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

Online Appendix

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*

November 2013

Contents

*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)

List of Tables

List of Figures

This appendix presents additional derivations and estimation results that are removed from the paper for space limitations.

In section A we present the effects of misspecified levels on posterior results in a standard New Keynesian Phillips Curve (NKPC). This analysis provides a straightforward motivation for the extended NKPC and HNKPC models in the paper. Specifically, we show that a priori demeaning and detrending of the data, without considering the short and long-run data properties obscure inference in these standard models.

In section B we elaborate and compare the inference of the NKPC model using structural form and unrestricted reduced form. This section illustrates the difficulty of inferring the unrestricted reduced form parameters and to obtain the main parameters of interest, the structural parameters, using these. This difficulty is based on the non-linear parameter transformations required to link the structural and the reduced form models. Through simulation examples, we show that flat prior distributions used in one of the model representations can be very informative in the other model representation. This section motivates the structural parameter estimation approach we follow throughout the paper.

Sections C and D provide the details of the posterior sampling algorithms for the extended NKPC and HNKPC models proposed in the main paper. In these sections, the state space representations of the extended models and the appropriate sampling scheme are explained in detail. We further report the exact prior parameters used for the results in the paper and present a sketch of a prior sensitivity analysis based on prior-predictive likelihood comparisons.

Sections E, F and G provide posterior and predictive results for the extended NKPC and HNKPC models which are not included in the paper due to space constraints. In section E we present additional posterior and predictive results for the extended NKPC models. Main conclusions from these models are similar to the extended HNKPC model results discussed in the paper. Nonetheless, we provide these results for clarity and the ease of comparison. In section F we present additional results for the HNKPC models which are in line with the main conclusions of the paper. Section G presents the entire distribution of the inflation predictions for extended NKPC and HNKPC models we propose.

Section H presents the results of the prior-predictive likelihood analysis for the proposed models. The main conclusion of this section is that the adopted priors in the paper do not dominate the results. The data information is the main factor favoring the extended models we propose.

Section I presents the posterior and predictive results of the alternative NKPC and HNKPC models, considered for robustness checks, in detail. Several alternative models are compared with the extended models in the paper. We show that our main conclusions on the improved model performance through modeling the trends and levels in the data, and the use of survey data hold. We further disentangle the predictive gains from these two sources of extensions.

In section J presents a further alternative HNKPC model to the proposed HNKPC models in our paper. This model aims at accounting for the possibility of measurement errors in survey expectations. The results obtained from this alternative model are very similar to the corresponding results of in the paper, thus, we conclude that the effect of the measurement errors in survey expectations is negligible.

Section K presents a straightforward cointegration analysis for inflation and marginal cost series, based on the time-varying NKPC model structure. This analysis is performed to justify an implicit assumption in the proposed models namely the assumption that there is no stable long-run relationship between the inflation and marginal cost series. The results of this cointegration analysis are in line with the implicit assumption we make in the proposed NKPC model structures.

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

The linear NKPC 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 NKPC.

The linear NKPC model can be written as

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

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 costs 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, the slope of the Phillips curve, is restricted as $0 \leq \lambda < 1$ which is in line with previous evidence on the slope of the NKPC.

Since NKPC in (1) specifies the relation between the *short-run* stationary fluctuations in real marginal costs and inflation, $\tilde{\pi}_t$ and $\tilde{z}_t$ can be interpreted as the *transitory* components of inflation and marginal costs, 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

$$\begin{aligned} \pi_t &= \tilde{\pi}_t + c_{\pi,t}, \\ z_t &= \tilde{z}_t + c_{z,t}, \end{aligned} \tag{2}$$

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 in order to mimic the high inflationary period during the 1970s. For modelling the permanent component of the real marginal cost series, we use a trend specification mimicking the declining real marginal cost levels in the U.S. over the sample starting from the 1960s. This specification can be formulated as follows

$$\begin{aligned} c_{\pi,t} &= c_{\pi,t-1} + \kappa_t \eta_{t-1}, \;\; c_{z,t} = c_{z,t-1} + \mu_{z,t-1}, \\ \mu_{z,t} &= \mu_{z,t-1}, \;\; \eta_t \sim NID(0, \omega^2), \end{aligned} \tag{3}$$

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 costs, 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 costs in (3), we specify a deterministic trend for this series.

We simulate three sets of data from the model in (1)–(3). 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.

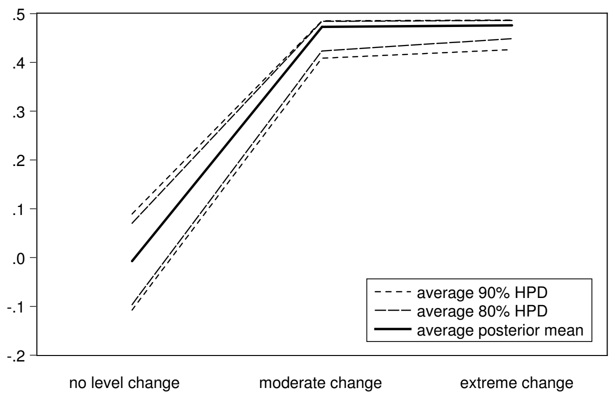
In order to capture the effect of model misspecification on posterior inference, when computing the transitory component, we ignore level shifts in the simulated inflation series and simply demean the series. For the marginal cost series, we remove the linear trend prior to the analysis and only focus 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 (1) using flat priors on restricted parameter regions:

$$p(\phi_1, \phi_2, \gamma_b, \lambda) \propto \begin{cases} 1, & \text{if } |\phi_1| + \phi_2 < 1, \ |\phi_2| < 1, \ 0 \leq \gamma_b < 1, \ 0 \leq \lambda < 1 \\ 0, & \text{otherwise} \end{cases}. \tag{4}$$

Given that model (1) 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 (2013). 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 (1). 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 1 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) for this overestimation.

Figure 1: Overestimation illustration for the backward looking NKPC model



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

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

This section presents the unrestricted reduced form inference (URF) of the NKPC model, and the inference of the corresponding structural form (SF) model parameters. The structural form (SF) representation for the basic NKPC model derived from the firm's price setting for filtered data is given as

$$\begin{aligned} \tilde{\pi}_t &= \lambda \tilde{z}_t + \gamma_f E_t(\tilde{\pi}_{t+1}) + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{5}$$

where $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID\,(0, \Sigma)$ and standard stationary restrictions hold for ϕ_1, ϕ_2.

We show that the posterior draws from the structural form parameters can be obtained using the reduced form representation of (5):

$$\begin{aligned} \tilde{\pi}_t &= \alpha_1 \tilde{z}_{t-1} + \alpha_2 \tilde{z}_{t-2} + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{6}$$

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 (5):

$$\alpha_1 = \tfrac{\lambda(\phi_1 + \gamma\phi_2)}{1 - \gamma(\phi_1 + \gamma\phi_2)}, \;\; \alpha_2 = \tfrac{\lambda\phi_2}{1 - \gamma(\phi_1 + \gamma\phi_2)}. \tag{7}$$

Finally, the model in (5) is related to an Instrumental Variables (IV) model with exact identification. Bayesian estimation of the unrestricted reduced form model in (6) is straightforward under flat or conjugate priors. Given the posterior draws of reduced form parameters, posterior draws of structural form parameters in (5) can be obtained using the transformation in (7). 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 $\lambda \neq 0$.[1]

Figure 2 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 2 shows the transformations 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 2 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 $\alpha_2 = 0$.

