Online Appendix

Regime-Switching Density Forecasts Using Economists' Scenarios

Graziano Moramarco*

University of Bologna

Appendix A Estimation and forecasting

A.1 Bayesian estimation of MSAR with multiple views

Let us define here $\mathbf{y} = (y_1, \dots, y_T)$ and $\mathbf{S} = (S_1, \dots, S_T)$. Also, let $\boldsymbol{\vartheta} = (\beta_1, \dots, \beta_K, \sigma_1, \dots, \sigma_K, \alpha_1, \dots, \alpha_p, \boldsymbol{\xi})$ and $\boldsymbol{\theta} = (\beta_1, \dots, \beta_K, \sigma_1, \dots, \sigma_K, \alpha_1, \dots, \alpha_p,)$. The posterior distribution $p(\boldsymbol{\vartheta}|\mathbf{y})$ for model (1) in the paper is obtained using Bayes' theorem:

$$p(\boldsymbol{\vartheta}|\mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\vartheta})p(\boldsymbol{\vartheta}) \tag{1}$$

where $p(\boldsymbol{\vartheta})$ is the prior on the parameters and $p(\mathbf{y}|\boldsymbol{\vartheta})$ is the likelihood function, which in this case is a Markov mixture of normals (Frühwirth-Schnatter 2006). Treating the state vector $\mathbf{S}$ as data, the Markov mixture likelihood can be expressed as the sum of the complete-data likelihood $p(\mathbf{y}, \mathbf{S}|\boldsymbol{\vartheta})$ over all possible values of $\mathbf{S}$:

$$\begin{aligned} p(\mathbf{y}|\boldsymbol{\vartheta}) &= \sum_{\mathbf{S} \in S_K} p(\mathbf{y}, \mathbf{S}|\boldsymbol{\vartheta}) \\ &= \sum_{\mathbf{S} \in S_K} p(\mathbf{y}|\mathbf{S}, \boldsymbol{\theta}_1 \dots, \boldsymbol{\theta}_K)p(\mathbf{S}|\boldsymbol{\xi}) \end{aligned} \tag{2}$$

As shown in Frühwirth-Schnatter (2006), expression (2) factors in a convenient way that makes estimation easier. In particular, if the prior assumes (i) the independence of the parameter vector $\boldsymbol{\theta}$ across regimes and (ii) the independence between parameters $\boldsymbol{\theta}$ and the transition matrix $\boldsymbol{\xi}$, i.e.,

$$p(\boldsymbol{\vartheta}) = \prod_{k=1}^{K} p(\boldsymbol{\theta}_k)p(\boldsymbol{\xi}) \tag{3}$$

then the complete-data posterior, i.e.,

$$p(\boldsymbol{\vartheta}|\mathbf{y}, \mathbf{S}) \propto \prod_{k=1}^{K} p(\boldsymbol{\theta}_k|\mathbf{y}, \mathbf{S})p(\boldsymbol{\xi}|\mathbf{S}) \tag{4}$$

*University of Bologna, Department of Economics - Piazza Scaravilli 2, 40126 Bologna, Italy. E-mail: graziano.moramarco@unibo.it.

factors in the same way as the complete-data likelihood $p(\mathbf{y}, \mathbf{S}|\boldsymbol{\vartheta})$. This facilitates the application of conventional Markov Chain Monte Carlo (MCMC) methods used for Bayesian estimation, in a context where, due to the Markov-switching behavior, the prior $p(\boldsymbol{\vartheta})$ and the posterior $p(\boldsymbol{\vartheta}|\mathbf{y})$ are not conjugate, and the posterior does not assume any convenient analytical form.

The posterior $p(\boldsymbol{\vartheta}|\mathbf{y})$ can be expressed as the sum of the posterior for the augmented parameter vector $(\mathbf{S}, \boldsymbol{\vartheta})$ over all possible realizations of $\mathbf{S}$:

$$p(\boldsymbol{\vartheta}|y) = \sum_{\mathbf{S}\in S_K} p(\mathbf{S}, \boldsymbol{\vartheta}|\mathbf{y}) \tag{5}$$

In practice, Bayesian estimation samples from the joint posterior $p(\mathbf{S}, \boldsymbol{\vartheta}|\mathbf{y})$, using:

$$p(\mathbf{S}, \boldsymbol{\vartheta}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{S}, \boldsymbol{\vartheta})p(\mathbf{S}|\boldsymbol{\vartheta})p(\boldsymbol{\vartheta}) \tag{6}$$

