Technical Appendix for Interest Rate Rules in Practice - the Taylor Rule or a Tailor-Made Rule?

Adam Check*

November 3, 2015

1 Details of Full Sample Estimation

To perform full sample estimation with known break dates, I use an independent Normal-Inverse Gamma prior. That is, I assume:

$$p(\beta, \sigma_1^2, \sigma_2^2, \sigma_3^2) = p(\beta)p(\sigma_1^2)p(\sigma_2^2)p(\sigma_3^2)$$

I assume that $p(\beta)$ follows a normal distribution, so that:

$$\beta \sim N(0, V_{pri})$$

I assume that $p(\sigma_1^2)$, $p(\sigma_2^2)$, and $p(\sigma_3^2)$ follow independent inverse-gamma distributions, so that:

$$\sigma_1^2 \sim IG(\alpha_{1,pri}, \beta_{1,pri})$$
$$\sigma_2^2 \sim IG(\alpha_{2,pri}, \beta_{2,pri})$$
$$\sigma_3^2 \sim IG(\alpha_{3,pri}, \beta_{3,pri})$$

Due to the fact that there are two known break dates, it will be helpful to partition the

*I can be contacted at ajc@uoregon.edu

X matrix and Y vector into three blocks. That is,

$$X = [X_1 \ X_2 \ X_3]'$$
$$Y = [Y_1 \ Y_2 \ Y_3]'$$

where X_1 contains all variables in X between the start date at the first break date, X_2 contains all variables in X between the first break date and the second break date, and X_3 contains all variables in X following the second break date, and likewise for Y_1, Y_2, and Y_3.

Because there are three variance regimes, it will also be helpful to define transformations of X and X_i for $i = 1, 2, 3$. Let Σ be a vector of length T that holds the variance parameter corresponding to each time period. That is, $\Sigma_i = [\sigma_i^2 \iota_{T_i}']'$ for $i = 1, 2, 3$, where ι_{T_i} is a vector of ones of length T_i, which is the number of periods that the i^{th} regime is in place. To construct Σ, we stack these partitions $[\Sigma_1' \Sigma_2' \Sigma_3']'$. Then, we define $X_{i,\sigma}$ as X_i divided element-wise by $\sqrt{\Sigma_i}$, and X_σ as X divided element-wise by $\sqrt{\Sigma}$.

Given the priors, the kernel of the unconditional posterior distributions do not correspond to known distributions. However, as shown in Koop (2003), chapter 4, the kernels of the *conditional* posterior distributions do correspond to known distributions. Working through the algebra, we find:

$$p(\beta | X, Y, \sigma_1^2, \sigma_2^2, \sigma_3^2) \propto \exp\left[\frac{-1}{2}(\beta - \beta_{post})' V_{post} (\beta - \beta_{post})\right]$$

i.e. $(\beta | X, Y, \sigma_1^2, \sigma_2^2, \sigma_3^2) \sim N(\beta_{post}, V_{post})$ where

$$V_{post} = (V_{pri}^{-1} + X_\sigma' X_\sigma)^{-1}$$
$$\beta_{post} = V_{post}(V_{pri}^{-1}\beta_{pri} + X_\sigma' Y)$$

The kernel of the conditional posterior distribution for σ_i^2 is given by:

$$p(\sigma_i^2|X,Y,\beta) \propto \beta_{i,post}^{\alpha_{i,post}} (\sigma_i^2)^{-\alpha-1} \exp\left(\frac{-\beta_{i,post}}{\sigma_i^2}\right)$$

for $i = 1, 2, 3$. Therefore, we have $(\sigma_i^2|X,Y,\beta) \sim IG(\alpha_{i,post}, \beta_{i,post})$ where:

$$\begin{aligned}\alpha_{i,post} &= \alpha_{i,pri} + \frac{T_i}{2} \\ \beta_{i,post} &= \beta_{i,pri} + \frac{(Y_i - X_{i,\sigma}\beta)'(Y_i - X_{i,\sigma}\beta)}{2}\end{aligned}$$

Summarizing, we have two blocks of parameters:

$$\begin{aligned}(\beta|X,Y,\sigma_1^2,\sigma_2^2,\sigma_3^2) &\sim N(\beta_{post}, V_{post}) \\ (\sigma_i^2|X,Y,\beta) &\sim IG(\alpha_{i,post}, \beta_{i,post})\end{aligned}$$

In order to perform inference, I utilize the Gibbs sampler. To initialize the sampler, I set

$$\begin{aligned}\beta &= \beta^{(0)} \\ \sigma_1^2 &= \sigma_1^{2,(0)} \\ \sigma_2^2 &= \sigma_2^{2,(0)} \\ \sigma_3^2 &= \sigma_3^{2,(0)}\end{aligned}$$

Then, I perform the following steps $G_0 + G$ times, where the first G_0 draws serve as burn-in, and the next G are used for inference:

1. Draw $(\beta^{(j)}|X,Y,\sigma_1^{2,(j-1)},\sigma_2^{2,(j-1)},\sigma_3^{2,(j-1)}) \sim N(\beta_{post}, V_{post})$.

2. Draw $(\sigma_i^{2,(j)}|X,Y,\beta^{(j)}) \sim IG(\alpha_{i,post}, \beta_{i,post})$ for $i = 1, 2, 3$

Running the Gibbs sampler simulates from the posterior distribution. We can then use these

Note that in the limit, as $|V_{pri}| \to \infty$ and $\beta_{pri} \to 0$, the mean of $\beta_{post} \to (X_\sigma' X_\sigma)^{-1} X_\sigma' Y$

posterior draws to compute most posterior features of interest, such as the central tendency of the parameter values or the central tendency or credible intervals for the fitted values, $\hat{Y}$.