C Bayesian inference of the extended NKPC model

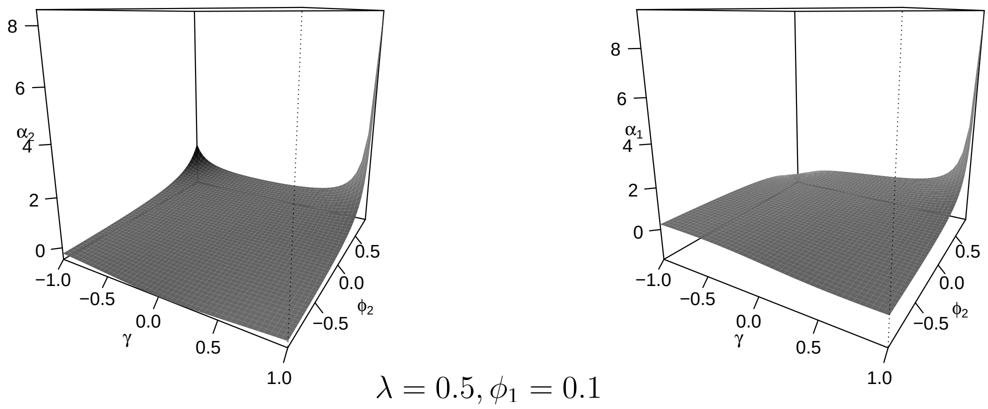
In this section we summarize the prior specifications, our use of prior predictive likelihoods, and the posterior sampling algorithms for the extended NKPC and HNKPC models. We further present a prior sensitivity analysis for the proposed models using a prior-predictive analysis.

C.1 Prior specification for parameters

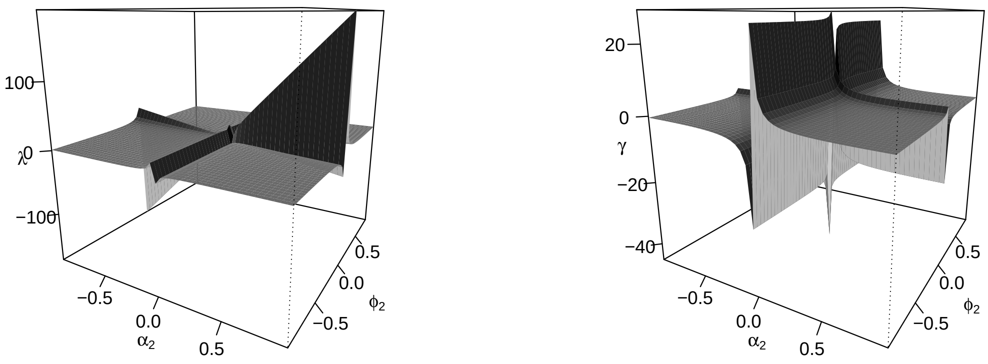
The extended NKPC and HNKPC models contain several additional parameters compared to the standard NKPC model. We classify the model parameters in five groups, and assign independent priors for each group. The first group includes the common parameters in the NKPC and HNKPC models, $\theta_N = \{\lambda, \gamma_f, \phi_1, \phi_2, \Sigma\}$, in

[1]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.

Figure 2: Nonlinear parameter transformations



$\lambda = 0.5, \phi_1 = 0.1$



$\alpha_1 = 0.5, \phi_1 = 0.1$

Note: The top panel presents the implied unrestricted reduced form parameters in (6) given structural form parameters in (5). The bottom panel presents implied structural form parameters in (5) given unrestricted reduced form parameters in (6). Parameter transformations are obtained using the RRF restrictions in (7).

(5). 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 and the (observation) variance priors are of inverse-Wishart type[2]

$$\begin{aligned} p(\lambda, \gamma_f, \phi_1, \phi_2|\Sigma) &\propto \text{ constant for } |\lambda| < 1,\ |\gamma_f| < 1,\ |\phi_1| + \phi_2 < 1,\ |\phi_2| < 1, \\ \Sigma &\sim IW(1, 20 \times \tilde{\Sigma}), \end{aligned} \tag{8}$$

where $IW(\nu, \Psi)$ is the inverse Wishart density with scale Ψ and degrees of freedom ν. It is possible to use economic theory or steady state relationships to construct priors for these parameters, see Del Negro and Schorfheide (2008). We do not follow this approach but let the data information dominate our relatively weak prior information. For the same reason, we perform a prior-predictive analysis and investigate the sensitivity of our posterior results with respect to the prior.

Note that the prior specifications of the observation and state covariances are important in this class of models and for 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 (the state variation) can be cumbersome, see Canova (2012). We therefore impose a data based prior on the observation covariances. We first estimate an unrestricted reduced form VAR model using demeaned inflation series and (linear) detrended (log) real marginal cost series, and base the observation variance prior on this covariance 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 priors for the state variances

$$\sigma_{\eta_1} \sim IG(20, 20 \times 10^{-2}), \quad \sigma_{\eta_2} \sim IG(20, 20 \times 10^{-3}), \quad \sigma_{\eta_3} \sim IG(1, 1 \times 10^{-5}), \tag{9}$$

[2]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.

where $IG(\alpha, \alpha\xi)$ is the inverse-Gamma distribution with shape α and scale $\alpha\xi$. Parameters α and ξ are the a priori number and variance of dummy observations.

Similar to the standard counterparts, the extended NKPC and HNKPC models may also suffer from flat likelihood functions. We therefore set weakly 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 much less than the number of observations to limit the prior information.

The fourth prior distribution we consider is applicable to the NKPC and HNKPC 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, this parameter can be estimated together with other model parameters. However, often the limited number of level shifts 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).

C.2 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 one relatively weak instrument (the second lag of the marginal cost series) the posterior will have very fat tails and the existence of the posterior distribution is ensured through priors defined on a bounded region, see Zellner et al. (2013) for a detailed analysis of a linear IV model with small numbers of weak instruments.

The MCMC sampler for the full conditional posterior distribution is based on Gibbs sampling with a Metropolis-Hastings step and data augmentation, combining the methodologies in Geman and Geman (1984); Tanner and Wong (1987); Gerlach et al. (2000) and Çakmaklı et al. (2011).

Together with the level specifications of the inflation and real marginal cost series the proposed extended NKPC model takes the following form

$$
\begin{aligned}
\pi_t - c_{\pi,t} &= \tfrac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\left(z_t - c_{z,t}\right) + \tfrac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\left(z_{t-1} - c_{z,t-1}\right) + \epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}, \\
c_{\pi,t+1} &= c_{\pi,t} + \kappa_t\eta_{1,t+1}, \\
c_{z,t+1} &= \mu_{z,t} + c_{z,t} + \eta_{2,t+1}, \\
\mu_{z,t+1} &= \mu_{z,t} + \eta_{3,t+1},
\end{aligned}
\tag{10}
$$

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 disturbances $(\epsilon_{1,t},\ \epsilon_{2,t})'$ and $(\eta_{1,t},\ \eta_{2,t},\ \eta_{3,t})'$ are independent for all t.

