TI 2013-090/III
Tinbergen Institute Discussion Paper

Posterior-Predictive Evidence on US Inflation using Extended Phillips Curve Models with Non-filtered Data

Nalan Basturk[1,3]
Cem Cakmakli[2]
Pinar Ceyhan[1,3]
Herman K. van Dijk[1,3,4]

[1] *Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam;*
[2] *Faculty of Economics and Business, University of Amsterdam;*
[3] *Tinbergen Institute;*
[4] *Faculty of Economics and Business Administration, VU University Amsterdam.*

Posterior-Predictive Evidence on US Inflation using Extended Phillips Curve Models with non-filtered Data*

Nalan Basturk[1,3], Cem Cakmakli[†,2], Pinar Ceyhan[1,3], and Herman K. van Dijk[1,3,4]

[1] *Econometric Institute, Erasmus University Rotterdam*
[2] *Department of Quantitative Economics, University of Amsterdam*
[3] *Tinbergen Institute*
[4] *Department of Econometrics, VU University Amsterdam*

July 2013

Abstract

Changing time series properties of US inflation and economic activity, measured as marginal costs, are modeled within a set of extended Phillips Curve (PC) models. It is shown that mechanical removal or modeling of simple low frequency movements in the data may yield poor predictive results which depend on the model specification used. Basic PC models are extended to include structural time series models that describe typical time varying patterns in levels and volatilities. Forward as well as backward looking expectation mechanisms for inflation are incorporated and their relative importance evaluated. Survey data on expected inflation are introduced to strengthen the information in the likelihood. Use is made of simulation based Bayesian techniques for the empirical analysis. No credible evidence is found on endogeneity and long run stability between inflation and marginal costs. Backward-looking inflation appears stronger that forward-looking one. Levels and volatilities of inflation are estimated more precisely using rich PC models. Estimated inflation expectations track nicely the observed long run inflation from the survey data. The extended PC structures compare favorably with existing basic Bayesian Vector Autoregressive and Stochastic Volatility models in terms of fit and prediction. Tails of the complete predictive distributions indicate an increase in the probability of disinflation in recent years.

Keywords: New Keynesian Phillips curve, unobserved components, time varying parameters, level shifts, inflation expectations, survey data
JEL Classification: C11, C32, E31, E37

*A preliminary version of this paper was presented at the Tinbergen Institute workshop 'Recent Theory and Applications of DSGE Models', the Federal Reserve Bank of Atlanta workshop on 'Methods and Applications for Dynamic Stochastic General Equilibrium Models', the European Seminar on Bayesian Econometrics in Vienna, the Ecares Seminar in Brussels, the 5th CSDA international conference on computational and financial econometrics (CFE'12, Oviedo) and the Tripartite Seminar at the Wharton School. We thank Gianni Amisano, Frank DiTraglia, Sylvia Fruehwirth-Schnatter, Domenico Giannone, James Heckman, Alejandro Justitiano, Frank Kleibergen, Michele Lenza, Richard Paap, Georgio Primicieri, Ricardo Reis and three anonymous referees and the guest-editor Frank Schorfheide for constructive discussions and very helpful comments on the preliminary version of the paper. All remaining errors are the authors responsibility. Nalan Basturk and Herman K. van Dijk are supported by the NWO Grant number 400–09–340. Cem Cakmakli is grateful to the AXA Research Fund for financial support.

†Nalan Basturk (basturk@ese.eur.nl), Cem Cakmakli (C.Cakmakli@uva.nl., corresponding author), Pinar Ceyhan (ceyhan@ese.eur.nl), Herman K. van Dijk (hkvandijk@ese.eur.nl)

1 Introduction

Modeling the relation between inflation and fluctuations in economic activity has been one of the building blocks of macroeconomic policy analysis. Often, the analysis of this relation, denoted as Phillips Curve (PC) models, is conducted using the short-run variations in inflation and economic activity[1]. The conventional method for extracting this short run variation in the observed series is to demean and detrend the data prior to analysis, see Galí and Gertler (1999); Nason and Smith (2008). However, mechanical removal of the low frequency movements in the data may lead to misspecification in the models, as suggested in Ferroni (2011); Canova (2012) for DSGE models. The existence of complex low frequency movements, such as potential structural breaks and level shifts in the observed series in particular in the inflation series, is well documented in the literature (McConnell and Perez-Quiros, 2000; Stock and Watson, 2008). For instance distinct periods with different patterns can be observed for the non-filtered inflation series. The period between the beginning of 1970s and beginning of 1980s is often labelled as a high inflationary period compared to the latter periods. A similar statement holds for economic activity as the real marginal cost series, often used as a proxy for the economic activity, see Galí and Gertler (1999), follows a negative trend which is amplified further in the recent decade. The importance of the joint analysis of such high and low frequency movements in macroeconomic data has recently been documented, see (Ferroni, 2011; Delle Monache and Harvey, 2011; Canova, 2012).

In this paper we aim to contribute to this literature in four ways. We illustrate and discuss possible effects that simple prior filtering of the low frequencies in the data may have on posterior and predictive inference using a basic PC model. The issue is that the observed inflation and marginal cost data have more complex low frequency structures than just a basic mean and/or a basic linear or HP trend. We show that this

[1] For notational convenience we use the abbreviation 'PC' instead of the common abbreviation of the New Keynesian Phillips Curve models.

misspecification affects posterior inference of the structural PC parameters and gives poor forecasting results depending on the model specification. In the Appendix A and Appendix H, we present extensive evidence using a set of simulated and real data and a range of PC model structures. Obviously, in well specified models and in series with simple means and linear trends the misspecification effects are not severe. However, from this outset the use of mechanical filters without properly examining the frequency features of the data is not advisable.

We extend the basic PC model by specifying structural time series models which allow for stochastic trends, structural breaks and stochastic volatility in the inflation and log marginal cost series and integrate these with the basic model. The more complex model structure enables the identification of the relation between macroeconomic variables inherent in the PC models, together with possible long and short run dynamics in each series.

Next, we enrich the extended PC models to include both forward and backward looking expectation mechanisms. There is a debate in the literature on the relative weights of these two components in explaining and forecasting inflation patterns in the U.S.. Our combined model structure can provide valuable inferential information on that point.

As a final contribution we make use of survey data on inflation expectations from University of Michigan Research Center, which provide quarterly one year ahead inflation expectations. It is well known that the class of PC models including complex time series features and rich expectation mechanisms is not easy to estimate given the usually weak data information. The proposed richer expectational mechanisms, making use of the survey data, strengthen the likelihood information and are expected to make inference more efficient and forecasting more accurate.

Several alternatives to structural time series models for efficiently combining the PC model with explicit low frequency movements in the data are available. One alternative is to focus only on the high frequencies rewriting the likelihood in frequency domain and maximizing the likelihood only over a portion of fluctuations, see e.g. Christiano and Vigfusson (2003). Another alternative is to utilize multiple prior filters, to capture the possibly incorrectly specified low frequency components (Canova and Ferroni, 2011). Here we aim to focus on explicitly modeling the low frequency movements to improve the predictive performances of the structural form models while we keep the theoretical model at a simple tractable level.

We apply the proposed set of models to quarterly U.S. data over the period 1960-I until 2012-I. For all models considered, posterior and predictive results are obtained using a simulation based Bayesian approach. Our results indicate that PC structures with three additional components (structural time series features, expectational mechanisms and inflation survey data) capture time variation in the low and high frequency movements of both inflation and marginal cost data. For the inflation series, the model identifies two distinct periods with different inflation levels. In terms of the marginal cost series, the trend specification accommodates the smoothly changing trend observed in the series, specifically after 2000. We also find improved forecasting performance of the extended PC model with three extra components included when this one is compared with basic PC models with demeaned and/or detrended data, with a standard stochastic volatility model proposed by Stock and Watson (2007) and, further, with an extended Bayesian vector autoregressive model which accounts for changing levels, trends and volatility in the data. The model comparison is based on predictive likelihood and out-of-sample Mean Squared Forecast Error (MSFE) comparisons. The Bayesian approach we adopt has several appealing features particularly for the models considered. In terms of inflation predictions, several measures of interest, such as disinflation probabilities obtained from the lower tail of the complete predictive densities, are obtained as a by-product of simulation based Bayesian inference. Furthermore, for the models with general trend and level structures, the non-existence of a stable long-run relationship between inflation and marginal cost series, can be easily assessed using the posterior draws of the trends and levels.

The structure of this paper is as follows: Section 2 presents the three extensions to the standard Phillips curve model structure. Section 3 summarizes the likelihood, prior and the posterior sampling algorithm. Section 4 provides the application of the proposed models and the standard PC model on U.S. inflation and marginal cost data. Section 5 concludes.

2 Extended Philips Curve models

We start with a standard PC model based on a priori filtered data. Next, we extend this model with a structural time series model in order to deal with low and high frequencies that are present in U.S. inflation and the low frequency property in the U.S. log marginal cost series. Thirdly, we extend the latter PC model by introducing a Hybrid PC model (HPC) with both backward and forward looking inflation expectations where the long-run expectations are anchored around observed values of inflation expectations obtained from survey data.

The standard PC can be derived by the approximation of the equilibrium conditions of the firms under staggered price setting using the Calvo formulation, (Calvo, 1983). The basic PC model derived from the firm's price setting is given as

$$\begin{aligned} \tilde{\pi}_t &= \lambda \tilde{z}_t + \gamma_f E_t(\tilde{\pi}_{t+1}) + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{1}$$

where $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID\,(0, \Sigma)$ and standard stationary restrictions hold for (ϕ_1, ϕ_2). Given the AR(2) dynamics for the steady state deviation of the marginal cost, the model can be solved for the inflation expectations by iterating the model forward. This implies that the entire stream of future inflation expectations is taken into account. The PC model takes the form of an instrumental variable model with two instruments and nonlinear parameters in the inflation equation[2]

$$
\begin{aligned}
\tilde{\pi}_t &= \tfrac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\tilde{z}_t + \tfrac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\tilde{z}_{t-1} + \epsilon_{1,t} \\
\tilde{z}_t &= \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + \epsilon_{2,t}.
\end{aligned}
\qquad (2)
$$

One way to estimate the structural parameters is to estimate the unrestricted reduced form model, and solve for the structural parameters, see the Appendix B and Kleibergen and Mavroeidis (2011) for details. However, this transformation, involving a complex Jacobian determinant, may seriously obscure the inference on the structural parameters. Hence we opt for estimating structural parameters directly.

Extended PC models: low frequency components, non-filtered data

We depart from the standard PC model by avoiding the a priori data filtering and emphasize that data filtering is an integral part of data modeling from an econometric point of view. Specifically, we make use of models with time varying levels as well as volatility for capturing both the low and high frequency changes in the U.S. inflation and marginal cost series. Furthermore, modeling data filters together with other model parameters concerns the uncertainty related to steady state specifications. Modeling the data filters explicitly incorporates this uncertainty into the model while the use of filtered data does not. In the latter case, levels and trends are assumed to be known prior to analysis. Finally, prior data filtering also has important effects on the predictive performance of the models as we will show in section 4.

From an economic point of view, Ascari (2004) among others, analyzes the implications of a (constant) trend inflation on the PC structure. Adding a trend inflation to standard PC assumptions, Ascari and Ropele (2007) show that the resulting PC coefficients depend on the trend inflation, thus the interpretation of the coefficients differs from the standard model. Schorfheide (2005) develops a DSGE model along the lines of Woodford (2003) and focuses on agents' learning of the discrete changes in

[2]The model in (2) can be written as a triangular simultaneous equations model: $\left(\begin{smallmatrix} 1 & -\alpha_1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} \pi_t \\ z_t \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & -\alpha_1 & -\alpha_2 & 0 \\ 0 & 1 & -\phi_1 & -\phi_2 \end{smallmatrix}\right)\left(\begin{smallmatrix} c_{\pi,t}, & c_{z,t}, & c_{z,t-1}, & c_{z,t-2} \end{smallmatrix}\right)' + \left(\begin{smallmatrix} 0 & \alpha_2 \\ 0 & \phi_1 \end{smallmatrix}\right)\left(\begin{smallmatrix} \pi_{t-1} \\ z_{t-1} \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 & 0 \\ 0 & \phi_2 \end{smallmatrix}\right)\left(\begin{smallmatrix} \pi_{t-2} \\ z_{t-2} \end{smallmatrix}\right) + \left(\begin{smallmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{smallmatrix}\right)$, and the following parameter restrictions hold: $\alpha_1 = \lambda/\left(1-(\phi_1+\phi_2\gamma_f)\gamma_f\right)$ and $\alpha_2 = \phi_2\gamma_f\alpha_1$.

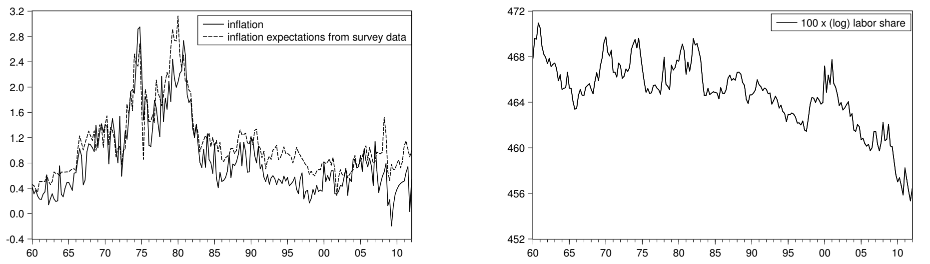
inflation target of the central bank under trend inflation. Cogley and Sbordone (2008) take one-step further and derive the PC model with time-varying trend inflation modeled as a driftless random walk. In these settings, the coefficients of the PC remain constant since a fraction of the firms adjust their price by the steady state inflation rate. Moreover, Nason and Smith (2008) provides empirical evidence in favor of stable structural parameters. In our extended PC models with non-filtered data we follow this assumption and keep the structural parameters constant focusing on low and high frequency movements in steady state levels.

The proposed joint modeling of data filters and other model parameters are also motivated by the stylized facts regarding the non-filtered U.S. inflation and log marginal cost data, shown in Figure 1 over the period between 1960-I and 2012-II[3]. The left panel displays two stylized facts. First, there exist distinct periods with differing patterns for the inflation series. The period between the beginning of the 1970s and the beginning of the 1980s can be labelled as a high inflationary period compared to the remaining periods. Existing evidence shows that the decline in level and volatility is due to credible monetary policy that stabilized inflationary expectations since the early eighties, see McConnell and Perez-Quiros (2000) and Stock and Watson (2007). We observe a temporary increase in the level of inflation during 1970s, while this increase in inflation switches back to the earlier levels after the second break in the first quarter of 1983.One way to model this changing behavior of the series is to allow for regime changes in parameters to capture the change in the structure of the series, see Sims and Zha (2006); Cogley and Sbordone (2008), among others. We consider two cases for the inflation process. In the first case, we assume continuous level shifts and we can model the changing inflation level using a random walk process as

$$c_{\pi,t+1} = c_{\pi,t} + \eta_{1,t+1}, \quad \eta_{1,t} \sim NID(0, \sigma^2_{\eta_1}). \tag{3}$$

[3]Inflation is computed as the continuously compounded growth rate of the implicit GDP deflator and for the real marginal cost series we use labor share in non-farm business sector obtained from http://research.stlouisfed.org/fred2/, see Galí and Gertler (1999) for details. The right panel in Figure 1 displays real marginal cost series, in natural logarithms and multiplied by 100.

Figure 1: Inflation, inflation expectations and log real marginal cost ($\times 100$) series over first quarter of 1960 and the first quarter 2012



Alternatively, we consider an inflation level subject to occasional and discrete shifts, allowing for permanent level shifts. Such level shifts are modeled as follows

$$c_{\pi,t+1} = c_{\pi,t} + \kappa_t \eta_{1,t+1}, \qquad \eta_{1,t} \sim NID(0, \sigma^2_{\eta_1}) \tag{4}$$

where κ_t is a binary variable taking the value of 1 with probability p_κ if there is level shift and the value 0 with probability $1 - p_\kappa$ if the level does not change. This model structure allows for level shifts depending on p_κ while preserving a parsimonious model structure with only a single additional parameter. Occasional and large level shifts correspond to low values of p_κ together with high values of σ_{η_1}. When p_κ is 1, the model becomes a local level model in (3). We use both specifications (3) and (4) in the empirical analysis. The attractive feature of this specification is that the implications on the resulting model is identical to the pure random walk case as the expectation of the future inflation levels is same as the current level, while it still allows for regime changes that are permanent until the next regime change.

