(\ln(\sqrt{\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})}), \widehat{\text{Var}}(\hat{\sigma}^2_{\hat{\varepsilon},t}) \cdot \frac{1}{2\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})})$

Our robustness checks revealed that the posterior distributions of the state parameters β_0, θ_0, ϕ_0 and h_0 are not very sensitive to the prior distribution: different starting values have different but temporary effects on the posterior distributions, only visible in the first to five years of the sample, after which the distribution converges to the same values.

The values of the priors for Σ_η and σ_v^2 do have an effect on the volatility of these state parameters, although not on the trend. As is commonly known in the literature, these values affect the signal-to-noise ratio's, and depending on the value of these ratio's the posterior distributions of the state parameters are more or less smooth. Given the non-linearity of the model, and the relatively short time series for some countries, we do not find this surprising. As we explained above, our decision to opt for a Bayesian approach was informed by this sensitivity.

2.3 Interpreting the width of β_t posterior distribution

In general the posterior distribution of β_t is wider than fixed parameter estimates of our model. Since the time-varying coefficients are estimated as locally weighted averages, the number of observations used for these averages can be small. For frequentist approaches, this implies increased uncertainty around the point estimates, and in the Bayesian context In the Bayesian context the posterior distributions become wider compared to fixed parameter estimates. This problem is not unique to our method: rolling regressions suffer from the same issue if the subsamples are chosen too small[6]. This problem is exacerbated as the number of time-varying parameters increases. Therefore, we limited the number of variables with time varying parameters in our specification to a maximum of five.

Since we know that the potentially wide posterior distribution of β_t is partially the result of a small local sample, we cannot evaluate the significance of the posterior values as is done in frequentist analyses without committing a Type II error. Therefore, when the majority of the posterior distribution β_t is below 0, throughout this paper we consider this to be evidence in favor of a negative Phillips curve slope.

[6] Also, state space specifications mitigate this issue somewhat by imposing more structure on the estimates for β_t by specifying a functional form for the data generating process, which is chosen to be a random walk in this paper.

2.4 Conditional model diagnostics

In this paper we present a simple method to help us asses to what extent the variables in equation (6) additional to those in equation (5) improve the fit of the model. We can also use this method to check standard model assumptions, such as serial correlation, skewness and kurtosis. Our method provides a quick and efficient, yet insightful alternative to computing the marginal likelihoods for our models (as is done in e.g. Chan, Clark, and Koop (2018)). Note that our method can only be used to evaluate the fit of *nested* models. We cannot reliably use our method to evaluate the empirical fit of short-term versus long-term expectation horizon variables.

Our method can be used to obtain *conditional* model diagnostics: we employ the Kalman filter-smoother to compute β_t and the one-step ahead prediction errors[7] v_t *conditional* on the posterior mean of the (hyper)parameters of equation (6). We can also use the conditional values of v_t to check standard model assumptions, such as serial correlation, skewness and kurtosis.

To gauge the gain in model fit from each variable additional to those in the accelerationist specification (5), we compute the Bayesian Information Criterion (BIC) based on the conditional likelihood from nested versions of the posterior values of equation (6). The likelihood, and therefore the BIC are functions of the prediction errors. The model with the lowest BIC is preferred.

We first compute the conditional likelihood of equation (6) with all time varying parameters except for β_t restricted to 0, conditional on the posterior mean[8] of σ_t^2 and $\sigma_{\eta_1}^2$ of equation (6). Next we relax the restriction on θ_t and compute the filtered and smoothed β_t and BIC with the restriction $\phi_t = \gamma_t 1, t = \gamma_{2,t} = 0$ conditional on the posterior means of σ_t^2, $\sigma_{\eta_1}^2$ and θ_t. In the next iteration we only restrict $\gamma_t 1, t = \gamma_{2,t} = 0$ and condition on the posterior moments of ϕ_t, σ_t^2, $\sigma_{\eta_1}^2$ and θ_t. The following iteration only restricts $\gamma_{2,t} = 0$ and condition on the means of $\gamma_{1,t}$ together with aforementioned variables. In the final iteration we condition on the posterior moments of all time varying parameters except for β_t.

[7]The one-step-ahead prediction error is defined as $\mathbb{E}(\pi_t - \hat{\pi}_t | Y_t)$, where $\hat{\pi}_t$ is the filtered estimate of π_t conditional on the data $Y_t = ((u_t - u_t^*), \pi_t^B, \pi_t^P, \pi_t^C, \boldsymbol{C_t})$.

[8]We also computed these diagnostics conditional on the median and mode. The differences in the resulting statistics are negligible.

For each iteration, the conditional smoothed β_t is virtually identical to the unconditional posterior mean obtained with the Gibbs sampler. Therefore, we are sufficiently confident that our conclusions based on the conditional diagnostics can be generalized to the unconditional posterior distributions described in the previous section. For the one-step ahead prediction errors we also compute conditional skewness, kurtosis, Jarque-Bera and Ljung-Box statistics for the general model (6).

3 Data

We provide the details of our data set for the United States (U.S.) and used in the empirical study to measure inflation, the unemployment gap and inflation expectations. We motivate our choices of variables in relation to earlier studies.

3.1 Inflation

We measure inflation as the quarter-on-quarter annualized change of the three-month average price index. Our sample includes both headline and core CPI and PCE inflation. Overall, CPI inflation in the United States decreased less than expected during the downturn and did not increase as much during the upturn, hence the "puzzle" of missing (dis)inflation, see Figure B.1 in the Appendix. While the CPI is the oldest measure of inflation, the target of the Federal Reserve Bank is actually 2% core PCE inflation. Ball and Mazumder (2011, 2019) explain how this implies a CPI inflation target of approximately 2.5%. Quarter-on-quarter annualized CPI headline and core inflation have decreased significantly since the 1980s and has been hovering around the target levels of 2-2.5% since the mid-1990s. The same trends are visible in PCE inflation.

3.2 The unemployment gap

Following Ball and Mazumder (2011, 2019), we measure the unemployment gap by subtracting estimates of the NAIRU from the unemployment rate. Here we rely on widely acknowledged exogenous measures. We take the NAIRU estimate of the Congressional Budget Office, see Figures B.2 and B.3 in the Appendix.

Naturally, since this is an unobserved entity, these estimates are never perfect. Also, since unemployment is an imperfect measure of real activity, we expect that estimates of the slope on the unemployment gap are biased. This provides us with another reason to employ Bayesian methods, so that we can use the prior to mitigate this problem.

Additionally, we test the robustness of our estimates to an alternative measure of the unemployment gap using the short-term instead of the total unemployment rate, as is done by Ball and Mazumder (2019). The short-term unemployment rate is the percentage of the labor force unemployed for less than 27 weeks, see Figure B.3 in the Appendix[9]. Ball and Mazumder follow Krueger, Cramer, and Cho (2014) in arguing that those unemployed for a longer period of time are are no longer appealing to potential employers or may stop searching for work. Consequentially, they no are longer part of the share of unemployed that form an excess supply of labor and hence no longer put downward pressure on wage growth and inflation during downturns. Figure B.2 in the Appendix shows that the level of slack in the economy during the Great Recession appears much less severe when measured by short-term unemployment, which could explain the lack of disinflation during this period.

We decided against endogenizing the unemployment gap, as is done in Matheson and Stavrev (2013), Blanchard, Cerutti, and Summers (2015), and Chan, Koop, and Potter (2016) mainly because it is difficult to identify β_t from u_t^* without any additional information in the state equations. Moreover, allowing for stochastic volatility in the innovations exacerbates this problem. This issue can be mitigated by imposing more structure on the equation for the unemployment gap[10]. However, since our focus is on comparing different Phillips curve specifications, we seek to keep the model parsimonious in other dimensions, and leave this extension for future research.

[9] There is no direct measure of short-term unemployment available in the database of the Bureau of Labor Statistics. Therefore, we follow Ball and Mazumder who compute the short-term unemployment rate by taking the number of people unemployed for more than 27 weeks and subtracting this number from the total number of unemployed. This number is then divided by the labor force.

[10] (This can be done by modelling it as an AR(2) or as an IS equation, relating it to a Taylor rule, yielding a small macroeconometric model of the economy, as is done in varying degrees by e.g. Chan, Koop, and Potter (2016) and Berger, Everaert, and Vierke (2016)).

Another way to mitigate measurement uncertainty is by evaluating the impact of alternative measures of real activity. These include the output gap, survey-based measures such as the capacity utilization rate, or measures of marginal costs. However, we focus on the unemployment rate and leave the analysis of other measures to future research. Our reasons for this are as follows: first, the unemployment rate is the classic measure (see Phillips (1958)) and also the most commonly used measure of real activity in the Phillips curve literature, which makes it easier to compare our results with other empirical studies. Second, given the Federal Reserve Bank's dual mandate of stable prices and maximum stable employment, specifically estimating the relationship between the unemployment gap and inflation is in itself a valuable analysis. Third, alternative measures also suffer from measurement problems: measures of marginal costs, such as labor's share of income or unit labor costs turned out to be poor proxies (see Coibion and Gorodnichenko (2015) for a summary of the literature on this). The output gap, as is the case for the unemployment gap, is not directly measured but must be estimated. Also, survey-based measures such as the capacity utilization rate are relatively short time series, making analyses of the long-term evolution of the Phillips curve impossible.

3.3 Professional and household expectations

Table 1 summarizes the expectations data[11] used in our study. The main empirical results in this paper are based on the average or median survey response, depending on the series.

3.3.1 Professional expectations

We primarily rely on the Survey of Professional Forecasters (SPF) for our measures of short- and long-term expectations of professionals. The SPF is the oldest *quarterly* survey of inflation forecasts. It was administered by the American Statistical Association and the National Bureau of Economic Research before the

[11]Note that this list not exhaustive: it does not include the household survey of the Federal Reserve Bank of New York. Since this survey was first administered in 2013, it cannot be used in this paper. Also, we do not use Federal Reserve Bank Greenbook forecasts, since forecasts of recent years are not yet disclosed.

Federal Reserve Bank of Philadelphia took over in 1990. The survey first recorded 1-year ahead (seasonally adjusted) CPI inflation expectations in 1981Q3.

To replicate the results of the papers mentioned above which start in the 1960s, we follow their approach and combine the SPF with a biannual series on 1-year ahead CPI inflation provided by the Livingston Survey. The Livingston survey (LS) was first administered in 1946 and contains the longest series on biannual CPI inflation expectations from economists from industry, government, banking and academia. Since 1990, the survey has been administered by the Federal Reserve Bank of Philadelphia[12]. We transform the biannual series into a quarterly series through simple interpolation by means of a piecewise cubic hermite interpolating polynomial (PCHIP).[13]

To measure long-term inflation expectations, we use the SPF and the LS, which have 10-years-ahead inflation forecasts from 1991Q4 and 1990Q2 onward, respectively. The 10-years-ahead inflation forecasts measure the annual average rate of headline CPI inflation over the next 10 years expected by economists and forecasters. To extend these series, the Federal Reserve Bank of Philadelphia refers to the biannual series of the Blue Chip Economic Indicators (BCEI). This survey is administered to business economists and starts in 1979Q4. Neither the SPF nor the LS has sufficiently long series on other measures of inflation, such as the PCE or measures of core inflation[14]. Therefore, we use the CPI expectation variable to approximate expectations for alternative measures of inflation.

An alternative to the SPF, LS and BCEI survey is the Consensus Forecast (CF) survey from Consensus Economics. These series are much shorter, as the data set starts in 1990. The CF was a biannual survey until 2014, when the survey was administered three times. From 2015 onward the survey is administered on a quarterly basis. The forecast window in this survey differs from the SPF, LS and BCEI in that respondents are asked to give a forecast of inflation for *next year* instead of in the next four consecutive quarters. Particularly for short-term forecasts this difference could make a difference to estimates and interpretations of the Phillips curve.

[12]Before, the survey was managed by J. Livingston, a columnist for the *Philadelphia Inquirer*.

[13]We use the Piecewise Cubic Hermite Interpolating Polynomial function from MATLAB.

[14]Expectations of these measures are available only from 2007 onward.

3.3.2 Household expectations

We measure household short- and long-term inflation expectations with The University of Michigan Surveys of households (MSC). Since January 1978 its monthly questionnaire includes expectations of 1-year ahead inflation, see Appendix B.1 for more detail. Since 1979 the MSC also collects inflation expectations for the "next 5 to 10 years", first on an annual basis, and since 1990 as a monthly series. Note here that households are not asked to give a forecast of the CPI or another index, but rather indicate the direction of household prices in general. These qualitative judgments are then converted into a forecast.

3.3.3 Comparing professional and household expectations

Since approximately 2000, professional one-year ahead expectations have closely tracked average core inflation in the United States, but household expectations, have not: since the early 2000s, household expectations as measured by the MSC have been higher than core inflation and professional expectations, and well above 2 and 2.5% for most of the sample. MSC expectations also appear to be more volatile, and appear to closer track headline inflation, suggesting their expectations of future inflation are mostly informed by current inflation. This divergence has been discussed extensively by Coibion and Gorodnichenko (2015). In their study, diverging household expectations is the main explanation for the missing disinflation.