The NKPC model in (10) can be cast into the state-space form as follows

$$
\begin{aligned}
Y_t &= HX_t + BU_t + \epsilon_{\mathbf{t}}, \quad \epsilon_{\mathbf{t}} \sim N(0, Q_t) \\
X_t &= FX_{t-1} + R_t\eta_{\mathbf{t}}, \qquad \eta_{\mathbf{t}} \sim N(0, I)
\end{aligned}
\tag{11}
$$

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 (11) 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 NKPC 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 level shift probability, p_κ, and by drawing 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 $p(X_t|\theta^{(m)}, Y_{1:T}, U_{1:T}, R_{1:T}, Q_{1:T})$ for $t = 1, 2, \ldots, T$.

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

5. Sample $\kappa_t^{(m)}$ from $p(\kappa_t^{(m)}|\theta^{(m)}, Y_{1:T}, h_{1:T}^{(m)}, U_{1:T}, R_{1:T}, \rho, \sigma^2_{\epsilon_2})$ for $t = 1, 2, \ldots, T$.

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

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

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

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).

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 (11) can be rewritten as

$$\begin{aligned} \tilde{\pi}_t &= \tfrac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\tilde{z}_t + \tfrac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\tilde{z}_{t-1} + \varepsilon_t \\ \tilde{z}_t &= \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + \epsilon_{2,t}. \end{aligned} \tag{12}$$

Posterior distributions of the structural parameters under flat priors are non-standard since z_t term also is on the right hand side of (12) and the model is highly non-linear in parameters. We therefore use two Metropolis Hastings steps to sample these structural parameters, see Metropolis et al. (1953) and 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 (12) 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 $t = 1, \ldots, T$

$$\begin{aligned}
X_{t|t-1} &= FX_{t-1|t-1} \\
P_{t|t-1} &= FP_{t-1|t-1}F' + R_t'R_t \\
\eta_{t|t-1} &= y_t - HX_{t|t-1} - BU_t \\
\zeta_{t|t-1} &= HP_{t|t-1}H' + Q_t \\
K_t &= P_{t|t-1}H'\zeta_{t|t-1}' \\
X_{t|t} &= X_{t|t-1} + K_t\eta_{t|t-1} \\
P_{t|t} &= P_{t|t-1} - K_tH'\zeta_{t|t-1}',
\end{aligned} \tag{13}$$

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 $t = T-1, \ldots, 1$

$$\begin{aligned}
\eta^*_{t+1|t} &= X_{t+1} - FX_{t|t} \\
\zeta^*_{t+1|t} &= FP_{t|t}F' + R_{t+1}'R_{t+1} \\
X_{t|t,X_{t+1}} &= X_{t|t} + P_{t|t}F'\zeta^{*-1}_{t+1|t}\eta^*_{t+1|t} \\
P_{t|t,P_{t+1}} &= P_{t|t} - P_{t|t}F'\zeta^{*-1}_{t+1|t}FP_{t|t}.
\end{aligned} \tag{14}$$

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 $P_{0|t,P_1} = F^{-1}(P_1 + R_1'R_1)F'^{-1}$ 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). For the estimation of the volatilities in the BVAR-TV-SV model we use the extension of the algorithm following Kastner and Frühwirth-Schnatter (2013) for improving the efficiency of the MCMC algorithm.

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^2_η

Using standard results from a linear regression model with a conjugate prior for the variances in (11), 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 $\Phi_{\eta_i} + \sum_{t=1}^{T} \eta^2_{i,t}$ 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 costs variance and correlation coefficient

To sample the variance of marginal costs 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

$$\prod_{t=1}^{T} f(\epsilon_{\mathbf{t}}) = \prod_{t=1}^{T} \frac{1}{\sigma_{\epsilon_{1,t}}} \phi\left(\frac{\epsilon_{1,t}}{\sigma_{\epsilon_{1,t}}}\right) \frac{1}{\sigma_{\epsilon_{2,t}} \sqrt{(1-\rho^2)}} \phi\left(\frac{\epsilon_{2,t} - \rho\epsilon_{1,t}}{\sigma_{\epsilon_{2,t}}(1-\rho^2)}\right), \tag{15}$$

Hence, together with prior for the variance in (11), variance of the marginal cost series can be sampled using (15) 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 (15). Conditional on the remaining parameters the posterior becomes

$$(1-\rho^2)^{-\frac{3}{2}} \prod_{t=1}^{T} \left(\frac{1}{\sqrt{(1-\rho^2)}} \phi\left(\frac{\epsilon_{2,t} - \rho\epsilon_{1,t}}{\sigma_{\epsilon_{2,t}}(1-\rho^2)}\right) \right). \tag{16}$$

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.

C.3 Prior-predictive likelihood analysis

In the proposed models, it is important to assess the effects of the specified prior distributions on the predictive likelihoods. Due to the nonlinear structure of the 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 NKPC and HNKPC model, we consider 1000 parameter draws from the joint prior distribution and compute the prior predictive likelihoods for 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.

D Bayesian inference of the extended HNKPC model

Posterior inference of the HNKPC models with time varying parameters follow similar to section C, using the Gibbs sampler with data augmentation. Together with the level specifications of the inflation and real marginal cost series the proposed extended HNKPC model takes the following form

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \tfrac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_t - c_{z,t}\right) + \tfrac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_{t-1} - c_{z,t-1}\right), \\
&+ \tfrac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\tfrac{\gamma_f^H}{1-\gamma_f^H\beta}\left(\mu_t - c_{\pi,t}\right) + \tfrac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\left(\pi_{t-1} - c_{\pi,t-1}\right) + \tfrac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}, \\
c_{\pi,t+1} &= c_{\pi,t} + \kappa_t\eta_{1,t+1}, \\
c_{z,t+1} &= \mu_{z,t} + c_{z,t} + \eta_{2,t+1}, \\
\mu_{z,t+1} &= \mu_{z,t} + \eta_{3,t+1}.
\end{aligned} \tag{17}$$

This can be cast into the state-space form as in (11)

$$\begin{aligned}
Y_t &= HX_t + BU_t + \epsilon_{\mathbf{t}}, \quad \epsilon_{\mathbf{t}} \sim N(0, Q_t) \\
X_t &= FX_{t-1} + R_t\eta_{\mathbf{t}}, \qquad \eta_{\mathbf{t}} \sim N(0, I)
\end{aligned} \tag{18}$$

using the following definitions

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

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

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

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

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

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

$$\alpha_3 = \frac{\gamma_b^H\gamma_f^H}{\left(1-\gamma_b^H\gamma_f^H\right)}\frac{\gamma_f^H}{\left(1-\gamma_f^H\beta\right)}, \quad \alpha_4 = \frac{\gamma_b^H}{\left(1-\gamma_b^H\gamma_f^H\right)}.$$

