

\*\*\*\*Exercise 1. Title: Using the AR(p) model to Understand the Properties of a Series\*\*\*\*

This exercise is loosely based on Geweke (1988). Let  $y_t$  for  $t = 1, \dots, T$  indicate observations on a time series variable.  $y_t$  is assumed to follow an AR(p) process:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \epsilon_t, \quad (1)$$

where  $\epsilon_t$  is i.i.d.  $N(0, h^{-1})$ . Many important properties of  $y_t$  depend on the roots of the polynomial  $1 - \sum_{i=1}^p \beta_i z^i$  which we will denote by  $r_i$  for  $i = 1, \dots, p$ . Geweke (1988) lets  $y_t$  be the log of real GDP and sets  $p = 3$  and, for this choice, focusses on the features of interest:  $C = \{\beta : \text{Two of } r_i \text{ are complex}\}$  and  $D = \{\beta : \min |r_i| < 1\}$  where  $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$ . Note that  $C$  and  $D$  are regions whose bounds are complicated nonlinear functions of  $\beta_1, \beta_2, \beta_3$ . If the AR coefficients lie in the region defined by  $C$  then real GDP exhibits an oscillatory response to a shock and if they lie in  $D$  then  $y_t$  exhibits an explosive response to a shock.

(a) Assuming a prior of the form:

$$p(\beta_0, \dots, \beta_p, h) \propto \frac{1}{h}, \quad (2)$$

derive the posterior for  $\beta$ . To simplify things, you may ignore the (minor) complications relating to the treatment of initial conditions. Thus, assume the dependent variable is  $y = (y_{p+1}, \dots, y_T)'$  and treat  $y_1, \dots, y_p$  as fixed initial conditions.

(b) Using an appropriate data set (e.g., the US real GDP data set provided on the website associated with this book), write a program which calculates the posterior means and standard deviations of  $\beta$  and  $\min |r_i|$ .

(c) Extend the program of part b) to calculate the probability that  $y_t$  is oscillatory (i.e.,  $\Pr(\beta \in C|y)$ ), the probability that  $y_t$  is explosive (i.e.,  $\Pr(\beta \in D|Data)$ ) and calculate these probabilities using your data set.

\*\*\*Solution to Exercise 1.\*\*\*\*\*

(a) This model can be written as a linear regression model in matrix notation as:

$$y = X\beta + \epsilon, \quad (3)$$

where  $\epsilon = (\epsilon_{p+1}, \dots, \epsilon_N)'$  and  $X$  is the  $(T - p) \times (p + 1)$  matrix containing an intercept and  $p$  lags of the dependent variable. For instance, if  $p = 3$

$$X = \begin{bmatrix} 1 & y_3 & y_2 & y_1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & y_{T-1} & y_{T-2} & y_{T-3} \end{bmatrix}.$$

Written in this way, it can be seen that (if we condition on the initial observations), the AR(p) model is simply the Normal linear regression model. The easiest way to answer this exercise is to note that the prior in (17.2) is the standard noninformative prior for the Normal linear regression model and is a limiting case of the natural conjugate prior. Thus, the answer to this exercise is exactly the same as the answer to Exercise 10.1 with particular prior hyperparameter values used. In particular, you can set  $\underline{\nu} = 0$  and  $\underline{Q}^{-1} = cI$ . The noninformative prior in (17.2) is obtained by letting the scalar  $c$  go to zero. Alternatively, multiplying the Normal linear regression model likelihood in (10.4) by the prior in (17.2) yields the relevant posterior. Manipulations exactly analogous to the solution to Exercise 10.1 can be done to confirm that the posterior is Normal-Gamma. To be precise, the posterior for  $\beta$  and  $h$  is  $NG(\bar{\beta}, \bar{Q}, \bar{s}^{-2}, \bar{\nu})$  where  $\bar{\nu} = T$ ,

$$\bar{Q} = (X'X)^{-1},$$

$$\bar{\beta} = (X'X)^{-1} X'y$$

and  $\bar{\nu}\bar{s}^2$  are the standard OLS residuals (denoted  $SSE$  in the introductory material of Chapter 10).