4 Empirical results

We first report the empirical results for the 'accelerationist' Phillips curve of equation (5) in Section 4.1. We discuss the effect of adding expectations and supply shock variables on the posterior distribution of β_t in section 4.2. In section 4.3 we review the fit of the model with the conditional model diagnostics introduced in section 2.4. We discuss the evidence regarding the anchoring of inflation expectations in the United States in section 4.4.

Table 1: Expectations data for the United States

Survey	Respondent	Measure	Forecast horizon	Starting	F
Survey of Profess. Forecasters	Professionals	CPI	1-year	1981Q3	Q
		GDP/GNP	1-year	1970Q2	Q
		CPI	10-years	1991Q4	Q
Livingston Survey	Professionals	CPI	1-year	1946M12,	B
			10-years	1990M06	B
Blue Chip Economic Indicators	Professionals	CPI	10-years	1979M12	B
Michigan Surveys of Households (HH)	HH	HH prices	1-year	1978M01	M
			5-years	1979M02[15]	Mix1
Consensus Forecasts	Professionals	CPI	next year, next 7 years	1990	Mix2

F - Frequency: Q - Quarterly; B - Biannual; M - Monthly.
Mix1 - Frequency Michigan Surveys is mixed: Annual (1979 - 1980), B (1980-1985), Q (1986-1987), M (1990-present).
Mix2 - Frequency Consensus Forecasts is mixed: B (until 2013), Triannual (2014), Q (2015-present).

4.1 Results for accelerationist Phillips curve specifications

Figure 4.1a and 4.1b show the posterior median, mean, 5^{th}, 16^{th}, 84^{th} and 95^{th} percentile of β_t for headline and core CPI inflation without stochastic volatility, which is equal to equation (5), or (6) with θ_t, ϕ_t and $\boldsymbol{\gamma}_t$ set to 0. Figures 4.1c and 4.1d display the same posterior moments for equation (5) with stochastic volatility. The volatilities are shown in Figure 4.1e and 4.1f. The results for median CPI and PCE inflation are comparable to those of headline and core CPI and are shown in Appendix C.2.

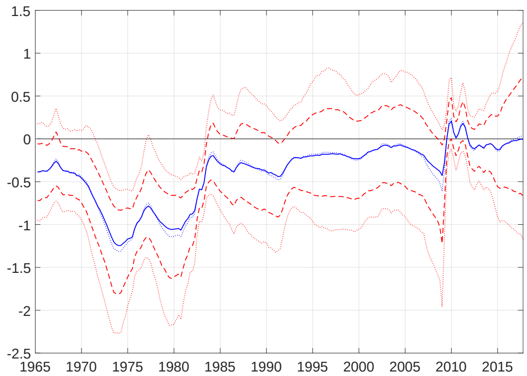
Based on the posterior distributions of β_t for the accelerationist Phillips curve we make four general observations. First, there is a decline in the slope since the 1970s, regardless of the measure of inflation used. For both headline and core CPI inflation the posterior median and mean of β_t hovered close to a value of -0.5 in the late 1960s, then decreased to around -0.7 to -1 in the 1970s and early 1980s during the 'Great Inflation' period when headline and core CPI q-o-q annualized ratcheted up to around 13-15%. During and after the 'Volcker' disinflation period, which started in 1980, inflation decreased rapidly. The Phillips curve relationship has weakened substantially since then: the slope deteriorates to the point where the posterior median and mean value of β_t are 0 or even slightly positive in the period after 2010. These results are in line with the main findings of Ball and Mazumder (2011), who estimate equation (5) without SV for median CPI inflation.

Second, allowing for stochastic volatility in the irregular component of equation (5) reduces the volatility and the width of the posterior distribution of β_t, since $\sigma^2_{\varepsilon,t}$ captures increased volatility during the high inflation of the 1970s and the subsequent Volcker disinflation as well as the Great Recession that followed the Financial Crisis of 2007-2008. Also, notice how the pattern of the posterior distribution of $\sigma^2_{\varepsilon,t}$ does not exhibit a peak in the aftermath of the Financial Crisis: increased volatility caused by the Financial Crisis mainly occurred in volatile markets such as those for food and energy.

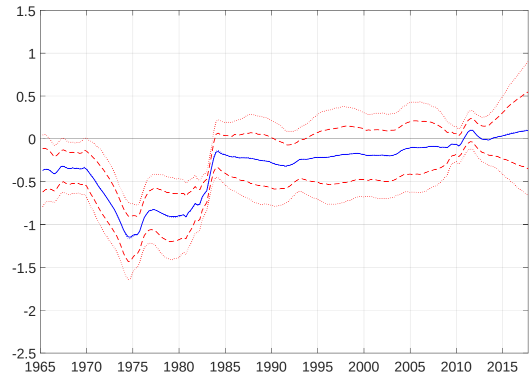
Third, we observe that β_t exhibits a cyclical pattern, particularly for core inflation. We find that periods with relatively rising inflation and a declining output gap, or decreasing inflation and a rising output gap correspond with more negative posterior values of β_t.

Figure 4.1: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t without SV (a) and with SV (b) from equation (5), and $\sigma^2_{\varepsilon,t}$ (c) from equation (4) for headline and core CPI inflation, 1965Q1 - 2017Q4

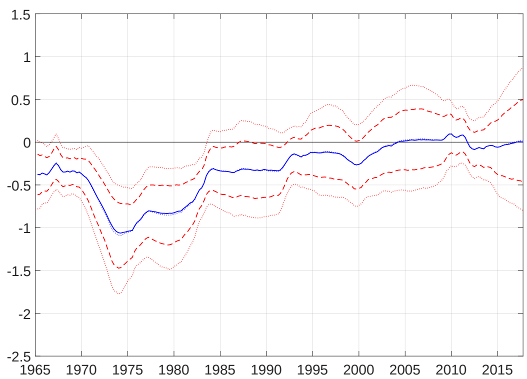
(a) β_t, eq. (5) without SV, headline CPI



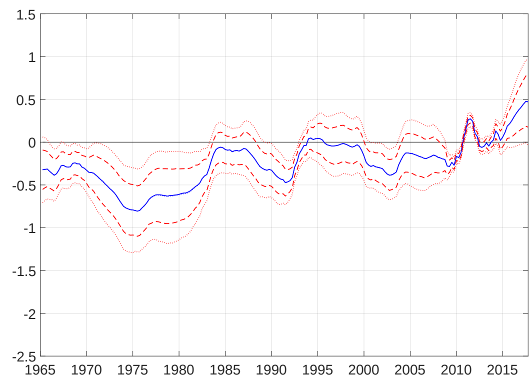
(b) β_t, eq. (5) without SV, core CPI



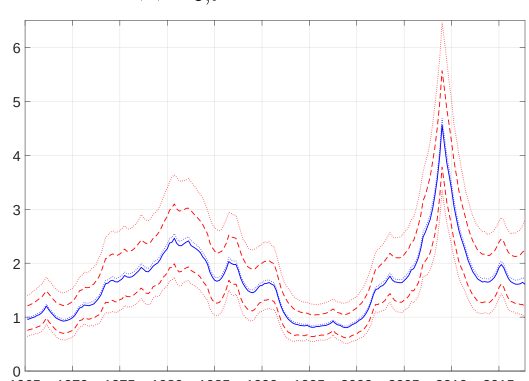
(c) β_t, eq. (5) with SV, headline CPI



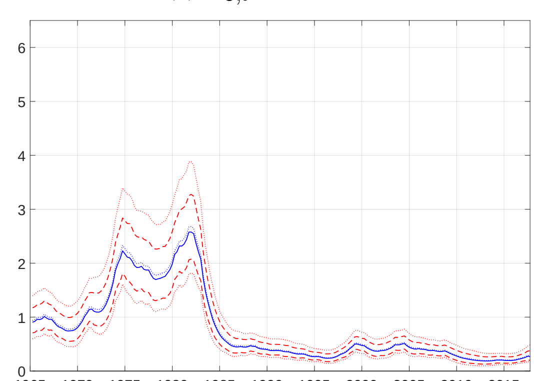
(d) β_t, eq. (5) with SV, core CPI



(e) $\sigma^2_{\varepsilon,t}$, headline CPI



(f) $\sigma^2_{\varepsilon,t}$, core CPI



4.2 Results for expectations-augmented Phillips curve

Next, we investigate whether our conclusions from the previous section are robust to specifications where in addition to the unemployment gap we include professional and household expectations and supply shock variables. Again, we allow for stochastic volatility in all models. Figures 4.2a to 4.2d show the posterior moments of β_t for headline and core CPI for different versions of equation (6). We compare these results with Figure 4.2e, which shows the posterior β_t values for a more commonly used Phillips curve specification with one-year-ahead professional expectations but no household expectations, and Figure 4.2f, which shows the posterior distribution of β_t for the complete model for headline inflation with one-year-ahead expectations and the measure for short-term unemployment discussed in section 3.2. The priors and a summary of the posterior distribution of the hyperparameters for the complete model with one-year-ahead expectations are reported in Table 4 in the Appendix.

Our findings are as follows. First, while adding one-year-ahead professional inflation expectations to the accelerationist Phillips curve does not alter the main conclusions drawn in the previous section (see Figure 4.2e), including household expectations from the Michigan Survey of Consumers reveals a posterior β_t that is more volatile: Figure 4.2a shows that the posterior distribution of β_t declines strongly from the start of the sample until the mid-1990s, when the slope is relatively stable but close to 0, and become more negative again for the period roughly between 2000 and 2015, and then weaken again afterwards. This corroborates some of the findings of Coibion and Gorodnichenko (2015), who explain the lack of disinflation in the United States by a rise in household expectations after 2008, coinciding with the rise in unemployment. However, we actually observe this rise much earlier on in the sample, in approximately 2000.

These fluctuations correspond to the dynamics in household expectations and ϕ_t: Figure 4.3a compares the dynamics in inflation, inflation expectations and their respective parameters β_t, θ_t and ϕ_t. The figure shows that professional and household expectations are closely aligned from the late 1980s until early 2000s, but household expectations deviate afterwards, as they appear to follow headline q-o-q inflation more closely. The relatively high correlation between inflation and household expectations results in a high median posterior value of ϕ_t compared to

θ_t until approximately 2008, when it declines again. From around 2011 onward we also see that household expectations and inflation are less aligned. This coincides with a slight decline in the posterior median of β_t. So, our results for a Phillips curve with short-term expectations supports the central thesis of Coibion and Gorodnichenko, but also show that the strength of the Phillips curve slope is still volatile, implying that the size and direction of the slope is time-dependent[16].

Second, if we replace short-term with long-term expectations, the posterior value of β_t becomes somewhat larger, but the decline in β_t is also more obvious, and more similar to the accelerationist Phillips curve. As we will see in the next section, long-term expectations are much more stable than short-term expectations, and consequently less strongly correlated with q-o-q inflation. As a result, the slope of an expectations-augmented Phillips curve with long-term expectations has more similar dynamics to an accelerationist Phillips curve.

Third, in contrast to headline inflation, for models for core inflation adding household expectations to the model for core inflation hardly has any effect on the distribution of β_t, which displays a declining trend as in the accelerationist Phillips curve. For both short-term and long-term expectations, ϕ_t decreases throughout the sample period, and the median value, shown in Figures 4.3c and 4.3d even becomes negative. By contrast θ_t remains relatively stable. The main explanation for this is that core inflation by definition is much less volatile than headline inflation. We already established that household expectations co-move more with headline than core inflation. Therefore, it has less of an effect on the fit of the model and on β_t.

Fourth, adding relative import and oil price inflation has the overall effect of smoothing the posterior distributions and reducing the previously observed cyclical dynamics in β_t. Including variables that control for supply shocks account for idiosyncratic shocks not captured by SV that would otherwise result in a volatile β_t. Figures 4.2c and 4.2d show that this is particularly true for core

[16] We investigated whether the shorter size of the series for household surveys has an impact on the difference between these two measures of expectations by estimating the model with professional 1-year-ahead expectations from 1978Q1, the start date of the series for 1-year-ahead household expectations. We found that the starting date only has a an impact on the posterior moments during the first three to five years of the sample, but then converges to the posterior moments generated by the longer sample.

inflation, providing evidence of the indirect effects of oil price dynamics.

Fifth, we checked whether replacing the unemployment gap with the short-term unemployment gap, as was suggested by Ball and Mazumder (2019) changes our results. Figure 4.2f shows that for headline cpi inflation in a model with short-term expectations, overall the posterior distribution of β_t becomes more negative, but the dynamics in the slope remain the same, albeit more smoothed. The same is true for core inflation and models with long-term expectations.

