Not-for-publication Appendix to the paper "A time-varying natural rate for the euro area"

J.S. Mésonnier and J.P. Renne - June 2006

1 Assessment of alternative proxies for inflation expectations

Alternative deflators can be considered for the computation of the real rate of interest from the nominal rate. To check for the acceptability of our model-consistent inflation expectations, we compare it here with two alternative common ways of proxying inflation expectations. A first method, as implemented e.g. by Laubach and Williams (2003), consists in deriving for each period t the expectation of the next quarter inflation $(E_t(\pi_{t+1}))$ from a univariate AR process estimated over the last 30 quarters[1]. A second method which also allows to continuously update the forecasting model is the univariate time-varying parameters procedure described by Stock and Watson (1996) and recently applied to the computation of real interest rates by Dotsey and Scholl (2003). Let us denote with k the lag length, and with β_t the $(k+1)\times 1$ vector of varying coefficients. Inflation is then assumed to follow

$$\pi_t = \beta_t'[\begin{array}{cccc} 1 & \pi_{t-1} & \cdots & \pi_{t-k} \end{array}]' + \varepsilon_t$$

and β_t is a random walk

$$\beta_t = \beta_{t-1} + \zeta_t$$

The variance-covariance matrix of ζ_t is diagonal. The Kalman filter is then used to estimate the unobserved varying coefficients. The variances of $\zeta_t^0, \zeta_t^1 \ldots, \zeta_t^k$, as well as the initial state vector β_0 and the lag length k are chosen to minimize the conditional predictive squared errors $\sum(\pi_t - E_{t-1}(\pi_t))^2$.

The root mean squared errors over the whole sample are 0.89 for the TVP method, 0.99 for the "moving AR" method and 0.91 for our model-consistent inflation expectations.

2 Computation of the information matrix

A state-space model can be defined by the two following equations:

$$\begin{aligned} Y_t &= \mu_t + G_t\rho_t + M_t\varepsilon_t && (1)\\ \rho_t &= \upsilon_t + H_t\rho_{t-1} + N_t\xi_t && (2) \end{aligned}$$

[1] The AR lag length is chosen so as to minimize the predictive squared errors over the whole period: using only one lag proves to be sufficient. As regards the TVP method, we tried four values of k (from one to four) and the longest lag is found to be the most efficient.

where Y_t is a n-vector of observed variables, ρ_t is an unobserved state vector of dimension p, ε_t and ξ_t are independent gaussian white noises with zero mean and identity covariance matrices, μ_t, G_t, M_t, υ_t, H_t, N_t are functions of an unknown vector of parameters θ and of the past values of Y_t. θ is finite dimensional and therefore, the model is parametric. Equation (1) is referred to as the measurement equation, and (2) as the transition equation. The Kalman filter and smoother provide a simple recursive way of recovering optimally the state vector.

Let denote with $\rho_{t|\tau}$ the estimate of ρ_t upon information $Y^\tau = (Y_1, ..., Y_\tau)$, the output of the Kalman filter is $\rho_{t|t}$ and the output of the Kalman smoother is $\rho_{t|T}$, where T is the number of observations. Let $\Sigma_{t|\tau}$ denote the covariance matrix of ρ_t based upon information Y^τ. The filtering procedure consists of the prediction and updating equations. Whereas the prediction equations provides $\rho_{t|t-1}$, $\Sigma_{t|t-1}$, $Y_{t|t-1}$ and $\Omega_{t|t-1}$, where $\Omega_{t|t-1} = Var(Y_t|Y^{t-1})$, the updating equations provide $\rho_{t|t}$, $\Sigma_{t|t}$, $Y_{t|t}$ and $\Omega_{t|t}$.