Using the properties of the Normal-Gamma distribution (see Theorem 8 of the Appendix), it follows that the marginal posterior for  $\beta$  is multivariate t:

$$\beta|y \sim t(\bar{\beta}, \bar{s}^2\bar{Q}, \bar{\nu}) \tag{4}$$

(b) From part a), the posterior for  $\beta$  has a convenient form and means and standard deviations can be calculated using analytical formulae based on (17.4). That is, using the properties of the t distribution, we have  $E(\beta|y) = \bar{\beta}$  and  $var(\beta|y) = \frac{\bar{\nu}\bar{s}^2}{\bar{\nu}-2}\bar{Q}$ . However,  $\min|r_i|$  is a complicated nonlinear function of  $\beta$  and no analytical results exist for its posterior mean and variance. However, Monte Carlo integration (see Exercise 11.3) can be done by randomly drawing from (17.4), using a subroutine to calculate the solutions to

$1 - \sum_{i=1}^p \beta_i z^i = 0$  and, from these, evaluating  $\min|r_i|$  at each of the draws. This will produce a random sample from the posterior of  $\min|r_i|$  which can be used to calculate its posterior mean and standard deviation (or any other function of its posterior). A program which does this is available on the website associated with this book.

Using the log of the quarterly US real GDP series provided on the website, which runs from 1947Q1 through 2005Q2, we obtain the results in Table 17.1. The Monte Carlo integration program was run for 10,000 replications.

Table 17.1: Posterior Results for AR(3) Model		
Parameter	Mean	Standard Dev.
$\beta_0$	0.010	0.004
$\beta_1$	1.300	0.066
$\beta_2$	-0.196	0.108
$\beta_3$	-0.104	0.066
$\min r_i $	1.002	0.002

(c)  $\Pr(\beta \in C|y)$  and  $\Pr(\beta \in D|Data)$  can also be calculated in the Monte Carlo integration program. Formally, since  $\Pr(\beta \in C|y) = E[I(\beta \in C|y)]$  and  $\Pr(\beta \in D|y) = E[I(\beta \in D|y)]$ , where  $I(\cdot)$  is the indicator function, the probabilities can be written as expected values of functions of the model parameters. In our computer code, calculating these expected values is equivalent to simply calculating the proportion of draws from the posterior which are in  $C$  (or  $D$ ). Our Monte Carlo integration program, using 10,000 replications, calculates  $\Pr(\beta \in C|y) = 0.021$  and  $\Pr(\beta \in D|y) = 0.132$ . Thus, it is very unlikely that real GDP exhibits oscillatory behavior and only slightly more likely that it is explosive.

\*\*\*\*Exercise 2. Title: The Threshold Autoregressive Model\*\*\*\*

Dynamics of many important macroeconomic variables can potentially vary over the business cycle. This has motivated the development of many models where different autoregressive representations apply in different regimes. Threshold autoregressive (TAR) models are a class of simple and popular regime-switching models. Potter (1995) provides an early exposition of these models in macroeconomics and Geweke and Terui (1993) is an early Bayesian treatment. This exercise asks you to derive Bayesian methods for a simple variant of a TAR model.

Consider a two regime TAR for a time series variable  $y_t$  for  $t = p + 1, \dots, T$  (where  $t = 1, \dots, p$  are used as initial conditions):

$$\begin{aligned} y_t &= \beta_{10} + \beta_{11}y_{t-1} + \dots + \beta_{1p}y_{t-p} + \epsilon_t \text{ if } y_{t-1} \leq \tau \\ y_t &= \beta_{20} + \beta_{21}y_{t-1} + \dots + \beta_{2p}y_{t-p} + \epsilon_t \text{ if } y_{t-1} > \tau \end{aligned} \quad (5)$$

where  $\epsilon_t$  is i.i.d.  $N(0, h^{-1})$ . We will use the notation  $\beta = (\beta'_1, \beta'_2)'$  where  $\beta_j = (\beta_{j0}, \beta_{j1}, \dots, \beta_{jp})'$  for  $j = 1, 2$ . For all parts of this question, you may proceed conditionally on the first  $p$  observations and, thus, ignore the (minor) complications caused by initial conditions [see Exercise 17.1 (a) for more detail].

