# Online Appendix for *Evaluating DSGE Model Forecasts of Comovements*

**Edward Herbst and Frank Schorfheide**

## A Prior Distribution for Small-Scale Model

The prior distribution for the small-scale model is summarized in Table A-1. Priors for the autocorrelations and standard deviations of the exogenous processes, the steady-state parameters $\bar{\gamma}$, $\pi^{(A)}$, and $r^{(A)}$, as well as the standard deviation of the monetary policy shock, are quantified based on regressions on pre-1984 observations of output growth, inflation, and nominal interest rates. The priors for the policy rule coefficients $\psi_2$ and $\rho_R$ are loosely centered around Taylor (1993)'s values. The prior for the parameter that governs price stickiness is chosen based on micro-evidence on price setting-behavior provided. More formal methods for the elicitation of priors for DSGE model parameters are discussed in Del Negro and Schorfheide (2008).

Table A-1: Prior Distribution for Small-Scale Model

| | Density | Para (1) | Para (2) |
|---|---|---|---|
| $\tau$ | Gamma | 2.00 | 0.50 |
| $\kappa$ | Gamma | 0.20 | 0.10 |
| $\psi_2$ | Gamma | 0.50 | 0.25 |
| $\rho_R$ | Beta | 0.50 | 0.20 |
| $\rho_G$ | Beta | 0.80 | 0.10 |
| $\rho_Z$ | Beta | 0.66 | 0.15 |
| $r^{(A)}$ | Gamma | 0.50 | 0.50 |
| $\pi^{(A)}$ | Gamma | 7.00 | 2.00 |
| $\bar{\gamma}$ | Normal | 0.40 | 0.20 |
| $\sigma_R$ | InvGamma | 0.40 | 4.00 |
| $\sigma_G$ | InvGamma | 1.00 | 4.00 |
| $\sigma_Z$ | InvGamma | 0.50 | 4.00 |

*Notes:* The following parameter is fixed: $\psi_1 = 1.70$. Para (1) and Para (2) list the means and the standard deviations for Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; and $s$ and $\nu$ for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu, s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region.

# B The Smets-Wouters Model

The equilibrium conditions of the Smets and Wouters (2007) model take the following form:

$$\begin{aligned}
\hat{y}_t &= c_y\hat{c}_t + i_y\hat{i}_t + z_y\hat{z}_t + \varepsilon_t^g && \text{(A.1)}\\
\hat{c}_t &= \frac{h/\gamma}{1+h/\gamma}\hat{c}_{t-1} + \frac{1}{1+h/\gamma}E_t\hat{c}_{t+1} + \frac{wl_c(\sigma_c-1)}{\sigma_c(1+h/\gamma)}(\hat{l}_t - E_t\hat{l}_{t+1}) && \text{(A.2)}\\
&\quad -\frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}(\hat{r}_t - E_t\hat{\pi}_{t+1}) - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}\varepsilon_t^b\\
\hat{i}_t &= \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}\hat{i}_{t-1} + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}E_t\hat{i}_{t+1} + \frac{1}{\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})}\hat{q}_t + \varepsilon_t^i && \text{(A.3)}\\
\hat{q}_t &= \beta(1-\delta)\gamma^{-\sigma_c}E_t\hat{q}_{t+1} - \hat{r}_t + E_t\hat{\pi}_{t+1} + (1-\beta(1-\delta)\gamma^{-\sigma_c})E_t\hat{r}_{t+1}^k - \varepsilon_t^b && \text{(A.4)}\\
\hat{y}_t &= \Phi(\alpha\hat{k}_t^s + (1-\alpha)\hat{l}_t + \varepsilon_t^a) && \text{(A.5)}\\
\hat{k}_t^s &= \hat{k}_{t-1} + \hat{z}_t && \text{(A.6)}\\
\hat{z}_t &= \frac{1-\psi}{\psi}\hat{r}_t^k && \text{(A.7)}\\
\hat{k}_t &= \frac{(1-\delta)}{\gamma}\hat{k}_{t-1} + (1-(1-\delta)/\gamma)\hat{i}_t + (1-(1-\delta)/\gamma)\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})\varepsilon_t^i && \text{(A.8)}\\
\hat{\mu}_t^p &= \alpha(\hat{k}_t^s - \hat{l}_t) - \hat{w}_t + \varepsilon_t^a && \text{(A.9)}\\
\hat{\pi}_t &= \frac{\beta\gamma^{(1-\sigma_c)}}{1+\iota_p\beta\gamma^{(1-\sigma_c)}}E_t\hat{\pi}_{t+1} + \frac{\iota_p}{1+\beta\gamma^{(1-\sigma_c)}}\hat{\pi}_{t-1} && \text{(A.10)}\\
&\quad -\frac{(1-\beta\gamma^{(1-\sigma_c)}\xi_p)(1-\xi_p)}{(1+\iota_p\beta\gamma^{(1-\sigma_c)})(1+(\Phi-1)\varepsilon_p)\xi_p}\hat{\mu}_t^p + \varepsilon_t^p\\
\hat{r}_t^k &= \hat{l}_t + \hat{w}_t - \hat{k}_t^s && \text{(A.11)}\\
\hat{\mu}_t^w &= \hat{w}_t - \sigma_l\hat{l}_t - \frac{1}{1-h/\gamma}(\hat{c}_t - h/\gamma\hat{c}_{t-1}) && \text{(A.12)}\\
\hat{w}_t &= \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}(E_t\hat{w}_{t+1} + E_t\hat{\pi}_{t+1}) + \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}(\hat{w}_{t-1} - \iota_w\hat{\pi}_{t-1}) && \text{(A.13)}\\
&\quad -\frac{1+\beta\gamma^{(1-\sigma_c)}\iota_w}{1+\beta\gamma^{(1-\sigma_c)}}\hat{\pi}_t - \frac{(1-\beta\gamma^{(1-\sigma_c)}\xi_w)(1-\xi_w)}{(1+\beta\gamma^{(1-\sigma_c)})(1+(\lambda_w-1)\epsilon_w)\xi_w}\hat{\mu}_t^w + \varepsilon_t^w\\
\hat{r}_t &= \rho\hat{r}_{t-1} + (1-\rho)(r_\pi\hat{\pi}_t + r_y(\hat{y}_t - \hat{y}_t^*)) && \text{(A.14)}\\
&\quad + r_{\Delta y}((\hat{y}_t - \hat{y}_t^*) - (\hat{y}_{t-1} - \hat{y}_{t-1}^*)) + \varepsilon_t^r.
\end{aligned}$$

The exogenous shocks evolve according to

$$\varepsilon_t^a = \rho_a \varepsilon_{t-1}^a + \eta_t^a \quad \text{(A.15)}$$
$$\varepsilon_t^b = \rho_b \varepsilon_{t-1}^b + \eta_t^b \quad \text{(A.16)}$$
$$\varepsilon_t^g = \rho_g \varepsilon_{t-1}^a + \rho_{ga}\eta_t^a + \eta_t^g \quad \text{(A.17)}$$
$$\varepsilon_t^i = \rho_i \varepsilon_{t-1}^i + \eta_t^i \quad \text{(A.18)}$$
$$\varepsilon_t^r = \rho_r \varepsilon_{t-1}^r + \eta_t^r \quad \text{(A.19)}$$
$$\varepsilon_t^p = \rho_r \varepsilon_{t-1}^p + \eta_t^p - \mu_p \eta_{t-1}^p \quad \text{(A.20)}$$
$$\varepsilon_t^w = \rho_w \varepsilon_{t-1}^w + \eta_t^w - \mu_w \eta_{t-1}^w. \quad \text{(A.21)}$$