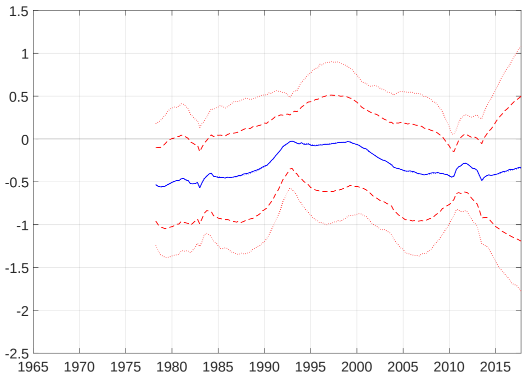
For PCE inflation we draw similar conclusions as for CPI inflation, except that overall β_t is weaker. We expected this: since there are no sufficiently long time series available fro PCE inflation expectations, so we included CPI inflation expectations instead. These are necessarily less closely correlated with PCE inflation, and hence have a smaller impact on β_t, resulting in a smaller difference between the accelerationist and expectations-augmented Phillips curves. The results are in Appendix C.2.

4.3 Conditional model diagnostics

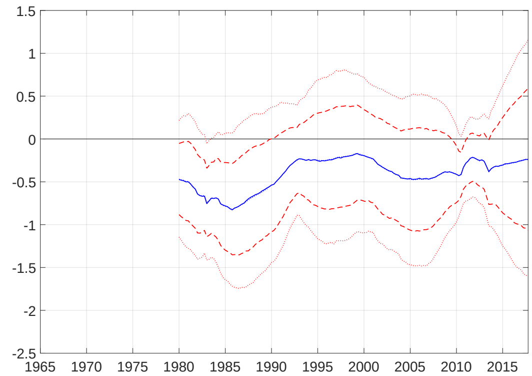
The conditional BICs described in section 2.4 are displayed in the first section of Table 2 below. The conditional BICs for the accelerationist model are in row 1. Row 5 shows the BICs for the unrestricted model, i.e. equation (6). The unrestricted model results in the lowest BIC for all specifications. Note that we cannot use our method to compare non-nested models. Hence, we cannot draw any conclusions on whether short or long-term inflation expectations are more empirically suitable. As was mentioned above, using short-term horizons is theoretically more correct, but given the number of papers that use long-term horizons instead, we also computed our conditional diagnostics for these specifications.

Figure 4.2: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t for (restricted versions of) equation(6) for headline CPI inflation, 1965Q1 - 2017Q4

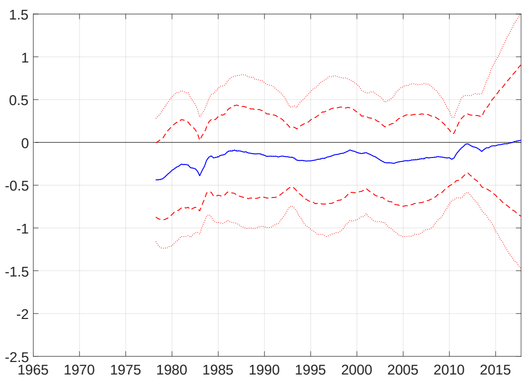
(a) Headline CPI, 1-year ahead expectations



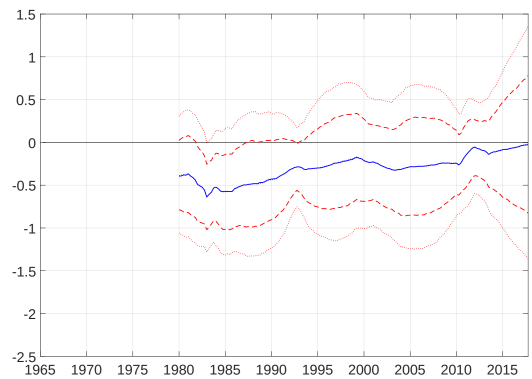
(b) Headline CPI, 10-year ahead expectations



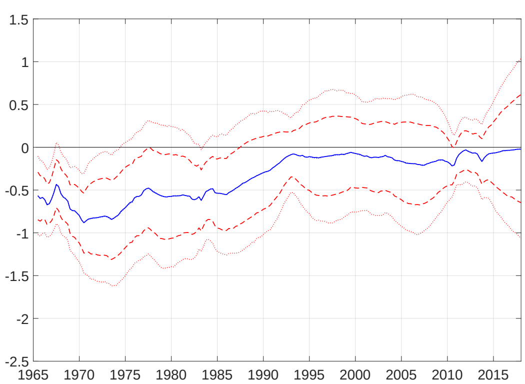
(c) Core CPI, 1-year ahead expectations



(d) Core CPI, 10-year ahead expectations



(e) Headline CPI, $\phi_t = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(f) Headline CPI, short-term unemployment, 1-year ahead expectations

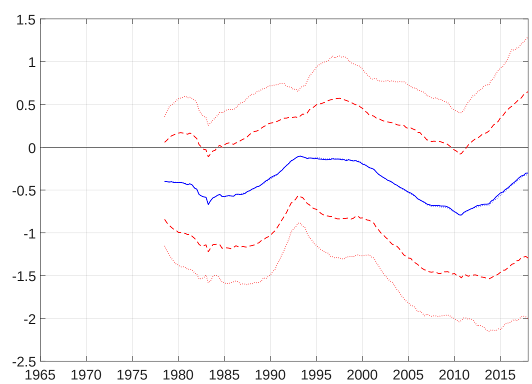
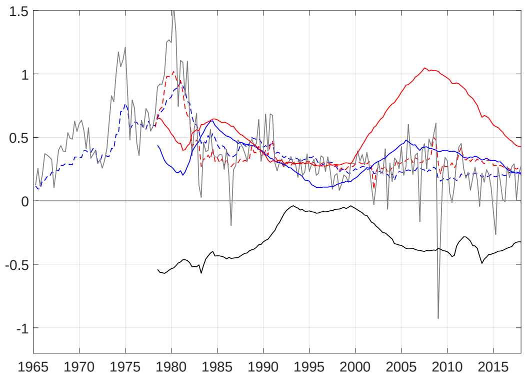
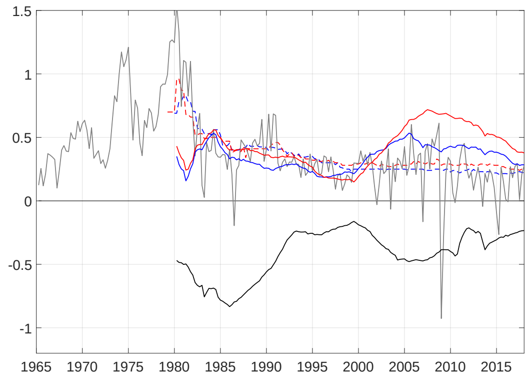


Figure 4.3: Posterior median of β_t (black), θ_t (blue), ϕ_t (red), professional (dashed blue) and household expectations (dashed red) and inflation (gray) for (restricted versions of) equation (6) for headline and core CPI inflation, 1965Q1 - 2017Q4

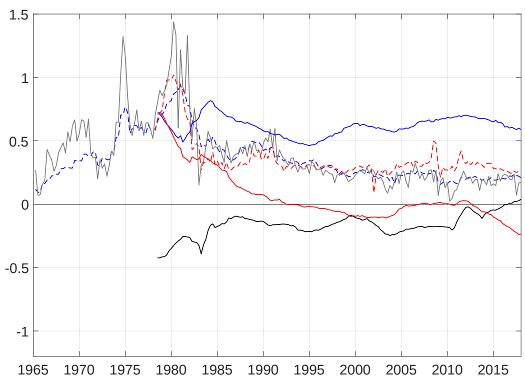
(a) Headline CPI, 1-year ahead expectations



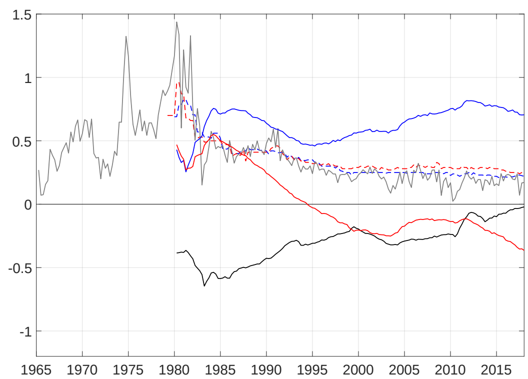
(b) Headline CPI, 10-year ahead expectations



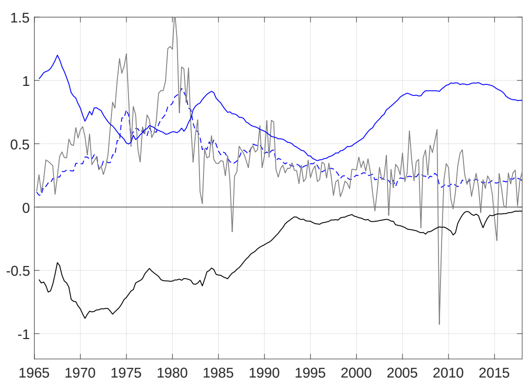
(c) Core CPI, 1-year ahead expectations



(d) Core CPI, 10-year ahead expectations



(e) Headline CPI, $\phi_t = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(f) Headline CPI, $\phi_t = 0$, $\pi_t^{e,P}$ is 10-year-ahead L-SPF

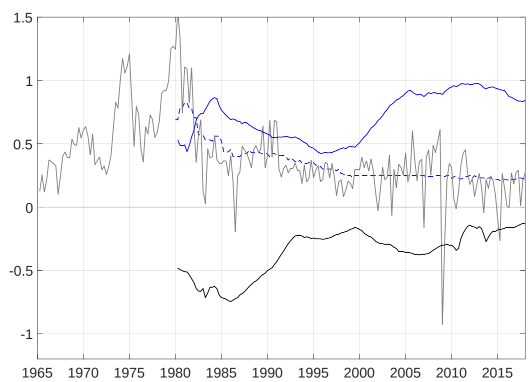


Table 2: Conditional Mean Estimates of the BIC and Other Model Diagnostics of equation (6)

	Conditional BIC			
	Short-term expectations		Long-term expectations	
	Headline CPI	Core CPI	Headline CPI	Core CPI
1. $\theta_t = \phi_t = \gamma_{1,t} = \gamma_{2,t} = 0$	4.525	2.848	4.497	2.831
2. $\phi_t = \gamma_{1,t} = \gamma_{2,t} = 0$	4.233	2.800	4.249	2.791
3. $\gamma_{1,t} = \gamma_{2,t} = 0$	3.833	2.781	4.092	2.806
4. $\gamma_{2,t} = 0$	3.530	2.724	3.786	2.822
5. unrestricted	2.711	2.472	2.704	2.374
difference row 1 and 5 (%)	-40.088	-13.202	-39.871	-16.143
	Relative contribution to BIC of unrestricted model			
$\pi_t^{e,P}$ (%)	-16.097	-12.766	-13.832	-8.753
$\pi_t^{e,H}$ (%)	-22.051	-5.053	-8.756	3.282
import inflation(%)	-16.703	-15.600	-17.066	3.501
oil price inflation (%)	-45.149	-67.021	-60.346	-98.031
	Conditional model diagnostics			
$Q(12)$	17.384	8.638	21.667**	10.991
$Q(20)$	26.178	29.161*	29.890*	22.464
Skewness	0.214	-1.216	-0.222	-0.674
Kurtosis	4.050	4.654	4.868	4.542
JB	8.519**	57.324***	23.335***	26.559***

Note: *, ** and *** indicate significance at 10, 5 and 1%-level, respectively.

The final row of the first section of the table shows the percentage difference between the conditional BICs of equation (6) and the accelerationist Phillips curve (5). For the models for headline inflation, the fit is improved by approximately 40%, regardless of whether one uses short- or long-term expectations. For the models for core inflation the expectation variables and supply shock variables only result in an improvement of 13 to 16%. These results confirms our earlier finding that since the variables additional to $u_t - u_t^*$ have less of an impact on the Phillips curve for core inflation, the posterior values of β_t of the full model are more similar to the accelerationist Phillips curve.

To compute the relative contribution of each additional variable to the BIC of this model we take the difference between two subsequent BICs and divide this by the difference between the BIC of the model row 1 and the model in row 5. Thus, each row in the second section of the table shows how much, percentage-wise, each additional variable contributes to the difference between the BIC of the accelerationist model (5) in row 1 and the largest model (6) in row 5.

In line with our findings in the previous section, for the model for headline inflation with short-term expectations, both expectation variables improve the fit. Household expectations have a larger impact than professional expectations. For a model with long-term expectations, the relative contribution of both expectation variables is smaller, and the relative contribution of household expectations is smaller than that of professional expectations. Again confirming our earlier findings, for the models for core inflation both types of professional and household expectations have a relatively small contribution. For the model with long-term expectation the difference is even positive, revealing the limited explanatory value of household expectations in explaining core inflation.

