Online appendix to "On the stability of the excess sensitivity of aggregate consumption growth in the US"

Gerdie Everaert[1], Lorenzo Pozzi*[2], and Ruben Schoonackers[3]

[1] *Ghent University & SHERPPA*
[2] *Erasmus University Rotterdam & Tinbergen Institute*
[3] *National Bank of Belgium & Ghent University*

August 30, 2016

Abstract

This online appendix provides technical details on the the estimation methodology used in the main paper, including an offset mixture representation for the stochastic volatility components, a general outline of the Gibbs sampler and details on the exact implementation for each of the Gibbs blocks.

A1 Offset mixture representation for the stochastic volatility components

Using $\nu_t \sim \mathcal{N}\left(0, e^{h_{\nu t}}\right)$ and $\mu_t \sim \mathcal{N}\left(0, e^{h_{\mu t}}\right)$ together with equation (18) in the main paper, we can write $\nu_t = e^{h_{\nu t}/2}\mu_{1t}$ and $\mu_t = e^{h_{\mu t}/2}\mu_{2t}$ with μ_{1t} and μ_{2t} *i.i.d.* error terms with unit variance. A key feature of these stochastic volatility components is that they are nonlinear but can be transformed into linear components by taking the logarithm of their squares

$$\ln\left(e^{h_{\nu t}/2}\mu_{1t}\right)^2 = h_{\nu t} + \ln\left(\mu_{1t}\right)^2, \qquad \ln\left(e^{h_{\mu t}/2}\mu_{2t}\right)^2 = h_{\mu t} + \ln\left(\mu_{2t}\right)^2, \tag{A-1}$$

where $ln\left(\mu_{1t}\right)^2$ and $ln\left(\mu_{2t}\right)^2$ are log-chi-square distributed with expected value -1.2704 and variance 4.93. Following Kim et al. (1998), we approximate the linear models in (A-1) by an offset mixture time series model as

$$g_{jt} = h_{jt} + \epsilon_{jt}, \tag{A-2}$$

*Corresponding author at: Department of Economics, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands, Tel:+31 (10) 4081256, Email: pozzi@ese.eur.nl, Website: http://people.few.eur.nl/pozzi.

for $j = \nu, \mu$, where $g_{\nu t} = \ln\left(\left(e^{h_{\nu t}/2}\mu_{1t}\right)^2 + c\right)$ and $g_{\mu t} = \ln\left(\left(e^{h_{\mu t}/2}\mu_{2t}\right)^2 + c\right)$, with $c = .001$ being an offset constant, and the distribution of ϵ_{jt} given by the following mixture of normals

$$f(\epsilon_{jt}) = \sum_{n=1}^{M} q_n f_N\left(\epsilon_{jt} | m_n - 1.2704, s_n^2\right), \tag{A-3}$$

with component probabilities q_n, means $m_n - 1.2704$ and variances s_j^2. Equivalently, this mixture density can be written in terms of the component indicator variable κ_{jt} as

$$\epsilon_{jt} | (\kappa_{jt} = n) \sim \mathcal{N}\left(m_n - 1.2704, s_n^2\right), \qquad \text{with} \qquad Pr(\kappa_{jt} = n) = q_n. \tag{A-4}$$

Following Kim et al. (1998), we use a mixture of $M = 7$ normal distributions to make the approximation to the log-chi-square distribution sufficiently good. Values for $\{q_n, m_n, s_n^2\}$ are provided by Kim et al. in their Table 4.

A2 General outline of the Gibbs sampler

Taken together, equations (14) and (26) in the main paper constitute the observation equations of a State Space (SS) model, with the unobserved states β_{it} and h_{jt} evolving according to the state equations (22)-(23) and (24)-(25) respectively. In a standard linear Gaussian SS model, the Kalman filter can be used to filter the unobserved states from the data and to construct the likelihood function such that the unknown parameters can be estimated using Maximum Likelihood (ML). However, the stochastic volatility components introduced in Section 3.1 and the stochastic model specification search outlined in Section 3.2 imply a non-regular estimation problem for which the standard approach via the Kalman filter and ML is not feasible. Instead, we use the Gibbs sampler which is a Markov Chain Monte Carlo (MCMC) method to simulate draws from the intractable joint and marginal posterior distributions of the unknown parameters and the unobserved states using only tractable conditional distributions. Intuitively, this amounts to reducing the complex non-linear model into a sequence of blocks for subsets of parameters/states that are tractable conditional on the other blocks in the sequence.