The real marginal cost series, shown in the right panel of Figure 1, does not exhibit discrete changes as the inflation series. This data instead has a continuously changing pattern around a negative trend, which can be attributed to technology shocks. Since this trend is more prominent in the second half of the sample period, we allow for a changing trend using a local linear trend specification

$$\begin{aligned} c_{z,t+1} &= \mu_{z,t} + c_{z,t} + \eta_{2,t+1}, \quad \eta_{2,t} \sim NID(0, \sigma^2_{\eta_2}) \\ \mu_{z,t+1} &= \mu_{z,t} + \eta_{3,t+1}, \qquad\qquad \eta_{3,t} \sim NID(0, \sigma^2_{\eta_3}). \end{aligned} \tag{5}$$

This specification is flexible enough to encompass many types of filters used for detrending, see Delle Monache and Harvey (2011), see also Canova (2012) for a similar specification in the more general context of DSGE models. When $\sigma^2_{\eta_3} = 0$, the level of the real marginal cost follows a random walk with a drift, μ_z. Additionally, when $\sigma^2_{\eta_2} = 0$, a deterministic trend is obtained. Note that, setting only $\sigma^2_{\eta_2} = 0$ but allowing $\sigma^2_{\eta_3}$ to be positive results in an integrated random walk process which can approximate nonlinear trends including the Hodrick-Prescott (HP) trend.

Together with the level specifications of the inflation and real marginal cost series the PC model in (2) using (4) and (5) takes the following form

$$\begin{aligned} \pi_t - c_{\pi,t} &= \tfrac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f} \left(z_t - c_{z,t}\right) + \tfrac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f} \left(z_{t-1} - c_{z,t-1}\right) + \epsilon_{1,t}, \\ z_t - c_{z,t} &= \phi_1 \left(z_{t-1} - c_{z,t-1}\right) + \phi_2 \left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}, \\ c_{\pi,t+1} &= c_{\pi,t} + \kappa_t \eta_{1,t+1}, \\ c_{z,t+1} &= \mu_{z,t} + c_{z,t} + \eta_{2,t+1}, \\ \mu_{z,t+1} &= \mu_{z,t} + \eta_{3,t+1}, \end{aligned} \tag{6}$$

where $(\epsilon_{1,t},\ \epsilon_{2,t})' \sim NID\left(0, \begin{pmatrix} \sigma^2_{\epsilon_1} & \rho\sigma_{\epsilon_1}\sigma_{\epsilon_2} \\ \rho\sigma_{\epsilon_1}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{pmatrix}\right)$, $(\eta_{1,t},\ \eta_{2,t},\ \eta_{3,t})' \sim NID\left(0, \begin{pmatrix} \sigma^2_{\eta_1} & 0 & 0 \\ 0 & \sigma^2_{\eta_2} & 0 \\ 0 & 0 & \sigma^2_{\eta_3} \end{pmatrix}\right)$ and the residuals $(\epsilon_{1,t},\ \epsilon_{2,t})'$ and $(\eta_{1,t},\ \eta_{2,t},\ \eta_{3,t})'$ are independent for all t.

Adding stochastic volatility as high frequency component

A further refinement in the PC model can be achieved allowing for time variation in residual variances. This extension is particularly appealing for the inflation series, as the variance of this series changes over time substantially, see e.g. Stock and Watson (2007) for a reduced form model with a stochastic volatility component. To extend the PC model with a stochastic volatility process in the inflation shocks, we add the following state equation to the system

$$h_{t+1} = h_t + \eta_{4,t+1},\ \eta_{4,t+1} \sim NID(0, \sigma^2_{\eta 4}), \tag{7}$$

where we specify a time-varying volatility, $\sigma_{\epsilon_{1,t}} = \exp(h_t/2)$, in the first equation in (6). We follow the practice in Stock and Watson (2007) by fixing the value of $\sigma^2_{\eta 4}$ prior to analysis to facilitate inference. We set $\sigma_{\eta 4} = 0.5$, which seems to work well for the U.S. inflation series.

An important estimation challenge in this extended model is the close relation between the changing inflation levels or level shifts and inflation fluctuations in the extended models. Changing data patterns can be captured by either of these model components, which makes it hard to identify these components unless one makes restrictions as we performed in our analysis. For this reason, we fix the value of $\sigma^2_{\eta 4}$ prior to analysis to facilitate inference and to impose smoothness in the volatility process. It is straightforward to extend the model so that parameter $\sigma^2_{\eta 4}$ is estimated together with the rest of the parameters. The estimation, however, is not trivial since the stochastic volatility component can capture all inflation behavior unless strong priors are imposed on this parameter.

Hybrid PC: forward and backward expectations using survey data

The PC model structure only allows for forward looking inflation expectations while the 'Hybrid' PC (HPC) model combines both backward and forward looking dynamics by including the first lag of inflation deviation in the model along with forward looking dynamics (Galí and Gertler, 1999). The HPC can be derived using an additional assumption on the firm's behaviour, where a fraction ω of the firms, that are unable to reset their prices, adjust their price by the lagged inflation rate rather than the steady state rate. The HPC model takes the form of

$$\begin{aligned} \tilde{\pi}_t &= \lambda^H \tilde{z}_t + \gamma_f^H E_t(\tilde{\pi}_{t+1}) + \gamma_b^H \tilde{\pi}_{t-1} + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{8}$$

where parameters of the HPC model, indicated by a superscript H are functions of the price stickiness parameter, a discount factor and the fraction of firms with backward looking pricing behavior. Iterating the first equation forward, the HPC implies the triangular simultaneous equations model which is nonlinear in parameters

$$\begin{aligned}
\tilde{\pi}_t &= \tfrac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\tilde{z}_t + \tfrac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\tilde{z}_{t-1} \\
&+ \tfrac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\textstyle\sum_{k=1}^{\infty}\gamma_f^{H,k}E_t(\tilde{\pi}_{t+k}) + \tfrac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\tilde{\pi}_{t-1} + \tfrac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t} \qquad (9)\\
\tilde{z}_t &= \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + \epsilon_{2,t}.
\end{aligned}$$

Unlike the PC solution, this system involves an infinite sum of expectations and a closed form solution only exists under certain assumptions such as rational expectations. Here, we do not follow this practice but model the inflation expectations using survey data instead. Specifically, let $S_t = E_t(\pi_{t+1})$ be the next period inflation expectation. We specify an adaptive rule in the sense that inflation expectations partially adjust to the survey expectations in each period:

$$S_t = \mu_t + \beta(S_{t-1} - \mu_{t-1}) + \eta_{5,t}, \qquad (10)$$

where $|\beta| < 1$. Given this restriction on the range of β, one can solve (10) for S_t and obtain $S_t = \mu_t + \sum_{j=1}^{\infty}\beta^j\eta_{5,t-j}$. This specification allows for the interpretation that expected inflation is equal to the survey values with a measurement error that is specified as an infinite moving average with declining weights. We make use of the inflation expectations data which strengthen the information content of the data.

We emphasize that alternative ways of specifying expectation mechanisms in the proposed set of extended PC models is possible and a topic of great interest. Further, their possible connections to survey based expectations are of equal interest. One possibility is to restrict μ to be the constant long run expectation which serves as an anchor. In that case one has a random walk model when $\beta = 1$. The observed series of survey data on inflation expectations and the posterior evidence on β reported in section 4 do not constitute credible evidence for this case. A more detailed comparison of alternative expectation mechanisms is, however, a topic beyond the scope of the present paper and left for future research.

Note that the model-implied expectation is for GDP inflation while the overlaid data is CPI inflation expectations. For this reason we subtract the average difference between CPI and GDP inflation from the survey data[4]. Finally, we note that the survey data provide four-steps-ahead (one-year) expectations. Assuming constant expectations over the year, we divide the survey data by 4 in order to get a comparable quarterly expectation data.

Specifying inflation expectations as in (10), the HPC model becomes

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \tfrac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_t - c_{z,t}\right) + \tfrac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_{t-1} - c_{z,t-1}\right), \\
&+ \tfrac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\tfrac{\gamma_f^H}{1-\gamma_f^H\beta}(S_t - c_{\pi,t}) + \tfrac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\left(\pi_{t-1} - c_{\pi,t-1}\right) + \tfrac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}.
\end{aligned} \tag{11}$$

Similar to the PC model, we consider three case of the HPC model with different specifications for inflation: (i) continuous level changes; (ii) discrete occasional level changes; and (iii) discrete occasional level changes and stochastic volatility.

3 Bayesian inference

In this section we summarize the prior specifications and the posterior sampling algorithms for the extended PC and HPC models.

Prior specification for parameters and prior predictive likelihood

The extended PC and HPC models contain several additional parameters compared to the standard PC model. We classify the model parameters in five groups, and assign independent priors for each group. The first parameter group includes the

[4]We thank an anonymous referee for pointing this out. Our approach of recalculating the inflation expectations is similar to Del Negro and Schorfheide (2012)

common parameters in the PC and HPC models, $\theta_N = \{\lambda, \gamma_f, \phi_1, \phi_2, \Sigma\}$, in (1). For the structural parameters $\{\lambda, \gamma_f, \phi_1, \phi_2\}$ we define flat priors on restricted regions, which also ensure that the autoregressive parameters, ϕ_1 and ϕ_2, are in the stationary region[5]. The (observation) variance priors are of inverse-Wishart type

$$\begin{aligned} p(\lambda, \gamma_f, \phi_1, \phi_2|\Sigma) &\propto 1 \text{ for } |\lambda| < 1,\ |\gamma_f| < 1,\ |\phi_1| + \phi_2 < 1,\ |\phi_2| < 1, \\ \Sigma &\sim IW(1, 20 \times \tilde{\Sigma}), \end{aligned} \tag{12}$$

where $IW(\nu, \Psi)$ is the inverse Wishart density with scale Ψ and ν degrees of freedom.

Note that the prior specifications of the observation and state covariance matrices are important in this class of models and for the case of macroeconomic data. Since the sample size is typically small, differentiating the short-run variation in series (the observation variances) from the variation in the long-run behavior (the state variation) can be cumbersome (Canova, 2012). For this reason, we impose a data based prior structure on the observation covariance matrices. We first estimate the implied unrestricted reduced form VAR model using demeaned inflation series and (linear) detrended real marginal cost series, and base the observation variance prior on this covariance matrix estimate, $\tilde{\Sigma}$. This specification imposes smoothness for the estimated levels and trends, and ensures that the state errors do not capture all variation in the observed variables. Second, prior distributions for the extra model parameters stemming from the hybrid models, $\theta_H = \{\gamma_b^H, \beta\}$ are defined as uniform priors on restricted regions $|\gamma_b^H| < 1$, $|\beta| < 1$. Third, we define independent inverse-Gamma prior densities for the variances of latent state variables: (3), (4) and (5)

$$\begin{aligned} &\sigma_{\eta_1} \sim IG(20, 20 \times 10^{-2}), \quad \sigma_{\eta_2} \sim IG(20, 20 \times 10^{-3}), \\ &\sigma_{\eta_3} \sim IG(1, 1 \times 10^{-5}), \qquad \sigma_{\eta_5} \sim IG(40, 40 * 10^{-1}), \end{aligned} \tag{13}$$

where $IG(\alpha, \alpha\xi)$ denotes the inverse-gamma distribution with shape $\alpha > 0$ and scale

[5] We experimented with wider truncated uniform densities for the λ and γ_f parameters. The prior truncation does not seem to have a substantial affect on the posterior results.

$\alpha\xi > 0$. Parameter α can be interpreted as the number of prior dummy observations while ξ is the a priori variance of a dummy observation.

Similar to the standard counterparts, the extended PC and HPC models may also suffer from flat likelihood functions. We therefore set only slightly informative priors for the state parameters, such that not all variation in inflation and marginal cost series are captured by the time-varying trends and levels. For example, the number of prior dummy observations for σ_{η_1} and σ_{η_2} is very small compared to the number of observations in the data, limiting the prior's information content. We use slightly informative priors on σ_{η_5}, to ensure that the implied inflation expectations of the models do not diverge substantially from the survey expectations.

The fourth prior distribution we consider is applicable to the PC and HPC models with level shifts. For these models, we consider a fixed level shift probability of 0.04. This choice leads to an a priori expected number of shifts of 8 for 200 observations in the sample. Alternatively, we could also estimate this parameter together with other model parameters. However, often the limited level shift observations plague the inference of this parameter. Hence, we set this value, obtained trough an extensive search over intuitive values of this parameter, prior to analysis.

Finally, for the stochastic volatility models, we specify an inverse-gamma prior for the marginal cost variances. For the correlation coefficient, ρ, we take an uninformative prior $p(\rho) \propto (1-\rho^2)^{-3/2}$, see Çakmaklı et al. (2011).

In the proposed models, it is important to assess the effects of the specified prior distributions on the predictive likelihoods of the proposed models. Due to the nonlinear structure of the proposed models, assessing the amount of prior information on the predictive results is not trivial. We present a prior-predictive analysis as in Geweke (2010). For each of the extended PC and HPC models, we consider 1000 parameter values drawn from their joint prior distributions and compute the predictive likelihoods for the data points for the period from the period between 1973-II and 2012-I. Hence a comparison of the resulting prior predictions will indicate which model is preferred by the priors. We provide these results in section 4.

Posterior existence and the sampling algorithm

We summarize the Bayesian inference for the proposed models. An important point regarding the posterior of the structural parameters is the existence of a posterior distribution and its moments, which depends on the number of instruments and the prior. Given two instruments (lagged values of marginal costs), the existence of the posterior distribution is ensured through priors defined on a bounded region although, see Zellner et al. (2012) for a detailed analysis of a linear IV model. Furthermore, due to the small number of instruments, there is a large posterior uncertainty in the PC models, irrespective of the instrument strength.

The MCMC algorithm to sample from the full conditional posterior distributions is based on Gibbs sampling with a Metropolis-Hastings step and data augmentation which combines the methodologies of Geman and Geman (1984); Tanner and Wong (1987); Gerlach et al. (2000); Çakmaklı et al. (2011). Details of the algorithm are provided in Appendices C and D.

4 Posterior and Predictive Evidence

In this section we present posterior and predictive evidence on several features of the extended PC models using U.S. data on inflation and marginal costs. We compare the results with those obtained from alternative reduced form models like Bayesian Vector Auto Regressive (BVAR) models and the stochastic volatility model from Stock and Watson (2007). Specifically, we estimate PC models with a linear trend and HP filter, labeled as PC-LT and PC-HP, respectively. In six PC models we make use of structural time series models to specify low and high frequencies. The first three of these models use the PC framework, allowing for continuous changes in the level of inflation (PC-TV), in addition for discrete occasional level shifts (PC-TV-LS), and in further addition allowing for stochastic volatility for inflation (PC-TV-LS-SV). The final three models use the hybrid form of the HPC framework with forward and backward looking expectations and using survey data. The corresponding extensions are denoted as HPC-TV, HPC-TV-LS and HPC-TV-LS-SV. A summary of the eight models used in this paper is given in Table 1. For a robustness check we ran experiments with several other model specifications and filter methods in order to obtain a smooth transition from a basic PC model with a mechanical filter to an extended PC model with low and high frequencies in levels and volatilities and with rich expectations and survey data. These results, together with a detailed discussion are reported in the Appendix H.

Table 1: Standard and extended Phillips curve models

model structure / low/high frequencies	Phillips curve	Hybrid Phillips curve
linear trend	PC-LT	HPC-LT*
Hodrick-Prescott filter	PC-HP	HPC-HP*
time varying levels	PC-TV (2)-(3)-(5)	HPC-TV (3)-(5)-(11)-(10)
switching and time varying levels	PC-TV-LS (2)-(4)-(5)	HPC-TV-LS (4)-(5)-(11)-(10)
switching, time varying levels and stochastic volatility	PC-TV-LS-SV (2)-(4)-(5)-(7)	HPC-TV-LS-SV (4)-(5)-(7)-(11)-(10)

Note: Results for the models indicated by (*) are provided in Appendix H.

Posterior evidence

We display the estimation results in Table 2 and focus on four features. First, the slope of the PC ($\lambda^{(H)}$) is estimated around 0.07 and 0.09 which is slightly higher than the conventional estimates of the Phillips curve slope, that indicate an almost flat curve (see e.g. Galí and Gertler (1999); Galí et al. (2005); Nason and Smith (2008)). When we model the levels of the series explicitly, $\lambda^{(H)}$ drops to values around 0.05 for both PC and HPC models. A possible explanation for this difference is the departure from the zero steady state inflation assumed in the traditional PC models. As shown in Ascari (2004); Ascari and Ropele (2007) among others, when the firms that cannot re-optimize their prices keep their prices fixed, trend inflation can affect the slope of the PC. In this case, this slope is a decreasing function of the trend inflation. Still, in both PC and HPC models, the estimated slopes are substantially different from zero as point 0 is outside the 95% Highest Posterior Density Interval (HPDI) for most cases.