For all specifications, the oil price variable has the largest relative contribution to the model fit, particularly for core inflation, serving as evidence for the indirect effect of supply shocks. We saw earlier that the the main effect of adding oil prices is to smooth the slopes and reduce overall volatility, but adding oil prices has no substantial effect on the trends in β_t.

The assumptions underlying our model are that the disturbances are normally distributed and serially independent. Under these assumptions the standardized one-step ahead prediction errors $v_{s,t}$ are also normally distributed and serially independent. Therefore, for the full model we also compute Ljung-Box Q-statistics for autocorrelation for 12 and 20 lags, as well as the skewness, kurtosis and the Jarque-Bera (JB) test for Normality of the standardized prediction error $v_{s,t}$. While there is evidence of kurtosis, the skewness and autocorrelation diagnostics support our assumptions.

To summarize, based on accelerationist Phillips curves specifications, the slope β_t has declined since 1965. The inclusion of professional expectations does not materially alter this result. Based on our results for models that also include household expectations, conclusions about the strength of the Phillips curve are dependent on whether headline or core inflation is a more suitable dependent variable, and to what extent one believes the accuracy of the inflation expectation measures available, and whether one should use short- or long-term expectation horizons. On the basis of our conditional diagnostics, we cannot conclude which expectation horizon results in a better fit. Overall, the posterior distributions support the notion of a negative but volatile Phillips curve for headline inflation, that at the end of the sample period declines again for all specifications. If we interpret core inflation as a measure of 'true' inflation free from transitory noise, we would conclude that the Phillips curve has weakened over time. However, the main explanation for this is that household expectations are explicitly about headline inflation household inflation expectations and therefore correlate less with core inflation. This is confirmed by our conditional diagnostics. We would therefore not conclude that the Phillips curve is in secular decline.

4.4 Have expectations become more anchored in the U.S.?

The decline in the slope of the Phillips curve is often explained by expectations which have become *anchored* (Blanchard 2016, Ball and Mazumder 2011, 2018, Bernanke (2007)). Here, 'anchoring' refers to two coinciding phenomena: first, as the parameter on the unemployment gap declines, the parameter on expectation variables and hence the importance of expectations in explaining inflation dynamics increases. Second, inflation expectations have over time come to depend less on their lagged values and supply shock variables such as the oil price, but instead converged to the target set by the central bank. If one wants to attribute the decline in β_t to increased anchoring of inflation expectations, these two 'conditions' need to hold. Blanchard (2016) and Ball and Mazumder (2011, 2019) find evidence for the first and second condition for a model for headline inflation with long-term professional expectations, although the results of the latter two authors are not robust to different measures of inflation.

Regarding the first condition, for models with only professional expectations, based on Figures 4.3e and 4.3f one would confirm that from the 1990s onward β_t declined while θ_t increased. Before that, as β_t declined, θ_t was more volatile, and during the 1980s a decline in the median of θ_t coincided with a decline in β_t. This suggests the shift to inflation targeting in the 1990s resulted in increased anchoring.

However, for models which also include household expectations the opposite can also be true: Figures 4.3a and 4.3b show that since the late 1970s, as the parameter on household expectations lost and gained strength throughout the sample period for the full model, so did the parameter on the unemployment gap. Hence, while expectations matter, their increased importance does not necessarily coincide with a decline in β_t.

For core inflation, the median of β_t has declined the most, but whether this is due to increased anchoring however, is not obvious from the observed patterns in the posterior distribution of θ_t. The trends in θ_t and β_t are much more comparable to models without household expectations, mainly because household expectations contribute far less to the fit of the model, resulting in the median of ϕ_t turning 0 or even negative.

Loosely following Blanchard (2016), we can test for the second condition by again exploiting the advantages of state space methods. We estimate equation (8), in which we let inflation expectations depend on a time-varying intercept, contemporaneous inflation and lagged inflation and oil price changes, that is

$$\pi_t^{e,i} = \alpha_t + \phi_{1,t}\pi_t + \phi_{2,t}\pi_{t-1} + \gamma_t \Delta oil_t + \varepsilon_t, \tag{8}$$

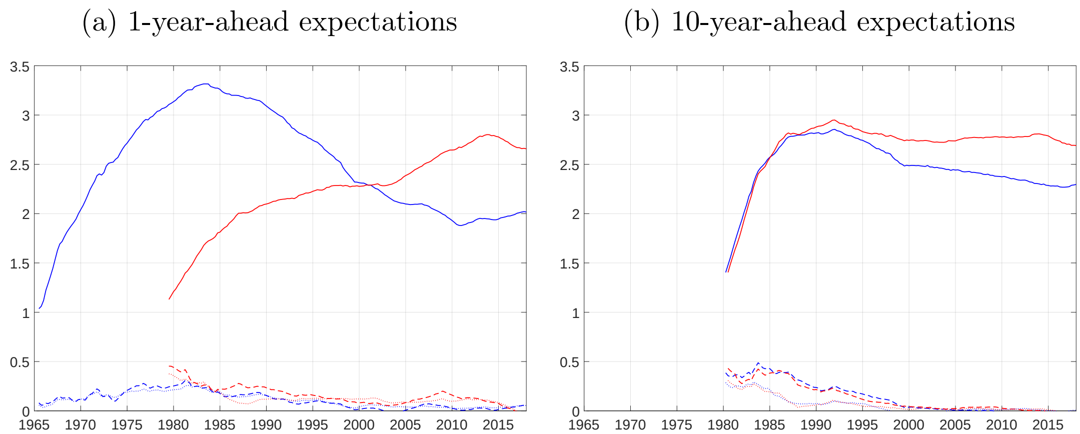
where $\pi_t^{e,i}$ again refers to a measure of inflation expectations i, where i can refer to either professional or household short- or long-term inflation expectations. The time-varying intercept α_t as well as the other time-varying parameters follow the stochastic process described in equation (3). Again, the natural logarithm of the variance of ε_t is assumed to follow a random walk process as in equation (4). We used the same priors for the hyperparameters as we did above.

The time-varying intercept α_t is a measure of convergence: if α_t converges to 2% or 2.5% (as was mentioned above, the target is either 2% or 2.5%, depending on the measure of inflation) while the posterior values of the other parameters decrease, one can conclude that that particular measure of inflation expectations has converged to the inflation target.

Figures 4.4a and 4.4b show the posterior median values of α_t, $\phi_{1,t}$ and $\phi_{2,t}$ for professional and household expectations for short- and long-term expectations.

The posterior median of the parameter on oil price changes hovers around 0, and is therefore not shown in the figures. Over time, both $\phi_{1,t}$ and $\phi_{2,t}$ have declined towards 0, implying strong anchoring of expectations, captured by the posterior median of intercept α_t. For one-year ahead professional expectations, this intercept has decreased since the early 1980s and has been hovering around 2% since 2005. For professional ten-year ahead expectations, the intercept is declining somewhat, but still in between 2 and 2.5%. This would indicate that professional expectations in the United Sates on average are anchored between the PCE target of 2% and the implicit CPI target of 2.5%.

Figure 4.4: Posterior median of α_t (solid line), $\phi_{1,t}$ (dashed line) and $\phi_{2,t}$ (dotted line) of equation (8), for professional (blue) and household expectations (red), 1965Q1 - 2017Q4



This is however not the case for household expectations: one-year ahead expectations have been trending upward since the beginning of the sample, and at the end of the sample exceed 2.5%. The posterior median estimate of ten-year ahead expectations are more stable, but appear to be anchored to a value above 2.5% since 1985. Hence, while the impact of transitory shocks on household expectations has diminished, they have not converged to the target set by the Federal Reserve Bank. We saw above that increased household inflation expectations explained some of the inflation dynamics we observed after the Financial crisis, when both household inflation expectations and unemployment increased. It hence appears that there is no disinflation because household expectations are *not* anchored to the target set by the Federal Reserve Bank. This finding echoes Coibion and Gorodnichenko's (2015) conclusions, although we find less of an explanatory role for oil price changes. Instead, the intercept of household expectations has increased, suggesting an overall higher level of household inflation expectations, regardless of temporary shocks.

5 Conclusions

This paper has employed Bayesian inference on unobserved components models to assess the dynamics in the slope of the Phillips curve over time. We have combined several Phillips curve specifications common in the literature, and allowed for stochastic volatility in the irregular component which reduced the overall width of the posterior distributions and the volatility of β_t. We have also proposed a simple method to evaluate the contribution of inflation expectations and supply shock variables, and to evaluate other relevant model diagnostics.

We systematically compared the outcomes for different measures of inflation, inflation expectations and the unemployment gap. We found that conclusions about the strength of the Phillips curve depend on which specification and which variables one includes to measure inflation, inflation expectations and the unemployment gap. The posterior results for Phillips curves for core inflation are in line with the nowadays common narrative that the slope of the Phillips curve is in decline; a specification with both short-term professional and household inflation expectations as well as short-term unemployment yielded the overall strongest slope, but this slope is also volatile. The key difference between these two outcomes is the inclusion of household expectations: these correlate more with headline inflation than core inflation, which explains the main difference in the posterior distribution of the slope. Our Bayesian model diagnostics support this finding. Overall, the posterior distributions support the notion of a Phillips curve slope that is negative throughout the sample but also volatile, declining again for all specifications at the end of the sample. We have not found convincing evidence that the decline in the Phillips curve since the 1960s can be attributed to an increased anchoring of inflation expectations.

References

Ball, L. M. and S. Mazumder (2011). "IInflation dynamics and the great recession". *Brookings Papers on Economic Activity* 42, pp. 337–405.

— (2019). "A Phillips Curve with Anchored Expectations and Short-Term Unemployment". *Journal of Money, Credit and Banking* 51(1), pp. 111–137.

Berger, T., G. Everaert, and H. Vierke (2016). "Testing for time variation in an unobserved components model for the US economy". *Journal of Economic Dynamics and Control* 69, pp. 179–208.

Bernanke, B. S. (July 2007). *Inflation Expectations and Inflation Forecasting.* Speech at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts.

Blanchard, O. (2016). "The Phillips Curve: Back to the '60s?" *American Economic Review* 106(5), pp. 31–34.

Blanchard, O., E. Cerutti, and L. Summers (2015). *Inflation and activity–two explorations and their monetary policy implications.* Tech. rep. National Bureau of Economic Research.

Chan, J. C. C., T. E. Clark, and G. Koop (2018). "A new model of inflation, trend inflation, and long-run inflation expectations". *Journal of Money, Credit and Banking* 50(1), pp. 5–53.

Chan, J. C. C., G. Koop, and S. M. Potter (2013). "A new model of trend inflation". *Journal of Business & Economic Statistics* 31(1), pp. 94–106.

— (2016). "A bounded model of time variation in trend inflation, NAIRU and the Phillips curve". *Journal of Applied Econometrics* 31(3), pp. 551–565.

Cogley, T. and T. J. Sargent (2005). "Drifts and volatilities: monetary policies and outcomes in the post WWII US". *Review of Economic dynamics* 8(2), pp. 262–302.

Coibion, O. and Y. Gorodnichenko (2015). "Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation". *American Economic Journal: Macroeconomics* 7(1), pp. 197–232.

Coibion, O., Y. Gorodnichenko, and R. Kamdar (2018). "The formation of expectations, inflation, and the phillips curve". *Journal of Economic Literature* 56(4), pp. 1447–91.

Del Negro, M. and G. E. Primiceri (2015). "Time Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum". *The Review of Economic Studies* 82(4), pp. 1342–1345.

Friedman, M. (1968). "The Role of Monetary Policy". *The American Economic Review* 58(1), pp. 1–17.

Gordon, R. J. (1982). "Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment". In: *Workers, Jobs, and Inflation*. Ed. by M. Baily. Washington, DC: The Brookings Institution, pp. 89–158.

— (2013). *The Phillips Curve is Alive and Well: Inflation and the NAIRU During the Slow Recovery*. Working Paper 19390. National Bureau of Economic Research.

Hall, R. (2013). "The routes into and out of the zero lower bound". *Proceedings-Economic Policy Symposium-Jackson Hole*. Federal Reserve Bank of Kansas City, pp. 1–35.

IMF (Apr. 2013). "Chapter 3: The Dog that Didn't Bark: Has Inflation Been Muzzled or Was it Just Sleeping?" *World Economic Outlook*.

Krueger, A. B., J. Cramer, and D. Cho (2014). "Are the long-term unemployed on the margins of the labor market?" *Brookings Papers on Economic Activity* 45(1), pp. 229–280.

Matheson, T. and E. Stavrev (2013). "The Great Recession and the inflation puzzle". *Economics Letters* 120(3), pp. 468–472.

Phillips, A. W. (1958). "The relation between unemployment and the Rate of change of money wage rates in the United Kingdom, 1861–1957". *Economica* 25(100), pp. 283–299.

Primiceri, G. E. (2005). "Time Varying Structural Vector Autoregressions and Monetary Policy". *The Review of Economic Studies* 72(3), pp. 821–852.