For notational convenience, define the non-centered time-varying parameter vector $\beta_t^* = (\beta_{0t}^*, \beta_{1t}^*)$, the non-centered stochastic volatilities vector $h_t^* = \left(h_{\nu t}^*, h_{\mu t}^*\right)$, the mixture indicators vector $\kappa_t = (\kappa_{\nu t}, \kappa_{\mu t})$, the unknown parameter vectors $\theta = (\theta_1, \ldots, \theta_q)$, $\phi_\beta = (\beta_{00},\ \beta_{10},\ \gamma,\ \rho,\ \sigma_{\eta_0},\ \sigma_{\eta_1})$, $\phi_h = (h_{\nu 0},\ h_{\mu 0},\ \sigma_{\upsilon_\nu},\ \sigma_{\upsilon_\mu})$, $\phi = (\delta, \theta, \phi_\beta, \phi_h)$ and the model indicators $\mathcal{M}_\theta = (\iota_{\theta_1}, \ldots, \iota_{\theta_q})$, $\mathcal{M}_{\phi_\beta} = (\iota_{\beta_{00}}, \iota_{\beta_{10}}, \iota_\gamma, \iota_\rho, \iota_{\beta_{0t}}, \iota_{\beta_{1t}})$, $\mathcal{M}_{\phi_h} = (\iota_{h_\nu},\ \iota_{h_\mu})$, $\mathcal{M} = (\mathcal{M}_\theta,\ \mathcal{M}_{\phi_\beta},\ \mathcal{M}_{\phi_h})$. Further let $D_t = (\Delta \ln C_t,\ \Delta \ln Y_t,\ Z_t)$ be the data vector. Stacking observations over time, we denote $D = \{D_t\}_{t=1}^T$ and similarly for β^*, h^* and κ. The posterior density of interest is then given by $f(\phi, \beta^*, h^*, \kappa, \mathcal{M}|D)$. Building on Frühwirth-Schnatter and Wagner (2010) for the stochastic model specification part, on Chib and Greenberg (1994) for the moving average (MA) part and on Kim et al. (1998) for the stochastic volatility part, our MCMC scheme is as follows:

1. Sample the binary indicators $(\mathcal{M}_\theta, \mathcal{M}_{\phi_\beta})$ and the constant parameters $(\delta, \theta, \phi_\beta)$ conditional on the time-varying parameters β^* and on the stochastic volatilities h^*. This is accomplished as:

(a) Sample the first step parameters δ from $f\left(\delta|\phi_\beta, h^*, D\right)$.

(b) Sample the binary indicators $\mathcal{M}_\theta$ and the MA coefficients θ:

i. Sample the binary indicators $\mathcal{M}_\theta$ from $f\left(\mathcal{M}_\theta|\delta, \phi_\beta, \phi_h, \beta^*, h^*, D\right)$ marginalizing over the parameters θ for which variable selection is carried out.

ii. Sample the unrestricted parameters in θ from $f\left(\theta|\delta, \phi_\beta, \phi_h, \beta^*, h^*, \mathcal{M}_\theta, D\right)$ while setting the restricted parameters in θ (for which the corresponding binary indicator in $\mathcal{M}_\theta$ is 0) equal to 0.

(c) Sample the binary indicators $\mathcal{M}_{\phi_\beta}$ and the parameters ϕ_β:

i. Sample the binary indicators $\mathcal{M}_{\phi_\beta}$ from $f\left(\mathcal{M}_{\phi_\beta}|\delta, \theta, \phi_h, \beta^*, h^*, D\right)$ marginalizing over the parameters ϕ_β for which variable selection is carried out.

ii. Sample the unrestricted parameters in ϕ_β from $f\left(\phi_\beta|\delta, \theta, \phi_h, \beta^*, h^*, \mathcal{M}_{\phi_\beta}, D\right)$ while setting the restricted parameters in ϕ_β (for which the corresponding binary indicator in $\mathcal{M}_{\phi_\beta}$ is 0) equal to 0.

2. Sample the unrestricted time-varying parameters in β^* from $f(\beta^*|\delta, \theta, \phi_\beta, h^*, \mathcal{M}_{\phi_\beta}, D)$. The restricted time-varying parameters in β^* (for which the corresponding binary indicator is 0) are sampled directly from their prior distribution in equation (23).

3. Sample the mixture indicators κ, the binary indicators $\mathcal{M}_{\phi_h}$, the constant parameters ϕ_h and the stochastic volatilities components h^* conditional on the constant parameters $(\delta, \theta, \phi_\beta)$ and on the time-varying parameters β^*. This is accomplished as:

(a) Sample the mixture indicators κ from $f\left(\kappa|\phi, \beta^*, h^*, D\right)$.[1]

(b) Sample the binary indicators $\mathcal{M}_{\phi_h}$ and the constant parameters ϕ_h conditional on the constant parameters $(\delta, \theta, \phi_\beta)$, on the time-varying parameters β^*, on the stochastic volatility components h^* and on the mixture indicators κ:

i. Sample the binary indicators $\mathcal{M}_{\phi_h}$ from $f\left(\mathcal{M}_{\phi_h}|\delta, \theta, \phi_\beta, \beta^*, h^*, \kappa, D\right)$ marginalizing over the parameters ϕ_h for which variable selection is carried out.

ii. Sample the unrestricted parameters in ϕ_h from $f\left(\phi_h|\delta, \theta, \phi_\beta, \beta^*, h^*, \kappa, \mathcal{M}_{\phi_h}, D\right)$ while setting the restricted parameters in ϕ_h (for which the corresponding binary indicator in $\mathcal{M}_{\phi_h}$ is 0) equal to 0.

(c) Sample the unrestricted stochastic volatilities in h^* from $f(h^*|\phi, \beta^*, \kappa, \mathcal{M}_{\phi_h}, D)$. The restricted stochastic volatilities in h^* (for which the corresponding binary indicator is 0) are sampled directly from their prior distribution in equation (25).