Because I am interested in model comparison, I also need to compute a measure of how well the model fits the data, for each model that I consider. As in most Bayesian model comparison exercises, I use the marginal likelihood for each model. Due to the use of priors that admit only *conditionally* conjugate posterior distributions, this value cannot be calculated analytically, and instead needs to be approximated. Using Bayes' rule, the marginal likelihood can be written in terms of the following decomposition:

$$\begin{aligned} p(\theta|y) &= \frac{p(Y|\theta)p(\theta)}{p(Y)} \\ \Rightarrow p(Y) &= \frac{p(Y|\theta)p(\theta)}{p(\theta|Y)} \end{aligned}$$

where $p(Y)$ is the marginal likelihood, θ is the vector of parameters of the model, $p(Y|\theta)$ is the likelihood, $p(\theta)$ is the prior, and $p(\theta|Y)$ is the posterior. Since the marginal likelihood is unconditional with respect to the value of the parameter vector, we can plug in any value, $\theta = \theta^*$, to evaluate the expression:

$$p(Y) = \frac{p(Y|\theta^*)p(\theta^*)}{p(\theta^*|Y)}$$

In our case, we know the analytic expressions for the prior and the likelihood function, so the values of $p(Y|\theta^*)$ and $p(\theta^*)$ can be computed analytically at any value of θ^*. However, as discussed earlier, we do not know the analytical form of the posterior distribution, $p(\theta|Y)$. Therefore, this value is approximated. First, I decompose the parameter vector into two blocks, $\theta = [\theta_1\ \theta_2]$, with $\theta_1 = \beta$ and $\theta_2 = [\sigma_1^2\ \sigma_2^2\ \sigma_3^2]'$, and write the posterior as $p(\theta_1, \theta_2|Y) = p(\theta_1|Y)p(\theta_2|Y, \theta_1)$. Then, I approximate $p(\theta_1^*|Y)$ and $p(\theta_2^*|Y, \theta_1^*)$ using the following steps adapted from the procedure in Chib (1995):

1. Set $\theta_1 = \theta_1^* = \beta^*$, where β^* is the posterior median of β computed from the output of

the Gibbs sampler.

2. Compute $p(\beta^*|Y) \approx \frac{1}{G}\sum_{j=1}^{G} p(\beta^*|Y, \sigma_1^{2,(j)}\ \sigma_2^{2,(j)}\ \sigma_3^{2,(j)})$

3. Set $\theta_2 = \theta_2^* = [\sigma_1^{2,*}\ \sigma_2^{2,*}\ \sigma_3^{2,*}]'$, where $\sigma_i^{2,*}$ is the posterior median of σ_i^2 computed from the output of the Gibbs sampler.

4. Compute $p(\theta_2^*|Y, \theta_1^*) = \prod_{i=1}^{3} p(\sigma_i^{2,*}|Y, \beta^*)$.

5. Compute $p(\theta^*|Y) = p(\theta_1^*, \theta_2^*|Y) = p(\theta_1^*|Y)p(\theta_2^*|Y, \theta_1^*)$.

Since I have analytical expressions for $p(\theta^*)$ and $p(Y|\theta^*)$, after completing this process I have all I need to compute the marginal likelihood.

Summarizing the complete process:

1. Choose a model, M_r from the feasible set that has not yet been chosen.

2. Set priors in this model according to a version of the g-prior adapted to the independent Normal-Inverse Gamma prior.

3. Draw from the posterior distributions by using the Gibbs Sampler.

4. Compute the marginal likelihood for model M_r, $p(Y|M_r)$, by using the method of Chib (1995) to approximate the posterior likelihood, $p(\theta^*|Y, M_r)$.

5. Repeat steps 1-4 2^p times, where p is the number of potentially included variables.

6. Compute the relative model probabilities using:

$$p(M_r) \propto p(Y|M_r)p(M_r)$$

   where $p(M_r)$ is the prior probability of model M_r. This simplifies to $p(M_r) \propto p(Y|M_r)$ in my baseline estimation, which assumes equal prior probability on each model.

2 Details of Estimation used in Forecasting Exercise

For forecasting, I only use the post-1983 sample. Therefore, I assume that there are no structural breaks. Because there are no structural breaks, I can use the natural conjugate prior for the linear regression model. This procedure is standard, and no longer requires any simulation, as all distributions, including the posterior, take known analytic forms. Analysis in this model is standard, and full details can be found in chapter 11 of Koop (2003)[1]

[1] There are errata in some of the formulae provided in the text, corrections can be found at http://www.wiley.com/legacy/wileychi/koopbayesian/

References

Chib, S. (1995). Marginal likelihood from the gibbs output. *Journal of the American Statistical Association*, 90:1313–1321.

Koop, G. (2003). *Bayesian Econometrics*. Wiley.