Sargan, J. D. and A. Bhargava (1983). "Maximum likelihood estimation of regression models with first order moving average errors when the root lies on the unit circle". *Econometrica* 51, pp. 799–820.

Shephard, N. G. and A. C. Harvey (1990). "On the probability of estimating a deterministic component in the local level model". *Journal of Time Series Analysis* 11(4), pp. 339–347.

Stock, J. H. and M. W. Watson (1998). "Median unbiased estimation of coefficient variance in a time-varying parameter model". *Journal of the American Statistical Association* 93(441), pp. 349–358.

— (2007). "Why has US inflation become harder to forecast?" *Journal of Money, Credit and Banking* 39(S1), pp. 3–32.

Appendices: An empirical assessment of the U.S. Phillips curve over time*

Marente Vlekke† Siem Jan Koopman‡ Martin Mellens§

November 2021

*The authors would like to thank Leon Bettendorf, Dennis Bonam, Adam Elbourne, Rob Euwals, Lennart Hoogerheide, Gerhard Ruenstler and Loes Verstegen for helpful comments and suggestions.

†Corresponding author. CPB Netherlands Bureau for Economic Policy Analysis, Vrije Universiteit Amsterdam. E-mail: m.vlekke@cpb.nl

‡Vrije Universiteit Amsterdam, CREATES Aarhus University, Tinbergen Institute.

§CPB Netherlands Bureau for Economic Policy Analysis.

Appendices

A Model summary and Bayesian inference

A.1 Model summary and state space representation

We can summarize the main model as follows. Let $X_t = [u_t - u_t^*, \pi_t^{e,P}, \pi_t^{e,H}, \boldsymbol{C_t}]$ and $B_t = [\beta_t, \theta_t, \phi_t, \boldsymbol{\gamma_t}']'$. Then equations (6) simplifies to

$$\pi_t = X_t B_t + \varepsilon_t, \tag{1}$$

where B_t evolves according to a random walk process, that is

$$B_t = B_{t-1} + \eta_t \qquad \eta_t \sim NID(0, \Sigma_\eta), \tag{2}$$

where

$$\Sigma_\eta = \begin{pmatrix} \sigma^2_{\eta,1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\eta,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2_{\eta,k} \end{pmatrix}, \tag{3}$$

where k denotes the number of time-varing coefficients in the model. We restrict Σ_η to be diagonal. Our robustness check showed that allowing Σ_η to be estimated freely did not change our main empirical results. As was mentioned above, we allow for the variance of ε_t to vary over time according to equation (4).

To cast equations (1) and (2) into state space form, recall the general linear Gaussian state space model

$$\begin{aligned} y_t &= Z_t \alpha_t + \varepsilon_t, & \varepsilon_t &\sim NID(0, H_t), \\ \alpha_t &= T_t \alpha_{t-1} + \eta_{t-1}, & \eta_t &\sim NID(0, Q_t), \end{aligned} \tag{4}$$

where the first equation, relating the $p\times1$ vector of dependent variable(s) y_t to the $k\times1$ vector of unobservables α_t, is generally referred to as the *measurement* or *observation* equation, and the second equation, which describes the dynamics of α_t, as the *state* or *transition* equation. Let $y_t = \pi_t - \pi_t^{e,B}$, $Z_t = [u_t - u_t^*, \pi_t^{e,P} - \pi_t^{e,B}, \pi_t^{e,H} - \pi_t^{e,B}, \boldsymbol{C_t}]$ and $\alpha_t = B_t = [\beta_t, \theta_t, \phi_t, \boldsymbol{\gamma_t}']'$. Here, $H_t = \sigma^2_{\varepsilon,t}$. The matrix $T_t = T$ is an identity matrix of size k, which equals 5 in the complete model. $Q_t = Q = \Sigma_\eta$.

A.2 Priors

As discussed in section 2.2, we used Maximum Likehelihood to estimate equation (7) on the first 40 observations of our dataset to inform our priors. Here we discuss in detail how obtain the the priors for the states B_0 and h_0, and hyperparameters Σ_η and σ^2_v.

A.2.1 $B_0 \sim N(b_{40}, P_{40})$

Earlier we defined $B_{ML,t} = [\beta_t, \theta_t, \gamma_t]'$ and $\tilde{y}_t = [\pi_1, \pi_2, \dots, \pi_t]'$. The elements in B_0 have a Normal conjugate prior for which we set the mean and variance equal to $\mathbb{E}(B_{ML,40}|\tilde{y}_{39}) = b_{40}$ and $\text{Var}(B_{ML,40}|\tilde{y}_{39}) = P_{40}$, respectively.

A.2.2 $\Sigma_\eta \sim IW(T_0, \Theta)$

We derive the underlying scale parameters of $\Sigma_{\eta,ML} = \text{diag}(\sigma^2_{\eta,ML,1}, \sigma^2_{\eta,ML,2}, \sigma^2_{\eta,ML,3})$ and use these to inform the prior instead of using the ML estimates directly as input for the priors of Σ_η (as is done in e.g. Primiceri (2005)). Σ_η has an inverse-Wishart conjugate prior distribution for which the degrees of freedom T_0 are set to 2 + the dimension of the matrix, and a diagonal scale matrix Θ derived from the ML point estimates $\hat{\sigma}^2_{\eta,ML,i}$. We assume that each $\hat{\sigma}^2_{\eta,ML,i}$ has an inverse-Gamma distribution with location parameter $\frac{T_0}{2}$ and scale parameter $\frac{\vartheta^\eta_i}{2}$, which we can derive by using that $\sigma^2_{\eta,ML,i} \sim IG(\frac{T_0}{2}, \frac{\vartheta^\eta_i}{2})$ with

$$\hat{\sigma}^2_{\eta,ML,i} = \widehat{\mathbb{E}}(\sigma^2_{\eta,ML,i}) = \frac{\frac{\vartheta^\eta_i}{2}}{\frac{T_0}{2} - 1},$$

which implies

$$\Big(\frac{T_0}{2} - 1\Big)\hat{\sigma}^2_{\eta,ML,i} = \frac{\vartheta^\eta_i}{2}.$$

We subsequently define $\Theta = \text{diag}(\frac{\vartheta^\eta_1}{2}, \frac{\vartheta^\eta_2}{2}, \frac{\vartheta^\eta_3}{2})$.

A.2.3 $h_0 \sim N(\ln(\sqrt{\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})}), \widehat{\text{Var}}(\hat{\sigma}^2_{\hat{\varepsilon},t}) \cdot \frac{1}{2\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})})$

For the prior for h_0 we use the smoothed disturbances $\hat{\varepsilon}_{ML,t} = \sigma_{\hat{\varepsilon},t}\epsilon_t$, $\epsilon_t \sim NID(0,1)$ to estimate a GARCH(1,1) model:

$$\sigma_{\hat{\varepsilon},t} = \omega_0 + \omega_1 \hat{\varepsilon}_{ML,t-1} + \xi \sigma^2_{\hat{\varepsilon},t-1}.$$

We use the estimates of of the parameters ω_0, ω_1 and ξ to compute $\mathbb{E}(\sigma^2_{\hat{\varepsilon},t}) = \mathbb{E}(\hat{\varepsilon}^2_{ML,t}) = \frac{\omega_0}{1-\omega_1-\xi}$ and $\text{Var}(\sigma^2_{\hat{\varepsilon},t}) = \frac{\omega_0^2 \cdot 2\omega_1^2}{(1-\omega_1-\xi)^2(1-3\omega_1^2-2\omega_1\xi-\xi^2)}$ to compute a prior for h_0. Recall that $h_t = \ln(\sigma_{\varepsilon,t})$. Hence, using the Delta method we derive $\widehat{\text{Var}}(\hat{h}_t) = \widehat{\text{Var}}(\ln\sqrt{\hat{\sigma}^2_{\hat{\varepsilon},t}}) = \text{Var}(\hat{\sigma}^2_{\hat{\varepsilon},t}) \cdot \frac{1}{2\hat{\sigma}^2_{\hat{\varepsilon},t}}$, where we set $\hat{\sigma}^2_{\hat{\varepsilon},t}$ equal to $\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})$. We thus obtain $h_0 \sim N(\ln(\sqrt{\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})}), \widehat{\text{Var}}(\hat{\sigma}^2_{\hat{\varepsilon},t}) \cdot \frac{1}{2\mathbb{E}(\hat{\sigma}^2_{\hat{\varepsilon},t})})$

A.2.4 $\sigma^2_v \sim IG(\frac{T_0}{2}, \frac{\vartheta^v}{2})$

Estimating the GARCH(1,1) model results in an estimated series of $\hat{\sigma}^2_{\hat{\varepsilon},t}$. To compute a prior for σ^2_v, analogous to the specification for h_t, we assume the logarithm of this series is generated by a random walk

$$\ln(\sqrt{\hat{\sigma}^2_{\hat{\varepsilon},t}}) = \ln(\sqrt{\hat{\sigma}^2_{\hat{\varepsilon},t-1}}) + v_{\hat{\varepsilon},t}, \qquad v_{\hat{\varepsilon},t} \sim NID(0, \sigma^2_v),$$

which we estimate with maximum likelihood. From this exercise we obtain $\hat{\sigma}_v^2$ and compute the the prior according to the same procedure as was described above, assuming an inverse-Gamma distribution for σ_v^2 with degrees of freedom T_0 and scale parameter ϑ^v.

A.3 The Gibbs sampler algorithm

Define $\tilde{y}_T = [y_1, \ldots, y_T]'$, $\tilde{Z}_T = [Z_1, \ldots, Z_T] = \tilde{\alpha}_T = [\alpha_1', \ldots, \alpha_T']'$ and $\tilde{h}_T = [h_1, \ldots, h_T]'$. The Gibbs sampler for the models evaluated in this paper consists of the following steps, which are repeated for $L + M$ iterations, where first L iterations are discarded and every j^{th} the last M iterations[1] are saved for inference:

1. Initialize $\tilde{h}_T^{(0)}$ and the hyperparameters $\Sigma_\eta^{(0)}$ and $\sigma_v^{2(0)}$.

   For $i = 1, \ldots, L + M$:

2. Draw $\tilde{\alpha}_T^{(i)}$ from $p(\tilde{\alpha}_T^{(i)} \mid \tilde{y}_T, \tilde{Z}_T, \tilde{h}_T^{(i-1)}, \Sigma_\eta^{(i-1)}, \sigma_v^{2(i-1)})$. We use the algorithm proposed by Carter and Kohn (1994).

3. Drawing $h_T^{(i)}$ requires equation (1) to be rewritten such that h_t appears on the right-hand side. For this, define $\sigma_t \epsilon_t = \varepsilon_t$ and continue as follows:

$$\pi_t - X_t B_t = \sigma_t \epsilon_t, \qquad \epsilon_t \sim NID(0, 1),$$

$$\log([\pi_t - X_t B_t]^2) = y_t^* = 2h_t + \log(\epsilon_t^2),$$

   where the $\log \chi^2(1)$ distribution of $\log(\epsilon_t^2)$ is approximated by a mixture of seven Normals as in S. Kim, Shephard, and Chib (1998). For each iteration, a vector $\tilde{s}_T = [s_1, \ldots, s_T]'$ is drawn, which for each t selects which of the seven components of the mixture of the Normal approximation is used for $\log(\epsilon_t^2)$. Hence, first $\tilde{s}_T^{(i)}$ is drawn from $p(\tilde{s}_T^{(i)} \mid \tilde{y}_T, \tilde{Z}_T, \tilde{\alpha}_T^{(i)}, \tilde{h}_T^{(i-1)}, \Sigma_\eta^{2(i-1)}, \sigma_v^{2(i-1)})$. Then $\tilde{h}_T^{(i)}$ can be drawn from $p(\tilde{h}_T^{(i)} \mid \tilde{y}_T, \tilde{Z}_T, \tilde{\alpha}_T^{(i)}, \tilde{s}_T^{(i)}, \Sigma_\eta^{(i-1)}, \sigma_v^{2(i-1)})$.