Table 2: Posterior results of alternative Phillips curve models

Model	$\lambda^{(H)}$	$\gamma_f^{(H)}$	γ_b^H	β	ρ	ϕ_1	ϕ_2
PC-LT	0.068 (0.028)	0.356 (0.242)	–	–	-0.007 (0.018)	0.840 (0.046)	0.076 (0.048)
PC-HP	0.095 (0.050)	0.426 (0.274)	–	–	-0.045 (0.037)	0.659 (0.045)	-0.009 (0.045)
PC-TV	0.059 (0.027)	0.387 (0.250)	–	–	-0.085 (0.057)	0.816 (0.052)	0.064 (0.051)
PC-TV-LS	0.054 (0.023)	0.362 (0.242)	–	–	-0.055 (0.050)	0.824 (0.052)	0.069 (0.052)
PC-TV-LS-SV	0.064 (0.021)	0.320 (0.227)	–	–	-0.018 (0.067)	0.871 (0.050)	0.098 (0.051)
HPC-TV	0.048 (0.025)	0.015 (0.025)	0.384 (0.137)	0.494 (0.276)	0.014 (0.059)	0.811 (0.052)	0.066 (0.050)
HPC-TV-LS	0.035 (0.019)	0.008 (0.009)	0.485 (0.106)	0.520 (0.179)	0.016 (0.012)	0.791 (0.089)	0.187 (0.078)
HPC-TV-LS-SV	0.059 (0.020)	0.040 (0.096)	0.217 (0.115)	0.435 (0.237)	-0.012 (0.005)	0.828 (0.031)	0.147 (0.043)

Note: The table presents posterior means and standard deviations (in parentheses) of parameters for the competing New Keynesian Phillips Curve (PC) type models estimated for quarterly inflation and real marginal cost over over the period from the first quarter of 1960 and the first quarter of 2012. λ (λ^H) and γ_f (γ_f^H) are the slope of the Phillips curve and the coefficient of inflation expectations in PC (HPC) model in (2) ((11)). γ_b^H is the coefficient of the backward looking component in the HPC model in (11). H denotes the parameters of the hybrid models while these parameters without H superscript correspond to the PC model counterparts. β is the autoregressive parameter for the deviation inflation expectations from the long-run trend, as defined in (10). ρ is the correlation coefficient of the residuals ϵ_1 and ϵ_2. ϕ_1 and ϕ_2 are the autoregressive parameters for the real marginal cost specification in (2). Posterior results are based on 40000 simulations of which the first 20000 are discarded for burn-in. Model abbreviations are as in Table 1.

Second, with respect to inflation expectations, it is shown in Table 2 that the coefficient of the short-run inflation expectations, $\gamma_f^{(H)}$ is much lower than the conventional estimates, which is above 0.9 in most of the cases. A potential reason for this finding is the methodology used for inference. Conventional analysis replaces inflation expectations by the real leading value of the inflation relying on the rational expectations hypothesis, see e.g. Galí and Gertler (1999); Sims (2002). However, we opt for explicitly solving for expectations resulting in a highly nonlinear system of simultaneous equations. We also notice a relatively higher posterior standard deviation for this parameter, hence another potential cause of this finding is the relatively low information content in the data about this parameter. This is in accordance with the discussion in the section 3 on the shape of the likelihood in these macro-models. Still more conventional values of this parameter are inside the 95% HPDI.

A striking result from Table 2 is related to the relative importance of the forward and backward looking components of the HPC, measured by parameters γ_f^H and γ_b^H. On the one hand, the evidence in Galí et al. (2005) suggests a dominant forward looking effect. Cogley and Sbordone (2008) document that the forward looking component of the HPC model dominates once the trend variation in inflation is taken into account. Similarly, Benati (2008) shows that, under stable monetary regimes with clearly defined nominal anchors, inflation appears to be (nearly) forward looking. On the other hand, many studies including Fuhrer and Moore (1995); Rudd and Whelan (2005) document a dominant backward looking effects in PC models. Our results favor the latter view since the effect of the backward looking component of inflation estimated by the HPC models in the bottom panel of Table 2 are substantially higher than those of the forward looking components. More specifically, Table 2 shows that the HPC and PC model results differ in terms of the forward looking components' coefficient $\gamma_f^{(H)}$. From an economic point of view, these results maybe driven by the model assumptions on firm behavior that differs from those of Cogley and Sbordone (2008) and Benati (2008). As argued before from an econometric point of view, the difference can stem from the weak data information, see Nason and Smith (2008) for further empirical results and a discussion on this topic.

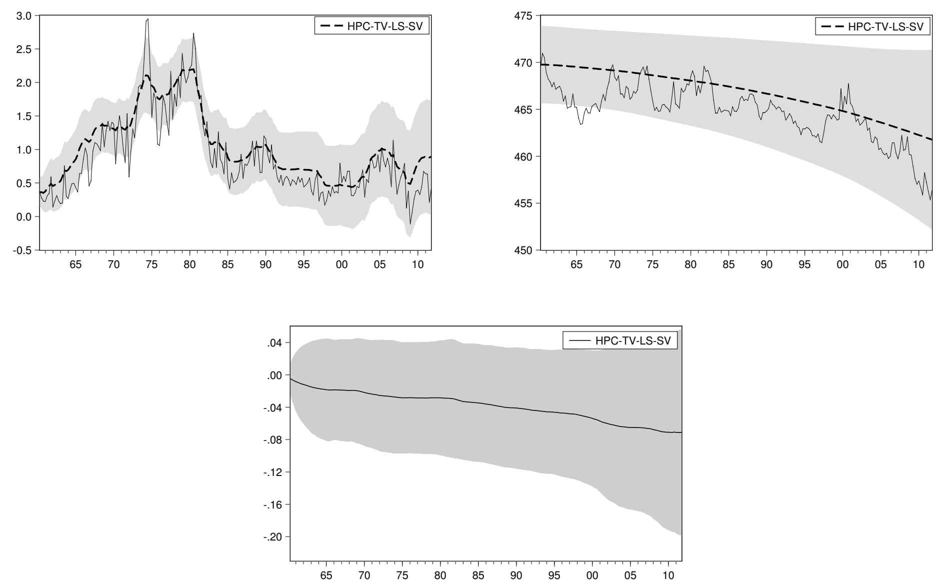
Third, the contemporaneous correlation between the observation disturbances determines the degree of endogeneity of the log real marginal cost in the PC. The estimates of this correlation parameter ρ are displayed in the fifth column of Table 2. Posterior means of ρ from all PC models are negative and close to 0, with high standard deviations. Consequently, 0 is inside the 95% HPDI. For the HPC models, posterior means of ρ are mostly positive with an even smaller magnitude. Therefore, the endogeneity problem does not seem to be severe and single equation inference may yield credible results for inflation and marginal costs. Still, we refrain from doing so since one neglects several cross-equations restrictions in that case.

A further consideration is the β parameter, which shows the persistence in measurement errors in survey inflation expectations. Posterior means of the β are given in the fifth column of Table 2. All HPC models indicate a mediocre persistence, as the posterior means are around 0.4, which implies that measurement errors in inflation expectations are systematic, albeit to a limited extend.

Estimated Levels, Volatilities, Breaks and Inflation Expectations

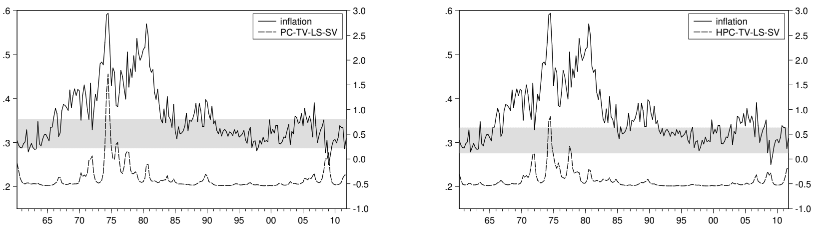
We present estimated levels, trends, inflation volatilities and break probabilities for the proposed HPC models in Figures 2, 3 and 4, respectively. Estimates for the PC counterparts are similar, and are provided in Appendix C.

Figure 2: Level, trend and slope estimates from the HPC-TV-LS-SV model



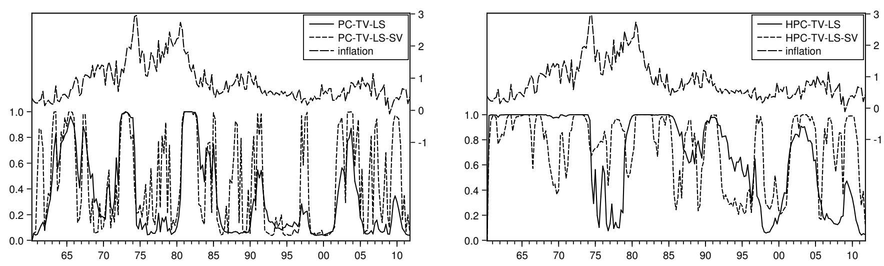
Note: The top-left panel exhibits estimated inflation levels. The top-right and bottom panels show estimated real marginal cost levels and the slopes of the levels, respectively. Grey shaded areas correspond to the 95% HPDI. Model abbreviations are as in Table 1. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Figure 3: Estimated inflation volatility from the (H)PC-TV-LS-SV models



Note: The dashed and solid lines show the posterior mean of the time varying inflation volatility and the observed inflation level. The shaded areas are the 90% HPDI of inflation volatility estimated by the equivalent models without the stochastic volatility components. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Figure 4: Estimated level shift probabilities for the PC and HPC models



Note: The solid and long-dashed lines are the posterior means of the estimated level shift probabilities from the (H)PC-TV-LS model and the (H)PC-TV-LS models, respectively. The dashed line is the observed inflation level. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

The top-left panel of Figure 2 shows estimated levels for the HPC-TV-LS-SV model. We first stress that models that only allow for discrete and occasional level shifts lead to smoother inflation levels compared to the model that allows for continuous level changes, especially in the second half of the sample period. Detailed results on this issue are provided in Appendix F. Furthermore, the model indicates frequent level shifts with a stable inflationary pattern between these level shifts. In DSGE models, mean inflation is generally connected to the inflation target in the central bank's policy rule. Hence movements in trend inflation reflect to a large extend changes in the monetary policy target (see also Schorfheide (2005); Cogley and Sbordone (2008)). Adding the stochastic volatility component to the model with level shifts cause more frequent discrete changes in the inflation level, possibly reflecting the uncertainty in monetary policy target captured by the changing volatility. Estimated marginal cost levels for the HPC-TV-LS-SV are given in the top-right panel of Figure 2. Marginal cost series follows a slightly nonlinear trend during the sample period.

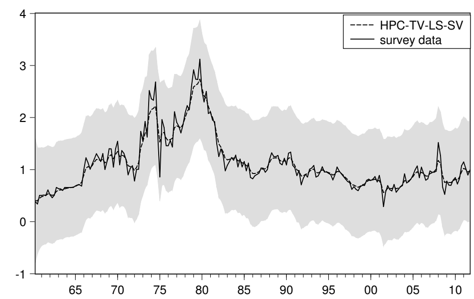
Figure 3 presents estimated volatility levels for the (H)PC model with level shifts and the stochastic volatility component. The stochastic volatility pattern coincides nicely with data features of the Great Moderation, which refers to the decline of the volatility of many U.S. macroeconomic series, see McConnell and Perez-Quiros (2000) among others. The period before the beginning of 1980s is characterized by high inflation levels accompanied by a high volatility, whereas inflation becomes more stable in the second half of the sample period. The decline in inflation volatility after 1980s is linked to credible monetary policy that stabilized inflationary expectations at a low level via commitment to a nominal anchor since the early eighties, see Ahmed et al. (2004); Stock and Watson (2007). The effect of this is also seen in the inflation levels presented in Figure 2. This period of low volatility is replaced by a highly volatile period after 2005 and during the recent financial crisis. A slight difference between PC and HPC models is related to the volatility peaks around 1975. It seems that the high volatility is distributed more evenly in the HPC model with stochastic volatility, whereas for the PC counterpart, high volatility is concentrated around 1975. Finally, the peak points of estimated volatilities coincide with rapid and substantial changes in inflation.

Estimated break probabilities for the PC and HPC models with and without the stochastic volatility component are given in Figure 4. On the one hand, estimated level shift probabilities from the PC-TV-LS model identify four major shifts in the inflation level around 1966, 1973, 1982 and 2005, which comprise the beginning and the end of the high inflationary periods. On the other hand, estimated shift probabilities in the PC-TV-LS-SV model demonstrate the complementarity of level shifts with the changing volatility. The probabilities follow a similar pattern with the PC-TV-LS model, however, the periods subject to level shifts are much longer. During the highly volatile periods of 1970s, the model produces quite clear signals of changing inflation levels, as high volatility levels cause rapid changes in inflation. Accordingly, low volatility periods are characterized by mild changes in inflation, leading to a stable inflation level. Still, for the low volatility periods, mild but significant changes in the inflation level are attributed to level shifts leading to higher level shift probabilities and more clear signals of level shifts.

Both models indicate subsequent level shifts from the beginning of the sample period until 1975, which corresponds to the period during which inflation increased from around 0.20% to around 3%. Unlike the PC model, HPC based models indicate continuous inflation changes during this period. This picture is reversed for the remaining sample period, as the level shift probabilities for both HPC models are considerably smaller. The model with only level shift signals a clear level change in the inflation at the beginning of 1980s, where inflation is subject to a rapid decrease. However, for the period of Great Moderation, the model implies a stable inflationary pattern with moderate signals of level shifts around 1990 and around 2005. As for the PC model with level shifts and stochastic volatility, the periods of level changes indicated by high break probabilities are longer and more clear compared to the counterpart without stochastic volatility. Again, this shows the complementarity of the stochastic volatility component to the level shifts.

Finally, we report implied inflation expectations, computed as the posterior mean of S_t, for the HPC-TV-LS-SV model in Figure 12. The shaded areas around the posterior mean represents the 95% HPDI for the estimated long-term inflation expectations, that track nicely the observed long-term inflation expectations. A noticeable difference between unobserved inflation expectations and the survey data is that the former are smoother than the latter, particularly around 1975 and 1980. In line with the volatility findings, these deviations become considerably smaller during the second half of the sample period. This indicates that inflation expectations are anchored nicely around the observed expectations, with values of parameter β around 0.45 indicating rapid convergence. We note that similar results hold for the remaining HPC models, and these results are included in Appendix F.

Figure 5: Implied inflation expectations by HPC-TV-LS-SV model



Note: The dashed line is the posterior mean of inflation expectations and the solid line is the survey data. Grey shaded areas are the 95% HPDI for estimated expectations. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Predictive Performance

Predictive performances of the models are reported using MSFEs, predictive likelihoods and predictive densities which enable us to report the disinflation probabilities.

The first metric we consider is the predictive likelihoods of all models in order to compare the density forecasts of the models. The one-step ahead predictive likelihood of the observation at $t_0 + 1$, y_{t_0+1}, conditional on the previous observations $y_{1:t_0}$, is given by

$$f(y_{t_0+1}|y_{1:t_0}) = \int p(y_{t_0+1}|X_{t_0+1}, \theta)p(X_{t_0+1}, \theta|y_{1:t_0})dX_{t_0+1}d\theta, \tag{14}$$

which can be computed by first generating $\{X_{t_0+1}\}_{m=1}^{M}$ for M posterior draws, using the corresponding state equations of the models. Next, the predictive likelihood of the observation at $t_0 + 1$ can be approximated by $\frac{1}{M}\sum_{m=1}^{M} p(y_{t_0+1}|X_{t_0+1}^{m}, \theta_{1:t_0}^{m})$, where $p(y_{t_0+1}|X_{t_0+1}^{m}, \theta_{1:t_0}^{m})$ is a multivariate normal density and M is a sufficiently large number.

A feature of the predictive likelihoods is that these can be used to compute the marginal likelihood as $p(y_{t_0+1:T}) = \prod_{t=t_0}^{T} f(y_{t+1}|y_{1:t})$, which provides a tool to analyze the contribution of each observation at time period t to the (log) marginal likelihoods, see Geweke and Amisano (2010). For the models without a priori demeaning and detrending, standard marginal likelihood calculations obtained by integrating out all model parameters using MCMC hold. For the models without a priori demeaning and detrending, standard marginal likelihood calculations obtained by integrating out all model parameters using MCMC hold. For the models with a priori demeaning and detrending, marginal likelihoods can be calculated in two ways. First, all model parameters can be integrated out. The parameter set then includes the mean and trend extracted from data at first place, and uncertainty in these parameters are also taken into account. Second, a priori mean and trend can be taken as 'constants', and the marginal likelihood calculation can be based on integrating out the remaining model parameters. The marginal likelihood in this case is the marginal likelihood of the demeaned and detrended data. The first approach includes the extra parameter uncertainty from a priori parameters, therefore the predictive power is likely to be lower than the one based on the second approach. Furthermore, most of the existing studies with demeaned and detrended data do not take into account the parameter uncertainty arising from this a priori step. We therefore choose the second approach to calculate the marginal likelihoods, which provides a strong alternative to the models we propose in terms of predictive power and is a fair replication of the literature.