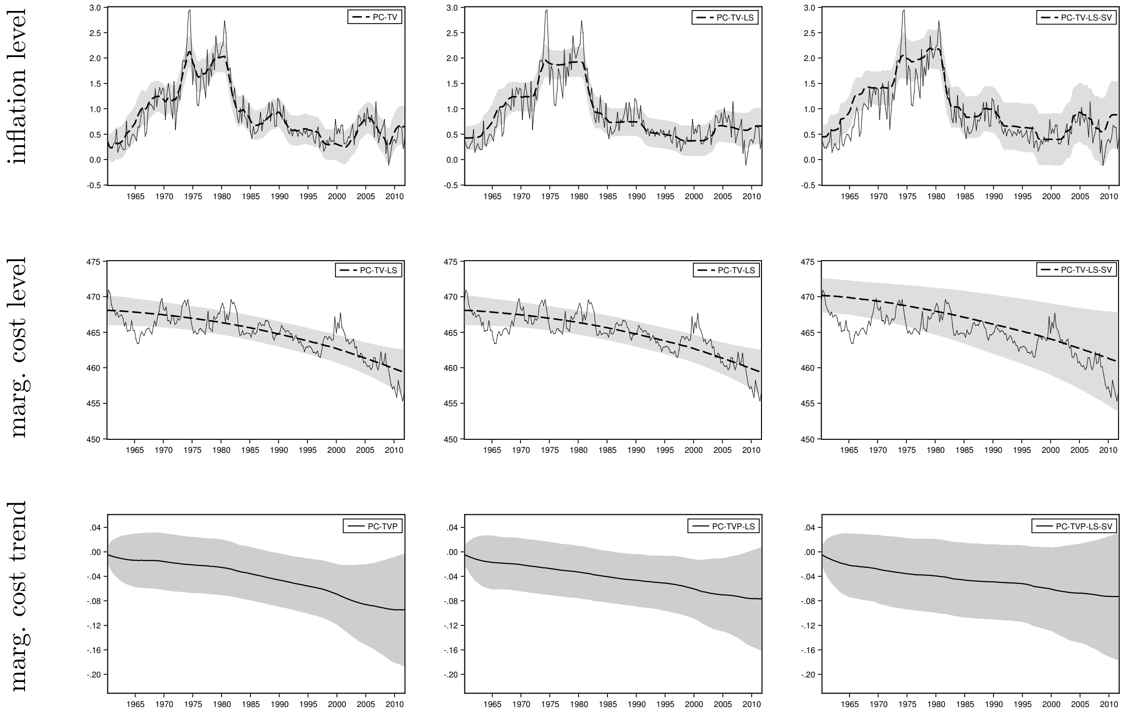
Given this setup, posterior inference can be carried out using the steps outlined in section C.

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

This section presents additional estimation results for the NKPC models with non-filtered time series. We summarize the estimated levels, volatilities, breaks and inflation expectations obtained from the NKPC-TV, NKPC-TV-LS and NKPC-TV-LS-SV models. Figure 3 shows the estimated levels from the three NKPC models. Estimated inflation levels, computed as the posterior mean of the smoothed states, are given in the first row of Figure 3. 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 3. 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 3 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

Figure 3: Level, trend and slope estimates from the NKPC models



Note: The top panel exhibits estimated inflation levels. The middle and the bottom panels show estimated real marginal cost levels and slopes, respectively. Grey shaded areas correspond to the 95% HPDI. NKPC-TV refers to the NKPC model with time varying levels and trends. NKPC-TV-LS refers to the NKPC model with time varying levels and trends. NKPC-TV-LS-SV refers to the NKPC model with time varying levels, trends and volatility. HNKPC-TV refers to the Hybrid NKPC model with time varying levels, trends and inflation expectations. HNKPC-TV-LS refers to the HNKPC model with time varying levels, trends and inflation expectations. HNKPC-TV-LS-SV refers to the HNKPC 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.

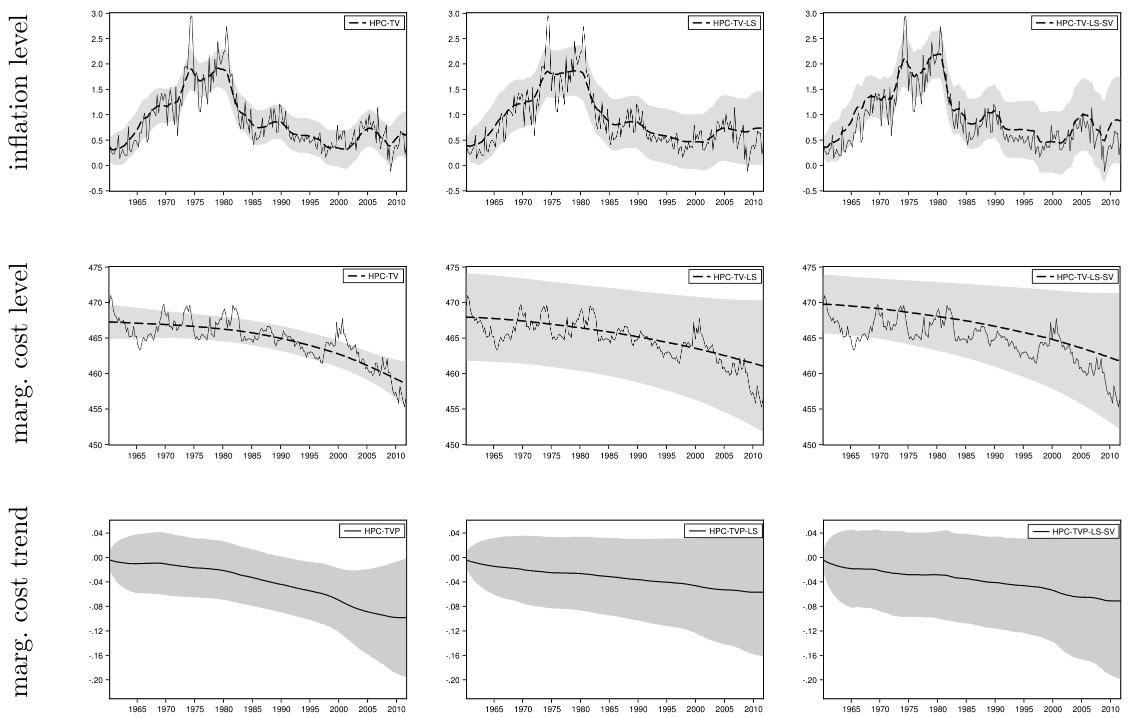
all models are given in the bottom panel of Figure 3, together with the 95% HPDIs. Nonlinearity of the negative trend is reflected in the negative values for the slope of the trend, with an increasing magnitude at the end of the sample. This change in the slope of the trend is accompanied by the increasing uncertainty about the slope. The difference between the models in terms of the estimated marginal cost structures is negligible.

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

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

Figure 4 presents the estimated inflation levels, together with estimated levels and trends of the marginal cost series.