The counterfactual no-rigidity prices and quantities evolve according to

$$\hat{y}_t^* = c_y \hat{c}_t^* + i_y \hat{i}_t^* + z_y \hat{z}_t^* + \varepsilon_t^g \quad \text{(A.22)}$$
$$\hat{c}_t^* = \frac{h/\gamma}{1+h/\gamma}\hat{c}_{t-1}^* + \frac{1}{1+h/\gamma}E_t\hat{c}_{t+1}^* + \frac{wl_c(\sigma_c-1)}{\sigma_c(1+h/\gamma)}(\hat{l}_t^* - E_t\hat{l}_{t+1}^*) - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}r_t^* - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}\varepsilon_t^b \quad \text{(A.23)}$$
$$\hat{i}_t^* = \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}\hat{i}_{t-1}^* + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}E_t\hat{i}_{t+1}^* + \frac{1}{\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})}\hat{q}_t^* + \varepsilon_t^i \quad \text{(A.24)}$$
$$\hat{q}_t^* = \beta(1-\delta)\gamma^{-\sigma_c}E_t\hat{q}_{t+1}^* - r_t^* + (1-\beta(1-\delta)\gamma^{-\sigma_c})E_t r_{t+1}^{k*} - \varepsilon_t^b \quad \text{(A.25)}$$
$$\hat{y}_t^* = \Phi(\alpha k_t^{s*} + (1-\alpha)\hat{l}_t^* + \varepsilon_t^a) \quad \text{(A.26)}$$
$$\hat{k}_t^{s*} = k_{t-1}^* + z_t^* \quad \text{(A.27)}$$
$$\hat{z}_t^* = \frac{1-\psi}{\psi}\hat{r}_t^{k*} \quad \text{(A.28)}$$
$$\hat{k}_t^* = \frac{(1-\delta)}{\gamma}\hat{k}_{t-1}^* + (1-(1-\delta)/\gamma)\hat{i}_t + (1-(1-\delta)/\gamma)\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})\varepsilon_t^i \quad \text{(A.29)}$$
$$\hat{w}_t^* = \alpha(\hat{k}_t^{s*} - \hat{l}_t^*) + \varepsilon_t^a \quad \text{(A.30)}$$
$$\hat{r}_t^{k*} = \hat{l}_t^* + \hat{w}_t^* - \hat{k}_t^* \quad \text{(A.31)}$$
$$\hat{w}_t^* = \sigma_l \hat{l}_t^* + \frac{1}{1-h/\gamma}(\hat{c}_t^* + h/\gamma\hat{c}_{t-1}^*). \quad \text{(A.32)}$$

The steady state (ratios) that appear in the measurement equation or the log-linearized equilibrium conditions are given by

$$\gamma = \bar{\gamma}/100 + 1 \tag{A.33}$$

$$\pi^* = \bar{\pi}/100 + 1 \tag{A.34}$$

$$\bar{r} = 100(\beta^{-1}\gamma^{\sigma_c}\pi^* - 1) \tag{A.35}$$

$$r_{ss}^k = \gamma^{\sigma_c}/\beta - (1-\delta) \tag{A.36}$$

$$w_{ss} = \left(\frac{\alpha^\alpha(1-\alpha)^{(1-\alpha)}}{\Phi {r_{ss}^k}^\alpha}\right)^{\frac{1}{1-\alpha}} \tag{A.37}$$

$$i_k = (1-(1-\delta)/\gamma)\gamma \tag{A.38}$$

$$l_k = \frac{1-\alpha}{\alpha}\frac{r_{ss}^k}{w_{ss}} \tag{A.39}$$

$$k_y = \Phi l_k^{(\alpha-1)} \tag{A.40}$$

$$i_y = (\gamma - 1 + \delta)k_y \tag{A.41}$$

$$c_y = 1 - g_y - i_y \tag{A.42}$$

$$z_y = r_{ss}^k k_y \tag{A.43}$$

$$wl_c = \frac{1}{\lambda_w}\frac{1-\alpha}{\alpha}\frac{r_{ss}^k k_y}{c_y}. \tag{A.44}$$

The prior distribution for the parameters of the SW model is summarized in Table A-2.