Accurate point predictions of inflation is of key importance for economic agents such as investors and central banks. Therefore, we also consider MSFE, computed as the mean of the sum of squares of the prediction errors. For inflation forecasts we use mean of the predictive distribution of inflation, consistent with a quadratic loss function. We consider MSFE for one and four period ahead forecasts in order to examine the forecasting ability of the models also for longer horizons.

As a third performance criteria, we report the disinflation risk indicated by each model. Typically, increased uncertainty about future inflation is penalized by the predictive likelihood comparisons. This uncertainty, however, may simply indicate the increasing inflationary risk. We include this criterion in order to gain insights on the inflationary risk implied by each model. Disinflation probabilities are computed as the tail probability of the predictive distributions such that the one step ahead predicted inflation values are below zero.

Apart from the models we considered so far, we also consider alternative reduced form models that are proven to have superior predictive abilities. The first model we include is the unobserved component model proposed by Stock and Watson (2007), henceforth denoted as SW2007. This model captures the unobserved trend in inflation where both the inflation and the trend volatility follow a stochastic process. We refer to Stock and Watson (2007) for the details of this model. The second model we consider is an unrestricted Bayesian VAR (BVAR-SV) model with two lags and with stochastic volatility for inflation. BVAR models are one of the workhorse models used for forecasting macroeconomic series. For the sake of brevity, we do not provide the details of this model, and refer to Del Negro and Schorfheide (2012). As for the structural models, we use the identical structural time series methods for modeling the level and the trends of the inflation and marginal cost series in the BVAR-SV. Both SW2007 and BVAR-SV models are strong competitors for the extended PC and HPC models we propose. In all considered models, the prior distributions in section 3, calculated using the full sample data, are used.

Marginal likelihoods and MSFE of the alternative models are presented in Table 6. The likelihood contribution of each observation and the corresponding cumulative predictive likelihoods are displayed in Figure 6. We present the (log) marginal likelihood of the competing models in the first column of Table 6. These values together with Figure 6 indicate three groups of models in terms of their predictive performances. The first group of models include BVAR-SV and the conventional PC models with demeaned and detrended data (PC-LT and PC-HP). The second group consists of the PC models with time variation in inflation levels (PC-TV, PC-TV-LS) and the SW2007 model. The models in the second group have much superior performance in terms of the marginal likelihood values. A second increase in the marginal likelihood values can be observed when we consider the models in the third group, namely the HPC models (HPC-TV, HPC-TV-LS, HPC-TV-LS-SV) and the PC model together with discrete level shifts and stochastic volatility for inflation (PC-TV-LS-SV).

A similar clustering of models is observed when we compare the models' performances using the one period ahead MSFE, with the exception of the BVAR-SV model. Unlike the model fit performance, measured by the marginal likelihood values, BVAR-SV model performs considerably better in terms of point prediction.

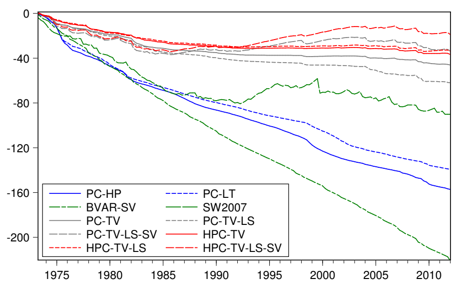
Three main conclusions can be drawn from these findings. First, the conventional PC models with demeaned and detrended data (PC-LT and PC-HP) perform worse than the competing models both in terms of MSFE and in terms of the marginal likelihood metric. However, the difference between HPC and PC models in terms of

Table 3: Predictive performance of Phillips curve models

Model	(Log) Marg. Likelihood	MSFE 1 period ahead	MSFE 4 period ahead
SW2007	-78.033	0.168	0.250
BVAR-SV	-220.710	0.091	0.195
PC-LT	-139.327	0.353	0.358
PC-HP	-157.195	0.458	0.367
PC-TV	-46.162	0.142	0.263
PC-TV-LS	-61.972	0.138	0.247
PC-TV-LS-SV	-33.476	0.126	0.213
HPC-TV	-36.683	0.109	0.220
HPC-TV-LS	-33.913	0.084	0.195
HPC-TV-LS-SV	-18.960	0.102	0.178

Note: The table reports the predictive performances of all competing models for the prediction sample over the second quarter of 1973 and the first quarter of 2012. '(Log) Marg. Likelihood' stands for the natural logarithm of the marginal likelihoods. 'MSFE' stands for the Mean Squared Forecast Error. Marginal likelihood values in the first column are calculated as the sum of the predictive likelihood values in the prediction sample. Results are based on 10000 simulations of which the first 5000 are discarded for burn-in. 'SW2007' stands for the model proposed by Stock and Watson (2007), and 'BVAR-SV' stands for the Bayesian VAR model with time varying levels and trends and a stochastic volatility component for the inflation equation. Remaining abbreviations are as in Table 1.

Figure 6: Predictive likelihoods from competing models



Note: The figure displays the evolution of the (log) predictive likelihoods for the computing models between the third quarter of 1973 and the first quarter of 2012. Model abbreviations are as in Table 1. Results are based on 5000 simulations of which the first 10000 are discarded for burn-in.

point forecasts is less pronounced compared to the increase in precision when switching from models using demeaned and detrended data to the models that use the raw data. This indicates the importance of estimating levels and trends together with the structural model parameters.

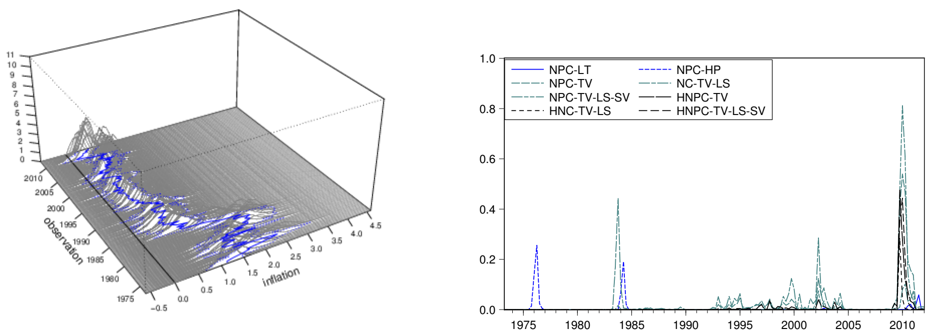
Second, the difference between the PC model with level shifts and stochastic volatility with the remaining PC models is considerably large. The performance of this model is comparable to the HPC models which perform superior both in terms of point forecasts and the model fit. On the one hand, models with level shifts and stochastic volatility deliver the most accurate point predictions considering MSFE and marginal likelihood values. These results pinpoint the importance of incorporating the high and low frequency movements in the structural models. On the other hand, this model performance can be increased further by incorporating the survey data and the backward looking component in the HPC models.

Third, structural models perform at least as well as the strong reduced form candidates, the SW2007 and BVAR-SV models. These findings are crucial since structural models deliver both structural macroeconomic information and predictive performance, whereas the reduced form models are solely designed for improving the predictive performance. Incorporating high and low frequency movements in structural models increase their predictive power substantially while still exploiting the macroeconomic information indicated by economic theory. These findings also hold for four period ahead forecasts, as shown in the last column of Table 6.

We next consider the evolution of the model performance over the forecast sample in detail, shown in Figure 6. An important finding from the figure is the increasing performance of the HPC models and the models with stochastic volatility components after mid 1980s. This period is characterized by a decrease in inflation volatility during the Great Moderation. The stochastic volatility component seems to capture this decrease in volatility accurately. Moreover, the effect of the level shifts can be observed when we compare the PC-TV-LS-SV model with the SW2007 model. Much of the difference in the performance of these models can be attributed to the changes in inflation levels. This shows that the inflation process exhibits rare regime changes and within each regime inflation follows a stable path.

The last metric we use for model comparison considers the implied disinflationary risk. The left panel in Figure 13 shows the entire distribution of the inflation predictions for the HPC-LS-SV model where the levels and trends are estimated together with the structural parameters. The mean predicted inflation is represented by the solid line, and the width of the predictive distribution is indicated by the white area under the inflation density. As expected, inflation predictions are concentrated around high (low) values during the high (low) inflationary periods. The uncertainty around the inflation predictions are also high for these periods, together with the periods when inflation is subject to a transition to low values around 1980s. When the observed inflation values are close to the zero bound, the predictive densities indicate disinflationary risk.

Figure 7: Predicted inflation densities from HPC-LS-SV model and disinflation probabilities implied by different Phillips curve models



Note: The left figure presents one period ahead predictive distribution of inflation from the HPC-LS-SV model, for the period between the third quarter of 1973 and the first quarter of 2012. The right figure presents disinflation probabilities computed using the one period ahead predictive distributions of inflation for the period between the third quarter of 1973 and the first quarter of 2012. Model abbreviation is based on Table 1. Results are based on 5000 simulations of which the first 10000 are discarded for burn-in.

The right panel in Figure 13 displays this disinflationary risk, which is of key importance especially for policy making. The figure shows that PC models with a priori demeaned and detrended data do not signal any pronounced disinflation risk except for the low disinflation probabilities during mid 1970s and mid 1980s. However, extended PC and HPC models exploiting the high and low frequency movements produce clear signals of disinflation risk and disinflationary pressure during the recent recession.

Note that actual disinflation only occurs around 2009 in this sample period and the models signal disinflationary risk slightly later than this period. This result can be explained by the agents' learning process. As indicated in Schorfheide (2005), if agents learn about the monetary policy changes later than the inflation level changes, the perceived target inflation in general equilibrium happens only gradually. In Schorfheide (2005), this is incorporated as Bayesian learning of the agents which is in line with the econometric assumption underlying our models. As the modelled state-space updating incorporates Bayesian learning, the changes in the inflation level occurs gradually and the inflationary risk signals are delayed. Our models are still able to capture this disinflationary pressure quite successfully.

Prior predictive likelihoods of proposed models

Due to the complex model structures in the proposed models, it is important to address the effects of the specified prior distributions on the predictive performances. We therefore perform the prior predictive analysis outlined in section 3 for the extended PC models. Table 4 presents the average and cumulative prior predictive likelihoods for the forecast sample, where we show that the adopted prior distributions clearly favor the less parameterized model, PC-TV. Moreover, the priors clearly do not favor the models with the stochastic volatility components. More importantly, the 'best performing model' according to the predictive results, HPC-TV-LS-SV, is the least favorable one according to the adopted prior distributions. We therefore conclude that data information is dominant, and the superior predictive performance of the HPC-TV-LS-SV model is not driven by the prior distribution.

Robustness analysis

Table 4: Prior-predictive results for the proposed models

Model	Average (Log) Pred. Likelihood	Cumulative (Log) Pred. Likelihood
PC-TV	-1.160	-180.883
PC-TV-LS	-1.361	-210.909
PC-TV-LS-SV	-1.449	-224.661
HPC-TV	-1.277	-199.219
HPC-TV-LS	-1.267	-197.677
HPC-TV-LS-SV	-2.043	-318.771

Note: The table reports the prior-predictive performances of all competing models for the prediction sample over the second quarter of 1973 and the first quarter of 2012. '(Log) Pred. Likelihood' stands for the natural logarithm of the predictive likelihoods. Results are based on 1000 simulations from the joint priors of model parameters. Model abbreviations are as in Table 1.

The proposed models extend the standard models in several ways and the predictive performance comparison in section 4 stems from these extensions jointly. However, it is important to address which extensions in the proposed model provide the largest predictive gains. For this reason, we estimate several alternative models, using which distinct gains from each contribution in the models can be identified. For space limitations, detailed results are provided in Appendix H and here we briefly summarize the main findings.

First, predictive gains solely from including the survey expectations in the models are substantial. Second, incorporating the low and high frequency data movements in the model is crucial. These extensions increase the predictive performances drastically in all models we consider. Finally, once survey data and time variation are included in the model, additive gains from the backward looking component in the hybrid models are negligible in terms of the prediction results. Moreover, the iterative solution of the inflation expectations regulates the posterior distributions of the parameters. We therefore conclude that the superiority of the most extensive model, HPC-TV-LS-SV, is not based on one of these extensions but is rather based on the combination of them.

We conclude this section with a remark on the possible existence of a long run stable relation between inflation and marginal costs. The models we considered so far rely on the implicit assumption of the absence of a long-run cointegrating relationship. We assess whether this assumption is plausible for the U.S. data considering the PC-TV model, and find credible evidence that such a cointegrating relationship is unlikely. Details are provided in Appendix I.

5 Conclusion

Phillips curve models constitute an integral part of macroeconomic models used for forecasting and policy analysis. These models are often estimated after demeaning and/or detrending the data. In this paper it is shown that mechanical removal of the low frequency movements in the data may lead to poor forecasts. Potential structural breaks and level shifts as well as changing volatility in the observed series require more complex models, which can handle these time variation together with the standard PC parameters. We have proposed a set of models where levels and trends of the series together with the volatility process are integrated with a structural PC model. Furthermore, we consider richer expectational mechanisms for the inflation series in enlarged Hybrid-PC models, where the inflation expectations are anchored around the inflation expectations obtained from survey data.

The proposed models capture time variation in the low frequency moments of both inflation and marginal cost data. For the inflation series we identify three distinct periods with high and low inflation. The high inflationary period corresponds to 1970s, following a low inflationary period of 1960s. The last period starting with 1980s is characterized by low inflation levels corresponding to an annual inflation level around 2%. When this model is blended with the stochastic volatility component, the level shifts can be identified even more precisely.

The use of macroeconomic information in the structural models together with the remaining high and low frequency movements in the data improves the predictive ability also compared to celebrated reduced form models, including the Bayesian VAR and the stochastic volatility model (Stock and Watson, 2007). Furthermore, modeling inflation expectations using survey data and adding stochastic volatility to the PC model structure improves in sample fit and out of sample predictive performance substantially. We also analyze the disinflation probabilities indicated by each competing model. The complete predictive densities, most notably from the enlarged models, indicate an increase in the probability of disinflation in the U.S. in recent years.

Modeling forward and backward looking components of inflation has important effects on empirical results. Endogeneity and persistence do not appear to be very important empirical issues in PC model structures. Finally, we analyze the existence of a long-run relation between the low frequency movements of both series. No evidence is found on such a long run stable cointegrating relation for the U.S. series.

Given that incorporating low and high frequency movements explicitly in macroeconomic models provides additional insights for both policy analysis and more accurate predictions, we plan to enlarge the proposed model to a more general DSGE framework in future work. Another interesting possibility of future research is to combine different PC models using their predictive performances, which seems to be time varying.

Appendix

A Effect of misspecified level shifts on posterior estimates of inflation persistence

The linear Backward Looking Phillips Curve (BLPC) captures the relation between real marginal cost $\tilde{z}_t$ and inflation $\tilde{\pi}_t$. We illustrate in this section that model misspecification resulting from ignoring level shifts in inflation data leads to overestimation of persistence in the inflation equation within a linear BLPC.

The linear BLPC model can be written as

$$\begin{aligned} \tilde{\pi}_t &= \lambda \tilde{z}_t + \gamma_b \tilde{\pi}_{t-1} + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{15}$$

with $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID(0, \Sigma)$. This model is a triangular simultaneous equations model and can also be interpreted as an instrumental variable model with two instruments. We specify an AR(2) model for the marginal cost in order to mimic for the cyclical behavior of the observed series, see Basistha and Nelson (2007); Kleibergen and Mavroeidis (2011) for a similar specification. The AR(2) parameters are restricted to the stationary region $|\phi_1| + \phi_2 < 1,\ |\phi_2| < 1$, and the lagged adjustment parameter in the inflation equation is restricted as $0 \leq \gamma_b < 1$. The structural parameter λ is restricted as $0 \leq \lambda < 1$ which is in line with previous evidence on the slope of the BLPC.

Since BLPC in (15) specifies the relation between the *short-run* stationary fluctuations in real marginal cost and inflation, $\tilde{\pi}_t$ and $\tilde{z}_t$ can be interpreted as the *transitory* components of inflation and marginal cost, in deviation from their long-run components. In fact, the observed non-filtered data can be decomposed into permanent and transitory components in a straightforward way as

$$\begin{aligned} \pi_t &= \tilde{\pi}_t + c_{\pi,t}, \\ z_t &= \tilde{z}_t + c_{z,t}, \end{aligned} \tag{16}$$

where π_t and z_t are the inflation and marginal cost data, respectively, and $c_{\pi,t}$ and $c_{z,t}$ are the permanent components of the series.