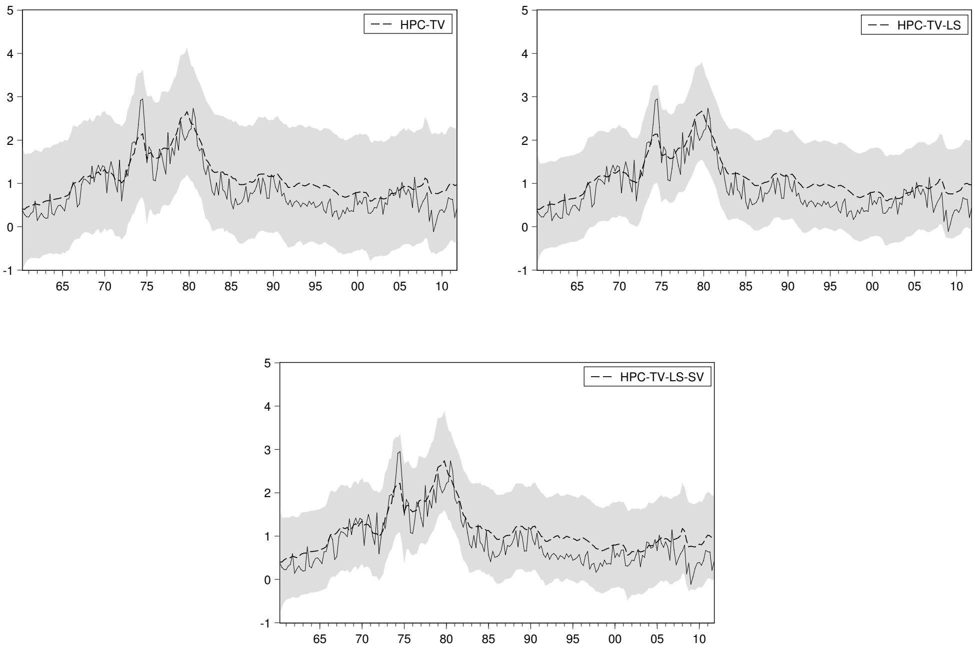
Figure 4: Level, trend and slope estimates from the HNKPC models



Note: The top panel exhibits estimated inflation levels. The middle and the bottom panels show estimated real marginal cost levels and slopes, 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.

Figure 5 presents the estimated inflation expectations together with observed survey based inflation expectations.

Figure 5: Implied inflation expectations by HNKPC models



Note: The thick solid lines are the posterior means of inflation expectations from the HNKPC 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.

G Predicted inflation densities from all proposed models

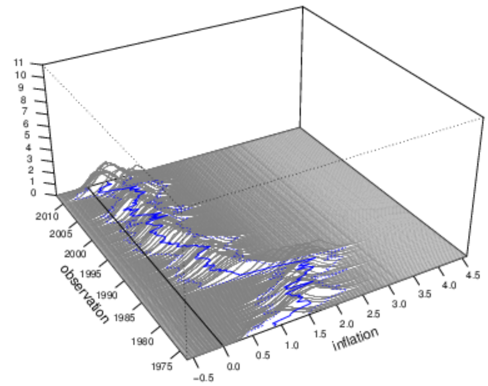
This section presents the entire distribution of the inflation predictions for all NKPC and HNKPC models. Predicted inflation densities from all proposed models are presented in Figure 6. In these figures, the solid lines represent the posterior mean of predicted inflation, and the white areas under 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 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.

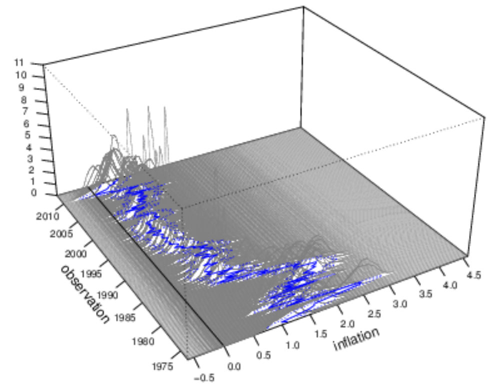
H Prior-predictive likelihoods of proposed models

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 C for the extended NKPC models, for the forecast sample analyzed earlier, covering the period between 1973-II and 2012-I. Table 1 presents the average and cumulative prior predictive likelihoods for the forecast sample. Prior predictive likelihoods, not using the data information and also using weak prior information, naturally perform worse than the predictive results reported in Table 3. Table 1 also shows that the adopted prior distributions clearly favor the less parameterized model, NKPC-TV. Moreover, the priors clearly do not favor models with stochastic volatility components. Most importantly, the 'best performing model' according to the predictive results in Table 3, HNKPC-TV-LS-SV, is the least favorable one according to the adopted prior distributions using the same forecast sample. We therefore conclude that data information is dominant, and the superior predictive performance of the HNKPC-TV-LS-SV model is not driven by the prior distribution.

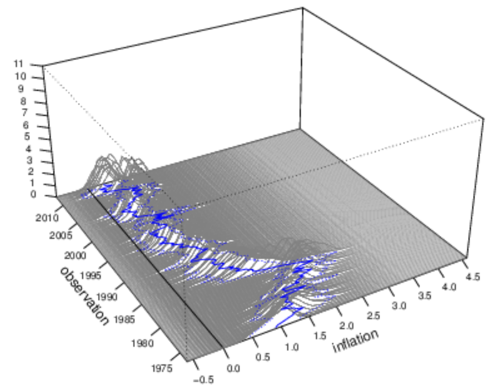
Figure 6: Predicted inflation densities from NKPC and HNKPC models



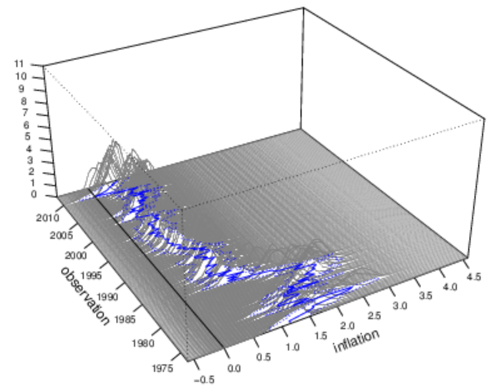
PC-TV



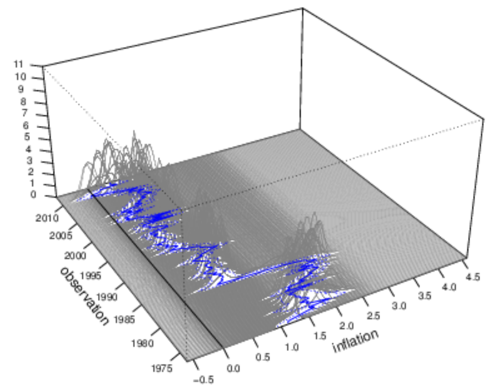
HPC-TV



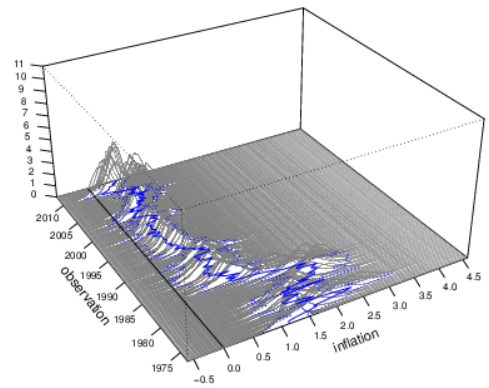
PC-TV-LS



HPC-TV-LS



PC-TV-LS-SV



HPC-TV-LS-SV

Note: The figure presents one period ahead predictive distributions of inflation from the NKPC and HNKPC models, for the period between the third quarter of 1973 and the first quarter of 2012. Model abbreviations are as in Figure 3 . Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Table 1: Prior-predictive results for the NKPC models

Model	Average (Log) Pred. Likelihood	Cumulative (Log) Pred. Likelihood
NKPC-TV	-1.16	-180.88
NKPC-TV-LS	-1.36	-210.91
NKPC-TV-LS-SV	-1.45	-224.66
HNKPC-TV	-1.28	-199.22
HNKPC-TV-LS	-1.27	-197.68
HNKPC-TV-LS-SV	-2.04	-318.77

Note: The table reports the prior-predictive performances of all competing models for the prediction sample over the period 1973-II until 2012-I. 'Average (Cumulative) Log Pred. Likelihood' stands for the average (sum) of the natural logarithms of predictive likelihoods. Results are based on 1000 simulations from the joint priors of model parameters. Model abbreviations are as in Table 1 in the paper.