[1] Note that the ordering of the Gibbs steps is in line with Del Negro and Primiceri (2015) who argue that the mixture indicators κ should be drawn after the components that do not condition on κ directly but before the components that do condition on κ.

4. Perform a random sign switch for σ_{η_i} and $\{\beta_{it}^*\}_{t=1}^T$ and for σ_{υ_j} and $\{h_{jt}^*\}_{t=1}^T$, e.g., σ_{η_1} and $\{\beta_{1t}^*\}_{t=1}^T$ are left unchanged with probability 0.5 while with the same probability they are replaced by $-\sigma_{\eta_1}$ and $\{-\beta_{1t}^*\}_{t=1}^T$.

Given an arbitrary set of starting values, sampling from these blocks is iterated J times and, after a sufficiently large number of burn-in draws B, the sequence of draws $(B+1, ..., J)$ approximates a sample from the virtual posterior distribution $f(\phi, \beta^*, h^*, \kappa, \mathcal{M}|D)$. Details on the exact implementation of each of the blocks can be found in Section A3. The results reported in the main paper are based on 15 000 iterations, with the first 5 000 draws discarded as a burn-in sequence.

A3 Details on the blocks in the Gibbs sampling algorithm

In this section we provide details on the exact implementation of the Gibbs sampling algorithm outlined in Section A2 to jointly sample the binary indicators $\mathcal{M}$, the constant parameters ϕ, the time-varying parameters β^*, the stochastic volatilities h^* and the mixture indicators κ.

For notational convenience, let us first define a general regression model

$$y = x^m b^m + \theta(L)\, e, \qquad e \sim \mathcal{N}(0, \Sigma), \tag{A-5}$$

where $y = (y_1, \ldots, y_T)'$ is a $(T \times 1)$ vector stacking observations on the dependent variable y_t, $x = (x_1', \ldots, x_T')'$ is a $(T \times k)$ unrestricted predictor matrix with typical row x_t, b is a $(k \times 1)$ unrestricted parameter vector, $\theta(L)$ is a lag polynomial of order q and Σ is a diagonal matrix with elements σ_{et}^2 on the diagonal that may vary over time to allow for heteroskedasticity of a known form. The restricted predictor matrix x^m and restricted parameter vector b^m exclude those elements in x and b for which the corresponding binary indicator in the m is zero, with m being a subset of the model $\mathcal{M}$.

The MA(q) errors in equation (A-5) imply a model that is non-linear in the parameters. As suggested by Ullah et al. (1986) and Chib and Greenberg (1994), conditional on θ a linear model can be obtained from a recursive transformation of the data. For $t = 1, \ldots T$ let

$$\widetilde{y}_t = y_t - \sum_{i=1}^{q} \theta_i \widetilde{y}_{t-i}, \qquad \text{with} \quad \widetilde{y}_t = 0 \quad \text{for } t \leq 0, \tag{A-6}$$

$$\widetilde{x}_t = x_t - \sum_{i=1}^{q} \theta_i \widetilde{x}_{t-i}, \qquad \text{with} \quad \widetilde{x}_t = 0 \quad \text{for } t \leq 0, \tag{A-7}$$

and further for $j = 1, \ldots, q$

$$\omega_{jt} = -\sum_{i=1}^{q} \theta_i \omega_{j,t-i} + \theta_{t+j-1}, \qquad \text{with} \quad \omega_{jt} = 0 \quad \text{for } t \leq 0, \tag{A-8}$$

where $\theta_s = 0$ for $s > q$. Equation (A-5) can then be transformed as

$$\widetilde{y} = \widetilde{x}^m b^m + \omega\lambda + e = \widetilde{w}^m \Phi^m + e, \tag{A-9}$$

where $\widetilde{w} = (\widetilde{x}, \omega)$ with $\omega = (\omega_1', \ldots, \omega_T')'$ and $\omega_t = (\omega_{1t}, \ldots, \omega_{qt})$, $\Phi^m = \left(b^{m'}, \lambda'\right)'$ and $\lambda = (e_0, \ldots, e_{-q+1})'$ initial conditions that can be estimated as unknown parameters.

Conditional on θ and σ_{et}^2, equation (A-9) is a standard linear regression with observed variables $\widetilde{y}$ and $\widetilde{w}^m$ and heteroskedastic normal errors with known covariance matrix Σ. A naive implementation of the Gibbs sampler would be to first sample the binary indicators in m from $f(m|\Phi, \Sigma, \widetilde{y}, \widetilde{w})$ and next Φ^m from $f(\Phi^m|m, \Sigma, \widetilde{y}, \widetilde{w})$. However, this approach does not result in an irreducible Markov chain as whenever an indicator in m equals zero, the corresponding coefficient in Φ is also zero which implies that the chain has absorbing states. Therefore, as in Frühwirth-Schnatter and Wagner (2010) we marginalize over the parameters Φ when sampling m and next draw Φ^m conditional on the binary indicators in m. The posterior distribution $f(m|\Sigma, \widetilde{y}, \widetilde{w})$ can be obtained using Bayes' Theorem as