In our simulation experiment, we model the steady state inflation as a constant level subject to regime shifts that mimic the high inflationary period during the 1970s. For modelling the permanent component of the real marginal cost series, we use a linear negative trend in order to mimic the declining real marginal cost levels in the U.S. over the sample starting from the 1960s. This specification can be formulated as follows

$$\begin{aligned} c_{\pi,t} &= c_{\pi,t-1} + \kappa_t \eta_{t-1}, \;\; c_{z,t} = c_{z,t-1} + \mu_{z,t-1}, \\ \mu_{z,t} &= \mu_{z,t-1}, \;\; \eta_t \sim NID(0, \omega^2), \end{aligned} \tag{17}$$

where κ_t is a binary variable indicating a level shift in the level series, $c_{\pi,t}$ and $c_{z,t}$ indicate the level value of inflation and real marginal cost, respectively, in period t and $\mu_{z,t}$ is the slope of the trend in the real marginal cost series. By excluding the stochastic component for the slope and the trend of the real marginal cost in (17), we specify a deterministic trend for this series.

We simulate three sets of data from the model in (15)–(17). For the first set, the inflation series show no level shifts, i.e. $\kappa_t = 0, \;\; \forall t$. For the other two sets of data, we impose different level shifts with moderate ($\omega^2 = 2.5$) and large ($\omega^2 = 5$) changes in the level values, respectively. For each specification we simulate 100 datasets with $T = 200$ observations, where two level shifts occur in periods $t = 50$ and $t = 150$. The observation error variance is set to $\left(\begin{smallmatrix} 1 & 0.01 \\ 0.01 & 0.01 \end{smallmatrix}\right)$, which leads to a correlation of 0.1 between the disturbances, and parameter λ is set to 0.1. Note that parameters $\phi_1 = 0.1$ and $\phi_2 = 0.5$ are chosen such that the transitory component of the series is stationary.

In order to capture the effect of model misspecification on posterior inference, when computing the transitory component, we ignore level shifts in the simulated inflation series and simply demean the series. For the marginal cost series, we remove the linear trend prior to the analysis and only focus on the effect of misspecification in the inflation series. This implies that for the simulated data with no level shifts, the model is correctly specified and the posterior results should be close to the true values. For each simulated data set we estimate the model in (15) using flat priors on restricted parameter regions:

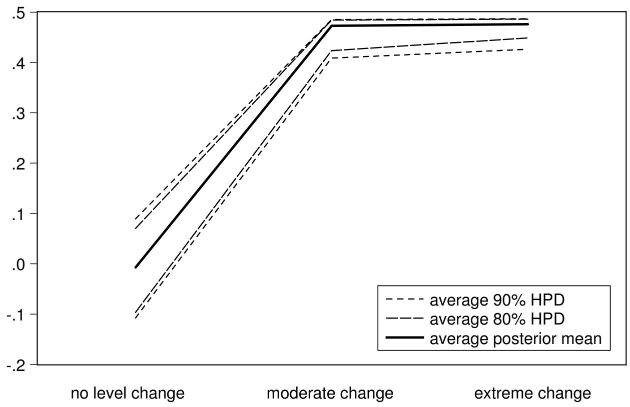
$$p(\phi_1, \phi_2, \gamma_b, \lambda) \propto \begin{cases} 1, & \text{if } |\phi_1| + \phi_2 < 1, \ |\phi_2| < 1, \ 0 \leq \gamma_b < 1, \ 0 \leq \lambda < 1 \\ 0, & \text{otherwise} \end{cases}. \tag{18}$$

Given that model (15) is equivalent to an instrumental variables model with 2 instruments, it can be shown that the likelihood function for such a model combined with the flat prior on a large space yields a posterior distribution that exists but it has no first or higher moments. Due to the bounded region condition on the parameters, where the structural parameter λ is restricted to the unit interval, all moments exist. For details, we refer to Zellner, Ando, Baştürk, Hoogerheide and Van Dijk (2012). We mention this existence result since it provides an econometric explanation why it is often difficult to estimate a structural model for macro-economic data such as (15). Indeed, the rather flat posterior surface plagues the inference, in particular, when ϕ_2 is close to zero. Posterior moments are in our case computed by means of standard Metropolis-Hastings method on ϕ_1 and ϕ_2 and λ and γ_b. Other Monte Carlo methods like Gibbs sampling are also feasible in this case.

Figure 8 presents the overestimation results from 100 different simulations for each setting we consider. We report the average overestimation in posterior γ_b estimates and 95% highest posterior density intervals (HPDI) intervals for this overestimation.

The persistence parameter γ_b is overestimated in all cases except for the correctly

Figure 8: Overestimation illustration for the backward looking PC model



Note: The figure presents overestimation probability of parameter γ_b for simulated data from the BLPC model with different structural breaks structures. We report average quantiles of overestimation based on 100 simulation replications for each parameter setting.

specified model. The degree of overestimation becomes larger with a larger shift in the level of inflation. Note that the average 95% HPDI of overestimation becomes tighter for data with extreme changes in levels. Hence the effect of model misspecification on the persistence estimates is more pronounced if the regime shifts are extreme.

In summary, our simulation experiments using BLPC show that when the shifts in the inflation level are not modelled, inference on model persistence parameters may be severely biased due to the model misspecification. This will also hold for predictive estimates.

We note that we focused on misspecification effects on persistence measures when level shifts in the series are ignored. Similar experiments can be set up for the BLPC with weak identification (or weak instruments) by setting $\phi_2 \approx 0$. The effect of misspecification on posterior and predictive estimates in the case of weak identification is a topic outside the scope of the present paper. We refer to Kleibergen and Mavroeidis (2011) for details on Bayesian estimation in case of weak identification.

B Structural and reduced form inference of the PC model

This section presents the unrestricted reduced form inference (URF) of the PC model, and the inference of the corresponding structural form (SF) model parameters. We show that the posterior draws from the structural form parameters can be obtained using the reduced form representation of (1):

$$\begin{aligned} \tilde{\pi}_t &= \alpha_1 \tilde{z}_{t-1} + \alpha_2 \tilde{z}_{t-2} + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{19}$$

where $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID\,(0, \Sigma)$, and the restricted reduced form (RRF) representation is obtained by introducing the following restrictions on parameters in (1):

$$\alpha_1 = \tfrac{\lambda(\phi_1+\gamma_f\phi_2)}{1-\gamma_f(\phi_1+\gamma_f\phi_2)}, \;\; \alpha_2 = \tfrac{\lambda\phi_2}{1-\gamma_f(\phi_1+\gamma_f\phi_2)}. \tag{20}$$

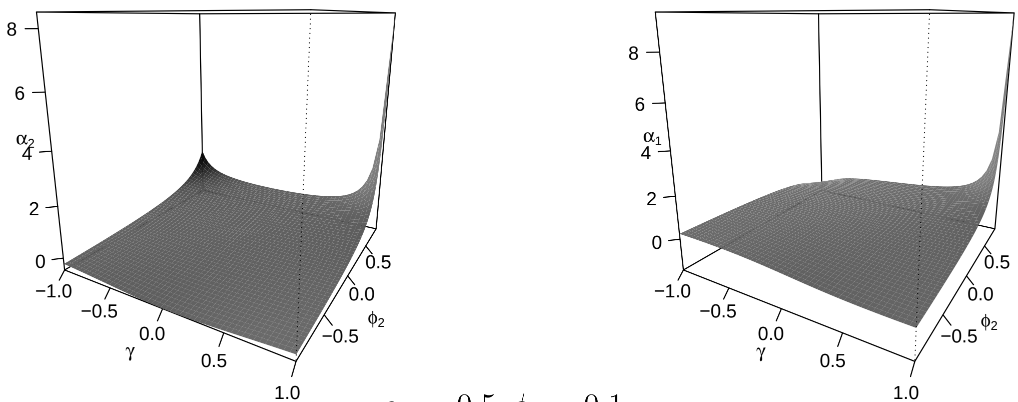
Finally, the model in (1) is related to an Instrumental Variables (IV) model with exact identification. Bayesian estimation of the unrestricted reduced form model in (19) is straightforward under flat or conjugate priors. Given the posterior draws of reduced form parameters, posterior draws of structural form parameters in (1) can be obtained using the transformation in (20). This nonlinear transformation, however, causes difficulties in setting the priors in an adequate way. The determinant of the Jacobian of this nonlinear transformation is $\mid J \mid = \frac{\lambda\phi_2^2}{(1-\gamma(\phi_1+\gamma\phi_2))^2}$, where the Jacobian is non-zero and finite if $\gamma(\phi_1 + \gamma\phi_2) \neq 1$, $\phi_2 \neq 0$ and $\lambda \neq 0$.[6]

Figure 9 illustrates the nonlinear transformation for the SF and RRF representations, for a grid of parameter values from SF representations, and plot the corresponding RRF parameter values, and vice versa. The top panel in Figure 9 shows the transfor-

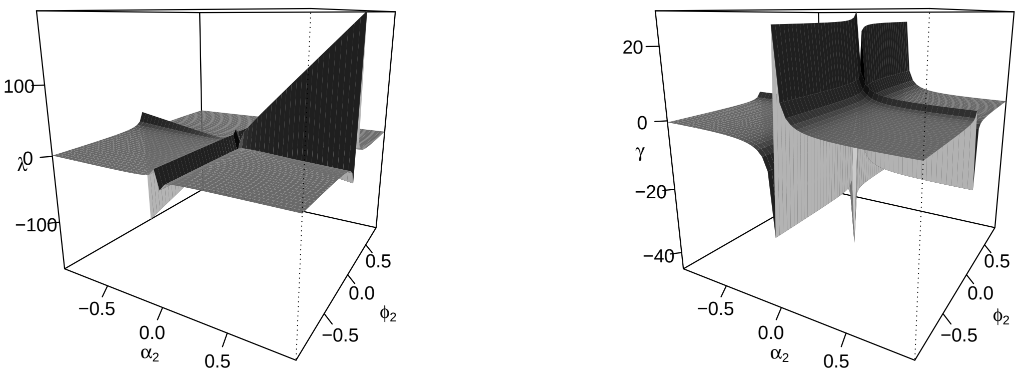
[6]We only consider the transformation from $\{\lambda, \gamma, \phi_1, \phi_2\}$ to $\{\alpha_1, \alpha_2, \phi_1, \phi_2\}$, i.e. variance parameters in the transformed model are left as free parameters.

mations from SF to RRF. Reduced form parameters α_1 and α_2 tend to infinity when persistence in inflation and marginal cost series are high, i.e. when the structural form parameters λ and $\phi_1 + \phi_2$ tend to 1. The bottom panel in Figure 9 shows the RRF to SF transformations. The corresponding SF parameters lead to an irregular shape, for example, when the instrument z_{t-2} has no explanatory power with $\phi_2 = 0$ or when $\alpha_2 = 0$.

Figure 9: Nonlinear parameter transformations



$\alpha_1 = 0.5, \phi_1 = 0.1$



$\lambda = 0.5, \phi_1 = 0.1$

Note: The top panel presents the implied unrestricted reduced form parameters in (19) given structural form parameters in (1). The top panel presents implied structural form parameters in (1) given unrestricted reduced form parameters in (19). Parameter transformations are obtained using the RRF restrictions in (20).

C Bayesian inference of the extended PC model

This section presents the MCMC scheme for the posterior inference of the PC model. Specifically, we use a Gibbs sampler together with data augmentation (see Geman and Geman, 1984; Tanner and Wong, 1987).

The PC model in (6) can be cast into the state-space form as follows

$$\begin{aligned} Y_t &= HX_t + BU_t + \epsilon_{\mathbf{t}}, \quad \epsilon_{\mathbf{t}} \sim N(0, Q_t) \\ X_t &= FX_{t-1} + R_t\eta_{\mathbf{t}}, \qquad \eta_{\mathbf{t}} \sim N(0, I) \end{aligned} \tag{21}$$

where

$$Y_t = \begin{pmatrix} \pi_t \\ z_t \end{pmatrix}, \quad X_t = \begin{pmatrix} c_{\pi,t}, & c_{z,t}, & \mu_{z,t}, & c_{z,t-1}, & c_{z,t-2} \end{pmatrix}', \quad U_t = \begin{pmatrix} z_t \\ z_{t-1} \\ z_{t-2} \end{pmatrix}, \quad \epsilon_{\mathbf{t}} = \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix},$$

$$H = \begin{pmatrix} 1 & -\alpha_1 & 0 & -\alpha_2 & 0 \\ 0 & 1 & 0 & -\phi_1 & -\phi_2 \end{pmatrix}, \quad B = \begin{pmatrix} \alpha_1 & \alpha_2 & 0 \\ 0 & \phi_1 & \phi_2 \end{pmatrix}, \quad Q_t = \begin{pmatrix} \sigma^2_{\epsilon_1,t} & \rho\sigma_{\epsilon_1,t}\sigma_{\epsilon_2} \\ \rho\sigma_{\epsilon_1,t}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{pmatrix},$$

$$F = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad R_t = \begin{pmatrix} \kappa_t\sigma_{\eta_1} & 0 & 0 \\ 0 & \sigma_{\eta_2} & 0 \\ 0 & 0 & \sigma_{\eta_3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \eta_{\mathbf{t}} = \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix},$$

where $\alpha_1 = \frac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}$ and $\alpha_2 = \frac{\lambda\gamma\phi_2}{1-(\phi_1+\phi_2\gamma)\gamma}$.

Once the state-space form of the model is set as in (21) standard inference techniques in state-space models can be carried out. Let $Y_{1:T} = (Y_1, Y_2, \ldots, Y_T)'$, $X_{1:T} = (X_1, X_2, \ldots, X_T)'$, $U_{1:T} = (U_1, U_2, \ldots, U_T)'$, $\sigma^2_{\epsilon_1,1:T} = (\sigma^2_{\epsilon_1,1}, \sigma^2_{\epsilon_1,2}, \ldots, \sigma^2_{\epsilon_1,T})'$ and $\theta = (\phi_1, \phi_2, \gamma_f, \lambda)'$. For the most general PC model with level shifts and stochastic volatility, the simulation scheme is as follows

1. Initialize the parameters by drawing κ_t using the prior for κ and unobserved states X_t, h_t for $t = 1, 2, \ldots, T$ from standard normal distribution and conditional on κ_t for $t = 0, 1, \ldots, T$. Initialize $m = 1$.

2. Sample $\theta^{(m)}$ from $p(\theta | Y_{1:T}, X_{1:T}, U_{1:T}, R_{1:T}, Q_{1:T})$.

3. Sample $X_t^{(m)}$ from $p(X_t | \theta^{(m)}, Y_{1:T}, h_{1:T}, U_{1:T}, R_{1:T}, Q_{1:T})$ for $t = 1, 2, \ldots, T$.

4. Sample $h_t^{(m)}$ from $p(h_t | X_{1:T}^{(m)}, \theta^{(m)}, Y_{1:T}, X_{1:T}, U_{1:T}, R_{1:T}, \rho^{m-1}, \sigma_{\epsilon_2}^{2,(m-1)}, \sigma_{\eta_4}^{2,(m-1)})$ for $t = 1, 2, \ldots, T$.

5. Sample $\kappa_t^{(m)}$ from $p(\kappa^{(m)} | \theta^{(m)}, Y_{1:T}, h_{1:T}, U_{1:T}, R_{1:T}, Q_{1:T})$ for $t = 1, 2, \ldots, T$.

6. Sample $\sigma_{\eta_i}^{2,(m)}$ from $p(\sigma_{\eta_i}^{2,(m)} | X_{1:T}^{(m)}, h_{1:T}^{(m)}, \kappa_{1:T}^{(m)})$ for $i = 1, 2, 3, 4$.

7. Sample $\rho^{(m)}$ from from $p(\rho^{(m)} | X_{1:T}^{(m)}, h_{1:T}^{(m)}, Y_{1:T}, X_{1:T}, U_{1:T}, \theta^{(m)}, \sigma_{\epsilon_2}^{2,(m-1)})$.

8. Sample $\sigma_{\epsilon_2}^{2,(m)}$ from from $p(\sigma_{\epsilon_2}^{2,(m)} | \rho^{(m)}, X_{1:T}^{(m)}, h_{1:T}^{(m)}, Y_{1:T}, X_{1:T}, U_{1:T}, \theta^{(m)})$.

9. Set $m = m + 1$, repeat (2)-(9) until $m = M$.

Steps (3)-(5) are common to many models in the Bayesian state-space framework, see for example Kim and Nelson (1999); Gerlach et al. (2000); Çakmaklı (2012). Note that parameter p_κ is set a priori using heuristics.