Let λ_t denote the innovation in Y_t (that is, $\lambda_t = Y_t - Y_{t|t-1}$), the log-likelihood can then be written as:

$$L(Y^T;\theta) = \frac{NT}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log\left|\Omega_{t|t-1}(\theta)\right| + \lambda_t'(\theta)\Omega_{t-t-1}^{-1}(\theta)\lambda_t(\theta)\right) \quad (3)$$

The computation of the information matrix is based on the results of Engle and Watson (1981). They use in particular the following expressions for derivatives of a symmetric matrix:

$$\frac{\partial |S|}{\partial x} = |S|\, tr\left(S^{-1}\frac{\partial S}{\partial x}\right) \quad (4)$$

$$\frac{\partial S^{-1}}{\partial x} = -S^{-1}\frac{\partial S}{\partial x}S^{-1} \quad (5)$$

And the considered estimate of the information matrix is:

$$\hat{I}_{F\,i,j} = -\frac{1}{T}\sum_{t=1}^{T}E\left(\frac{\partial^2 f_t(Y_t|Y^{t-1};\hat{\theta}_T)}{\partial\theta\partial\theta'}|Y^{t-1}\right) \quad (6)$$

Using (4) and (5), let differentiate (3):

$$\frac{\partial L_t}{\partial \theta_i} = -\frac{1}{2}tr\left(\Omega_{t|t-1}^{-1}\frac{\partial \Omega_{t|t-1}}{\partial\theta_i}\right) - \left(\frac{\partial\lambda_t}{\partial\theta_i}\right)'\Omega_{t|t-1}^{-1}\lambda_t + \frac{1}{2}\lambda_t'\Omega_{t|t-1}^{-1}\frac{\partial\Omega_{t|t-1}}{\partial\theta_i}\Omega_{t|t-1}^{-1}\lambda_t \quad (7)$$

Taking the trace of the last term gives:

$$\frac{\partial L_t}{\partial \theta_i} = -\frac{1}{2}tr\left(\left(\Omega_{t|t-1}^{-1}\frac{\partial \Omega_{t|t-1}}{\partial\theta_i}\right)\left(Id - \Omega_{t|t-1}^{-1}\lambda_t\lambda_t'\right)\right) - \left(\frac{\partial\lambda_t}{\partial\theta_i}\right)'\Omega_{t|t-1}^{-1}\lambda_t \quad (8)$$

that we write with obvious notations:

$$\frac{\partial L_t}{\partial \theta_i} = L_t^1 + L_t^2 \quad (9)$$

In order to get the second order derivative of the log-likelihood, we have to differentiate L_t^1:

$$\frac{\partial L_t^1}{\partial \theta_j} = -\frac{1}{2} tr \left(\frac{\partial}{\partial \theta_j} \left(\Omega_{t|t-1}^{-1} \frac{\partial \Omega_{t|t-1}}{\partial \theta_i} \right) \left(Id - \Omega_{t|t-1}^{-1} \lambda_t \lambda_t' \right) \right) \tag{10}$$

$$-\frac{1}{2} tr \left(\Omega_{t|t-1}^{-1} \frac{\partial \Omega_{t|t-1}}{\partial \theta_i} \Omega_{t|t-1}^{-1} \frac{\partial \Omega_{t|t-1}}{\partial \theta_j} \Omega_{t|t-1}^{-1} \lambda_t \lambda_t' \right)$$
$$+\frac{1}{2} tr \left(\Omega_{t|t-1}^{-1} \frac{\partial \Omega_{t|t-1}}{\partial \theta_i} \Omega_{t|t-1}^{-1} \left(\frac{\partial \lambda_t}{\partial \theta_j} \lambda_t' + \lambda_t \frac{\partial \lambda_t'}{\partial \theta_j} \right) \right)$$