I Posterior and predictive results from alternative models for robustness checks

The proposed NKPC and HNKPC 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 HNKPC models enables the use of more data information. Furthermore, extended and standard HNKPC models use the additional information from a backward looking component for the inflation series compared to the HNKPC counterparts. According to the predictive results, the most comprehensive model, HNKPC-TV-LS-SV is also the best performing model. However, a deeper analysis is needed in order to see the added predictive gains from each of 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 2 presents all NKPC and HNKPC model structures we compare to differentiate these effects.

Table 2: Standard and extended NKPC models

low/high frequencies \ model structure	iterated expectations solution		direct expectations data	
	NKPC	HNKPC	NKPC	HNKPC
linear trend	NKPC-LT	n/a ⋆	NKPCS-LT	HNKPCS-LT
Hodrick-Prescott filter	NKPC-HP	n/a ⋆	NKPCS-HP	HNKPCS-HP
time varying levels	NKPC-TV	HNKPC-TV	NKPCS-TV	HNKPCS-TV
time varying levels and switching	NKPC-TV-LS	HNKPC-TV-LS	NKPCS-TV-LS	HNKPCS-TV-LS
time varying levels and stochastic volatility	NKPC-TV-SV	HNKPC-TV-SV	NKPCS-TV-SV	HNKPCS-TV-SV
time varying levels, switching and stochastic volatility	NKPC-TV-LS-SV	HNKPC-TV-LS-SV	NKPCS-TV-LS-SV	HNKPCS-TV-LS-SV

Note: The first two columns present the standard and extended (H)NKPC models presented in the main paper, for which expectational mechanisms are solved explicitly. The last two columns present alternative model structures for (H)NKPC models. For these models, we do not iterate inflation expectations in the models, but instead replace them with survey data directly. NKPC(S)-LT (NKPC-HP(S)) refers to the NKPC model where the real marginal cost series is detrended using linear trend (Hodrick-Prescott) filter. For the remaining models real marginal cost series' trend is modeled using local linear trend model. NKPC(S)-TV refers to the NKPC model with time varying inflation levels. NKPC(S)-TV-LS refers to the NKPC model with time varying inflation levels together with level shifts. NKPC(S)-TV-SV refers to the NKPC model with time varying inflation levels and stochastic volatility. NKPC(S)-TV-LS-SV refers to the NKPC model with time varying inflation levels together with level shifts and stochastic volatility. HNKPC(S)-TV refers to the Hybrid NKPC model with time varying levels and inflation expectations. HNKPC(S)-TV-LS refers to the HNKPC model with time varying levels together with level shifts and inflation expectations. HNKPC(S)-TV-SV refers to the HNKPC model with time varying levels, inflation expectations and stochastic volatility. HNKPC(S)-TV-LS-SV refers to the HNKPC model with time varying levels together with level shifts, inflation expectations and stochastic 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 NKPC and HNKPC 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 2 and are abbreviated by NKPCS-LT, NKPCS-HP, HNKPCS-LT and HNKPCS-HP, according to linear detrending or HP detrending prior to analysis. The improved predictive performances of NKPCS-LT and NKPCS-HP models compared to the standard NKPC counterparts show predictive gains from incorporating survey expectations in the models. Furthermore, comparing the predictive performances of the HNKPCS-LT and HNKPCS-HP models with the time-varying hybrid models, such as the HNKPC-TV or HNKPC-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 2, are NKPC 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 (5) where the expectation term is replaced by survey expectations. We denote these models by NKPCS-TV, NKPCS-TV-LS and NKPCS-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 in inflation, respectively. Comparing the predictive results of these models to the HNKPC counterparts provide the predictive gains solely from the HNKPC 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 HNKPC models using the survey expectations directly, without solving for the expectational mechanisms. We denote these models by HNKPCS-TV, HNKPCS-TV-LS and HNKPCS-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 HNKPC models clarifies the predictive gains from solving for the inflation expectations iteratively in the hybrid models.

The final set of alternative models aim to pinpoint predictive gains from introducing level shifts in inflation in the models with a stochastic volatility component. The comparison of the predictive results of models with time-varying levels and stochastic volatility, (H)NKPC-TV-SV, and with level shifts and stochastic volatility, (H)NKPC-TV-LS-SV, highlights predictive gains solely from introducing level shifts when changes in inflation volatility are taken into account.

One period ahead MSFE and log marginal likelihoods of these models, together with the standard (H)NKPC models and the models proposed in the paper, are given in Table 3. The prediction results are based on the forecast sample, which covers the period between the second quarter of 1973 and the first quarter of 2012. Comparing the first block and the first two rows of the second block Table 3, we see that the gains from using survey data inflation are substantial even in the standard NKPC 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 specifically from introducing time-variation in the inflation and marginal cost series, by comparing the results of the HNKPCS-LT and HNKPCS-HP models with the HNKPC-TV or HNKPC-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 possible reason 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 3. The prediction results from the NKPC and HNKPC 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 NKPCS-TV, NKPCS-TV-LS and NKPCS-TV-LS-SV models with HNKPCS-TV, HNKPCS-TV-