Sampling of θ

Conditional on the states $c_{\pi,t}, c_{z,t}$ and h_t for $t = 1, 2, \ldots, T$, redefining the variables such that $\tilde{\pi}_t = \pi_t - c_{\pi,t}$, $\tilde{z}_t = z_t - c_{z,t}$ and $\varepsilon_t = \epsilon_t / \exp(h_t/2)$, the measurement equation in (21) can be rewritten as

$$\begin{aligned} \tilde{\pi}_t &= \tfrac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f} \tilde{z}_t + \tfrac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f} \tilde{z}_{t-1} + \varepsilon_t \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}. \end{aligned} \tag{22}$$

Posterior distributions of the structural parameters under flat priors are non-standard since z_t term also is on the right hand side of (22) and the model is highly non-linear in parameters. We therefore use two Metropolis Hastings steps to sample these structural parameters (Metropolis et al., 1953; Hastings, 1970). For sampling ϕ_1, ϕ_2 conditional on λ, γ_f and other model parameters, the candidate density is a multivariate student-t density on the stationary region with a mode and scale with the posterior mode and scale using only the second equation in (22) and 1 degrees of freedom. For sampling λ, γ_f conditional on ϕ_1, ϕ_2 and other model parameters, the candidate is a uniform density.

Sampling of states, X_t

Conditional on the remaining model parameters, drawing $X_{0:T}$ can be implemented using standard Bayesian inference. This constitutes running the Kalman filter first and running a simulation smoother using the filtered values for drawing smoothed states as in Carter and Kohn (1994) and Frühwirth-Schnatter (1994). We start the recursion for $t = 1, \ldots, T$

$$\begin{aligned}
X_{t|t-1} &= FX_{t-1|t-1} \\
P_{t|t-1} &= FP_{t-1|t-1}F' + R_t' R_t \\
\eta_{t|t-1} &= y_t - HX_{t|t-1} - BU_t \\
\zeta_{t|t-1} &= HP_{t|t-1}H' + Q_t \\
K_t &= P_{t|t-1}H'\zeta_{t|t-1}' \\
X_{t|t} &= X_{t|t-1} + K_t\eta_{t|t-1} \\
P_{t|t} &= P_{t|t-1} - K_tH'\zeta_{t|t-1}',
\end{aligned} \tag{23}$$

and store $X_{t|t}$ and $P_{t|t}$. The last filtered state $X_{T|T}$ and its covariance matrix $P_{T|T}$ correspond to the smoothed estimates of the mean and the covariance matrix of the states for period T. Having stored all the filtered values, simulation smoother involves the following backward recursions for $t = T-1, \ldots, 1$

$$\begin{aligned}
\eta^*_{t+1|t} &= X_{t+1} - FX_{t|t} \\
\zeta^*_{t+1|t} &= FP_{t|t}F' + R'_{t+1}R_{t+1} \\
X_{t|t,X_{t+1}} &= X_{t|t} + P_{t|t}F'\zeta^{*-1}_{t+1|t}\eta^*_{t+1|t} \\
P_{t|t,P_{t+1}} &= P_{t|t} - P_{t|t}F'\zeta^{*-1}_{t+1|t}FP_{t|t}.
\end{aligned} \tag{24}$$

Intuitively, the simulation smoother updates the states using the same principle as in the Kalman filter, where at each step filtered values are updated using the smoothed values obtained from backward recursion. For updating the initial states, using the state equation $X_{0|t,X_1} = F^{-1}(X_1)$ and $P_{0|t,P_1} = F^{-1}(P_1 + R'_1R_1)F'^{-1}$ can be written for the first observation. Given the mean $X_{t|t,X_{t+1}}$ and the covariance matrix $P_{t|t,P_{t+1}}$, the states can be sampled from $X_t \sim N(X_{t|t,X_{t+1}}, P_{t|t,P_{t+1}})$ for $t = 0, ..., T$.

Sampling of inflation volatilities, h_t

Conditional on the remaining model parameters, we can draw $h_{0:T}$ using standard Bayesian inference as in the case of X_t. One important difference, however, stems from the logarithmic transformation of the variance in the stochastic volatility model. As the transformation concerns the error structure, the square of which follows a χ^2 distribution, the system is not Gaussian but follows a log-χ^2 distribution. Noticing the properties of log-χ^2 distribution, Kim et al. (1998) and Omori et al. (2007) approximate this distribution using a mixture of Gaussian distributions. Hence, conditional on these mixture components the system remains Gaussian allowing for standard inference outlined above. For details, see Omori et al. (2007).

Sampling of structural break parameters, κ_t

Sampling of structural break parameters, κ_t relies on the conditional posterior of the binary outcomes, i.e. the posterior value in case of a structural break in period t and the posterior value of the case of no structural breaks. However, evaluating this posterior requires one sweep of filtering, which is of order $O(T)$. As this evaluation should be implemented for each period t the resulting procedure would be of order $O(T^2)$. When the number of sample size is large this would result in an infeasible scheme. Gerlach et al. (2000) propose an efficient algorithm for sampling structural break parameters, κ_t, conditional on the observed data, which is still of order $O(T)$. We implement this algorithm for estimation of the structural breaks and refer to Gerlach et al. (2000); Giordani and Kohn (2008) for details.

Sampling of state error variances, σ_η^2

Using standard results from a linear regression model with a conjugate prior for the variances in (21), it follows that the conditional posterior distribution of $\sigma^2_{\eta_i}$, with $i = 1, 2, 3, 4$ is an inverted Gamma distribution with scale parameter $\Phi_{\eta_i} + \sum_{t=1}^{T} \eta_{i,t}^2$ and with $T + \nu_{\eta_i}$ degrees of freedom for $i = 2, 3, 4$ where Φ_{η_i} and ν_{η_i} are the scale and degrees of freedom parameters of the prior density. For $i = 1$ the parameters of the inverted Gamma distribution becomes $\Phi_{\eta_1} + \sum_{t=1}^{T} \kappa_t \eta_{1,t}^2$ and $\sum_{t=1}^{T} \kappa_t + \nu_{\eta_1}$.

Sampling of marginal cost variance and correlation coefficient

To sample the variance of marginal cost and correlation coefficient, we decompose the multivariate normal distribution of $\epsilon_{\mathbf{t}}$ into the conditional distribution of $\epsilon_{2,t}$ given $\epsilon_{1,t}$ and the marginal distribution of $\epsilon_{1,t}$, as in Çakmaklı et al. (2011). This results in

$$\prod_{t=1}^{T} f(\epsilon_{\mathbf{t}}) = \prod_{t=1}^{T} \frac{1}{\sigma_{\epsilon_{1,t}}} \phi\left(\frac{\epsilon_{1,t}}{\sigma_{\epsilon_{1,t}}}\right) \frac{1}{\sigma_{\epsilon_{2,t}} \sqrt{(1-\rho^2)}} \phi\left(\frac{\epsilon_{2,t} - \rho\epsilon_{1,t}}{\sigma_{\epsilon_{2,t}}(1-\rho^2)}\right), \tag{25}$$

Hence, together with prior for the variance in (21), variance of the marginal cost series can be sampled using (25) by setting up a Metropolis-Hasting step using an inverted Gamma candidate density with scale parameter $\sum_{t=1}^{T} \epsilon_{2,t}^2$ and with T degrees of freedom. To sample ρ from its conditional posterior distribution we can again use (25). Conditional on the remaining parameters the posterior becomes

$$(1-\rho^2)^{-\frac{3}{2}} \prod_{t=1}^{T} \left(\frac{1}{\sqrt{(1-\rho^2)}} \phi\left(\frac{\epsilon_{2,t} - \rho\epsilon_{1,t}}{\sigma_{\epsilon_{2,t}}(1-\rho^2)}\right)\right). \tag{26}$$

We can easily implement the griddy Gibbs sampler approach of Ritter and Tanner (1992). Given that $\rho \in (-1, 1)$ we can setup a grid in this interval based on the precision we desire about the value of ρ.

D Bayesian inference of the extended HPC model

Posterior inference of the HPC models with time varying parameters follow similar to Appendix C, using the Gibbs sampler with data augmentation. The HPC models with time varying parameters (HPC-TV), with level shifts in inflation (HPC-TV-LS), and with level shifts and stochastic volatility in inflation (HPC-TV-LS-SV) and the inflation expectation specification presented in the paper can be cast into the state-space form in (21) using the following definitions

$$
Y_t = \begin{pmatrix} \pi_t \\ z_t \end{pmatrix}, \; X_t = \begin{pmatrix} c_{\pi,t} & c_{z,t} & \mu_{z,t}, & c_{z,t-1} & c_{z,t-2} & S_t & c_{\pi,t-1} \end{pmatrix}', \; \epsilon_{\mathbf{t}} = \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix},
$$

$$
U_t = \begin{pmatrix} z_t & z_{t-1} & z_{t-2} & \mu_t & \pi_{t-1} \end{pmatrix}', \; B_t = \begin{pmatrix} \alpha_1 & \alpha_2 & 0 & -\alpha_3 & \alpha_4 \\ 0 & \phi_1 & \phi_2 & 0 & 0 \end{pmatrix},
$$

$$
H_t = \begin{pmatrix} 1-\alpha_3 & -\alpha_1 & 0 & -\alpha_2 & 0 & \alpha_3 & -\alpha_4 \\ 0 & 1 & 0 & -\phi_1 & -\phi_2 & 0 & 0 \end{pmatrix}, \; Q_t = \begin{pmatrix} \sigma^2_{\epsilon_1,t} & \rho\sigma_{\epsilon_1,t}\sigma_{\epsilon_2} \\ \rho\sigma_{\epsilon_1,t}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{pmatrix},
$$

$$
F_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \beta & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad R_t = \begin{pmatrix} \kappa_t \sigma_{\eta_1} & 0 & 0 & 0 \\ 0 & \sigma_{\eta_2} & 0 & 0 \\ 0 & 0 & \sigma_{\eta_3} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\eta_5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \eta_{\mathbf{t}} = \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \\ \eta_{5,t} \end{pmatrix},
$$

where parameters $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are defined as functions of the structural form parameters

$$
\alpha_1 = \frac{\lambda^H}{\left(1 - (\phi_1 + \phi_2 \gamma_f^H)\gamma_f^H\right)\left(1 - \gamma_b^H \gamma_f^H\right)}, \quad \alpha_2 = \frac{\lambda^H \gamma_f^H \phi_2}{\left(1 - (\phi_1 + \phi_2 \gamma_f^H)\gamma_f^H\right)\left(1 - \gamma_b^H \gamma_f^H\right)},
$$
$$
\alpha_3 = \frac{\gamma_b^H \gamma_f^H}{\left(1 - \gamma_b^H \gamma_f^H\right)} \frac{\gamma_f^H}{\left(1 - \gamma_f^H \beta\right)}, \quad \alpha_4 = \frac{\gamma_b^H}{\left(1 - \gamma_b^H \gamma_f^H\right)}.
$$

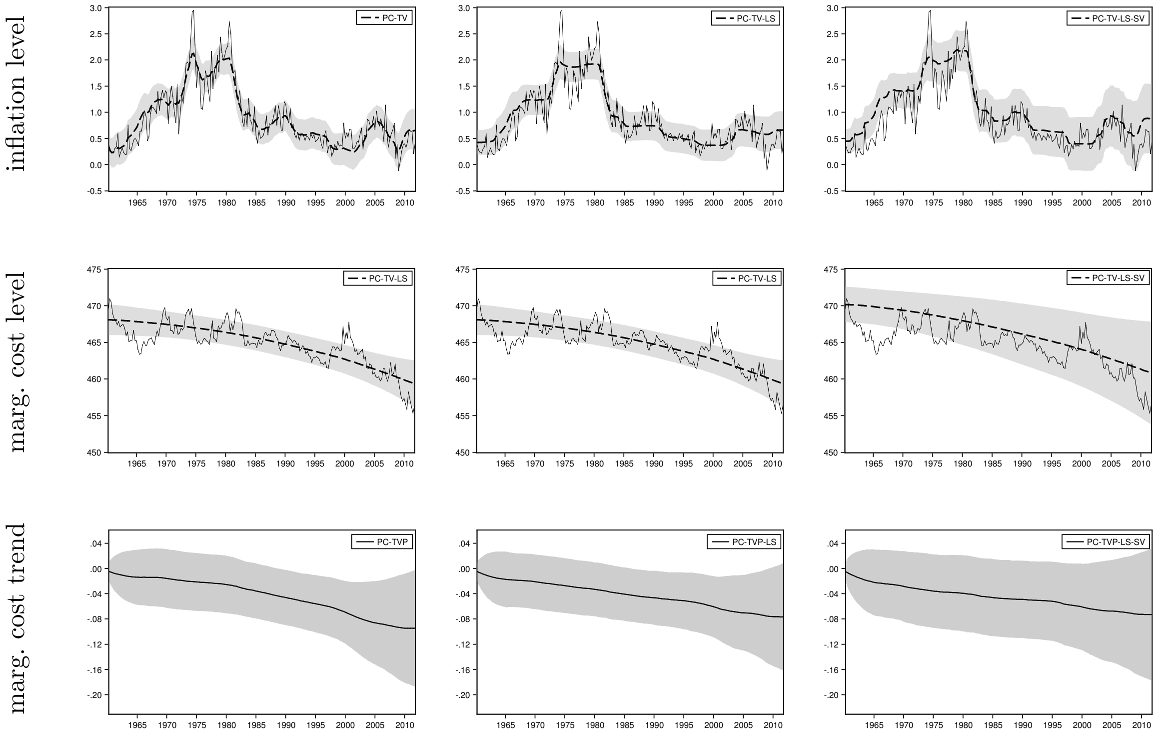
Given this setup, posterior inference can be carried out using the steps outlined in Appendix C.

E Posterior results for the PC models with non-filtered time series

This section presents additional estimation results for the PC models with non-filtered time series. We summarize the estimated levels, volatilities, breaks and inflation expectations obtained from the PC-TV, PC-TV-LS and PC-TV-LS-SV models. Figure 10 shows the estimated levels from the three PC models. Estimated inflation levels, computed as the posterior mean of the smoothed states, are given in the first row of Figure 10. Shaded areas around the posterior means represent the 95% HPDI for the estimated levels. For all three models, estimated inflation levels nicely track the observed inflation. Effects of the level specification are reflected in the estimates in various ways. First, when we model inflation level changes as discrete level shifts rather than continuous changes, we observe a relatively smoother pattern in estimated inflation levels. This effect can be seen by comparing the second and first graphs in the first row of Figure 10. While estimated inflation level in the first graph follows the observed inflation patterns closely, estimated inflation level in the second (and third to a less extent) graph mostly indicates three distinct periods. These periods are the high inflation periods capturing 1970s with a constant inflation level around 1.7% (quarterly inflation) following a low inflation period in 1960s, and the period after the beginning of 1980s with a stable inflation level around 0.5%, see Cecchetti et al. (2007) for similar findings. Second, adding the stochastic volatility together with level shifts results in discrete level shifts in inflation which are more frequent than the model with only level shifts.

The second panel in Figure 10 presents the estimated levels for the real marginal cost series for all models. A common feature of all these estimates is the smoothness of the estimated levels. In all models, marginal cost series follows a slightly nonlinear trend during the sample period. The estimated slopes of these trends for all models are given in the bottom panel of Figure 10, together with the 95% HPDIs. Nonlinearity of the negative trend is reflected in the negative values for the slope of the trend, with an increasing magnitude at the end of the sample. This change in the slope of the trend is accompanied by the increasing uncertainty about the slope. The difference between the models in terms of the estimated marginal cost structures is negligible.