We estimate the model using MCMC methods and assuming independence priors of the following form:

$$p(\alpha_1, \ldots, \alpha_p, , \beta_1, \ldots, \beta_K, \sigma_1^2, \ldots, \sigma_K^2) = \prod_{j=1}^{p} p(\alpha_j) \prod_{k=1}^{K} p(\beta_k) \prod_{k=1}^{K} p(\sigma_k^2) \tag{7}$$

If a gamma hyper-prior is used for C_0, as is the case in our empirical application, the MSAR independence prior becomes:

$$p(\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_K, \sigma_1^2, \ldots, \sigma_K^2, C_0) = \prod_{j=1}^{p} p(\alpha_j) \prod_{k=1}^{K} p(\beta_k) \prod_{k=1}^{K} p(\sigma_k^2)p(C_0)$$

When several views are considered, Bayesian averaging across different views can performed in the following way. Let $\boldsymbol{\vartheta}_{K,i}^0$ denote the generic i-th view assuming K states. First, the number of regimes is treated as uncertain. Accordingly, a discrete prior is defined for K, fixing a maximum number $\overline{K}$:

$$\pi_K^0 = Pr\,(K) \tag{8}$$

for $K = 1, \ldots, \overline{K}$, with $\sum_{K=1}^{\overline{K}} \pi_K^0 = 1$. Second, assuming that a number P_K of alternative priors (views) are available for any given number of states K, a prior probability $\pi(\boldsymbol{\vartheta}_{K,i}^0|K)$ is assigned to $\boldsymbol{\vartheta}_{K,i}^0$, such that $\sum_{i=1}^{P_K} \pi(\boldsymbol{\vartheta}_{K,i}^0\,|K) = 1$. In other words, a discrete hierarchical prior is defined with respect to $\boldsymbol{\vartheta}$. The unconditional prior probability of $\boldsymbol{\vartheta}_{K,i}^0$ is equal to the joint prior probability of $\boldsymbol{\vartheta}_{K,i}^0$ and the number K of regimes, i.e., $\pi(\boldsymbol{\vartheta}_{K,i}^0) = \pi(\boldsymbol{\vartheta}_{K,i}^0, K)$. Using $\pi_{K,i}^0$ to denote this unconditional probability, we have that:

$$\pi_{K,i}^0 \equiv \pi(\boldsymbol{\vartheta}_{K,i}^0) = \pi(\boldsymbol{\vartheta}_{K,i}^0\,|K)\,\pi_K^0 \tag{9}$$

Also, let $\boldsymbol{\pi}^0$ denote the vector of length $\sum_{K=1}^{\overline{K}} P_K$ containing the unconditional prior probabilities of all views, i.e., $\boldsymbol{\pi}^0 = (\pi_{1,1}^0, \ldots, \pi_{\overline{K},P_{\overline{K}}}^0)$. Finally, the posterior probabilities of the views depend on the prior vector $\boldsymbol{\pi}^0$, collecting the probabilities defined in (9), and on the marginal likelihood of the MSAR model under the different views. In particular, the posterior probability for view $\boldsymbol{\vartheta}_{K,i}^0$ is equal to the joint posterior probability of $\boldsymbol{\vartheta}_{K,i}^0$ and the number K of regimes, i.e., $\pi(\boldsymbol{\vartheta}_{K,i}^0|\mathbf{y}) = \pi(\boldsymbol{\vartheta}_{K,i}^0, K|\mathbf{y})$, and is given by:

$$\pi_{K,i} \equiv \pi(\boldsymbol{\vartheta}_{K,i}^0|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\vartheta}_{K,i}^0)\pi_{K,i}^0}{\sum_{K=1}^{\overline{K}} \sum_{j=1}^{P_K} p(\mathbf{y}|\boldsymbol{\vartheta}_{K,j}^0)\pi_{K,j}^0} \tag{10}$$

where $p(\mathbf{y}|\boldsymbol{\vartheta}^0_{K,i}) = p(\mathbf{y}|\boldsymbol{\vartheta}^0_{K,i}, K) = \int p(\mathbf{y}|\boldsymbol{\vartheta}_K, \boldsymbol{\vartheta}^0_{K,i}, K)p(\boldsymbol{\vartheta}_K|\boldsymbol{\vartheta}^0_{K,i}, K)d\boldsymbol{\vartheta}_K$, with $\boldsymbol{\vartheta}_K$ denoting the parameter vector in the MSAR model with K regimes. From (10), we have that $\sum_{K=1}^{\overline{K}} \sum_{i=1}^{P_K} \pi_{K,i} = 1$.