$$f(m|\Sigma, \widetilde{y}, \widetilde{w}) \propto f(\widetilde{y}\,|m, \Sigma, \widetilde{w})\, p(m), \tag{A-10}$$

with $p(m)$ being the prior probability of m and $f(\widetilde{y}\,|m, \Sigma, \widetilde{w})$ the marginal likelihood of the regression model (A-9) where the effect of the parameters Φ has been integrated out. Under the normal conjugate prior

$$p(\Phi^m) = \mathcal{N}(\Phi_0^m, V_0^m), \tag{A-11}$$

the closed form solution of the marginal likelihood is given by

$$f(\widetilde{y}\,|m, \Sigma, \widetilde{w}) \propto \frac{|\Sigma|^{-0.5}\,|V_T^m|^{0.5}}{|V_0^m|^{0.5}} \exp\left(-\frac{\widetilde{y}'\Sigma^{-1}\widetilde{y} + (\Phi_0^m)'(V_0^m)^{-1}\Phi_0^m - (\Phi_T^m)'(V_T^m)^{-1}\Phi_T^m}{2}\right), \tag{A-12}$$

where the posterior moments Φ_T^m and V_T^m can be calculated as

$$\Phi_T^m = V_T^m\left((\widetilde{w}^m)'\Sigma^{-1}\widetilde{y} + (V_0^m)^{-1}\Phi_0^m\right), \tag{A-13}$$

$$V_T^m = \left((\widetilde{w}^m)'\Sigma^{-1}\widetilde{w}^m + (V_0^m)^{-1}\right)^{-1}. \tag{A-14}$$

Following George and McCulloch (1993) we use a single-move sampler in which the binary indicators ι_k in m are sampled recursively from the Bernoulli distribution with probability

$$p(\iota_k = 1\,|\iota_{-k}, \Sigma, \widetilde{y}, \widetilde{w}) = \frac{f(\iota_k = 1\,|\iota_{-k}, \Sigma, \widetilde{y}, \widetilde{w})}{f(\iota_k = 0\,|\iota_{-k}, \Sigma, \widetilde{y}, \widetilde{w}) + f(\iota_k = 1\,|\iota_{-k}, \Sigma, \widetilde{y}, \widetilde{w})}, \tag{A-15}$$

for k depending on the particular subset m of $\mathcal{M}$. We further randomize over the sequence in which the binary indicators are drawn.

Conditional on m, the posterior distribution of Φ^m is given by

$$p(\Phi^m|m, \Sigma, \widetilde{y}, \widetilde{w}) = \mathcal{N}(\Phi_T^m, V_T^m), \tag{A-16}$$

with posterior moments b_T^m and V_T^m given in equations (A-13)-(A-14).

Block 1: Sampling $(\mathcal{M}_\theta, \mathcal{M}_{\phi_\beta})$ and $(\delta, \theta, \phi_\beta)$

Block 1(a): Sampling the first step parameters δ

Conditional on the stochastic volatility component $h_{\nu t}$, equation (14) can be written in the general notation of equation (A-5) as: $y_t = \Delta \ln Y_t$, $x_t^m = x_t = Z_t$, $b^m = b = \delta$ and $\theta(L) = 1$ such that $\theta(L)e_t = \nu_t$ and the elements in Σ are given by $\sigma_{et}^2 = e^{h_{\nu t}}$. In the notation of equation (A-9) we further have $\widetilde{y}_t = y_t$, $\widetilde{w}_t^m = \widetilde{w}_t = x_t$ and $\Phi^m = \Phi = \delta$. The parameters in δ can then be sampled from the posterior distribution in equation (A-16) and used to calculate $E_{t-1}(\Delta \ln Y_t) = Z_t\delta$, $\nu_t = \Delta \ln Y_t - Z_t\delta$ and

$$\nu_t^* = \frac{\theta(L)\,\nu_t e^{h_{\mu t}/2}}{\sqrt{1-\rho^2}e^{h_{\nu t}/2}},$$

conditional on θ, ρ, $h_{\nu t}$ and $h_{\mu t}$.

Block 1(b): Sampling the binary indicators $\mathcal{M}_\theta$ and the MA coefficients θ

Conditional on the parameters δ, ϕ_h and ϕ_β, on the time-varying coefficients β_t^*, on the stochastic volatility $h_{\mu t}^*$ and on the binary indicators $\mathcal{M}_{\phi_\beta}$, equation (16) can be written in the general notation of equation (A-5) as: $y_t = \Delta \ln C_t$, $x_t = (1,\ Z_t\delta,\ \beta_{0t}^*,\ \beta_{1t}^* Z_t\delta,\ \Delta \ln C_{t-1})$, $b = (\beta_{00},\ \beta_{10},\ \sigma_{\eta_0},\ \sigma_{\eta_1},\ \gamma)$ and $\theta(L)e_t = \theta(L)\varepsilon_t$, such that the elements in Σ are given by $\sigma_{et}^2 = \sigma_{\tilde{\varepsilon}t}^2 = e^{h_{\mu t}}/(1-\rho^2)$. The values of the binary indicators in the subset $\mathcal{M}_{\phi_\beta}$ of $\mathcal{M}$ then imply the restricted $x_t^{\mathcal{M}_{\phi_\beta}}$ and $b^{\mathcal{M}_{\phi_\beta}}$.