Figure 10: Level, trend and slope estimates from the PC models



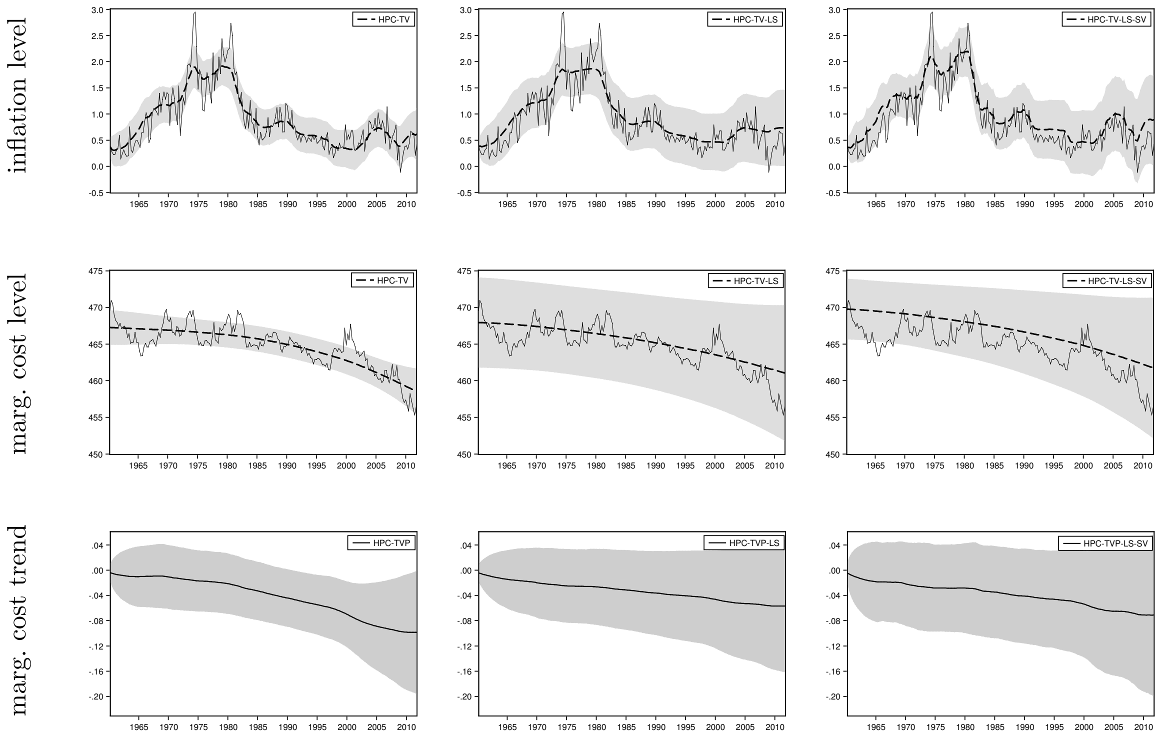
Note: The top panel exhibits estimated inflation levels. The middle and the bottom panels show estimated real marginal cost levels and the slopes of the levels, respectively. Grey shaded areas correspond to the 95% HPDI. PC-LT (PC-HP) refers to the PC model where the real marginal cost series is detrended using linear trend (Hodrick-Prescott) filter. PC-TV refers to the PC model with time varying levels and trends. PC-TV-LS refers to the PC model with time varying levels and trends. PC-TV-LS-SV refers to the PC model with time varying levels, trends and volatility. HPC-TV refers to the Hybrid PC model with time varying levels, trends and inflation expectations. HPC-TV-LS refers to the HPC model with time varying levels, trends and inflation expectations. HPC-TV-LS-SV refers to the HPC model with time varying levels, trends, inflation expectations and volatility. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

F Posterior results for the HPC models with non-filtered time series

This section presents additional estimation results for the HPC models with non-filtered time series. We summarize the estimated levels, volatilities, breaks and inflation expectations obtained from the HPC-TV, HPC-TV-LS and HPC-TV-LS-SV models.

Figure 11 presents the estimated inflation levels, together with estimated levels and trends of the marginal cost series.

Figure 11: Level, trend and slope estimates from the HPC models

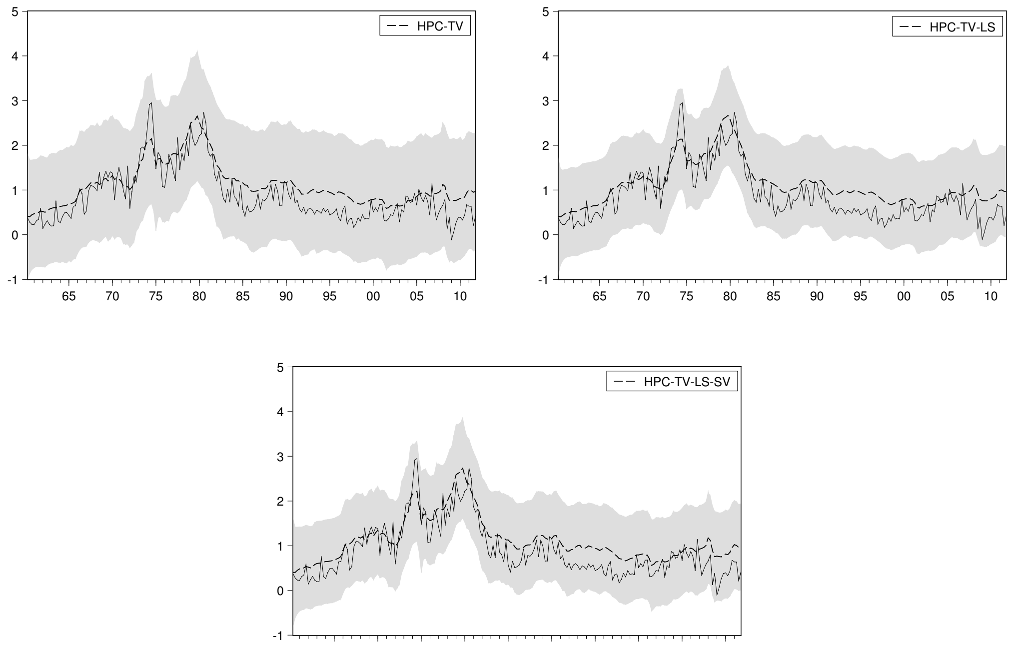


Note: The top panel exhibits estimated inflation levels. The middle and the bottom panels show estimated real marginal cost levels and the slopes of the levels, respectively. Grey shaded areas correspond to the 95% HPDI. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

G Predicted inflation densities from all proposed models

This section presents the entire distribution of the inflation predictions for all PC and HPC models. The solid lines represent the posterior mean of predicted inflation, and the white areas under the inflation densities show the inflation levels with non-zero posterior probability. For all models we propose, inflation predictions are concentrated around high (low) values during the high (low) inflationary periods. The uncertainty around the inflation predictions are also high for these periods, together with the periods when inflation is subject to a transition to low values around 1980s. When the

Figure 12: Implied inflation expectations by HPC models



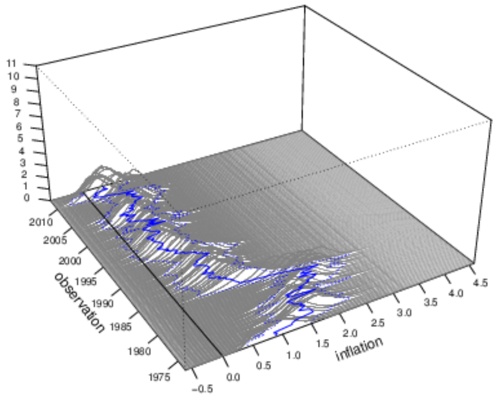
Note: The thick solid lines are the posterior means of inflation expectations from the HPC models. The thin solid lines are the observations of inflation expectations from survey data. Grey shaded areas are the 95% HPDI for estimated inflation expectations. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

observed inflation values are close to the zero bound, the predictive densities indicate disinflationary risk, computed as the fraction of the predictive distribution below zero.

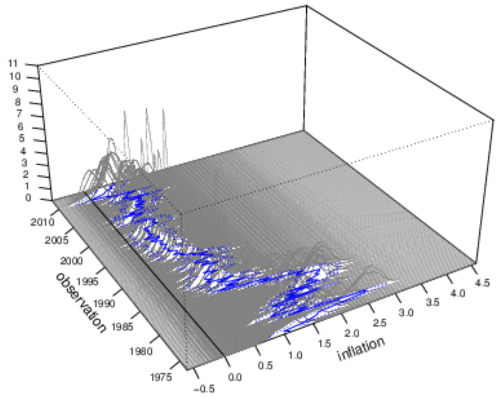
H Posterior and predictive results from alternative models for robustness checks

The proposed PC and HPC models extend the standard models in several ways. First, both model structures introduce time variation in the long and short run dynamics of inflation and marginal cost series. Second, the introduction and the iterative solution of the expectational mechanisms and the survey data in the extended HPC models enables the use of more data information. Furthermore, extended and standard HPC models use the additional information from a backward looking component for the inflation series compared to the HPC counterparts. According to the predictive results, the most comprehensive model, HPC-TV-LS-SV is also the best performing model. However, a deeper analysis is needed in order to see the added predictive gain from all these extensions. In this section we consider several alternative models and their predictive performances to separately address the predictive gains from each of these extensions in the model structure. Table 5 presents all PC and HPC model structures we compare to differentiate these effects.

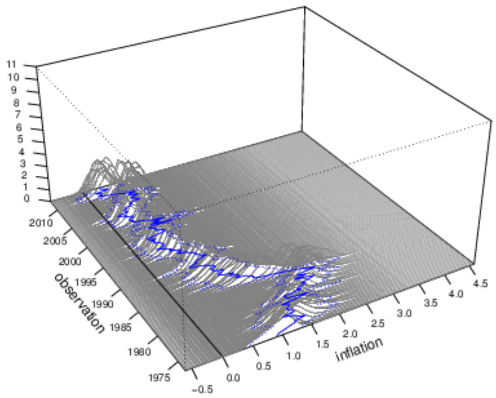
Figure 13: Predicted inflation densities from PC and HPC models



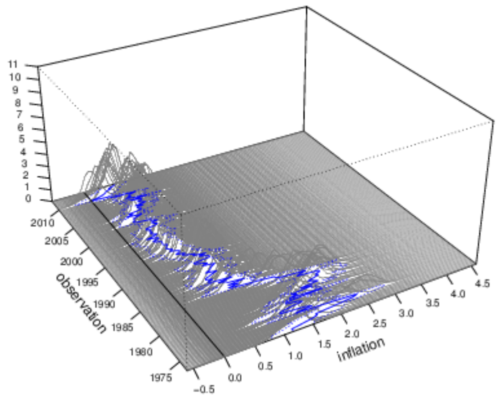
PC-TV



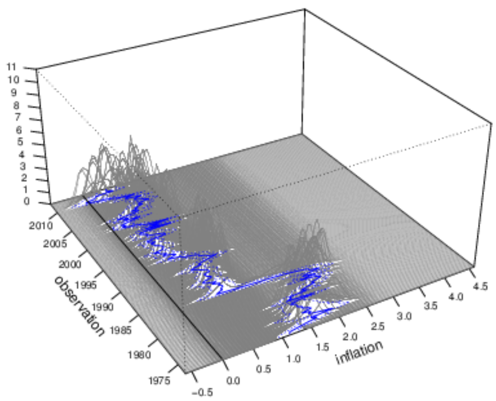
HPC-TV



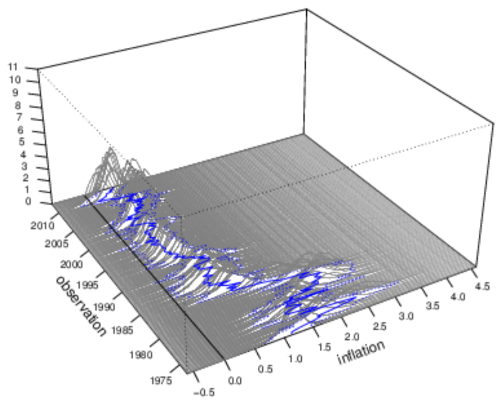
PC-TV-LS



HPC-TV-LS



PC-TV-LS-SV



HPC-TV-LS-SV

Note: The figure presents one period ahead predictive distributions of inflation from the PC and HPC models, for the period between the third quarter of 1973 and the first quarter of 2012. Model abbreviations are as in Figure 10 . Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Table 5: Standard and extended PC models

model structure / low/high frequencies	iterated expectations solution		direct expectations data	
	Phillips Curve	Hybrid Phillips Curve	Phillips Curve	Hybrid Phillips Curve
linear trend	PC-LT	n/a ⋆	PCS-LT	HPCS-LT
Hodrick-Prescott filter	PC-HP	n/a ⋆	PCS-HP	HPCS-HP
time varying levels	PC-TV	HPC-TV	PCS-TV	HPCS-TV
time varying levels and switching	PC-TV-LS	HPC-TV-LS	PCS-TV-LS	HPCS-TV-LS
time varying levels and stochastic volatility	PC-TV-SV	HPC-TV-SV	PCS-TV-SV	HPCS-TV-SV
time varying levels, switching and stochastic volatility	PC-TV-LS-SV	HPC-TV-LS-SV	PCS-TV-LS-SV	HPCS-TV-LS-SV

Note: The first two columns present the standard and extended (H)PC models presented in the main paper, for which expectational mechanisms are solved explicitly. The last two columns present alternative model structures for (H)PC models. For these models, we do not iterate inflation expectations in the models, but instead replace them with survey data directly. PC(S)-LT (PC-HP(S)) refers to the PC model where the real marginal cost series is detrended using linear trend (Hodrick-Prescott) filter. PC(S)-TV refers to the PC model with time varying levels and trends. PC(S)-TV-LS refers to the PC model with time varying levels and trends. PC(S)-TV-LS-SV refers to the PC model with time varying levels and volatility. PC(S)-TV-LS-SV refers to the PC model with time varying levels, trends and volatility. HPC(S)-TV refers to the Hybrid PC model with time varying levels, trends and inflation expectations. HPC(S)-TV-LS refers to the HPC model with time varying levels, trends and inflation expectations. HPC(S)-TV-LS-SV refers to the HPC model with time varying levels and volatility. HPC(S)-TV-LS-SV refers to the HPC model with time varying levels, trends, inflation expectations and volatility.

⋆ Iterative solution of these models without using the survey data does not exist.

The first set of alternative models we consider are the standard PC and HPC models combined with data from survey expectations, without introducing explicit time variation in the low frequency structure of data but instead demeaning the inflation series, and detrending the marginal cost series prior to analysis. These models are given in the first two rows of the right panel of Table 5 and are abbreviated by PCS-LT, PCS-HP, HPCS-LT and HPCS-HP, according to linear detrending or HP detrending prior to analysis. The improved predictive performances of PCS-LT and PCS-HP models compared to the standard PC counterparts show predictive gains from incorporating survey expectations in the models. Furthermore, comparing the predictive performances of the HPCS-LT and HPCS-HP models with the time-varying hybrid models, such as the HPC-TV or HPC-TV-LS models show the gains from incorporating time variation alone, since all these models use survey data and the backward looking component for inflation.

The second set of alternative models we consider, on the right panel of Table 5, are PC models with time-varying levels, where we incorporate the survey expectations in the model directly rather than solving the model iteratively. These models correspond to (1) where the expectation term is replaced by survey expectations. We denote these models by PCS-TV, PCS-TV-LS and PCS-TV-LS-SV, for the time-varying levels, time-varying levels with regimes shifts in inflation and time-varying levels with regime shifts and stochastic volatility component, respectively. Comparing the predictive results of these models to the HPC counterparts provide the predictive gains solely from the HPC extension, i.e. they separate the gains from incorporating the backward looking inflation component in the model from the other model extensions.

The third set of alternative models we consider are the HPC models using the survey expectations directly, without solving for the expectational mechanisms. We denote these models by HPCS-TV, HPCS-TV-LS and HPCS-TV-LS-SV, for the time-varying levels, time-varying levels with regimes shifts in inflation and time-varying levels with regime shifts and stochastic volatility component, respectively. Comparing the predictive performance of these models with the proposed HPC models clarifies the predictive gains from solving for the inflation expectations iteratively in the hybrid models.

The final set of alternative models aim to separate the predictive gains from the stochastic volatility component in the time-varying level models without level shifts. The comparison of the predictive results of these models, (H)PC-TV-SV with the models with level switching (H)PC-TV-LS-SV highlights predictive gains solely from introducing level shifts.

One period ahead MSFE and log marginal likelihoods of these models, together with the standard (H)PC models and the models proposed in the paper, are given in Table 6. The prediction results are based on the forecast sample, which covers the period between the second quarter of 1973 quarter and the first quarter of 2012. Comparing the first block and the first two rows of the second block Table 6, we see that the gains from using survey data inflation is substantial even in the standard PC models. In terms of predictive gains, the biggest improvement in predictive likelihoods and the MSFE are achieved with this contribution in the models. However, the predictive performances of these improved models are still far from the more involved models. Hence the gains from the proposed models do not only stem from the inclusion of the survey data information alone.

We also report the predictive gains resulting solely from introducing time-variation in the inflation and marginal cost series, by comparing the results of the HPCS-LT and HPCS-HP models with the HPC-TV or HPC-TV-LS models in the table. The more involved models with time variation clearly perform better according to the predictive results. Especially the difference in marginal likelihoods of these models enables us to conclude that incorporating time variation in the data is also important.