Table A-2: Prior Distribution for Smets-Wouters Model

| | Density | Para (1) | Para (2) | | Density | Para (1) | Para (2) |
|---|---|---|---|---|---|---|---|
| $\psi$ | Beta | 2.00 | 0.50 | $\Phi$ | Normal | 1.25 | 0.12 |
| $\rho$ | Beta | 0.75 | 0.10 | $r_y$ | Normal | 0.12 | 0.05 |
| $r_{\Delta y}$ | Normal | 0.12 | 0.05 | $\bar{\pi}$ | Gamma | 0.62 | 0.10 |
| $100(\beta^{-1} - 1)$ | Gamma | 0.25 | 0.10 | $\bar{l}$ | Normal | 875 | 10.0 |
| $\bar{\gamma}$ | Normal | 0.40 | 0.10 | $\sigma_a$ | Invgamma | 0.10 | 2.00 |
| $\sigma_b$ | Invgamma | 0.10 | 2.00 | $\sigma_g$ | Invgamma | 0.10 | 2.00 |
| $\sigma_I$ | Invgamma | 0.10 | 2.00 | $\sigma_r$ | Invgamma | 0.10 | 2.00 |
| $\sigma_p$ | Invgamma | 0.10 | 2.00 | $\sigma_w$ | Invgamma | 0.10 | 2.00 |
| $\rho_a$ | Beta | 0.50 | 0.20 | $\rho_b$ | Beta | 0.50 | 0.20 |
| $\rho_g$ | Beta | 0.50 | 0.20 | $\rho_I$ | Beta | 0.50 | 0.20 |
| $\rho_r$ | Beta | 0.50 | 0.20 | $\rho_p$ | Beta | 0.50 | 0.20 |
| $\rho_w$ | Beta | 0.50 | 0.20 | $\mu_p$ | Beta | 0.50 | 0.20 |
| $\mu_w$ | Beta | 0.50 | 0.20 | $\rho_{ga}$ | Beta | 0.50 | 0.20 |

*Notes:* The following parameters are fixed in Smets and Wouters (2007): $\delta = 0.025$, $g_y = 0.18$, $\lambda_w = 1.50$, $\varepsilon_w = 10.0$, and $\varepsilon_p = 10$. In addition, we fix: $\varphi = 5.00$, $\sigma_c = 1.5$, $h = 0.7$, $\xi_w = 0.7$, $\sigma_l = 2$, $\xi_p = 0.7$, $\iota_w = 0.5$, $\iota_p = 0.5$, $r_\pi = 2$, $\alpha = 0.3$. Para (1) and Para (2) list the means and the standard deviations for Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; and $s$ and $\nu$ for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu, s) \propto \sigma^{-\nu-1} e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region.

# C Computational Details

## C.1 Posterior Simulator

To implement the posterior predictive checks, we need to generate draws from a sequence of posterior distributions $p(\theta|Y_{1:t})$ for $t = R, \ldots, T-1$. For $t = R$ we use the Random-Walk Metropolis (RWM) algorithm in An and Schorfheide (2007). Draws for $\tau > 0$ are also generated with the RWM algorithm. However, the covariance matrix of the proposal density is constructed by re-weighting the draws from $p(\theta|Y_{1:t-1})$ with the importance weights constructed from $p(y_t|\theta, Y_{1:t-1})$.[9] We proceed in the same manner to obtain draws from the posteriors associated with the synthetic samples $(Y_{1:R}, Y^*_{R+1:t})$.

For Step 1(a) of Algorithm 1 we use the RWM algorithm to obtain 250,000 draws from $p(\theta|Y_{1:R})$ and take a subsample of $N = 500$ draws to generate 500 trajectories $Y^{*(j)}_{R+1:T}$ in Step 1(b). Since our MCMC methods have better mixing properties for posterior distributions created from simulated data and since the repeated recursive estimation of the DSGE model is computationally quite costly, we reduce the number of draws from the RWM algorithm in Step 1(c)(ii) from 250,000 to $25,000$ and take a subsample of $L = 2,500$ equally-spaced draws. For Step 1(c)(iii) we set $M = 10$ to simulate a total of $LM = 25,000$ trajectories $\tilde{y}^{(l,m)}_{t+h}$, $h = 1, \ldots, 8$. When we execute Step 2, that is, estimate the DSGE model recursively on the actual data, we increase the number of MCMC draws to 250,000 and use every 10'th, which leads to $L = 25,000$. We keep $M = 10$, which means that we are using a total of $LM = 250,000$ simulated trajectories from which point and density forecasts are computed.

