ANNALS OF ECONOMICS AND FINANCE **11-1**, 185–208 (2010)

Fluctuations of Real Interest Rates and Business Cycles

Yongli Zhang *

CEMA, Central University of Finance and Economics
E-mail: ayzhang@gmail.com

This paper investigates whether the occurrences of business cycles have caused the fluctuations of real interest rates in the US. Based on a standard consumption-based asset pricing model, the model incorporates a new feature that investors have to learn about the unobservable alternations between expansions and recessions. The model captures the qualitative property that real interest rates increase with expected future consumption growth. The simulation technique of the Gibbs Sampling is used to estimate and calibrate the model. It is discovered that the conditional variances of consumption growth are too small to be modeled as a time-vary volatility process. This finding casts doubt on Weitzman (2007). Furthermore, the model largely duplicates the dynamics of real interest rates prior to Year 1980. However, it fails to yield the drastic increase in the real interest rates during the 1981-1982 Recession, which was mainly caused by the quick tightening of monetary policy by the Federal Reserve. It is concluded that the consumption-based asset pricing models without a monetary perspective are difficult to fully capture the *dynamics* of real interest rates in the US data.

Key Words: Consumption-Based asset pricing; Bayesian learning; Gibbs sampling; Markov chain Monte Carlo.
JEL Classification Number: G12.

1. INTRODUCTION

Real interest rates in the US always experience troughs during recessions. This is evident in Figure 1, where the annualized real interest rates[1] of the

*I thank Siddhartha Chib, David Hinkley, John Hsu, Steve LeRoy and Jessica Wachter for helpful suggestions. All errors are mine. The research is supported by the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (No: 708015).

[1]The annualized real interest rates are approximated by the differences between the annualized nominal yields on three-month Treasury bills and realized inflation rates. The implicit assumption imposed here is that expected inflations coincide with actual inflations. Although real interest rates are better represented by the real yields of inflation-

1529-7373/2010

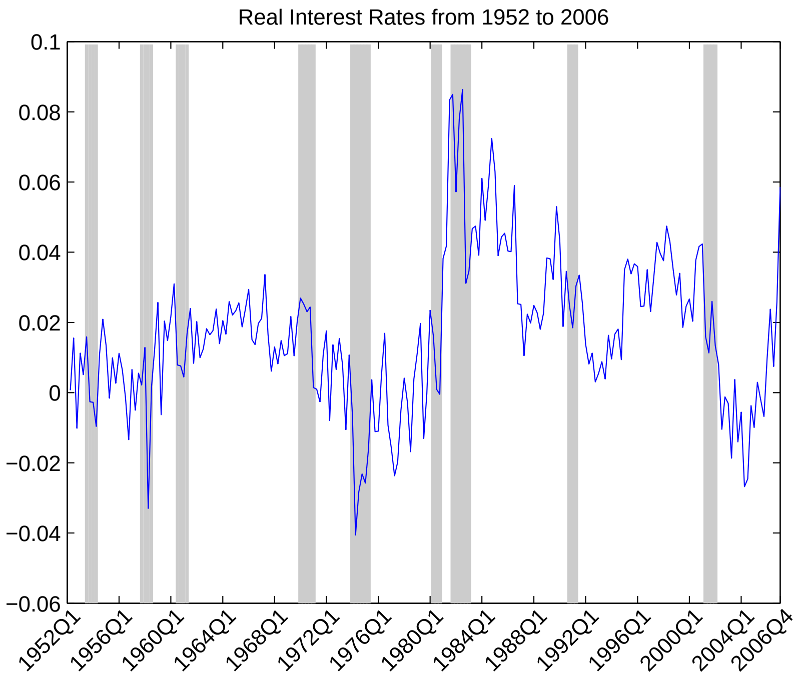

FIG. 1. The quarterly (annualized) real interest rates are approximated by subtracting the actual inflation rates from the nominal yields on three-month Treasury bills. Data on the yields of T-bills are from the Federal Reserve Statistical Release. The inflation rates are converted from the CPI that is provided by the Bureau of Labor Statistics. The shaded areas denote the periods of recessions defined by the NBER.