4. Conditional on $\tilde{y}_T$, $\tilde{Z}_T$, $\tilde{\alpha}_t$ and $\tilde{h}_T$, the variances Σ_η and σ_v^2 are independently distributed and hence $\Sigma_\eta^{(i)}$ and $\sigma_v^{2(i)}$ can be drawn from $p(\Sigma_\eta^{(i)}, \sigma_v^{2(i)} \mid \tilde{y}_T, \tilde{Z}_T, \tilde{\alpha}_T^{(i)}, \tilde{h}_T^{(i)}) = p(\Sigma_\eta^{(i)} \mid \tilde{y}_T, \tilde{Z}_T, \tilde{\alpha}_T^{(i)}, \tilde{h}_T^{(i)}) \cdot p(\sigma_v^{2(i)} \mid \tilde{y}_T, \tilde{Z}_T, \tilde{\alpha}_T^{(i)}, \tilde{h}_T^{(i)})$

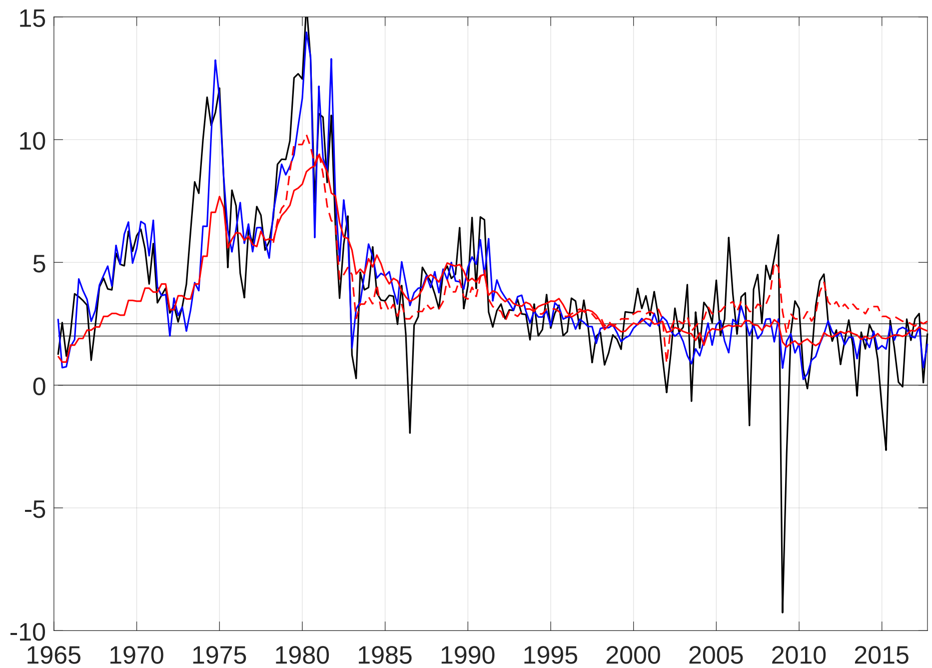
5. Go to step 2.

The algorithm for the model without stochastic volatility simplifies to a procedure where step 3 is skipped and where $\sigma_{\varepsilon,t}^{2(i)} = \sigma_\varepsilon^{2(i)}$ is drawn in step 4 together with $\Sigma_\eta^{(i)}$.

[1] In this paper, $L = 10{,}000$, $M = 50{,}000$ and $j = 5$.

B Data

Figure B.1: Annualized q-o-q headline CPI (black) and core CPI inflation (blue), short-term professional (solid red) and household inflation expectations (dashed red)[2]



[2]The horizontal black lines signify the 2% and 2.5% inflation target.

Figure B.2: Unemployment rate (solid) and NAIRU estimates (dashed), short-term unemployment and NAIRU in gray

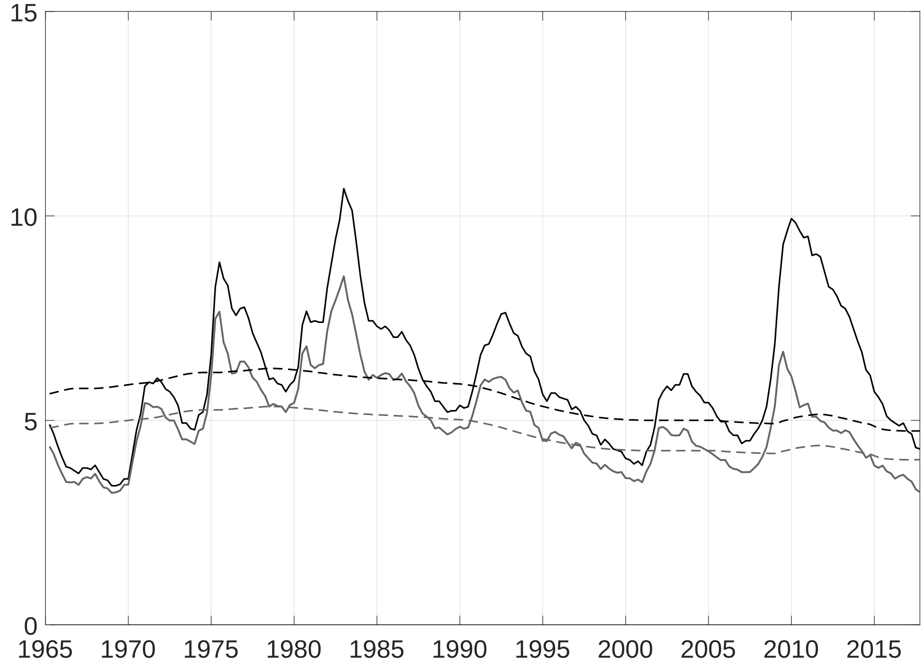
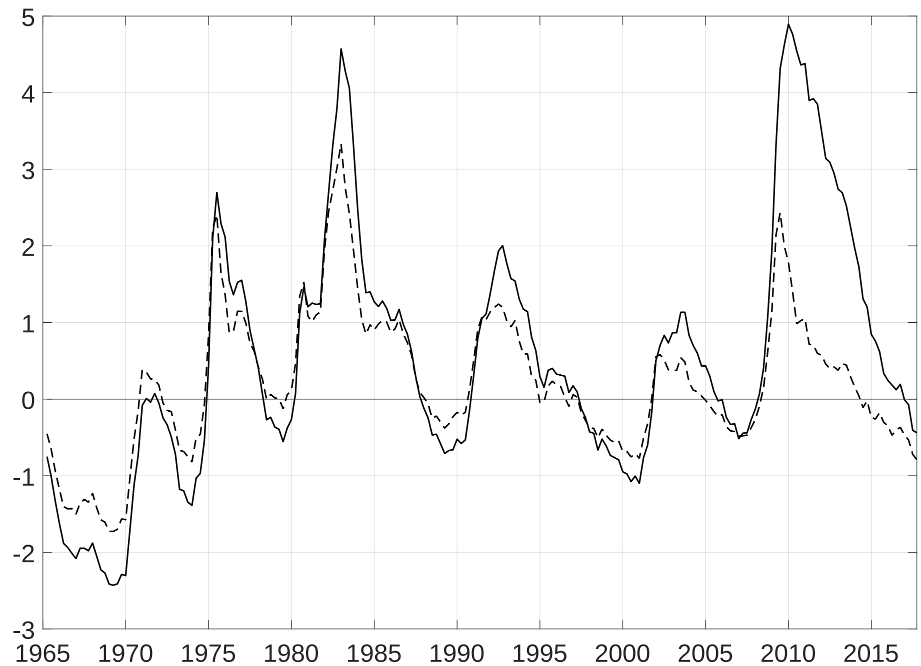


Figure B.3: Unemployment gap estimates (dashed line is short-term gap)



B.1 The University of Michigan Surveys of Consumers

In this survey, respondents are asked how much prices will change both in the short run (one-year-ahead) and the longer run (5 to 10 years ahead). The exact questions are:

- "During the next 12 months do you think that prices in general will go up, go down, or stay where they are now?"
- "By about what percent do you expect prices to go (up/down) on the average during the next 12 months?"
- "What about the outlook for prices over the next 5 to 10 years? Do you think prices will be higher, about the same, or lower, 5 to 10 years from now?"
- "By about what percent per year do you expect prices to go (up/down) on the average, during the next 5 to 10 years?"

After correcting for interpretation errors and missing information, the responses are transformed to percentage changes, of which the median is directly available. See Curtin (1996) for a detailed description of the procedure.

B.1.1 The disadvantages of alternative expectation measures

An alternative to surveys is to use market-based measures, such as the difference between the extra yield investors require to hold nominal assets that are exposed to inflation risk and assets that offer an inflation-adjusted return, such as Treasury inflation protected securities (TIPS). This difference, also known as 'inflation compensation' is often used to measure inflation expectations of investors. The three main objections against using them in this study are as follows. First, inflation compensation not only encompasses inflation expectations but also inflation risk premiums and potentially other factors, see e.g. Chen, Engstrom, and Grishchenko (2016). Identifying inflation expectations from risk premia is not straightforward, making inflation compensation a noisy measure of expectations, see e.g. Chernov and Mueller (2012), Grishchenko and Huang (2013) and D'Amico, D. H. Kim, and Wei (2018). Second, TIPS only became available only in 1997, which makes the times series too short for the goals of our analysis. Third, TIPS are not fully protected against inflation, because their payments are linked to the CPI three months prior to the date of payment. Inflation swaps suffer from the same disadvantages.

Another alternative is to endogenize expectations by assuming that long-run inflation expectations are implicitly captured by estimates of trend inflation. In unobserved components models with a trend-cycle decomposition it is then possible to interpret a time-varying trend, typically modelled as a random walk, as a measure of long-run inflation expectations. Papers adopting this approach to investigate the Phillips curve include Stock and Watson (2007, 2010), Harvey (2011), Chan, Koop, and Potter (2016), Berger, Everaert, and Vierke (2016), and Hindrayanto, Samarina, and Stanga (2019). The main advantage of this approach is that one is no longer limited by lack of data availability of survey measures. Another argument is that implicit long-run measures of expectations are not affected by sampling biases in surveys. While we acknowledge both advantages, we would argue that estimated trend inflation does not necessarily capture long-run inflation expectations either: estimated trend inflation in an unobserved components model for the Phillips curve can also be interpreted as a time-varying intercept, and can therefore also include shocks not captured by the rest of the model. Trend inflation is hence is not only a function of inflation expectations, but other factors as well. Therefore, even if surveyed inflation expectations and estimated trend inflation are not equal, this does not imply that there is a bias in the surveyed expectations. The supposed equivalence of trend inflation and long-run inflation forecasts was examined by Chan, Clark, and Koop (2018). They find that while long-run surveyed professional inflation expectations cannot be equated with trend inflation, surveys still have a valuable contribution to model fit and forecasting performance. Therefore, in this paper, we opt to explore the merits of surveyed expectations.

C Empirical results

C.1 Convergence diagnostics

Table 1: 20^{th}order sample correlation of the hyperparameter draws of model (6) for headline CPI inflation with 1-year ahead expectations

$\sigma^2_{\eta,\beta}$	$\sigma^2_{\eta,\theta}$	$\sigma^2_{\eta,\phi}$	σ^2_{η,γ_1}	σ^2_{η,γ_2}	σ^2_{v}
0.001	0.030	0.005	-0.002	-0.009	0.004

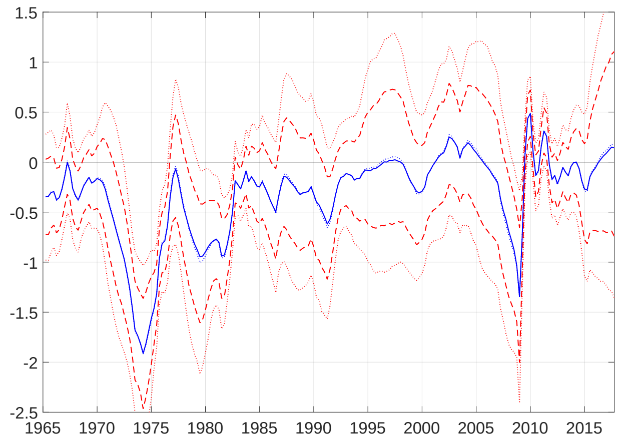
C.2 Posterior distributions of hyperparameters and β_t

Table 2: Priors and posterior distribution of the hyperparameters of model (6) for headline CPI inflation with 1-year ahead expectations

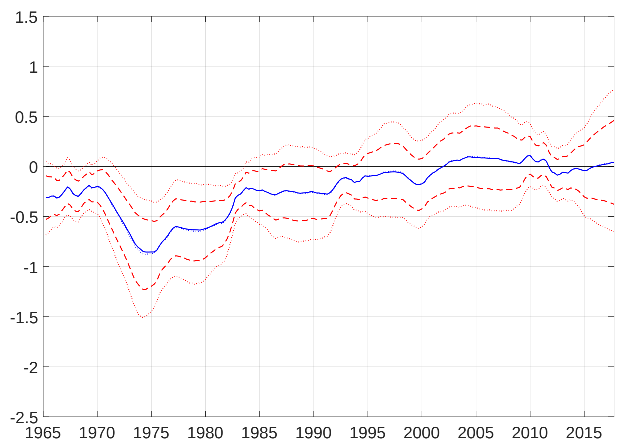
	Prior Value		Posterior Distribution (percentiles)				
Parameter	ϑ^η	T_0	5%	16%	median	84%	95%
$\sigma^2_{\eta,\beta}$	0.268	$\dim(B_t) + 2$	0.018	0.023	0.034	0.052	0.069
$\sigma^2_{\eta,\theta}$	0.268	$\dim(B_t) + 2$	0.019	0.026	0.040	0.064	0.090
$\sigma^2_{\eta,\phi}$	0.268	$\dim(B_t) + 2$	0.018	0.023	0.035	0.056	0.076
$\sigma^2_{\eta,\gamma}$	0.268	$\dim(B_t) + 2$	0.009	0.011	0.014	0.018	0.022
$\sigma^2_{\eta,\gamma}$	0.268	$\dim(B_t) + 2$	0.003	0.003	0.004	0.004	0.005
σ^2_{v}	0.454	$\dim(B_t) + 2$	0.021	0.025	0.034	0.048	0.060