[9] Strictly speaking, we are reweighting draws from the posterior distribution $p(\theta|Y^{t-1}_{1:t-1})$ with importance weights obtained from $p(y^t_t|\theta, Y^t_{1:t-1})$. In the presence of data revisions, this does not generate a posterior covariance matrix conditional on $Y^t_{1:t}$. However, the approximation appears to be sufficient to construct a good proposal density for the RWM algorithm.

## C.2 Generating Point Forecasts and PITs Based on the Output of a Posterior Simulator

In Step 1(c)(iii) of Algorithm 1 we generate draws $y_{t+h}^{(l,m)}$ from the predictive distribution that conditions on time $t$ information. To simplify the notation, we drop the $m$ superscript and assume that we have $L$ draws $y_{t+h}^{(l)}$, $l = 1, \ldots, L$. The (unconditional) point forecasts considered in this paper are defined as $\int y_{i,t+h} p(y_{i,t+h}|Y_{1:t}) dy_{i,t+h}$ and approximated by

$$\hat{y}_{i,t+h|t} = \frac{1}{L} \sum_{l=1}^{L} y_{i,t+h}^{(l)}.$$

The probability integral transformations are defined as $\int_{-\infty}^{y_{i,t+h}} p(\tilde{y}_{i,t+h}|Y_{1:t}) d\tilde{y}_{i,t+h}$ and approximated by

$$\hat{z}_{i,h,t} = \frac{1}{L} \sum_{l=1}^{L} \mathcal{I}\{\tilde{y}_{i,t+h}^{(l)} \leq y_{i,t+h}\}.$$

A Kernel approach is used to compute conditional point forecasts and PITs. The conditional point forecasts considered in this paper are defined as $\int y_{i,t+h} p(y_{i,t+h}|y_{j,t+h}, Y_{1:t}) dy_{i,t+h}$ and approximated by

$$\hat{y}_{i,t+h|j,t} = \frac{1}{L} \sum_{l=1}^{L} K\left(\frac{\tilde{y}_{j,t+h}^{(l)} - y_{j,t+h}}{b}\right) y_{i,t+h}^{(l)},$$

where $b$ is the bandwidth and $K(\cdot)$ is a normalized kernel. In our application we are using a Gaussian kernel and with the bandwidth equal to $1.06\hat{\sigma}_{j,t+h}L^{-1/5}$, where $\sigma_{j,t+h} = std(y_{j,t+h})$, the optimal bandwidth for the Gaussian kernel with Gaussian data. Adjusting the bandwidth does not qualitatively alter the results and only has a minor impact on the quantitative results. The probability integral transforms based on the conditional density forecasts are computed according to

$$\hat{z}_{i|j,h,t} = \frac{1}{L} \sum_{l=1}^{L} K\left(\frac{\tilde{y}_{j,t+h}^{(l)} - y_{j,t+h}}{b}\right) \mathcal{I}\{\tilde{y}_{i,t+h}^{(l)} \leq y_{i,t+h}\}.$$

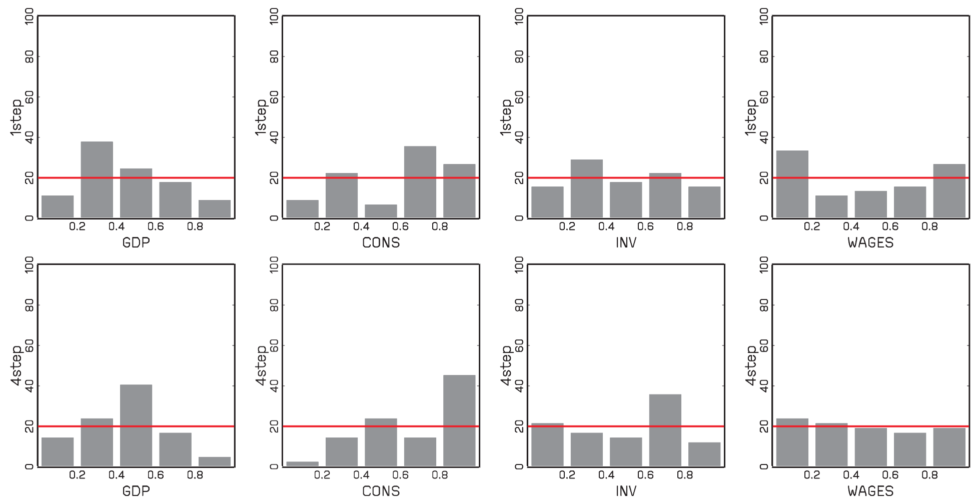
# D Estimation of SW Model on Detrended Data