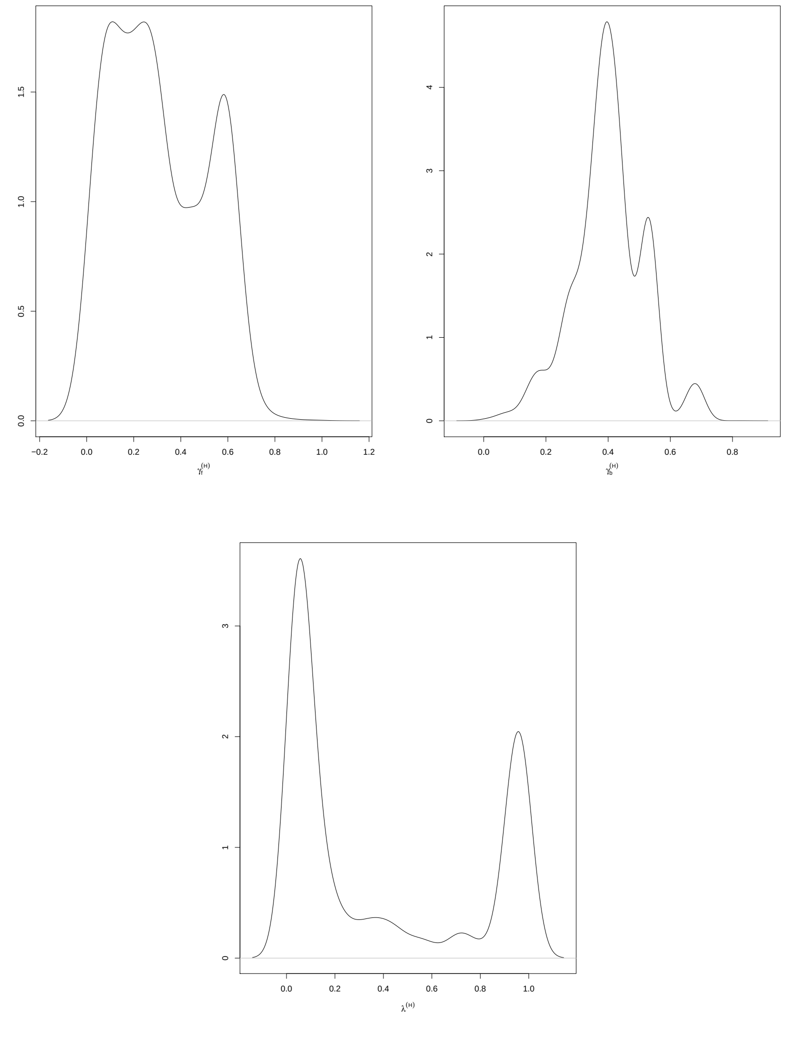
LS and HNKPCS-TV-LS-SV models, respectively. In all these comparisons, the models without the backward looking component performs slightly better (worse) in terms of MSFE (marginal likelihood), hence the backward looking component does not seem to improve predictive results in general and the improvements 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 (NKPCS-TV-LS-SV and HNKPCS-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 3, 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 section. Table 4 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 7 shows this irregularity for the HNKPCS-TV model, parameters $\lambda^{(H)}$, $\gamma_b^{(H)}$ and $\gamma_f^{(H)}$. The bimodality problem in posterior densities is most apperant in the NKPC 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)NKPC models with survey expectations deteriorate posterior inference compared to the iterative solution of these expectational terms.

Figure 7: Posterior density of $\lambda^{(H)}$, $\gamma_b^{(H)}$ and $\gamma_f^{(H)}$ from the HNKPCS-TV model



Note: The figure presents posterior densities of parameters from the HNKPCS-TV model. Model abbreviations are based on Table 2. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Table 3: Predictive performance of additional NKPC models

Model	(Log) Marg. Likelihood	MSFE 1 period ahead
NKPC-LT	-139.327	0.353
NKPC-HP	-157.195	0.458
NKPCS-LT	-79.141	0.105
NKPCS-HP	-85.397	0.130
HNKPCS-LT	-81.047	0.105
HNKPCS-HP	-85.200	0.119
NKPC-TV	-46.162	0.142
NKPC-TV-LS	-61.972	0.138
NKPC-TV-SV	-22.761	0.134
NKPC-TV-LS-SV	-33.476	0.126
HNKPC-TV	-36.385	0.123
HNKPC-TV-LS	-35.052	0.105
HNKPC-TV-SV	-19.695	0.106
HNKPC-TV-LS-SV	-18.150	0.091
NKPCS-TV	-34.407	0.129
NKPCS-TV-LS	-32.004	0.099
NKPCS-TV-LS-SV	-15.390	0.092
HNKPCS-TV	-40.465	0.176
HNKPCS-TV-LS	-38.082	0.297
HNKPCS-TV-LS-SV	-12.977	0.139
BVAR (constant)	-166.226	0.085
BVAR-TV-SV	-97.980	0.100
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 2. BVAR (constant) denotes the BVAR model with 2 lags and with constant parameters. 'BVAR-TV-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 4: Posterior results of alternative NKPC models

Model	$\lambda^{(H)}$	$\gamma_f^{(H)}$	γ_b^H	ρ	ϕ_1	ϕ_2
NKPCS-LT	0.011 (0.051)	0.611 (0.055)	–	-0.016 (0.021)	0.824 (0.047)	0.075 (0.044)
NKPCS-HP	0.064 (0.051)	0.627 (0.081)	–	-0.045 (0.064)	0.681 (0.096)	0.014 (0.081)
HNKPCS-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)
HNKPCS-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)
NKPCS-TV	0.057 (0.028)	0.142 (0.086)	–	-0.034 (0.061)	0.815 (0.052)	0.067 (0.052)
NKPCS-TV-LS	0.049 (0.023)	0.430 (0.125)	–	-0.027 (0.050)	0.821 (0.054)	0.072 (0.052)
NKPCS-TV-LS-SV	0.058 (0.025)	0.307 (0.165)	–	-0.015 (0.068)	0.826 (0.052)	0.078 (0.053)
HNKPCS-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)
HNKPCS-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)
HNKPCS-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 2.

To conclude, predictive gains obtained from including the survey expectations in the models are substantial and incorporating the low and high frequency data movements in the model is crucial. These two conclusions are in line with Faust and Wright (2013), who consider a large set of alternative models for inflation forecasting, including unrestricted reduced form models, and compare their forecast performances based on MSFE. Our model incorporates both these features in the NKPC model structure. Third, once survey data and time variation are included in the model, there are still additional predictive gains from the backward looking component in the hybrid models.

J Modeling inflation expectations using unobserved components

The HNKPC models implicitly assume that survey based inflation expectations capture 'real' inflation expectations for the next period accurately. However, survey expectations are likely to reflect real inflation expectations with a measurement error. In this section we extend the HNKPC model by including a latent variable for unobserved inflation expectations, aiming to account for the possibility of measurement errors in survey expectations. Specifically, we propose an adaptive rule under which inflation expectations partially adjust to survey expectations at each period:

$$S_{t+1} = \mu_{t+1} + \beta_S(S_t - \mu_t) + \eta_{S,t+1}, \tag{19}$$

where $|\beta_S| < 1$ and μ_t is the survey observation for inflation expectation at time t. This adaptive rule implies that unobserved inflation expectations converge to the survey based expectations in the long run. Given the restriction on parameter β_S, one can solve (19) for S_t and obtain $S_t = \mu_t + \sum_{j=0}^{\infty} \beta_S^j \eta_{S,t-j}$. 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 next consider the HNKPC model given the specified adaptive rule for the unobserved inflation expectations. Notice that we can factorize the expectation term in equation (9) in the main text of the paper, $E_t\left(\tilde{\pi}_{t+k}\right)$, into two parts related to the measurement error and the relation between survey based expectations and long run expectations, as $E_t\left(\tilde{\pi}_{t+k}\right) = E_t\left(S_{t+k-1} - \mu_{t+k-1}\right) + E_t\left(\mu_{t+k-1} - c_{\pi,t+k}\right)$. Then the weighted sum of expectations in equation (9) in the paper becomes

$$\sum_{k=1}^{\infty} \gamma_f^k E_t\left(\tilde{\pi}_{t+k}\right) = \sum_{k=1}^{\infty} \gamma_f^k E_t\left(S_{t+k-1} - \mu_{t+k-1}\right) + \sum_{k=1}^{\infty} \gamma_f^k E_t\left(\mu_{t+k-1} - c_{\pi,t+k}\right). \quad (20)$$

The first part of the summation, $\sum_{k=1}^{\infty} \gamma_f^k E_t\left(S_{t+k-1} - \mu_{t+k-1}\right)$, is related to the measurement error and can be computed from (19). For the second part of the summation, $\sum_{k=1}^{\infty} \gamma_f^k E_t\left(\mu_{t+k-1} - c_{\pi,t+k}\right)$, we specify a similar partial adjustment process as the process specified in the paper $\mu_t - c_{\pi,t+1} = \beta_\mu(\mu_{t-1} - c_{\pi,t}) + \eta_{\mu,t+1}$. The partial adjustment mechanism implies that the further one gets into the future the smaller will be the difference between short and long run inflation expectations. Estimates of β_μ will indicate the empirical speed of adjustment. For instance, for a value of the posterior mean of β_μ equal to 0.5 it follows that within a few periods one has almost complete adjustment.