A.2 Computing density forecasts

Computing density forecasts from the MSAR model requires three steps. Let $\mathbf{y}_t = (y_0, y_1, \ldots, y_t)$. Also, let us assume that the current time period is T and the forecast horizon is one period. The first step consists in using the MCMC algorithm to sample both the current unobserved regime S_T and the MSAR parameters $\boldsymbol{\vartheta}$ from the posterior distribution $p(\mathbf{S}, \boldsymbol{\vartheta}|\mathbf{y}_T)$. Let $(\boldsymbol{\vartheta}^{(d)}, S_T^{(d)})$ denote a generic MCMC draw. Next, each draw is used to forecast the future state of the economy. Taking $S_T^{(d)}$ as the starting value, a stochastic forecast $S_{T+1}^{(d)}$ is computed using the matrix of transition probabilities $\boldsymbol{\xi}^{(d)}$, i.e., based on equation (2) in the paper. Third, $y_{T+1}^{(d)}$ is sampled from the normal predictive density $p(y_{T+1}, |\mathbf{y}_T, \boldsymbol{\vartheta}^{(d)}, S_{T+1}^{(d)})$. In particular,

$$y_{T+1}|\mathbf{y}_T, \boldsymbol{\vartheta}^{(d)}, S_{T+1}^{(d)} = k \sim \mathcal{N}\left(\sum_{j=1}^{p} \alpha_j^{(d)} y_{T+1-j} + \beta_k^{(d)}, \sigma_k^{(d)2}\right) \tag{11}$$

Conditional on knowing the state of the economy in the future period $T+1$, the predictive distribution of y_{T+1} is a Normal for any given parameter vector. However, since the future state of the economy is unknown, the density forecast of y_{T+1} produced by the MSAR will be a mixture of the different regime-specific normals, where the mixture weights are given by the probabilities of the economy ending up in the different possible regimes at $T+1$. As a result, the MSAR generally produces non-normal forecast distributions. Also, the predictive densities are non-linear in y_T and heteroskedastic (Frühwirth-Schnatter 2006). More specifically, assuming a known number of regimes K and a known parameter vector $\boldsymbol{\vartheta}$, the one-step-ahead density forecast at time T is the following finite mixture of K normal components:

$$p\left(y_{T+1}|\mathbf{y}_T, \boldsymbol{\vartheta}\right) = \sum_{k=1}^{K} p\left(y_{T+1}|\mathbf{y}_T, \boldsymbol{\theta}_k\right) Pr\left(S_{T+1} = k|\mathbf{y}_T, \boldsymbol{\vartheta}\right) \tag{12}$$

Next, as a result of Bayesian estimation, the density forecast for any given view integrates out parameter uncertainty:

$$p\left(y_{T+1}|\mathbf{y}_T, \boldsymbol{\vartheta}^0_{K,i}\right) = \int p\left(y_{T+1}|\mathbf{y}_T, \boldsymbol{\vartheta}_K, \boldsymbol{\vartheta}^0_{K,i}\right) p(\boldsymbol{\vartheta}_K|\mathbf{y}_T, \boldsymbol{\vartheta}^0_{K,i}) d\boldsymbol{\vartheta}_K \tag{13}$$

where $\boldsymbol{\vartheta}_K$, as before, denotes the parameter vector when K regimes are assumed.

A.3 Prior on the regime-switching variance in the empirical application

Based on the properties of the gamma and inverse gamma distributions (see, e.g., Frühwirth-Schnatter 2006), it holds that:

$$\mathrm{E}(\sigma_k^2|C_0) = \frac{C_0}{c_0 - 1} \tag{14}$$

$$\mathrm{Var}(\sigma_k^2|C_0) = \frac{C_0^2}{(c_0-1)^2(c_0-2)} \tag{15}$$

$$\mathrm{E}(C_0) = \frac{g_0}{G_0} \tag{16}$$

$$\mathrm{Var}(C_0) = \frac{g_0}{G_0^2} \tag{17}$$

$$\mathrm{E}(C_0^2) = \left(\frac{g_0}{G_0}\right)^2 + \frac{g_0}{G_0^2} \tag{18}$$