The SW model imposes a common linear trend on real output, real consumption, real investment, and real wages. This trend is rejected by the data. We eliminated this trend from the model and treated the data by individually HP-filtering each series. For each forecast sample, we applied the HP-filter (with smoothing parameter 1600) individually to each of the four series over both the estimation and forecast period collectively. Using this treated data, we repeated the analysis in the paper without the trend in the model ($\bar{\gamma} = 0$), but we have omitted the calculation of credible bands because it is computationally very costly.

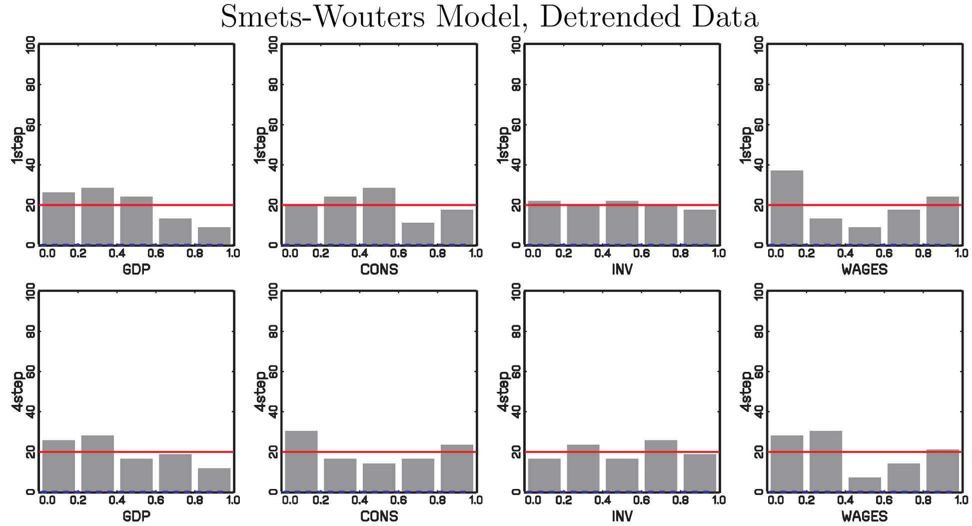
Figure A-1 shows the PITs from the unconditional forecasts for output, consumption, investment, and wages. The top panel shows the PIT histograms associated with the original version of the model, while the bottom panel shows the PIT histograms associated with the detrended model. Recall that for output, the predictive distribution implied by the SW model was too diffuse. The bottom panel indicates that the linear trend in output in the SW model may be driving this result. The PITs for output under the detrended do not cluster in the center of the distribution, especially at longer horizons. In the original model, consumption was underpredicted, as the PITs are overrepresented in the bins on the right side of the histogram. Once detrended, however, the PITs look much more uniform. A similar improvement (albeit from a smaller deficiency) is seen in the predictive distribution for investment. On the other hand, the calibration of the four step ahead predictive distribution for wages has deteriorated by using detrended data. Taken collectively, though, the histograms in Figure A-1 seem to indicate that the counterfactual common trend in output, consumption, investment, and wages deteriorates density forecasts in the SW model.

Figure A-1: PIT Histograms – Unconditional Forecasts

Smets-Wouters Model, Original Data




Smets-Wouters Model, Detrended Data




*Notes:* Probability integral transforms for one and four-step ahead forecasts of output growth (*GDP*), inflation (*INF*), and interest rates (*FFR*). *Bars* correspond to actuals, and *dashed bands* indicate 90% credible intervals obtained from the predictive distribution.