US economy are plotted from the first quarter of 1952 to the last quarter of 2006. Before 1981, there were two large decreases in real interest rates. One happened in 1974, when the OPEC cut the supply of oil and caused a dramatic increase in inflation. The US economy went into a recession. Six years after that, the other sharp drop occurred and the real interest rate reached the historically low of -4.56%. This was due to a record-high inflation rate triggered by the 1979 Oil Crisis. The economy went into a recession in 1980, and another one in 1981. Similar patterns can be seen during other recessions.

Motivated by the above evidence, I investigate in this paper whether the occurrences of business cycles cause the fluctuations of real interest

---

indexed bonds, it is a common practice to use this approximation, because the Treasury Inflation-Indexed Securities (also known as the TIPS), were not introduced until the 1990s.

rates. The consumption-based asset pricing model considered by Lucas (1978) is a good starting point. It considers an economy with a long-lived representative agent, who receives aggregate endowment in every period. The agent maximizes his lifetime expected utility by allocating his wealth to consumption and investment. The asset market has only a riskless bond that will return one unit of consumption for sure in the next period.

One important feature of this model is that the representative agent has no incentive to trade bonds and consumes all his endowment in equilibrium. This is because of the representative-agent setting where there is no one to take the opposite trade position. Aggregate consumption therefore equals aggregate endowment. Real interest rates are such that the no-trade condition is guaranteed. They are determined by the agent's expected marginal rates of substitution. In particular, under conditions that will be specified shortly, real interest rates become lower (higher) when expected consumption growth rates are lower (higher). If investors expect tomorrow's consumption growth to be lower, they would save more and consume less today because their marginal utilities tomorrow are higher. It takes a lower interest rate to dissuade the investors from saving today. Similarly, when future consumption growth is expected to be higher, investors would prefer consuming more and saving less today due to tomorrow's lower marginal utilities. A higher interest rate is needed to remove their incentives to dissave today.

Why do real interest rates first drop at the beginning and then rise near the end of a recession? It is important to note that recessions typically last for several quarters, which correspond to our model periods since we use quarterly data. At the beginning of a recession, investors expect consumption growth in the next quarter to be lower, thus a lower real interest rate. But when the recession is going to bottom out, investors will become optimistic and expect future consumption growth to be higher. Therefore a higher interest rate will result and shape a trough right before the end of a recession.

Despite the ultra-stark structure where money and production are abstracted away, the model works surprisingly well in a qualitative sense to capture the relation between business cycles and real interest rates. But how would it fare with the data quantitatively? This is the question that I will answer.

In order for the model to produce time-varying interest rates, I introduce Bayesian learning to the benchmark model. In particular, I assume that recessions and expansions are unobservable. Investors have to learn about the true state of the world by observing realized data on consumption growth. In addition, they also have to learn about the (conditional) means and (conditional) variances of the process that governs the dynamics of consumption growth. Since real interest rates depend on the trade-offs between current marginal utilities and expected future marginal utilities, they will start to fluctuate and reflect people's changing expectations as investors update their information.

These assumptions are realistic. Although a variety of macroeconomic data have been made easier to access and updated in a timely manner, investors are still not a hundred percent certain about whether a most recent expansion has ended or a new recession is starting. Even economists have to examine several more quarters' data to decisively determine whether a new business cycle has occurred in a previous time period. Therefore, it would be unrealistic to assume that rational investors possess perfect information about the true states of business cycles. Furthermore, no investors would have the ability of perfectly foreseeing the growth prospect and risks as represented by conditional means and conditional variances, respectively. It is therefore reasonable to assume that these parameters are not readily known.