Under the normal conjugate prior $p(\theta) = \mathcal{N}\left(b_0^\theta, V_0^\theta\right)$, the exact conditional distribution of θ is

$$p\left(\theta|\Phi^{\mathcal{M}_{\phi_\beta}}, \Sigma, y, x\right) \propto \prod_{t=1}^{T} \exp\left(-\frac{e_t(\theta)^2}{2\sigma_{et}^2}\right) \times \exp\left(-\frac{1}{2}\left(\theta - b_0^\theta\right)'\left(V_0^\theta\right)^{-1}\left(\theta - b_0^\theta\right)\right), \tag{A-17}$$

where $e_t(\theta) = \widetilde{y}_t(\theta) - \widetilde{w}_t^m(\theta)\,\Phi^{\mathcal{M}_{\phi_\beta}}$ is calculated from the transformed model in equation (A-9) further conditioning on the initial conditions λ to obtain $\Phi^{\mathcal{M}_{\phi_\beta}} = \left(b^{\mathcal{M}_{\phi_\beta}'}, \lambda'\right)'$.

Direct sampling of θ using equation (A-17) is not possible, though, as $e_t(\theta)$ is a non-linear function of θ. To solve this issue, Chib and Greenberg (1994) propose to linearize $e_t(\theta)$ around θ^* using a first-order Taylor expansion

$$e_t(\theta) \approx e_t(\theta^*) - \Psi_t(\theta - \theta^*), \tag{A-18}$$

where $\Psi_t = (\Psi_{1t}, \ldots, \Psi_{qt})$ is a $1 \times q$ vector including the first-order derivatives of $e_t(\theta)$ evaluated at θ^* obtained using the following recursion

$$\Psi_{it} = -e_{t-i}(\theta^*) - \sum_{j=1}^{q} \theta_j^* \Psi_{i,t-j}, \tag{A-19}$$

where $\Psi_{it} = 0$ for $t \leq 0$. An adequate approximation can be obtained by choosing θ^* to be the non-linear

least squares estimate of θ conditional on the other parameters in the model, which can be obtained as

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \sum_{t=1}^{T} \left(e_t\left(\theta\right)\right)^2, \tag{A-20}$$

Conditioning on θ^*, equation (A-18) can then be rewritten as an approximate linear regression model

$$e_t\left(\theta^*\right) + \Psi_t\theta^* \approx \Psi_t\theta + e_t\left(\theta\right), \tag{A-21}$$

with dependent variable $e_t\left(\theta^*\right)+\Psi_t\theta^*$ and explanatory variables Ψ_t. As such, the approximate expression in (A-21) can be written in the general notation of equation (A-5) by setting $y_t = e_t\left(\theta^*\right)+\Psi_t\theta^*$, $x_t = \Psi_t$, $b = \theta$ and $\theta(L) = 1$ such that $\theta(L)e_t = \varepsilon_t$ and $\sigma_{et}^2 = \sigma_{\varepsilon t}^2 = e^{h_{\mu t}}/(1-\rho^2)$. The values of the binary indicators in the subset $\mathcal{M}_\theta$ of $\mathcal{M}$ imply the restricted $x_t^{\mathcal{M}_\theta}$ and $b^{\mathcal{M}_\theta}$. In the notation of equation (A-9) we further have $\widetilde{y}_t = y_t$, $\widetilde{w}_t^m = x_t^{\mathcal{M}_\theta}$ and $\Phi^m = b^{\mathcal{M}_\theta}$.

The binary indicators in $\mathcal{M}_\theta$ can then be sampled from the posterior distribution in equation (A-10) with marginal likelihood calculated from equation (A-12). Next, we sample $\theta^{\mathcal{M}_\theta}$ using a Metropolis-Hastings (MH) step. Suppose $\theta^{\mathcal{M}_\theta}(i)$ is the current draw in the Markov chain. To obtain the next draw $\theta^{\mathcal{M}_\theta}(i+1)$, first draw a candidate $\theta^{\mathcal{M}_\theta}(c)$ using the following normal proposal distribution

$$q\left(\theta^{\mathcal{M}_\theta}|\theta^*, \Sigma, y, x\right) \sim \mathcal{N}\left(\Phi_T^{\mathcal{M}_\theta}, V_T^{\mathcal{M}_\theta}\right), \tag{A-22}$$

with posterior moments $\Phi_T^{\mathcal{M}_\theta}$ and $V_T^{\mathcal{M}_\theta}$ calculated using equations (A-13)-(A-14). The MH step then implies a further randomization which amounts to accepting the candidate draw $\theta^{\mathcal{M}_\theta}(c)$ with probability

$$\alpha\left(\theta^{\mathcal{M}_\theta}(i), \theta^{\mathcal{M}_\theta}(c)\right) = \min\left\{\frac{p\left(\theta^{\mathcal{M}_\theta}(c)|\Phi^{\mathcal{M}_\theta}, \Sigma, y, x\right)}{p\left(\theta^{\mathcal{M}_\theta}(i)|\Phi^{\mathcal{M}_\theta}, \Sigma, y, x\right)} \frac{q\left(\theta^{\mathcal{M}_\theta}(i)|\theta^*, \Sigma, y, x\right)}{q\left(\theta^{\mathcal{M}_\theta}(c)|\theta^*, \Sigma, y, x\right)}, 1\right\}. \tag{A-23}$$