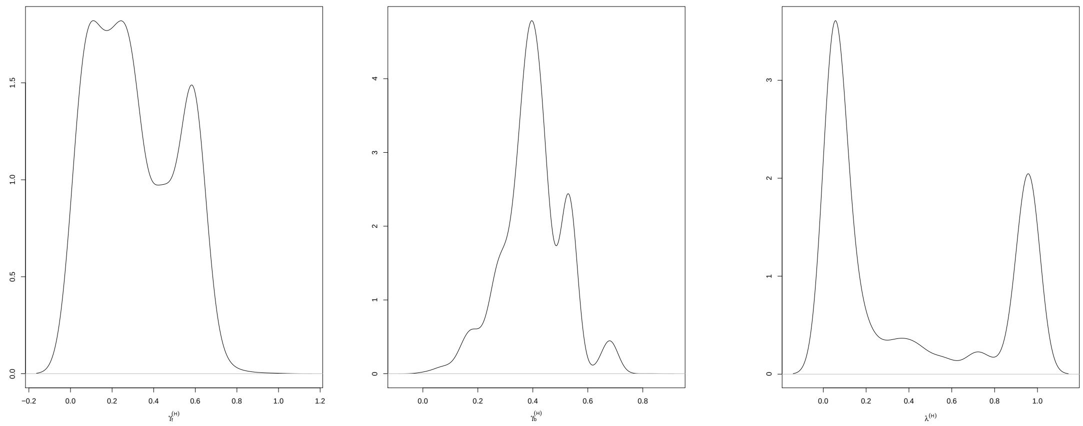
As a third possibility for predictive gains, we focus on the models with backward looking components. One way to separate the added value from this component is to consider the second block of Table 6. The prediction results from the PC and HPC models in this block are very similar, with slight improvements in the hybrid models, where the backward looking component is incorporated. Another way to see the effect of the backward looking component is to compare the PCS-TV, PCS-TV-LS and PCS-TV-LS-SV models with HPCS-TV, HPCS-TV-LS and HPCS-TV-LS-SV models, respectively. In all these comparison, the models without the backward looking component performs slightly better (worse) in terms of MSFE (marginal likelihood), hence the backward looking component does not seem to improve predictive results in general and the improvements in the hybrid models mainly stem from incorporating the survey expectations.

From the considered alternative models, time-varying level models with a stochastic volatility component using survey data directly (PCS-TV-LS-SV and HPCS-TV-LS-SV) clearly perform best. In terms of the predictive likelihoods, these models are also comparable to the 'best performing' model we propose.

A final source of possible predictive gains in the proposed models is the iterative solution of inflation expectations. This comparison is based on the comparison of the models in the third (fourth) block and the fifth (sixth) block of Table 6, where only the third (fourth) block uses the iterative solution. According to the MSFE, predictive results deteriorate slightly when we solve the system. We find this result rather counterintuitive since the iterative solution is based on the complete model structure. As we show briefly, despite this slight increase in the predictive performances, models without the iterative solutions suffer from identification issues.

We next focus on changes in parameter estimates for the alternative models proposed in this appendix. Table 7 presents the parameter estimates for all alternative models. Despite the predictive gains from these alternative models, parameter estimates are rather different from those obtained from the proposed models. Specifically for the hybrid models considered, uncertainty in posterior distributions increase substantially if the iterative model solution is not used. Furthermore, posterior densities of some parameters are quite irregular in most of these models which use expectations data directly. Figure 14 shows this irregularity for the HPCS-TV model, parameters $\lambda^{(H)}$, $\gamma_b^{(H)}$ and $\gamma_f^{(H)}$. The bimodality problem in posterior densities is most apperant in the PC slope, $\lambda_b^{(H)}$. Furthermore, the backward looking component $\gamma_b^{(H)}$ is spread over a wide region with multiple modes. Similar results hold for the remaining alternative models which make use of the survey expectations data directly. We therefore conclude that replacing the expectational term in the (H)PC models with survey expectations deteriorate posterior inference compared to the iterative solution of these expectational terms.

Figure 14: Posterior density of $\lambda^{(H)}$, $\gamma_b^{(H)}$ and $\gamma_f^{(H)}$ from the HPCS-TV model



Note: The figure presents posterior densities of parameters from the HPCS-TV model. Model abbreviations are based on Table 5. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Table 6: Predictive performance of additional PC models

Model	(Log) Marg. Likelihood	MSFE 1 period ahead
PC-LT	-139.327	0.353
PC-HP	-157.195	0.458
PCS-LT	-79.141	0.105
PCS-HP	-85.397	0.130
HPCS-LT	-81.047	0.105
HPCS-HP	-85.200	0.119
PC-TV	-46.162	0.142
PC-TV-LS	-61.972	0.138
PC-TV-SV	-22.761	0.134
PC-TV-LS-SV	-33.476	0.126
HPC-TV	-36.683	0.109
HPC-TV-LS	-33.913	0.084
HPC-TV-SV	-20.738	0.097
HPC-TV-LS-SV	-18.960	0.102
PCS-TV	-34.407	0.129
PCS-TV-LS	-32.004	0.099
PCS-TV-LS-SV	-15.390	0.092
HPCS-TV	-40.465	0.176
HPCS-TV-LS	-38.082	0.297
HPCS-TV-LS-SV	-12.977	0.139
BVAR (constant)	-166.226	0.085
BVAR-SV	-220.710	0.091
SW2007	-78.033	0.168

Note: The table reports the predictive performances of alternative models for the period between the second quarter of 1973 and the first quarter of 2012. '(Log) Marg. Likelihood' stands for the natural logarithm of the marginal likelihoods. 'MSFE' stands for the Mean Squared Forecast Error. Marginal likelihood values in the first column are calculated as the sum of the predictive likelihood values in the prediction sample. Results are based on 10000 simulations of which the first 5000 are discarded for burn-in. Model abbreviations are based on Table 5. BVAR (constant) denotes the BVAR model with 2 lags and with constant parameters. 'BVAR-SV' denotes the 'BVAR' model with 2 lags, time varying levels for both series and stochastic volatility for inflation. 'SW2007 stands for the model proposed by Stock and Watson (2007).

Table 7: Posterior results of alternative PC models

Model	$\lambda^{(H)}$	$\gamma_f^{(H)}$	γ_b^H	ρ	ϕ_1	ϕ_2
PCS-LT	0.011 (0.051)	0.611 (0.055)	–	-0.016 (0.021)	0.824 (0.047)	0.075 (0.044)
PCS-HP	0.064 (0.051)	0.627 (0.081)	–	-0.045 (0.064)	0.681 (0.096)	0.014 (0.081)
HPCS-LT	0.154 (0.205)	0.350 (0.236)	0.408 (0.202)	-0.114 (0.155)	0.823 (0.058)	0.069 (0.057)
HPCS-HP	0.234 (0.235)	0.333 (0.180)	0.472 (0.154)	-0.216 (0.197)	0.614 (0.079)	-0.018 (0.057)
PCS-TV	0.057 (0.028)	0.142 (0.086)	–	-0.034 (0.061)	0.815 (0.052)	0.067 (0.052)
PCS-TV-LS	0.049 (0.023)	0.430 (0.125)	–	-0.027 (0.050)	0.821 (0.054)	0.072 (0.052)
PCS-TV-LS-SV	0.058 (0.025)	0.307 (0.165)	–	-0.015 (0.068)	0.826 (0.052)	0.078 (0.053)
HPCS-TV	0.383 (0.395)	0.308 (0.197)	0.401 (0.111)	-0.322 (0.349)	0.593 (0.314)	0.007 (0.100)
HPCS-TV-LS	0.557 (0.432)	0.375 (0.196)	0.393 (0.094)	-0.468 (0.367)	0.432 (0.328)	-0.031 (0.101)
HPCS-TV-LS-SV	0.151 (0.178)	0.216 (0.161)	0.368 (0.149)	-0.024 (0.095)	0.871 (0.027)	0.112 (0.032)

Note: Posterior results are based on 40000 simulations of which the first 20000 are discarded for burn-in. Model abbreviations are based on Table 5.

I Analysis of cointegration in inflation and marginal cost levels

The models in the paper considered rely on the implicit assumption of the absence of a long-run cointegrating relationship between the inflation and marginal cost series. We assess whether this assumption is plausible for the U.S. data. For this reason, we consider the PC-TV model that provides the unobserved levels of both series at each posterior draw. For each of these obtained posterior draws, we perform a simple two-step analysis to check the existence of the cointegrating relationship, which can be seen as a Bayesian extension of the method of Engle and Granger (1987).

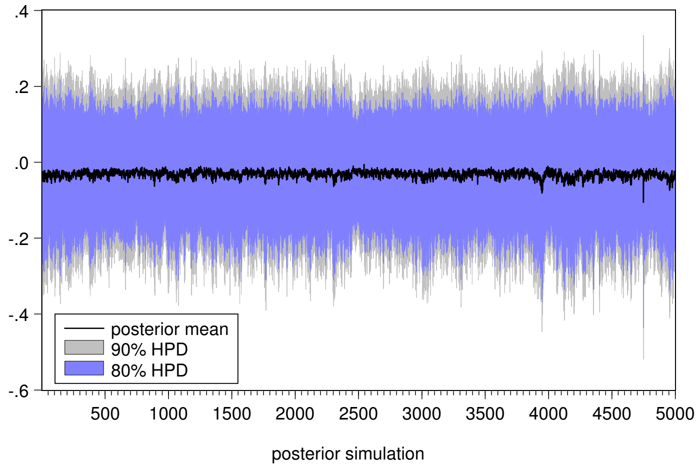
We perform a two step analysis, where in the first step we obtain the residuals from the regression of the estimated level of inflation on a constant and the estimated level of marginal cost, for each posterior draw. This implies that we take the estimation uncertainty in the analysis into account. Next, we obtain the posterior distribution of the autoregressive parameter, ρ, for each set of residuals from the following regression using flat priors on the identified region $\rho \in [-1, 1]$

$$\Delta\hat{\epsilon}_t = \rho\hat{\epsilon}_{t-1} + \eta_t, \;\; \eta_t \sim NID(0, \sigma^2), \tag{27}$$

where $\hat{\epsilon}_t$ denotes the residuals from the first stage, and $\rho = 0$ implies that there is no cointegrating relationship between the series. An HPDI including the value of 0 indicates that a cointegrating relation between inflation and marginal cost is unlikely.

We compute the mean and the quantiles of these individual densities using 5000 posterior draws, and report the average values of the mean and the quantiles of ρ based on 3000 simulations. These results are presented in Figure 15. Posterior means of parameter ρ are around 0 for all posterior draws of inflation and marginal cost levels, and the 80% an 90% percent quantiles of the distribution are around 0 as well. Hence this simulation experiment does not indicate a cointegrating relationship between the inflation and marginal cost levels. This pattern is also found for other TV-PC models we considered for the U.S. data, but these results are not reported for the sake of brevity. We conclude that the underlying assumption of 'no cointegrating relationship' is found to be feasible for the PC models we consider.

Figure 15: Cointegration analysis for the marginal cost and inflation series



Note: The figure presents the posterior means and quantiles of the ρ parameter from 5×10^3 posterior draws from the PC-TV models, where for each draw, the the reported values are calculated using 3000 simulations. $\rho = 0$ implies that there is no cointegrating relationship between the series.

References

Ahmed S, Levin A, Wilson BA. 2004. Recent U.S. macroeconomic stability: Good policies, good practices, or good luck? *The Review of Economics and Statistics* **86**: 824–832.

Ascari G. 2004. Staggered prices and trend inflation: Some nuisances. *Review of Economic Dynamics* **7**: 642–667.

Ascari G, Ropele T. 2007. Optimal monetary policy under low trend inflation. *Journal of Monetary Economics* **54**: 2568–2583.

Basistha A, Nelson CR. 2007. New measures of the output gap based on the forward-looking new Keynesian Phillips curve. *Journal of Monetary Economics* **54**: 498–511.

Benati L. 2008. Investigating inflation persistence across monetary regimes. *The Quarterly Journal of Economics* **123**: 1005–1060.

Çakmaklı C. 2012. Bayesian semiparametric dynamic nelson-siegel model. Working Paper Series 59-12, The Rimini Centre for Economic Analysis.

Çakmaklı C, Paap R, Van Dijk D. 2011. Modeling and estimation of synchronization in multistate Markov-switching models. Working Papers 11-002/4, Tinbergen Institute.

Calvo G. 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* **12**: 383–398.

Canova F. 2012. Bridging DSGE models and the raw data. Working Papers 635, Barcelona Graduate School of Economics.

Canova F, Ferroni F. 2011. Multiple filtering devices for the estimation of cyclical DSGE models. *Quantitative Economics* **2**: 73–98.

Carter CK, Kohn R. 1994. On Gibbs sampling for state space models. *Biometrika* **81**: 541–553.

Cecchetti SG, Hooper P, Kasman BC, Schoenholtz KL, Watson MW. 2007. Understanding the evolving inflation process. Paper written for the U.S. Monetary Policy Forum 2007, Washington, DC.

Christiano LJ, Vigfusson RJ. 2003. Maximum likelihood in the frequency domain: the importance of time-to-plan. *Journal of Monetary Economics* **50**: 789–815.

Cogley T, Sbordone AM. 2008. Trend inflation, indexation, and inflation persistence in the New Keynesian Phillips curve. *The American Economic Review* **98**: 2101–2126.

Del Negro M, Schorfheide F. 2012. DSGE model-based forecasting. Staff Reports 554, Federal Reserve Bank of New York.

Delle Monache D, Harvey AC. 2011. The effect of misspecification in models for extracting trends and cycles. Working Papers EWP 2011/013, Euroindicators.

Engle RF, Granger CWS. 1987. Co-integration and error correction: representation, estimation, and testing. *Econometrica* : 251–276.

Ferroni F. 2011. Trend agnostic one-step estimation of DSGE models. *The B.E. Journal of Macroeconomics* **11**: 1–36.

Frühwirth-Schnatter S. 1994. Data augmentation and dynamic linear models. *Journal of Time Series Analysis* **15**: 183–202.

Fuhrer J, Moore G. 1995. Inflation persistence. *The Quarterly Journal of Economics* **110**: 127–59.

Galí J, Gertler M. 1999. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics* **44**: 195–222.

Galí J, Gertler M, David Lopez-Salido J. 2005. Robustness of the estimates of the hybrid new keynesian phillips curve. *Journal of Monetary Economics* **52**: 1107–1118.

Geman S, Geman D. 1984. Stochastic relaxations, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence* **6**: 721–741.

Gerlach R, Carter C, Kohn R. 2000. Efficient Bayesian inference for dynamic mixture models. *Journal of the American Statistical Association* : 819–828.

Geweke J. 2010. *Complete and Incomplete Econometric Models (The Econometric and Tinbergen Institutes Lectures)*. Princeton University Press.

Geweke J, Amisano G. 2010. Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting* **26**: 216–230.

Giordani P, Kohn R. 2008. Efficient Bayesian inference for multiple change-point and mixture innovation models. *Journal of Business & Economic Statistics* **26**: 66–77.

Hastings WK. 1970. Monte Carlo sampling using Markov chains and their applications. *Biometrika* **57**: 97–10.

Kim CJ, Nelson CR. 1999. *State-space models with regime switching: Classical and Gibbs-Sampling approaches with applications*, volume 1. The MIT Press, 1 edition.

Kim S, Shephard N, Chib S. 1998. Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies* **65**: 361–93.

Kleibergen F, Mavroeidis S. 2011. Identification robust priors for Bayesian analysis in DSGE models. Manuscript.

McConnell MM, Perez-Quiros G. 2000. Output fluctuations in the United States: What has changed since the early 1980's? *American Economic Review* **90**: 1464–1476.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. 1953. Equations of state calculations by fast computing machines. *Journal of Chemical Physics* **21**: 1087–1092.

Nason JM, Smith GW. 2008. The New Keynesian Phillips curve: lessons from single-equation econometric estimation. *Economic Quarterly* **94**: 361–395.

Omori Y, Chib S, Shephard N, Nakajima J. 2007. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics* **140**: 425–449.

Ritter C, Tanner MA. 1992. Facilitating the Gibbs sampler: The Gibbs stopper and the griddy-Gibbs sampler. *Journal of the American Statistical Association* **87**: 861–868.

Rudd J, Whelan K. 2005. New tests of the new-Keynesian Phillips curve. *Journal of Monetary Economics* **52**: 1167–1181.

Schorfheide F. 2005. Learning and monetary policy shifts. *Review of Economic Dynamics* **8**: 392–419.

Sims CA. 2002. Solving linear rational expectations models. *Computational Economics* **20**: 1–20.

Sims CA, Zha T. 2006. Were there regime switches in U.S. monetary policy? *American Economic Review* **96**: 54–81.

Stock JH, Watson MW. 2007. Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking* **39**: 3–33.

Stock JH, Watson MW. 2008. Phillips Curve inflation forecasts. Working Paper 14322, National Bureau of Economic Research.

Tanner MA, Wong WH. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* **82**: 528–550.

Woodford M. 2003. Optimal interest-rate smoothing. *Review of Economic Studies* **70**: 861–886.

Zellner A, Ando T, Baştürk N, Hoogerheide L, Van Dijk HK. 2012. Bayesian analysis of instrumental variable models: Acceptance-rejection within Direct Monte Carlo. *Econometric Reviews*, accepted.