With both structural uncertainty (unobservable states) and parameter uncertainty (unknown conditional means and variances), the model has the potential of yielding low short-term interest rates that are comparable with the data. Mehra and Prescott (1985) point out that the consumption-based price model with CRRA preferences yields real interest rates that are too high to match the data. Structural and parameter uncertainty, as an important source of risk that affects investors' decision making, fundamentally changes investors' perception and expectation about the future. Without perfect information, the investors living in this world are more uncertain about the future. Being risk averse, they would want to save more today. Real interest rates are therefore lower to prevent them from saving in the world with imperfect information. Moreover, Zhang (2007) calibrates the same type of model but with the assumption of only one unknown parameter. The main finding is that parameter uncertainty is not enough to account for the low interest rates. By introducing unobservable state variables that govern the occurrences of business cycles, together with more unknown parameters, we can reasonably expect that the model will better duplicate the data.

Structural uncertainty has already been prescribed by Weitzman (2007) as a resolution to three asset pricing puzzles. While Weitzman (2007) focuses on the steady-state implications of learning, this paper studies the interest rate dynamics that are implied by Bayesian learning.

To incorporate learning about economic expansions and recessions, we model the exogenous process of consumption growth as an autoregressive process of order one with two-state Markov-switching conditional means and variances. That is, we assume that the conditional mean and the conditional variance of the AR process at any point in time can each take one of two possible values: high or low. Although it is natural to associate recessions (expansions) with low (high) conditional means and high (low) conditional variances, whether this association is valid is ultimately an empirical issue.

Indeed, our estimation results suggest that the conditional means are typically higher and the conditional variances lower in expansions than in recessions. However, the differences in the conditional variances are much less significant. This is because the data of consumption growth are much less volatile than asset prices, whose volatilities are often (appropriately) modeled as a stochastic process. The implication is that it is unnecessary to model the conditional volatilities of consumption growth as a stochastic process, since the estimates in different states would not be distinguished anyways. This finding casts doubt on Weitzman (2007), who claims that all three asset pricing puzzles can be resolved. His results crucially depend on modeling the precisions (defined as the reciprocals of variances) of consumption growth as a stochastic process.

Since the introduction of the Markov processes significantly complicates the inference of unknown parameters and hidden states, we utilize the Gibbs Sampling, a technique of the more general family of simulation methods called the Markov Chain Monte Carlo, to simulate samples that approximate posterior moments and asset prices.

The results show that investors with Bayesian learning have been informative about the occurrences of business cycles. In other words, their subjective beliefs about which economic states they live in match the alternations of expansions and recessions in the data. However, the results of interest rates are mixed. Although the model's results largely follow the trend of the real interest rates in the data before 1980, it misses the volatile episode from 1980 to 1982. As can be observed from Figure 1, a prominent feature is that the 1980-1982 period has seen dramatic changes in the real interest rates. First, real interest rates started to rise sharply from the fourth quarter of 1980 and then reached an unprecedented level one year after that. Then the rates briefly dropped and rose again to the all-time high of 9.55% (per annum) in the second quarter of 1982. This huge increase in the real interest rates has been associated with the shift from an expansionary to a contractionary monetary policy by the Federal Reserve chaired by Paul Volcker. We therefore conclude that, despite the success of using structural uncertainty to match the average level of real interest rates in the data, incorporating Bayesian learning into the consumption-based asset pricing models will not explain fluctuations of real interest rates in the US economy. The model is missing a monetary perspective that may be fruitful in duplicating the interest-rate dynamics. This approach is left for research.

Wachter (2006) uses a consumption-based asset pricing model to explain the term structure of nominal interest rates. The driving force of her model is a time-varying price of risk generated by external habit. In comparison, without deviating from CAAR preference, my model is driven by Bayesian learning about the hidden state and unknown parameters. I do not consider the term structure and only focus on the quarterly real interest rates in this paper.

The rest of the chapter is organized as follows. Section 2 describes the consumption and inflation data and their sources. A model with Bayesian learning is studied in Section 3. Section 4 estimates the model and reports the results. The conclusion is in the last section.