If $\theta^{\mathcal{M}_\theta}(c)$ is accepted, $\theta^{\mathcal{M}_\theta}(i+1)$ is set equal to $\theta^{\mathcal{M}_\theta}(c)$ while if $\theta^{\mathcal{M}_\theta}(c)$ is rejected, $\theta^{\mathcal{M}_\theta}(i+1)$ is set equal to $\theta^{\mathcal{M}_\theta}(i)$. Note that the unrestricted θ is restricted to obtain $\theta^{\mathcal{M}_\theta}$ by excluding θ_l when $\iota_{\theta_l} = 0$. In this case θ_l is not sampled but set equal to zero.

Block 1(c): Sampling the binary indicators $\mathcal{M}_{\phi_\beta}$ and the second step parameters ϕ_β

Conditional on β_t^*, δ and ν_t^*, equation (26) can be written in the general notation of equation (A-5) setting $y_t = \Delta \ln C_t$, $x_t = (1,\ Z_t\delta,\ \beta_{0t}^*,\ \beta_{1t}^* Z_t\delta,\ \Delta \ln C_{t-1},\ \nu_t^*)$, $b = (\beta_{00},\ \beta_{01},\ \sigma_{\eta_0},\ \sigma_{\eta_1},\ \gamma,\ \rho)$ and $\theta(L)e_t = \theta(L)\mu_t$, such that the elements in Σ are given by $\sigma_{et}^2 = e^{h_{\mu t}}$. Further conditioning on the MA parameters θ, the unrestricted transformed variables $\widetilde{y}_t$ and $\widetilde{w}_t$ in equation (A-9) are obtained, with corresponding unrestricted extended parameter vector $\Phi = (b', \lambda')'$. The values of the binary indicators in $\mathcal{M}_{\phi_\beta}$ then imply the restricted $\widetilde{w}_t^{\mathcal{M}_{\phi_\beta}}$ and $\Phi^{\mathcal{M}_{\phi_\beta}}$.

First, the binary indicators in $\mathcal{M}_{\phi_\beta}$ are sampled from the posterior distribution in equation (A-10) with marginal likelihood calculated from equation (A-12). Second, conditional on $\mathcal{M}_{\phi_\beta}$, ϕ_β is sampled, together with λ, by drawing $\Phi^{\mathcal{M}_{\phi_\beta}}$ from the general expression in equation (A-16). Note that the unrestricted $\Phi = (\beta_{00},\ \beta_{10},\ \sigma_{\eta_0},\ \sigma_{\eta_1},\ \gamma,\ \rho,\ \lambda)$ is restricted to obtain $\Phi^{\mathcal{M}_{\phi_\beta}}$ by excluding those coefficients for which the

corresponding binary indicator in $\mathcal{M}_{\phi_\beta}$ is zero. These restricted coefficients are not sampled but set equal to zero.

Block 2: Sampling β^*

In this block we use the forward-filtering and backward-sampling approach of Carter and Kohn (1994) and De Jong and Shephard (1995) to sample the time-varying parameters β^*. Conditional on the coefficients ϕ_β, on the stochastic volatility $h_{\mu t}$, on the first block results $Z_t\delta$ and ν_t^* and on the binary indicators $\mathcal{M}_{\phi_\beta}$, equation (26) can be rewritten as

$$y_t = \iota_{\beta_{0t}}\sigma_{\eta_0}\beta_{0t}^* + \iota_{\beta_{1t}}\sigma_{\eta_1}\beta_{1t}^* x_{1t} + \theta(L)\,\mu_t, \tag{A-24}$$

with $y_t = \Delta \ln C_t - \beta_{00} - \beta_{10} Z_t\delta - \gamma\Delta \ln C_{t-1} - \rho\nu_t^*$ and $x_{1t} = Z_t\delta$. Using the recursive transformation suggested by Ullah et al. (1986) and Chib and Greenberg (1994), the model in equation (A-24) can be transformed to

$$\widetilde{y}_t = \iota_{\beta_{0t}}\sigma_{\eta_0}\widetilde{\beta}_{0t} + \iota_{\beta_{1t}}\sigma_{\eta_1}\widetilde{\beta}_{1t} + \omega_t\lambda + \mu_t, \tag{A-25}$$

where $\widetilde{y}_t$ and $\omega_t = (\omega_{1t}, \ldots, \omega_{qt})$ are calculated (conditional on θ) from equations (A-6) and (A-8) and similarly

$$\widetilde{\beta}_{0t} = \beta_{0t}^* - \sum_{i=1}^{q}\theta_i\widetilde{\beta}_{0,t-i}, \qquad \text{with} \quad \widetilde{\beta}_{0t} = 0 \quad \text{for } t \le 0, \tag{A-26}$$

$$\widetilde{\beta}_{1t} = \beta_{1t}^* x_{1t} - \sum_{i=1}^{q}\theta_i\widetilde{\beta}_{1,t-i}, \qquad \text{with} \quad \widetilde{\beta}_{1t} = 0 \quad \text{for } t \le 0. \tag{A-27}$$