(a) Assume that  $\tau$  is known (e.g.,  $\tau = 0$ ) and Normal-Gamma priors are used (i.e., the joint prior for  $\beta$  and  $h$  is  $NG(\underline{\beta}, \underline{Q}, \underline{s}^{-2}, \underline{\nu})$ ), derive the posterior for the model defined by (17.5).

(b) Using an appropriate data set (e.g., the real GDP data set available on the website associated with this book), write a program and carry out Bayesian inference in the TAR model using your results from part (a). Note: When working with GDP, TAR models are usually specified in terms of GDP growth. Hence, if you are working with a GDP series, you should define  $y_t$  as its first difference.

(c) Repeat parts (a) and (b) assuming that  $\tau$  is an unknown parameter.

\*\*\*\*Solution to Exercise 2

(a) The solution to this question follows immediately from noting that the TAR with known  $\tau$  can be written as a Normal linear regression model. That is, analogously to the solution to Exercise 17.1(a), we can write the TAR as:

$$y = X\beta + \epsilon,$$

where  $y = (y_{p+1}, \dots, y_T)'$ ,  $\epsilon = (\epsilon_{p+1}, \dots, \epsilon_N)'$  and  $X$  is the  $(T - p) \times 2(p + 1)$  matrix with  $t^{th}$  row given by  $[D_t, D_t y_{t-1}, \dots, D_t y_{t-p}, (1 - D_t), (1 - D_t) y_{t-1}, \dots, (1 - D_t) y_{t-p}]$ , where  $D_t$  is a dummy variable which equals 1 if  $y_{t-1} \leq \tau$  and equals 0 if  $y_{t-1} > \tau$ . Thus, we can immediately use the results from Exercise 10.1 to find that the joint posterior for  $\beta$  and  $h$  is  $NG(\bar{\beta}, \bar{Q}, \bar{s}^{-2}, \bar{\nu})$ , where  $\bar{\nu} = T - p + \underline{\nu}$ ,

$$\bar{Q} = (\underline{Q}^{-1} + X'X)^{-1},$$

$$\bar{\beta} = \bar{Q} (\underline{Q}^{-1} \underline{\beta} + X'X \hat{\beta})$$

and

$$\bar{s}^2 = \frac{\nu s^2 + SSE + (\hat{\beta} - \underline{\beta})' X' X \bar{Q} Q^{-1} (\hat{\beta} - \underline{\beta})}{\bar{\nu}}.$$

Note that, in the preceding equation, we are using the same notation defined at the beginning of Chapter 10 where  $\hat{\beta}$  is the standard OLS estimate and  $SSE$  the accompanying sum of squared residuals.

(b) A MATLAB program which provides the solution to this part of the exercise is available on the website associated with this book. The basic structure of this program is that it loads in the data, specifies prior hyperparameters then evaluates the formulae provided in the solution to part (a). For US real GDP growth from 1947Q2 through 2005Q2, we obtain the following results relating to the posteriors of  $\beta$  and  $\sigma^2 = \frac{1}{h}$ . We set  $p = 2$  and use the noninformative variant of the Normal-Gamma prior. Furthermore, we set  $\tau$  to the mean of  $y_{t-1}$  and thus the two regimes can be interpreted as "below average growth" and "above average growth" regimes.