Replacing the infinite sum of expectations of inflation deviations using the two specifications for the measurement error and for the deviation of the survey expectations from the long run inflation expectations in (20), the HNKPC model becomes

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \tfrac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_t - c_{z,t}\right) + \tfrac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_{t-1} - c_{z,t-1}\right) \\
&+ \tfrac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\left(\tfrac{\gamma_f^H}{1-\gamma_f^H\beta_S}(S_t - \mu_t) + \tfrac{\gamma_f^H}{1-\gamma_f^H\beta_\mu}(\mu_t - c_{\pi,t})\right) \\
&+ \tfrac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\left(\pi_{t-1} - c_{\pi,t-1}\right) + \tfrac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}.
\end{aligned} \tag{21}$$

Notice that if the speed of adjustment for both specifications are equal, i.e. $\beta_S = \beta_\mu$, then the HNKPC reduces to

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \tfrac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_t - c_{z,t}\right) + \tfrac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_{t-1} - c_{z,t-1}\right) \\
&+ \tfrac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\tfrac{\gamma_f^H}{1-\gamma_f^H\beta_S}(S_t - c_{\pi,t}) + \tfrac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\left(\pi_{t-1} - c_{\pi,t-1}\right) + \tfrac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}.
\end{aligned} \tag{22}$$

We next compare the models specified in (21) and in (22) with a HNKPC-TV parametrization in terms of their forecast performances. For the forecast sample considered in the paper, the cumulative predictive likelihood for the HNKPC-TV model in (21) is -36.19 while for the model in (22) this value is -36.44. The cumulative predictive likelihood values for the HNKPC-TV model with and without the restriction $\beta_S = \beta_\mu$ indicate that this restriction is statistically valid as the difference between the likelihood values are very small. Following this evidence we display the parameter estimates of all extended HNKPC models using the expectation specification in (22) in Table 5. We further report the cumulative predictive likelihood values and 1 step ahead MSFE for these models in Table 6.

Results are very similar to the corresponding table in the paper (Table 2), thus, we conclude that the effect of the measurement errors in survey expectations is negligible.

Table 5: Posterior results of HNKPC models with unobserved inflation expectations

Model	λ^H	γ_f^H	γ_b^H	β_S	ρ	ϕ_1	ϕ_2
HNKPC-TV	0.05 (0.03)	0.02 (0.03)	0.38 (0.14)	0.49 (0.28)	0.01 (0.06)	0.81 (0.05)	0.07 (0.05)
HNKPC-TV-LS	0.04 (0.02)	0.01 (0.01)	0.49 (0.11)	0.52 (0.18)	0.02 (0.01)	0.79 (0.09)	0.19 (0.08)
HNKPC-TV-LS-SV	0.06 (0.02)	0.04 (0.10)	0.22 (0.12)	0.44 (0.24)	-0.01 (0.01)	0.82 (0.05)	0.15 (0.04)

Note: The table presents posterior means and standard deviations (in parentheses) of parameters for the competing HNKPC type models estimated for quarterly inflation and real marginal costs over the period 1960-I until 2012-I. λ^H and γ_f^H are the slope of the Phillips curve and the coefficient of inflation expectations in HNKPC model in (22). γ_b^H is the coefficient of the backward looking component in the HNKPC model in (22). β_S is the autoregressive parameter for the deviation of inflation expectations, as used in (22). 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. Posterior results are based on 40000 simulations of which the first 20000 are discarded for burn-in. Model abbreviations are as in Table 1 in the paper.

Table 6: Predictive performance of HNKPC models with unobserved inflation expectations

Model	Cumulative (Log) Pred. Likelihood	MSFE 1 period ahead
HNKPC-TV	-36.44	0.12
HNKPC-TV-LS	-35.77	0.09
HNKPC-TV-LS-SV	-17.96	0.09

Note: The table reports the predictive performances of competing models for the prediction sample over the period 1973-II until 2012-I. 'Cumulative (Log) Pred. Likelihood' stands for the sum of the natural logarithms of predictive likelihoods. 'MSFE' stands for the Mean Squared Forecast Error. Results are based on 10000 simulations of which the first 5000 are discarded for burn-in. Remaining abbreviations are as in Table 1 in the paper.

K 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 NKPC-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).

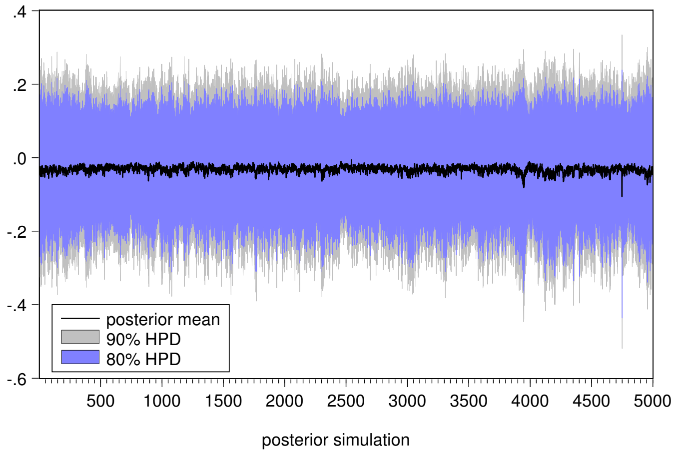
We perform a two step analysis, where in the first step we obtain the residuals from the regression of the estimated level of inflation on a constant and the estimated level of marginal costs, for each posterior draw. This implies that we take the estimation uncertainty in the analysis into account. 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 $\rho \in [-1, 1]$

$$\Delta\hat{\epsilon}_t = \rho\hat{\epsilon}_{t-1} + \eta_t, \;\; \eta_t \sim NID(0, \sigma^2), \tag{23}$$

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 8. 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-NKPC 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 NKPC models we consider.

Figure 8: Cointegration analysis for the marginal costs and inflation series



Note: The figure presents the posterior means and quantiles of the ρ parameter from 5×10^3 posterior draws from the NKPC-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

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.

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.

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

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. In *US Monetary Policy Forum*, volume 8.

Del Negro M, Schorfheide F. 2008. Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics* **55**: 1191–1208.

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

Faust J, Wright JH. 2013. Forecasting inflation. In Elliott G, Timmermann A (eds.) *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*. Elsevier, 2–56.

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

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.

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**: 97–109.

Kastner G, Frühwirth-Schnatter S. 2013. Ancillarity-sufficiency interweaving strategy (asis) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics & Data Analysis* .

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**: 361–393.

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

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.

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.

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

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

Zellner A, Ando T, Baştürk N, Hoogerheide L, Van Dijk HK. 2013. Bayesian analysis of instrumental variable models: Acceptance-rejection within Direct Monte Carlo. *Econometric Reviews* **33**: 3–35.