Substituting equation (23) in (A-26)-(A-27) yields

$$\widetilde{\beta}_{0,t+1} = \beta_{0t}^* - \sum_{i=1}^{q}\theta_i\widetilde{\beta}_{0,t+1-i} + \eta_{0t}^*, \tag{A-28}$$

$$\widetilde{\beta}_{1,t+1} = \beta_{1t}^* x_{1,t+1} - \sum_{i=1}^{q}\theta_i\widetilde{\beta}_{1,t+1-i} + x_{1,t+1}\eta_{1t}^*, \tag{A-29}$$

such that the state space representation of the model in equations (A-25), (23) and (A-28)-(A-29) is given by

$$\widetilde{y}_t - \omega_t\lambda = \overbrace{\begin{bmatrix} (0 & \sigma_{\eta_0} & 0 & \ldots & 0) & (0 & \sigma_{\eta_1} & 0 & \ldots & 0) \end{bmatrix}}^{\sigma_\eta} \overbrace{\begin{bmatrix} \alpha_{0t} \\ \alpha_{1t} \end{bmatrix}}^{\alpha_t} + \mu_t, \tag{A-30}$$

$$\underbrace{\begin{bmatrix} \alpha_{0,t+1} \\ \alpha_{1,t+1} \end{bmatrix}}_{\alpha_{t+1}} = \underbrace{\begin{bmatrix} T_{0t} & 0 \\ 0 & T_{1t} \end{bmatrix}}_{T_t}\underbrace{\begin{bmatrix} \alpha_{0t} \\ \alpha_{1t} \end{bmatrix}}_{\alpha_t} + \underbrace{\begin{bmatrix} K_{0t} & 0 \\ 0 & K_{1t} \end{bmatrix}}_{K_t}\underbrace{\begin{bmatrix} \eta_{0t}^* \\ \eta_{1t}^* \end{bmatrix}}_{\eta_t}, \tag{A-31}$$

with $\alpha_{i,t+1}$ given by

$$\underbrace{\begin{bmatrix} \beta^*_{i,t+1} \\ \widetilde{\beta}_{i,t+1} \\ \widetilde{\beta}_{it} \\ \vdots \\ \widetilde{\beta}_{i,t-(q-2)} \end{bmatrix}}_{\alpha_{i,t+1}} = \underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ x_{i,t+1} & -\theta_1 & \dots & -\theta_{q-1} & -\theta_q \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}}_{T_{it}} \underbrace{\begin{bmatrix} \beta^*_{it} \\ \widetilde{\beta}_{it} \\ \vdots \\ \widetilde{\beta}_{i,t-(q-2)} \\ \widetilde{\beta}_{i,t-(q-1)} \end{bmatrix}}_{\alpha_{it}} + \underbrace{\begin{bmatrix} 1 \\ x_{i,t+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{K_{it}} \begin{bmatrix} \eta^*_{it} \end{bmatrix}, \quad \text{(A-32)}$$

for $i = 0, 1$ and where $x_{0t} = 1 \ \forall t$, $\mu_t \sim \mathcal{N}\left(0, e^{h_{\mu t}}\right)$ and $\eta^*_{it} \sim \mathcal{N}(0, 1)$. In line with equations (23) and (A-26)-(A-27), each of the states in α_t is initialized at zero.

Equations (A-30)-(A-31) constitute a standard linear Gaussian state space model, from which the unknown state variables α_t can be filtered using the standard Kalman filter. Sampling α_t from its conditional distribution can then be done using the multimove simulation smoother of Carter and Kohn (1994) and De Jong and Shephard (1995). Using β_{i0}, σ_{η_i} and β^*_{it}, the centered time-varying coefficients β_{it} in equation (20) can easily be reconstructed from equation (22). Note that in a restricted model with $\iota_{\beta_{it}} = 0$, σ_{η_i} is excluded from σ_η and α_{it} is dropped from the state vector α_t. In this case, no forward-filtering and backward-sampling for β^*_{it} is needed as this can be sampled directly from its prior distribution using equation (23).

Block 3: Sampling κ, $\mathcal{M}_{\phi_h}$, ϕ_h and h^*

Block 3(a): Sampling the mixture indicators κ

Conditional on h_{jt} and on ϵ_{jt}, calculated from equation (A-2) as $\epsilon_{jt} = g_{jt} - h_{jt}$, we first sample the mixture indicators κ_{jt} from the conditional probability mass

$$p\left(\kappa_{jt} = n | h_{jt}, \epsilon_{jt}\right) \propto q_n f_{\mathcal{N}}\left(\epsilon_{jt} | h_{jt} + m_n - 1.2704, s^2_n\right), \quad \text{(A-33)}$$

for $j = \nu, \mu$.