Parameter	Mean	Standard Dev.
$\beta_{10}$	0.572	0.100
$\beta_{11}$	0.289	0.129
$\beta_{12}$	0.008	0.092
$\beta_{20}$	0.213	0.232
$\beta_{21}$	0.394	0.130
$\beta_{22}$	0.218	0.093
$\sigma^2$	0.879	0.083

(c) When  $\tau$  is treated as an unknown parameter, then we need methods of working with the posterior:  $p(\beta, h, \tau|y)$ . This can be done by noting that the rules of conditional probability imply:

$$p(\beta, h, \tau|y) = p(\beta, h|\tau, y) p(\tau|y).$$

The first component on the right hand side of this equation,  $p(\beta, h|\tau, y)$ , can be analyzed using the methods of parts (a) and (b). That is, conditioning upon  $\tau$  is equivalent to assuming it known and the answer to part (a) shows how to analyze the posterior for  $\beta$  and  $h$  for known  $\tau$ . Hence, we can focus on the second component:  $p(\tau|y)$ . But, the intuition that conditioning on  $\tau$  is equivalent to assuming it known, can be carried further to provide us

with a way of obtaining  $p(\tau|y)$ . Note first that, for a given value for  $\tau$ , the marginal likelihood can be calculated in the standard way:

$$p(y|\tau) = \int \int L(\beta, h|\tau) p(\beta, h|\tau) d\beta dh, \quad (6)$$

where  $L(\beta, h|\tau)$  is the likelihood function for a given value of  $\tau$ . However, as we have seen from part (a), the TAR model (conditional on  $\tau$ ) is simply a Normal linear regression model. Thus, using the Normal-Gamma natural conjugate prior, standard textbook expressions exist for (17.6) – conditional on  $\tau$  [see, e.g., Exercise 10.4 or Koop (2003), page 41 or Poirier (1995), page 543]. If we use the notation from the solution to part (a), then we can draw on these earlier derivations to write (17.6) as:

$$p(y|\tau) = \frac{\Gamma\left(\frac{\nu}{2}\right) (\nu s^2)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right) \pi^{\frac{T-p}{2}}} \left(\frac{|\overline{Q}|}{|Q|}\right)^{\frac{1}{2}} (\nu \overline{s}^2)^{-\frac{\nu}{2}}, \quad (7)$$

although we stress that  $X$  depends on  $\tau$  (and, thus, so will  $\overline{Q}$  and  $\overline{s}^2$ ) and, hence, (17.7) will be different for every possible value of  $\tau$ .

Since Bayes' theorem implies:

$$p(\tau|y) \propto p(y|\tau) p(\tau), \quad (8)$$

we can combine (17.7) with a prior for  $\tau$  in order to obtain the marginal posterior  $p(\tau|y)$ . Of course, any prior for  $\tau$  can be used. A common choice is a restricted noninformative one. This treats every possible value for  $\tau$  as, a priori, equally likely which implies each regime contains a minimum number of observations (e.g., 15% of the observations).

Note that, even though  $y_{t-1}$  is a continuous variable,  $\tau$  will be a discrete random variable since there are a finite number of ways of dividing a given data set into two regimes. Hence, in practice, posterior inference in the TAR with an unknown threshold can be performed by evaluating (17.7) and, thus, (17.8) for every allowable threshold value. This provides us with  $p(\tau|y)$ . If interest centers on  $\beta$  and/or  $h$ , then we can obtain their posterior properties by taking a weighted average over all possible thresholds, using  $p(\tau|y)$  as weights. The validity of this procedure follows from the fact that, if  $\tau \in \{\tau_1, \dots, \tau_{T^*}\}$  denotes the set of possible threshold values, then

$$p(\beta, h|y) = \sum_{i=1}^{T^*} p(\beta, h|\tau = \tau_i, y) p(\tau = \tau_i|y).$$