Figure C.1: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t without SV (a) and with SV (b) from equation (5), and $\sigma^2_{\varepsilon,t}$ (c) from equation (4) for headline PCE inflation, 1965Q1 - 2017Q4

(a) $\hat{\beta}_t$ from equation (5) without SV



(b) $\hat{\beta}_t$ from equation (5) with SV



(c) $\hat{\sigma}^2_{\varepsilon,t}$ from equation (4)

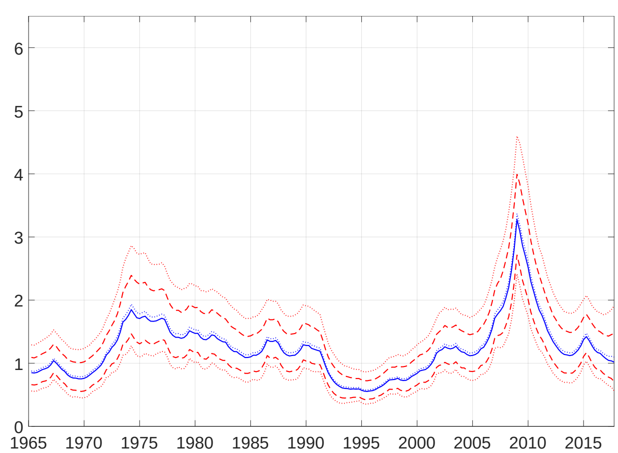
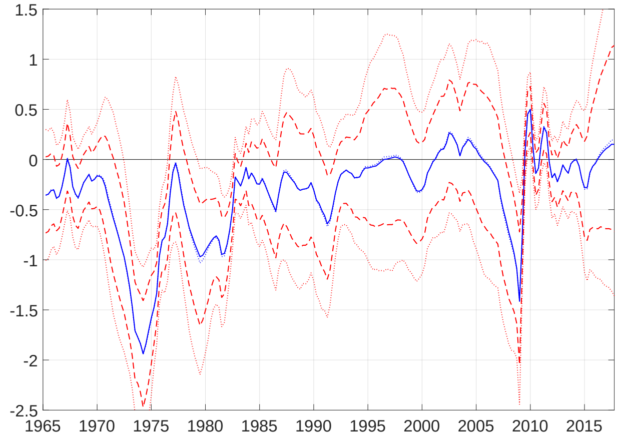
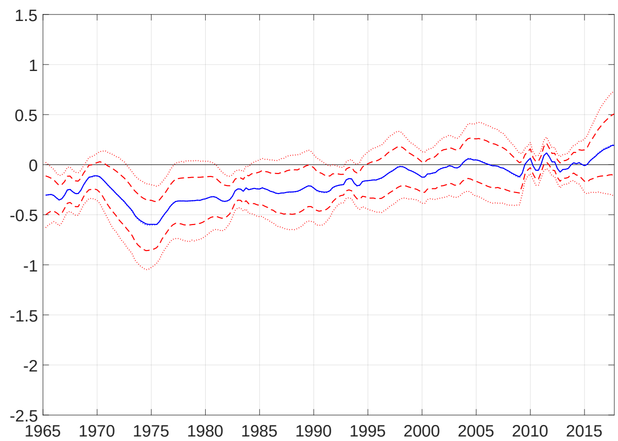


Figure C.2: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t without SV (a) and with SV (b) from equation (5), and $\sigma^2_{\varepsilon,t}$ (c) from equation (4) for PCE inflation excluding food and energy, 1965Q1 - 2017Q4

(a) $\hat{\beta}_t$ from equation (5) without SV



(b) $\hat{\beta}_t$ from equation (5) with SV



(c) $\hat{\sigma}^2_{\varepsilon,t}$ from equation (4)

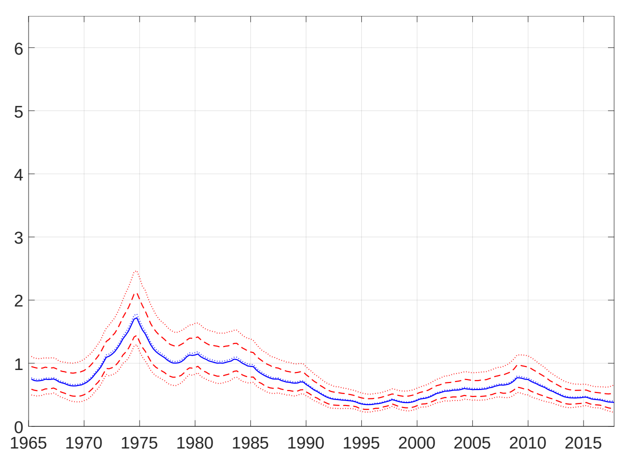
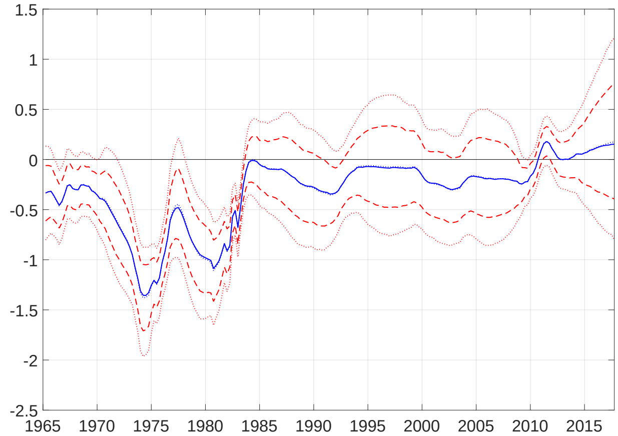
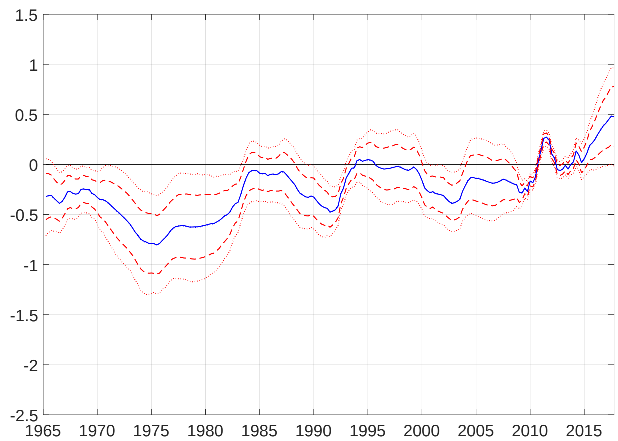


Figure C.3: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t without SV (a) and with SV (b) from equation (5), and $\sigma^2_{\varepsilon,t}$ (c) from equation (4) for median CPI inflation, 1965Q1 - 2017Q4

(a) $\hat{\beta}_t$ from equation (5) without SV



(b) $\hat{\beta}_t$ from equation (5) with SV



(c) $\hat{\sigma}^2_{\varepsilon,t}$ from equation (4)

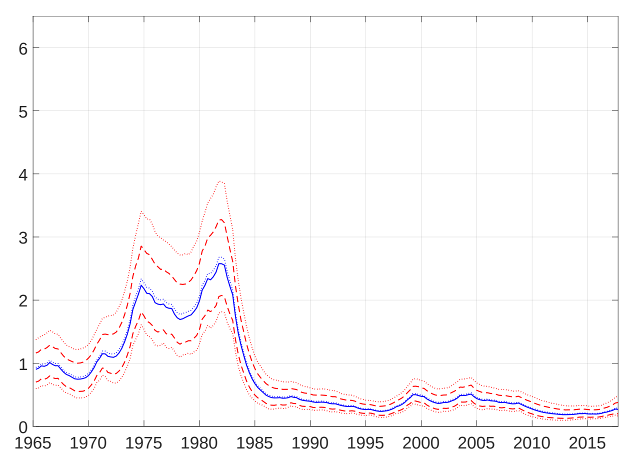
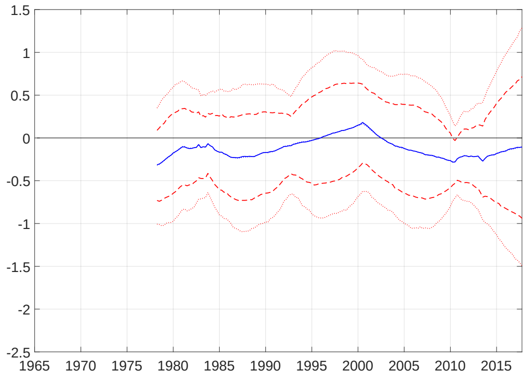
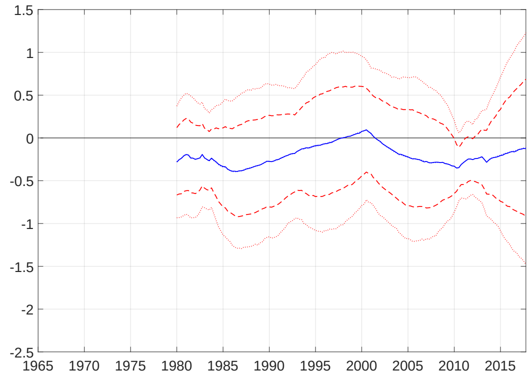


Figure C.4: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t for (restricted versions of) equation (6) for headline PCE inflation, 1965Q1 - 2017Q4

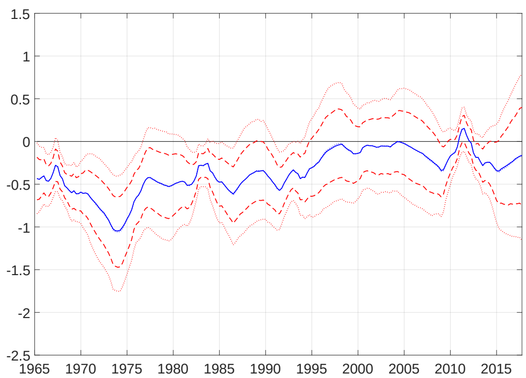
(a) 1-year ahead L-SPF and MSC



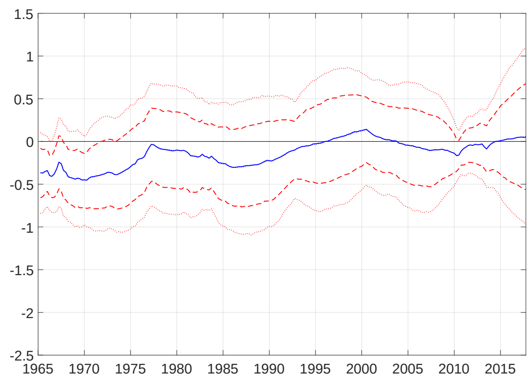
(b) 10-year ahead L-SPF and MSC



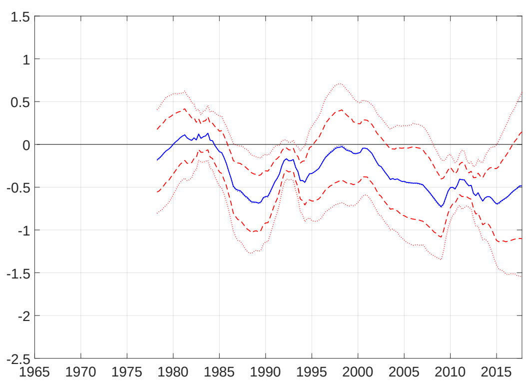
(c) $\phi_t = \boldsymbol{\gamma_t} = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(d) $\phi_t = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(e) $\theta_t = \boldsymbol{\gamma_t} = 0$, $\pi_t^{e,H}$ is 1-year-ahead MSC