2. THE DATA

I use quarterly data from the first quarter of 1952 to the fourth quarter of 2006 for all the series in this paper. The real growth rates of consumption of nondurable goods and services are used as indicators of business cycles. They were no doubt at lower levels near the ends of recessions, as can be seen from Figure 2. The average duration of these recessions is ten months, which amounts to at least three quarters. These recessions are defined by the National Bureau of Economic Research through examining a comprehensive combination of macroeconomic time series such as the real GDP and personal income excluding transfer payments[2]. Typically, the NBER defines a recession as the one that starts from the peak of a business cycle and ends at the trough. The series of consumption growth of non-durable goods and services mostly follows this pattern[3] and tracks the historical US business cycles pretty well. Furthermore, the sample correlation coefficient between the real interest rates and the recession-tracking consumption growth is .4, which is significantly positive.

Figure 2 plots the quarterly consumption growth rates, which are denoted by $\{y_t\}$ and exponentially compounded:

$$y_t = \ln \frac{C_t}{C_{t-1}} \tag{1}$$

where C_t denotes real aggregate consumption of nondurable goods and services at time t. The sample mean and the sample standard deviation are 0.0237 and 0.0193 (per annum), and the sample autocorrelations of the first four orders are 0.31, 0.20, 0.19 and 0.06, respectively. So one would not

[2]See the NBER's official website http://www.nber.org/cycles/recessions.html for the details about how a business cycle is determined.

[3]The exceptions are the recession in 1970 and the one in 2001. This is mainly due to the difference in measurement frequencies. While the recession periods defined by the NBER are precise up to months, the series of consumption of non-durables and services provided by the Bureau of Economic Analysis (BEA) is of quarterly frequency.

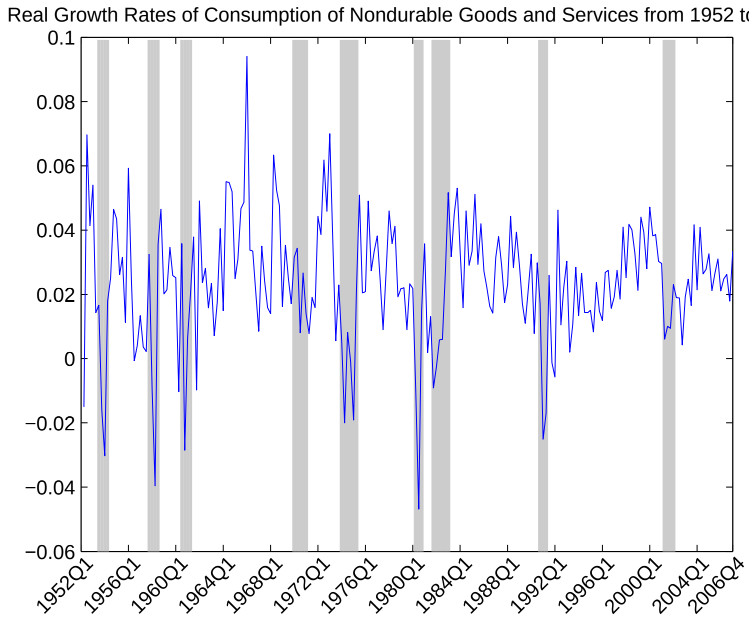

FIG. 2. Quarterly (annualized) real growth rates of consumption of non-durable goods and services from the BEA. The series are obtained by subtracting the growth rates of CPI from the nominal growth rates of consumption of non-durable goods and services. The shaded areas are recession periods defined by the NBER.

defend too strongly that the quarterly consumption growth series follows a Gaussian random walk with white noise.

The quarterly real interest rates are obtained by deflating the nominal rates by the inflation measure of CPI and are plotted in Figure 1. The sample mean and the sample standard deviation are 0.0168 and 0.0213 (per annum), respectively. As mentioned in the Introduction, anecdotal evidence seems to suggest that the fluctuations of real interest rates are closely related to business cycles. In the next section, we investigate whether the asset pricing models of the Lucas type, with the added feature that investors learn about business cycles using the Bayes rule, are promising to capture dynamics of the real interest rates seen in the data.