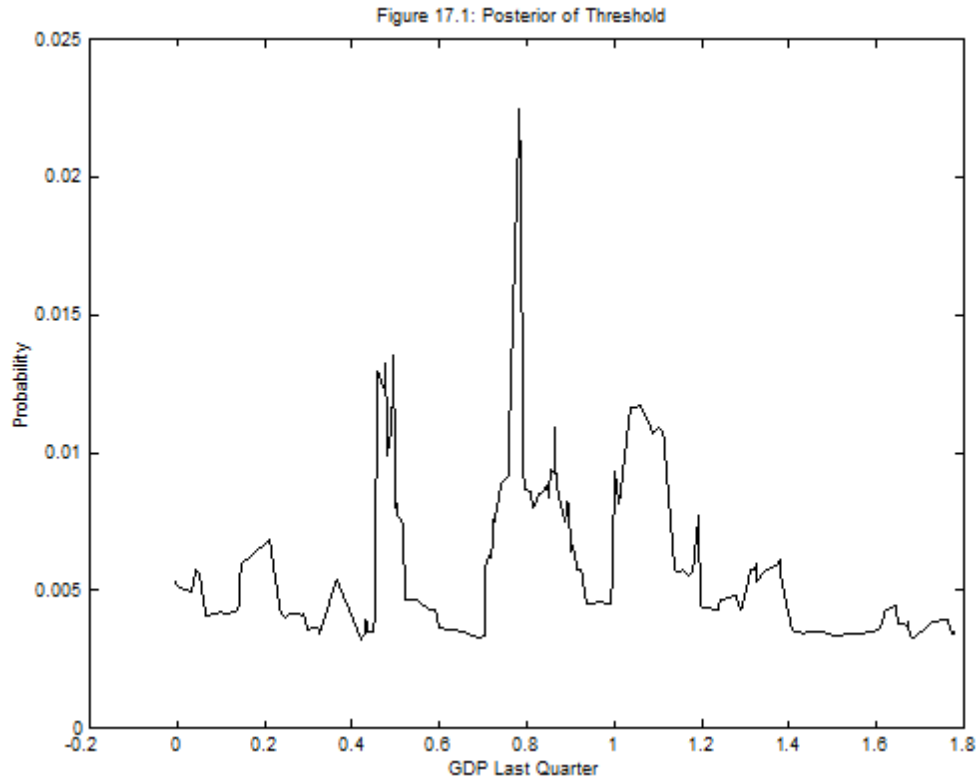
Thus, insofar as we are willing to use the Normal-Gamma natural conjugate prior, Bayesian inference in the TAR model can be performed using only textbook analytical results for the Normal linear regression model and no posterior simulation is required.

A MATLAB program which extends the program part (b) to allow for an unknown threshold is provided on the website associated with this book. Here we note one aspect of our program which is commonly used with regime-switching models. The assumption about the errors implies that we have added enough lags to make the errors independent of one another. Thus,  $y_t|X_t, \beta, h$  and  $y_s|X_s, \beta, h$  are independent of one another for  $t \neq s$ . This implies that the ordering of rows in our regression model is irrelevant. Thus, we can order all our variables (dependent and explanatory variable) by  $y_{t-1}$  without changing the posterior. That is, variables can be ordered so that the first observation of all variables is the one with the lowest value for GDP growth last quarter, the second observation has the second highest value for GDP growth last quarter, etc.. If we order our data in this way, then the TAR is equivalent to a structural break (also called changepoint) model (see Exercises 11.11 and 11.12). Such an ordering makes the step "evaluate (17.7) for every allowable threshold" particularly easy.