(f) $\theta_t = 0$, $\pi_t^{e,H}$ is 1-year-ahead MSC

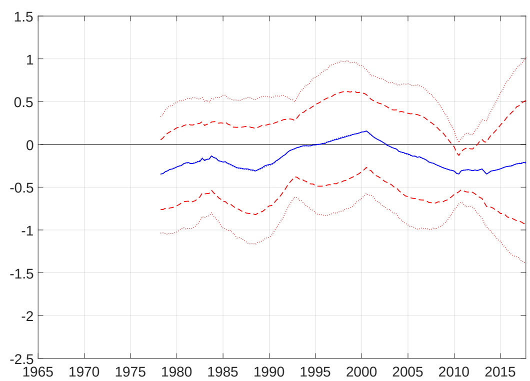
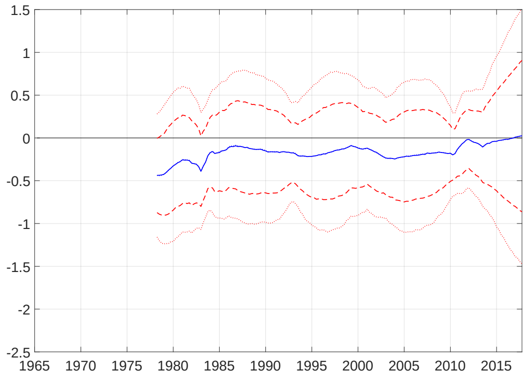
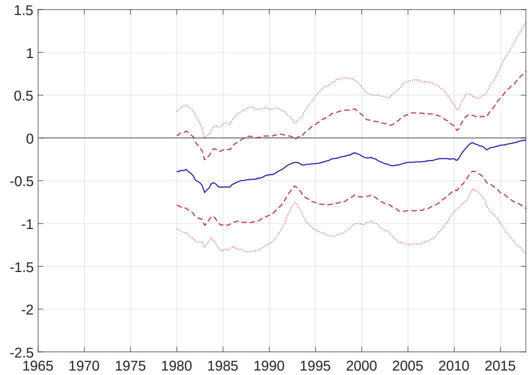


Figure C.5: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t for (restricted versions of) equation (6) for CPI inflation excluding food and energy, 1965Q1 - 2017Q4

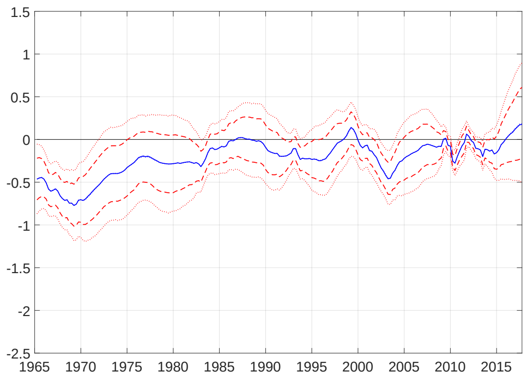
(a) 1-year ahead L-SPF and MSC



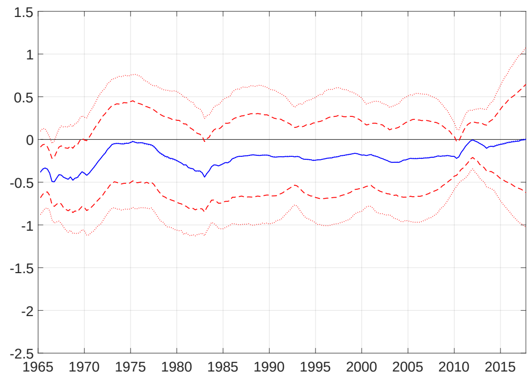
(b) 10-year ahead L-SPF and MSC



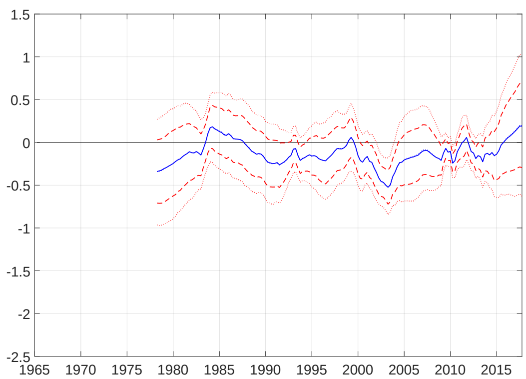
(c) $\phi_t = \boldsymbol{\gamma_t} = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(d) $\phi_t = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(e) $\theta_t = \boldsymbol{\gamma_t} = 0$, $\pi_t^{e,H}$ is 1-year-ahead MSC



(f) $\theta_t = 0$, $\pi_t^{e,H}$ is 1-year-ahead MSC

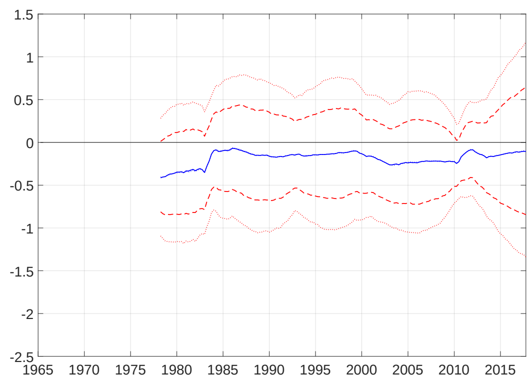
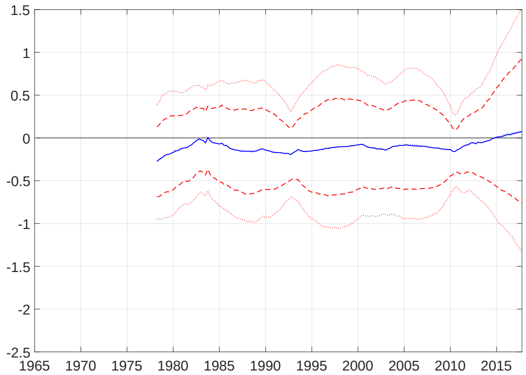
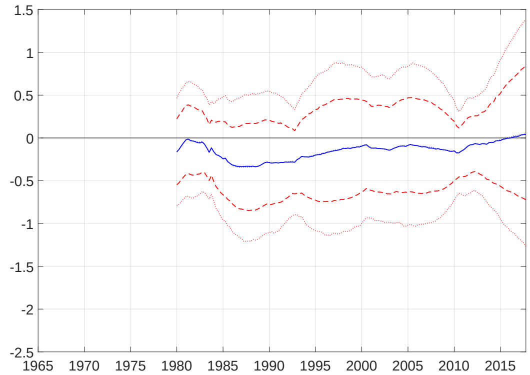


Figure C.6: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t for (restricted versions of) equation (6) for PCE inflation excluding food and energy, 1965Q1 - 2017Q4

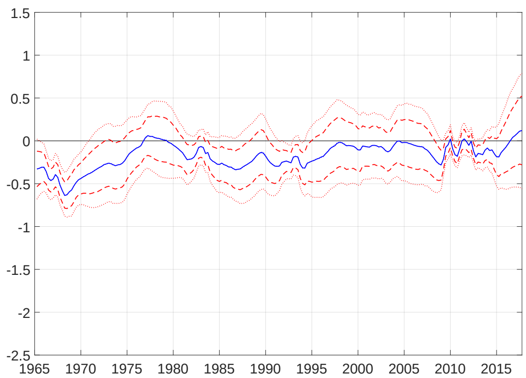
(a) 1-year ahead L-SPF and MSC



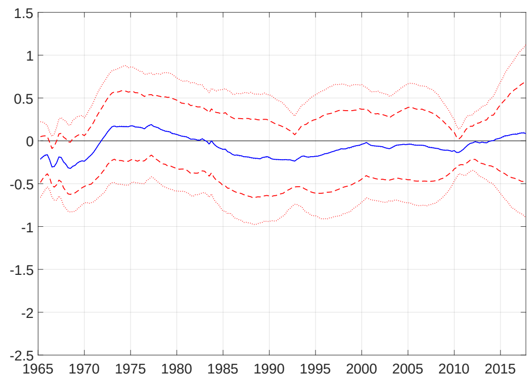
(b) 10-year ahead L-SPF and MSC



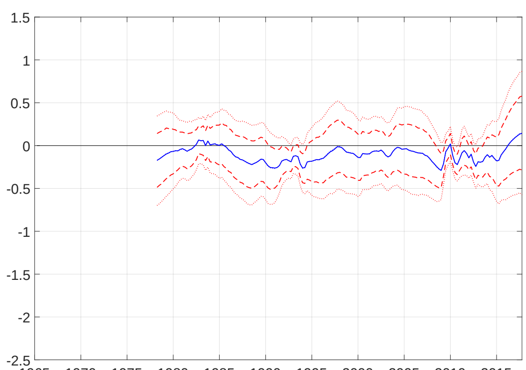
(c) $\phi_t = \boldsymbol{\gamma_t} = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(d) $\phi_t = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(e) $\theta_t = \boldsymbol{\gamma_t} = 0$, $\pi_t^{e,H}$ is 1-year-ahead MSC



(f) $\theta_t = 0$, $\pi_t^{e,H}$ is 1-year-ahead MSC

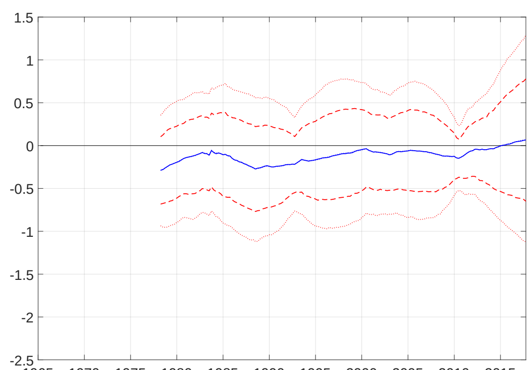
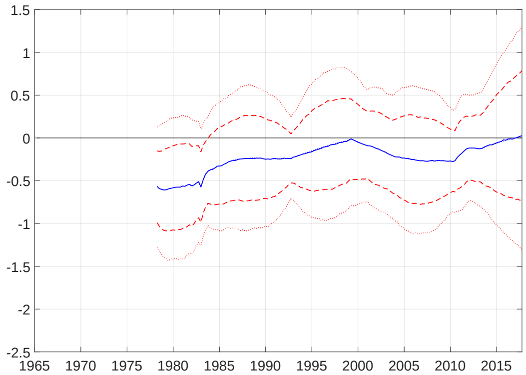
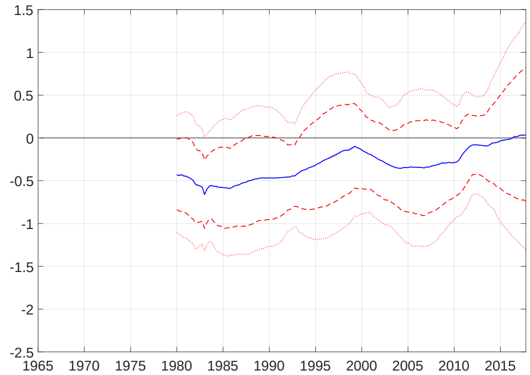


Figure C.7: Posterior median (blue), mean (dotted blue), 5^{th} and 95^{th} percentiles (dotted red) and 16^{th} and 84^{th} percentiles (dashed red) of β_t for (restricted versions of) equation (6) for median CPI inflation, 1965Q1 - 2017Q4

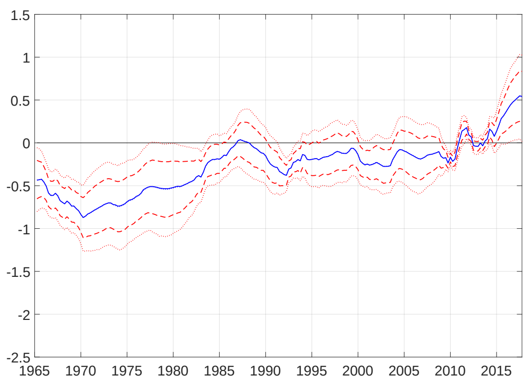
(a) 1-year ahead L-SPF and MSC



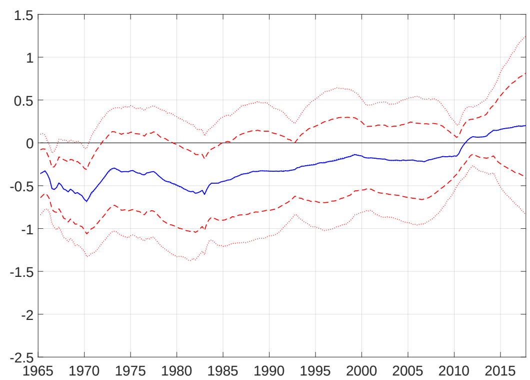
(b) 10-year ahead L-SPF and MSC



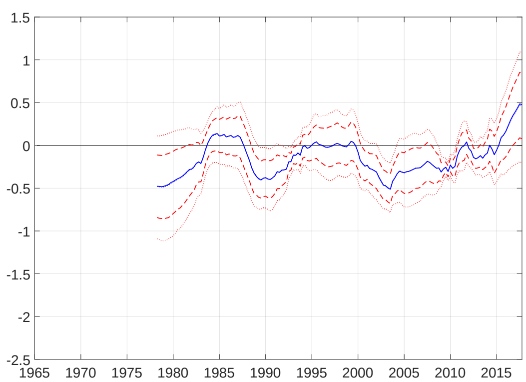
(c) $\phi_t = \boldsymbol{\gamma_t} = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(d) $\phi_t = 0$, $\pi_t^{e,P}$ is 1-year-ahead L-SPF



(e) $\theta_t = \boldsymbol{\gamma_t} = 0$, $\pi_t^{e,H}$ is 1-year-ahead MSC



(f) $\theta_t = 0$, $\pi_t^{e,H}$ is 1-year-ahead MSC