3. THE MODEL

Consider an exchange economy populated by a long-lived representative agent, who receives aggregate endowment in terms of the perishable consumption goods in every period. The agent maximizes his lifetime expected utility by allocating his wealth to consumption and investment. His preference is represented by the standard time-additive utility with Constant Relative Risk Aversion: $C^{1-\gamma}/1-\gamma$. The agent maximizes his life-time expected utility:

$$E\{\sum_{t=0}^{\infty}\beta^t\frac{C_t^{1-\gamma}}{1-\gamma}\} \tag{2}$$

where $\beta < 1$ measures his time preference and $\gamma > 0$ is the parameter of relative risk aversion. In the asset market, the only tradable asset is a one-period riskless zero-coupon bond, which – if bought at time t – will pay its holder one unit of the consumption good at time $t+1$ for sure.

In equilibrium, the asset market clears such that no one holds a positive net position on the risk-free asset and the agent consumes all his endowment:

$$C_t = W_t, \text{ for all } t \tag{3}$$

Then the growth rate y_t defined in (1) also represents the growth of aggregate endowment process. It follows some exogenous stochastic process which is yet to be specified.

The equilibrium first order condition, also known as the Euler equation, is:

$$E_t\{M_{t+1}R_t^f\} = 1 \tag{4}$$

where E_t is the expectation operator conditional on investors' information up to period t, R_t^f is the (gross) risk-free rate of return from time t to $t+1$, and M_{t+1} is the pricing kernel, which is the stochastic marginal rate of substitution

$$u'(C_{t+1})/u'(C_t) = \beta C_{t+1}^{-\gamma}/C_t^{-\gamma} \tag{5}$$

Using (1), (4) and the market clearing condition (3), we have that the (gross) return of the risk-free asset is:

$$R_t^f = \frac{1}{\beta E_t\{W_{t+1}^{-\gamma}/W_t^{-\gamma}\}} = \frac{1}{\beta E_t \exp(-\gamma y_{t+1})} \tag{6}$$

which can be computed if the stochastic process for the growth rate is specified.

I assume that the consumption growth process is an autoregressive process of order one:

$$y_t - \mu_{s_t} = \rho\cdot(y_{t-1}-\mu_{s_{t-1}}) + \sigma_{v_t}\varepsilon_t,\ \epsilon_t \overset{iid}{\sim} N(0,1) \tag{7}$$

To capture the alternation of business cycles, we introduce two hidden-state variables: s_t and v_t, each of which can take one of the two possible values: 0 and 1. The conditional mean μ_{s_t} and the conditional variance $\sigma^2_{v_t}$ at time t can be either high or low depending on the values of s_t and v_t:

$$\mu_{s_t} = \begin{cases} \lambda, & \text{if } s_t = 0 \\ \lambda + \xi, & \text{if } s_t = 1 \end{cases}, \ \xi > 0.$$

$$\sigma^2_{v_t} = \begin{cases} \frac{1}{\theta}, & \text{if } v_t = 0 \\ \frac{1}{\eta}, & \text{if } v_t = 1 \end{cases}, \ \theta > \eta.$$

In other words, λ and $\lambda + \xi$ are the conditional means in the low-mean and high-mean state, respectively. And $1/\theta$ and $1/\eta$ respectively represent the variance in the low-variance and high-variance state. Alternatively, the conditional mean and variance can be expressed in a more compact form:

$$\mu_{s_t} = \lambda + \xi s_t, \ \xi > 0 \tag{8}$$

$$\sigma^2_{v_t} = \frac{1}{\theta} + (\frac{1}{\eta} - \frac{1}{\theta})v_t, \ \text{provided that } \eta < \theta \tag{9}$$

We further assume that $\{s_t\}$ and $\{v_t\}$ follow two independent two-state Markov processes with transition matrices:

$$\begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \text{ and } \begin{pmatrix} 1-p_v & p_v \\ q_v & 1-q_v \end{pmatrix},$$

respectively. All parameters and states are unknown. Therefore the representative agent has to learn about the values of these parameters and the unobservable states $\{s_t\}$ and $\{v_t\}$ in which they have lived. Since we are only concerned with the states within the span of our sample period, we can treat the unknown states in the same way with the unknown parameters because the unknown states during our sample period have all been realized and are therefore fixed[4]. Note that this econometric model is essentially a mixture of four AR(1) processes with Gaussian innovations. The