Posterior properties of  $\beta$  and  $\sigma^2$ , using the US real GDP growth data, are provided in Table 17.3. Figure 17.1 plots the posterior of  $\tau$ . Since we are calculating marginal likelihoods, an informative prior is required. We use a weakly informative prior which reflects our belief that AR coefficients are likely to be well within the stationary region and the error variance is likely to be approximately 1 (since our data is measured as a percentage). However, we are quite uncertain about these prior means and, hence, attach moderately large prior variances to them. Prior hyperparameter values consistent with these beliefs are  $\underline{\beta} = 0, \underline{Q} = .25I, \underline{s}^2 = 1, \underline{\nu} = 5$ . In a serious empirical study, the researcher would likely elicit the prior in a more sophisticated manner and/or carry out a prior sensitivity analysis. The posterior properties of  $\beta$  and  $h$  are not that different from those found in part (b), despite the fact that we are now averaging over all the thresholds (rather than selecting a single value for the threshold and treating it as known). However, the posterior for the thresholds (see Figure 17.1), is very irregular. Appreciable posterior probability is found for many of possible thresholds. The posterior mode is roughly the same as the average quarterly growth rate of GDP (0.85%). This threshold value would divide the data into a "below average growth" regime and an "above average growth" regime, as was done in part (a). But many

other threshold values also receive some posterior support from the data.

Parameter	Mean	Standard Dev.
$\beta_{10}$	0.542	0.119
$\beta_{11}$	0.269	0.139
$\beta_{12}$	0.373	0.101
$\beta_{20}$	0.249	0.242
$\beta_{21}$	0.384	0.127
$\beta_{22}$	0.196	0.100
$\sigma^2$	0.896	0.083



\*\*\*Exercise 3. Title: Extensions of the Basic Threshold Autoregressive Model 1: Other Threshold Triggers

There are many extensions of the TAR which have been found useful in empirical work. In Exercise 17.2, we assumed that the first lag of the



dependent variable (last quarter's GDP growth) triggered the regime switch. However, in general, it might be another variable,  $z$ , that is the *threshold trigger* and it may take longer than one period to induce the regime switch. Thus, we use the same assumptions and definitions as in Exercise 17.2, except we replace (17.5) by:

$$\begin{aligned} y_t &= \beta_{10} + \beta_{11}y_{t-1} + \dots + \beta_{1p}y_{t-p} + \epsilon_t \text{ if } z_{t-d} \leq \tau \\ y_t &= \beta_{20} + \beta_{21}y_{t-1} + \dots + \beta_{2p}y_{t-p} + \epsilon_t \text{ if } z_{t-d} > \tau \end{aligned} \quad (9)$$

where  $d$  is the *delay parameter* and  $z_{t-d}$  is either an exogenous variable or a function of the lags of the dependent variable.

(a) Assume that  $d$  is an unknown parameter with a noninformative prior over  $1, \dots, p$  (i.e.,  $\Pr(d = i) = \frac{1}{p}$  for  $i = 1, \dots, p$ ) and Normal-Gamma priors are used (i.e., the joint prior for  $\beta$  and  $h$  is  $NG(\underline{\beta}, \underline{Q}, \underline{s}^{-2}, \underline{\nu})$ ), derive the posterior for the model defined by (17.9).

(b) Using an appropriate data set (e.g., the real GDP growth data set available on the website associated with this book), write a program and carry out Bayesian inference for this model using your results from part (a). Set  $p = 4$  and

$$z_{t-d} = \frac{\sum_{d=1}^p y_{t-d}}{d},$$

so that (if you are using quarterly real GDP growth data), the threshold trigger is average GDP growth over the last  $d$  quarters.

\*\*\*\*Solution to Exercise 3\*\*\*\*\*

(a) The solution to this exercise is a straightforward extension of the solution to Exercise 17.2 parts (a) and (c). That is, it follows immediately from noting that, if we condition on  $\tau$  and  $d$ , then this TAR model can be written as a Normal linear regression model. Note that, since  $z_{t-d}$  is either exogenous or a function of lags of the dependent variable, it can also be conditioned on in the same manner as we condition on  $X$ . As in Exercise 17.2 part (c), we can write the TAR as:

$$y = X\beta + \epsilon,$$

where  $X$  is the  $(T - p) \times 2(p + 1)$  matrix with  $t^{\text{th}}$  row given by

$[D_t, D_t y_{t-1}, \dots, D_t y_{t-p}, (1 - D_t), (1 - D_t) y_{t-1}, \dots, (1 - D_t) y_{t-p}]$ , where  $D_t$  is now a dummy variable which equals 1 if  $z_{t-d} \leq \tau$  and equals 0 if  $z_{t-d} > \tau$ .