$$\tag{19}$$

Given the hyperparameter values $c_0 = 3$, $g_0 = 0.5$ and $G_0 = 0.5$, it follows that:

$$\mathrm{E}(\sigma_k^2) = \frac{\mathrm{E}(C_0)}{c_0 - 1} = 0.5 \tag{20}$$

$$\mathrm{Var}(\sigma_k^2) = \mathrm{E}(\mathrm{Var}(\sigma_k^2|C_0)) + \mathrm{Var}(\mathrm{E}(\sigma_k^2|C_0)) = \tag{21}$$

$$= \frac{\mathrm{E}(C_0^2)}{(c_0-1)^2(c_0-2)} + \frac{\mathrm{Var}(C_0^2)}{(c_0-1)^2} = \tag{22}$$

$$= \frac{3}{4} + \frac{1}{2} = 1.25 \tag{23}$$

Appendix B Fed supervisory scenarios

Table A1: Fed stress tests 2015-2018: scenarios of GDP growth

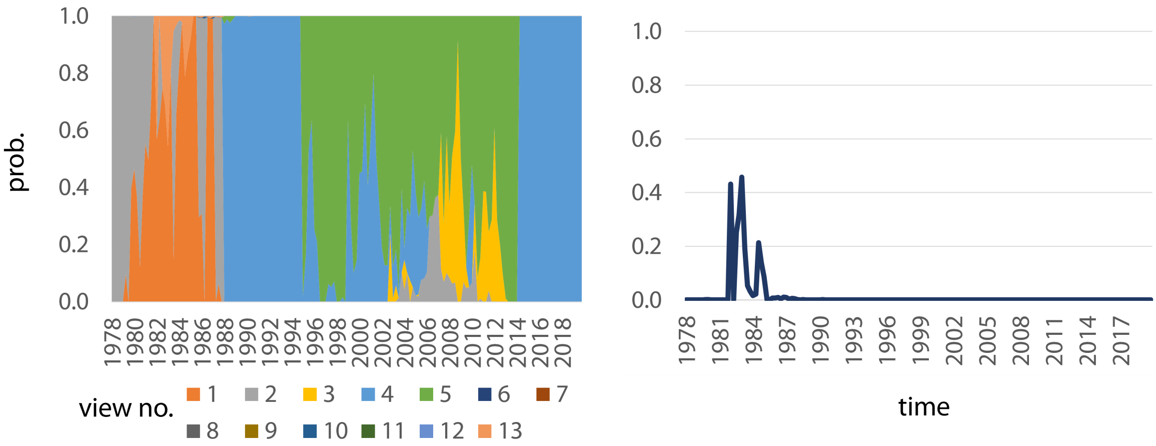
	2015			2016			2017			2018		
time	base	adv.	sev.	base	adv.	sev.	base	adv.	sev.	base	adv.	sev.
2014Q4	3	-0.6	-3.9									
2015Q1	2.9	-1.3	-6.1									
2015Q2	2.9	-0.2	-3.9									
2015Q3	2.9	0.2	-3.2									
2015Q4	2.9	0.3	-1.5									
2016Q1	2.9	0.8	1.2	2.5	-1.5	-5.1						
2016Q2	2.9	1.2	1.2	2.6	-2.8	-7.5						
2016Q3	2.9	1.7	3	2.6	-2	-5.9						
2016Q4	2.9	1.8	3	2.5	-1.1	-4.2						
2017Q1	2.7	1.8	3.9	2.4	0	-2.2	2.2	-1.5	-5.1			
2019Q4	2.7	1.9	3.9	2.5	1.3	0.4	2.3	-2.8	-7.5			
2017Q3	2.6	2	3.9	2.3	1.7	1.3	2.4	-2	-5.9			
2017Q4	2.6	2.2	3.9	2.3	2.6	3	2.3	-1.5	-5.1			
2018Q1				2.6	2.6	3	2.4	-0.5	-3	2.5	-1.3	-4.7
2018Q2				2.4	3	3.9	2.4	1	0	2.8	-3.5	-8.9
2018Q3				2.3	3	3.9	2.4	1.4	0.7	2.6	-2.4	-6.8
2018Q4				2.3	3	3.9	2.3	2.6	3	2.5	-1.3	-4.7
2019Q1				2.1	3	3.9	2	2.6	3	2.3	-0.7	-3.6
2019Q2							2.1	3	3.9	2.3	0.4	-1.3
2019Q3							2.1	3	3.9	2.1	1	-0.2
2019Q4							2	3	3.9	2	2.5	2.8
2020Q1							2	3	3.9	2.1	2.8	3.5
2020Q2										2.1	3	4
2020Q3										2.1	3.2	4.2
2020Q4										2.1	3.3	4.5
2021Q1										2.1	3.3	4.5