Conditionally on Y_t , the only random terms of the latter equation are the λ_t, which are zero mean. The first term hence vanishes when taking the expected value of the equation. Moreover, recall that $\lambda_t = Y_t - \mu_t - G_t \rho_{t|t-1}$ and then that $\partial \lambda_t / \partial \theta_i$ only depends on the information at $t-1$, hence, the expected value of the third term is zero. All this leads to

$$E \left(\frac{\partial L_t^1}{\partial \theta_j} | Y^{t-1} \right) = -\frac{1}{2} tr \left(\Omega_{t|t-1}^{-1} \frac{\partial \Omega_{t|t-1}}{\partial \theta_i} \Omega_{t|t-1}^{-1} \frac{\partial \Omega_{t|t-1}}{\partial \theta_j} \right) \tag{11}$$

Regarding L_t^2:

$$\frac{\partial L_t^2}{\partial \theta_j} = -\frac{\partial^2 \lambda_t}{\partial \theta_i \partial \theta_j} \Omega_{t|t-1}^{-1} \eta_t - \left(\frac{\partial \lambda_t}{\partial \theta_i} \right)' \frac{\partial \Omega_{t|t-1}^{-1}}{\partial \theta_j} \lambda_t - \left(\frac{\partial \lambda_t}{\partial \theta_i} \right)' \Omega_{t|t-1}^{-1} \frac{\partial \lambda_t}{\partial \theta_j} \tag{12}$$

For the same reasons as above, the first two terms vanishes when taking the conditional expected value. Since the third depends only on the past innovations, its conditional expected value is equal to itself:

$$E \left(\frac{\partial L_t^2}{\partial \theta_j} | Y^{t-1} \right) = - \left(\frac{\partial \lambda_t}{\partial \theta_i} \right)' \Omega_{t|t-1}^{-1} \frac{\partial \lambda_t}{\partial \theta_j} \tag{13}$$

Finally, the ijth element of the information matrix is the negative of the sum of (11) and (13), that is:

$$\hat{I}_{F\,i,j} = \sum_{t=1}^{T} \left[\left(\frac{\partial \lambda_t}{\partial \theta_i} \right)' \Omega_{t|t-1}^{-1} \frac{\partial \lambda_t}{\partial \theta_j} + \frac{1}{2} tr \left(\Omega_{t|t-1}^{-1} \frac{\partial \Omega_{t|t-1}}{\partial \theta_i} \Omega_{t|t-1}^{-1} \frac{\partial \Omega_{t|t-1}}{\partial \theta_j} \right) \right] \tag{14}$$

3 State-space form of the model

In order to use the Kalman filter, equations (2) to (7) in the paper have to be written in the state-space form. (15) is the measurement equation and (16) is the corresponding transition equation.

$$\begin{bmatrix} \Delta y_t \\ \pi_t \end{bmatrix} = G \begin{bmatrix} a_t \\ a_{t-1} \\ z_t \\ z_{t-1} \end{bmatrix} + M \begin{bmatrix} \pi_{t-1} \\ \pi_{t-2} \\ \pi_{t-3} \end{bmatrix} + \begin{bmatrix} \mu_y \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^\pi \end{bmatrix} \tag{15}$$

$$\begin{bmatrix} a_t \\ a_{t-1} \\ z_t \\ z_{t-1} \end{bmatrix} = H \begin{bmatrix} a_{t-1} \\ a_{t-2} \\ z_{t-1} \\ z_{t-2} \end{bmatrix} + N \begin{bmatrix} i_{t-1} \\ i_{t-2} \\ \pi_{t-1} \\ \pi_{t-2} \\ \pi_{t-3} \\ \pi_{t-4} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -\lambda\mu_r \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^a \\ 0 \\ \varepsilon_t^z \\ 0 \end{bmatrix} \tag{16}$$

with $G = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & \beta \end{bmatrix}$, $M = \begin{bmatrix} 0 & 0 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix}$, $H = \begin{bmatrix} \psi & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -\lambda\theta & \Phi & -\beta\lambda \\ 0 & 0 & 1 & 0 \end{bmatrix}$

and $N = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & -\lambda\alpha_1 & -\lambda\alpha_2 & -\lambda\alpha_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.