Thus, as in the solution to Exercise 17.2 (a), the posterior for  $\beta$  and  $h$  (now conditional on  $\tau$  and  $d$ ) is  $NG(\bar{\beta}, \bar{Q}, \bar{s}^{-2}, \bar{\nu})$  where  $\bar{\beta}, \bar{Q}, \bar{s}^{-2}$  and  $\bar{\nu}$  are defined as in that solution except  $X$  and, thus, the posterior hyperparameters now depend on  $\tau$  and  $d$ .

Analogous to the solution to Exercise 17.2 (c), we can write:

$$p(\beta, h, \tau, d|y) = p(\beta, h|\tau, d, y) p(\tau, d|y).$$

The preceding paragraph describes methods of posterior inference for  $p(\beta, h|\tau, d, y)$  and we can focus on  $p(\tau, d|y)$ . But, also as in Exercise 17.2 (c), we can use the facts that:

$$p(y|\tau, d) = \int \int L(\beta, h|\tau, d) p(\beta, h|\tau, d) d\beta dh$$

and  $p(\tau, d|y) \propto p(y|\tau, d) p(\tau, d)$  to note that, by calculating the marginal likelihoods for the Normal linear regression model (conditional on  $\tau$  and  $d$ ) for every possible value of  $\tau$  and  $d$ , we can build up the posterior  $p(\tau, d|y)$ . The relevant formula for the marginal likelihood is given in (17.7).

(b) A MATLAB program which extends the program of Exercise 17.4 part (c) to allow for an unknown delay (and general threshold trigger) is provided on the website associated with this book. Since we are calculating marginal likelihoods, we use the same weakly informative prior described in the solution to Exercise 17.4 (c). As we stressed in that solution, in a serious empirical study, the researcher would likely elicit the prior in a more sophisticated manner and/or carry out a prior sensitivity analysis. Posterior properties of  $\beta$  and  $\sigma^2$ , using the US real GDP growth data, are provided in Table 17.4. Figure 17.2 plots the posterior of  $d$ . Note that, since different values of  $d$  imply different threshold triggers (with different interpretations), the interpretation of the threshold differs across  $d$ . Hence, it would make little sense to plot the posterior of  $\tau$  and, hence, we do not do so (although one could plot  $p(\tau|y, d)$  for a choice of  $d$ ). Interestingly, the posterior for  $d$  allocates most of the probability to  $d = 2$ , indicating that it is not last quarter's GDP growth which triggers a regime shift (as the model of Exercise 17.2 implies), but rather GDP growth averaged over the past two quarters. The finding that GDP growth two quarters ago plays an important role in regime shifting in US GDP has been found by others [e.g., Potter (1995)].

Note also that the solution to this question can be interpreted as a standard Bayesian posterior analysis where  $d$  is an unknown parameter whose

posterior we have derived. Alternatively (and equivalently) it can be as an example of Bayesian model averaging where we are presenting results (e.g., in Table 17.4) which are averaged over different models where the different models are defined by different threshold triggers. Koop and Potter (1999) is an empirical paper which focusses on the second interpretation.

Parameter	Mean	Standard Dev.
$\beta_{10}$	0.255	0.226
$\beta_{11}$	0.080	0.174
$\beta_{12}$	-0.330	0.196
$\beta_{13}$	-0.340	0.167
$\beta_{14}$	-0.264	0.166
$\beta_{20}$	0.557	0.159
$\beta_{21}$	0.253	0.087
$\beta_{22}$	0.270	0.080
$\beta_{23}$	-0.110	0.076
$\beta_{24}$	-0.008	0.072
$\sigma^2$	0.806	0.076