Notes: For each year between 2015 and 2018 the table reports the baseline, adverse and severely adverse supervisory scenarios for U.S. GDP growth (annualized quarter-on-quarter, in percentage) included in the annual stress test conducted by the Federal Reserve (see Federal Reserve Board 2014, 2016, 2017, 2018).

Appendix C Additional Figures

Figure A1: Optimal log-score-based posterior probabilities (prior $\boldsymbol{\pi}_1^{0*}$) over time

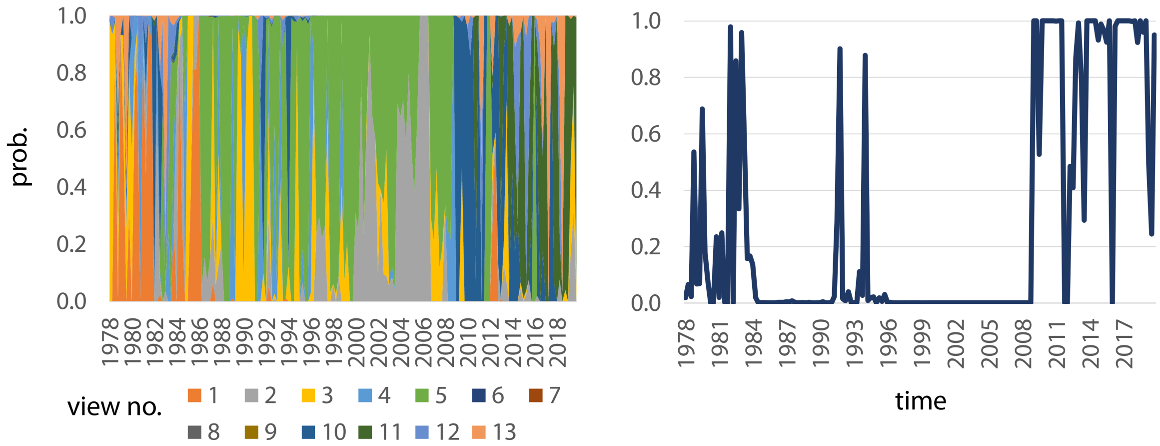
(a) All views (b) Fed-based views (cumulative)



Notes: The area chart in the left panel shows the time-varying Bayesian posterior probabilities for all the 13 views used to estimate the Markov-switching AR model. The chart goes from 1978Q1 to 2019Q4. The underlying prior probabilities $\boldsymbol{\pi}_1^{0*}$ are obtained using the log-score-based optimization procedure described in the paper. The right panel plots the cumulative weight assigned to the views derived from Fed supervisory scenarios (views 6-13). See Table 1 in the paper for the list of views.

Figure A2: Optimal PIT-based posterior probabilities (prior $\boldsymbol{\pi}_2^{0*}$) over time

(a) All views (b) Fed-based views (cumulative)



Notes: The area chart in the left panel shows the time-varying Bayesian posterior probabilities for all the 13 views used to estimate the Markov-switching AR model. The chart goes from 1978Q1 to 2019Q4. The underlying prior probabilities $\boldsymbol{\pi}_2^{0*}$ are obtained using the PIT-based optimization procedure described in the paper, where PIT stands for probability integral transform. The right panel plots the cumulative weight assigned to the views derived from Fed supervisory scenarios (views 6-13). See Table 1 in the paper for the list of views.

References

Federal Reserve Board (2014), "2015 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule," Federal Reserve Board report.

Federal Reserve Board (2016), "2016 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule," Federal Reserve Board report.

Federal Reserve Board (2017), "2017 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule," Federal Reserve Board report.

Federal Reserve Board (2018), "2018 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule," Federal Reserve Board report.

Frühwirth-Schnatter, S. (2006), *Finite Mixture and Markov Switching Models*, Springer Series in Statistics, Springer, New York.