Block 3(b): Sampling the binary indicators $\mathcal{M}_{\phi_h}$ and the constant parameters ϕ_h

Conditional on h^*_{jt}, g_{jt} and κ_{jt} and using the non-centered parameterization in equation (27), equation (A-2) can be rewritten in the general linear regression format of (A-5) setting $y_t = g_{jt} - (m_{\kappa_{jt}} - 1.2704)$, $x_t = (1, h^*_{jk})$, $b = (h_{j0}, \sigma_{\upsilon_j})$ and $\theta(L)e_t = \widetilde{\epsilon}_{jt} = \epsilon_{jt} - (m_{\kappa_{jt}} - 1.2704)$ for $j = \nu, \mu$. Given the mixture distribution of ϵ_{jt} defined in equation (A-4), the centered error term $\widetilde{\epsilon}_{jt}$ has a heteroskedastic variance $s^2_{\kappa_{jt}}$ such that $\Sigma = \text{diag}(s^2_{\kappa_{j1}}, \dots, s^2_{\kappa_{jT}})$. In the notation of equation (A-9) we further have $\widetilde{y}_t = y_t$ and the unrestricted $\widetilde{w}_t = x_t$ and $\Phi = b$. The values of the binary indicators in $\mathcal{M}_{\phi_h}$ then imply the restricted $\widetilde{w}_t^{\mathcal{M}_{\phi_h}}$ and $\Phi^{\mathcal{M}_{\phi_h}}$. Hence, the binary indicators ι_{h_j}, for $j = (\nu, \mu)$, in $\mathcal{M}_{\phi_h}$ can be sampled from the posterior distribution in equation (A-10) with marginal likelihood calculated from equation (A-12).

Next conditioning on $\mathcal{M}_{\phi_h}$, the parameters $(h_{j0}, \sigma_{\upsilon_j})$, for $j = \nu, \mu$ in ϕ_h are sampled by drawing $\Phi^{\mathcal{M}_{\phi_h}}$ using the general expression in equation (A-16). Note that the unrestricted $\Phi = (h_{j0}, \sigma_{\upsilon_j})$ is restricted

to obtain $\Phi^{\mathcal{M}_{\phi_h}}$ by excluding σ_{υ_j} when $\iota_{h_j} = 0$. In this case σ_{υ_j} is not sampled but set equal to zero.

Block 3(c): Sampling the stochastic volatilities h^*

Conditional on the transformed errors g_{jt} defined under equation (A-2), on the mixture indicators κ_{jt} and on the parameters ϕ_h, equations (A-2) and (24)-(25) can we written in the following conditional state space representation

$$\begin{bmatrix} g_{jt} - \left(m_{\kappa_{jt}} - 1.2704\right) - h_{j0} \end{bmatrix} = \begin{bmatrix} \sigma_{\upsilon_j} \end{bmatrix} \begin{bmatrix} h^*_{jt} \end{bmatrix} + \begin{bmatrix} \widetilde{\epsilon}_{jt} \end{bmatrix}, \tag{A-34}$$

$$\begin{bmatrix} h^*_{j,t+1} \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} h^*_{jt} \end{bmatrix} + \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} \upsilon^*_{jt} \end{bmatrix}, \tag{A-35}$$

for $j = \nu, \mu$, where $\widetilde{\epsilon}_{jt} = \epsilon_{jt} - \left(m_{\kappa_{jt}} - 1.2704\right)$ is ϵ_{jt} centered around zero with $\text{var}(\widetilde{\epsilon}_{jt}) = s^2_{\kappa_{jt}}$ and $\text{var}(\upsilon^*_{jt}) = 1$. In line with equation (25), the random walk component h^*_{jt} is initialized at zero.

As equations (A-34)-(A-35) constitute a standard linear Gaussian state space model, h^*_{jt} can be filtered using the Kalman filter and sampled using the multimove simulation smoother of Carter and Kohn (1994) and De Jong and Shephard (1995). Note that in a restricted model with $\iota_{h_j} = 0$, and hence $\sigma_{\upsilon_j} = 0$, no forward-filtering and backward-sampling for h^*_{jt} is needed as this can be sampled directly from its prior distribution in equation (25). Using h_{j0}, σ_{υ_j} and h^*_{jt}, the centered stochastic volatility component h_{jt} in equation (21) can easily be reconstructed from equation (24).

References

Carter, C. and Kohn, R. (1994). On gibbs sampling for state space models. *Biometrika*, 81:541–53.

Chib, S. and Greenberg, E. (1994). Bayes inference in regression models with ARMA(p,q) errors. *Journal of Econometrics*, 64:183–206.

De Jong, P. and Shephard, N. (1995). The simulation smoother fro time series models. *Biometrika*, 82:339–50.

Del Negro, M. and Primiceri, G. (2015). Time varying structural vector autoregressions and monetary policy: a corrigendum. *The review of economic studies*, 82:1342–1345.

Frühwirth-Schnatter, S. and Wagner, H. (2010). Stochastic model specification search for Gaussian and partial non-Gaussian state space models. *Journal of Econometrics*, 154(1):85–100.

George, E. and McCulloch, R. (1993). Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88:881–889.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies*, 65(3):361–93.

Ullah, A., Vinod, H., and Singh, R. (1986). Estimation of linear models with moving average disturbances. *Journal of Quantitative Economics*, 2:137–152.