[4]It is worth emphasizing that the latent state variables can be quite different conceptually depending on whether they have been realized or not. If they have been realized, then they are fixed but not observable by definition. If they have not been realized, then they are random variables objectively. The realized ones are intrinsically different from the unrealized ones, although they are all unobservable. The realized states are like unknown parameters. The researchers (investors and econometricians alike) know that the realized states never change their values. But they just do not know which value (0 or 1) they take. In a Bayesian world, researchers treat the realized states as subjective random variables, just like the way they treat unknown parameters. The unrealized states, however, are objective random variables that are governed by some distributions. Since we are dealing with the realized states because we are investigating the history but not predicting the future, it is in this sense that the states are indifferent from parameters.

motivation for introducing two Markov switching processes is discussed in the Introduction and is therefore not repeated.

Our stochastic structure is similar to the one studied in Hamilton (1989), Albert and Chib (1993), Kim and Nelson (1996). These authors document the evidence that the mean and the variance of the growth rates of macroeconomic time series such as GNP and GDP in expansions are different from those in recessions. However, none of them examine the time series of consumption of non-durable goods and services. The consumption series is much less volatile than GDP or GNP. As will be seen in the next section, our results show that the variance in the high-variance state differs little from that in the low-variance state. This implies that it is perhaps redundant to model the variance as a Markov-switching process. More details about model selection are provided in the next section.

Our model can be reduced to three special cases:

Model M: If $v_t = 0$ for all t and $p_v = q_v = 0$, then only the mean is switching between two states and the model is the same as the one studied in Hamilton (1989).

Model V: If $s_t = 0$ for all t and $p = q = 0$, then only the variance is switching between two states.

Model MVS: If $s_t = 1 - v_t$ for all t and $p = q_v$ and $q = p_v$, then the mean and the variance is governed by the same Markov process and the model collapses into the one in Albert and Chib (1993).

We have four models — the three special cases plus our model (**Model MV**) in (7)— to compare and choose from. This is the so-called "Bayesian model selection" problem in the Bayesian decision theory. The criterion of selecting a best Bayesian statistical model is similar to the maximum likelihood criterion by which competing estimators are chosen. In a Bayesian model selection framework, the true model that generates the data is not observable and therefore can be treated as an unknown state. The prior subjective belief about a model's authenticity can be characterized and shall be interpreted as a decision maker's[5] degree of confidence about whether the data are generated by that model. For example, suppose the prior beliefs about two competing models indexed by 1 and 2, $\pi(m = 1)$ and $\pi(m = 2)$, are fifty and fifty. It means that the decision maker has no information about which model better describes the data and therefore assigns equal probabilities to each model a priori. After the decision maker

[5]The decision maker refers to a researcher, who should be interpreted as a sophisticated investor. This is to maintain the presumption in the previous chapter that rational investors are as smart as econometricians and use the Bayes rule to learn about unknown parameters and unobservable states. Uncertainty about a model, as discussed above, is essentially similar to uncertainty about an unobservable state. Therefore, we are still modeling the uncertainty of the investors even when model uncertainty is incorporated.

obtains $\{Y_T\}$ data, she can update her subjective belief about the validity of each model using the Bayes rule:

$$\pi(m=i|Y_T)=\frac{\pi(m=i)f(Y_T|m=i)}{f(Y_T)},\ i=1,2. \tag{10}$$

This is the posterior belief about whether a model is the data generating model. The true model state m is updated and learned in a similar fashion to those unknown parameters. The *posterior odds*, which incorporates the information of both the marginal likelihood function (given a model) and the prior belief about a model, dictates which model is better. It is defined as the ratio:

$$\frac{\pi(m=1|Y_T)}{\pi(m=2|Y_T)}=\frac{\pi(m=1)}{\pi(m=2)}\frac{f(Y_T|m=1)}{f(Y_T|m=2)} \tag{11}$$

If this ratio is greater (less) than one, then Model 1 is more (less) favorable than Model 2. In the case where prior beliefs about two models are the same, the posterior odds reduces to the ratio of two marginal likelihood functions. This ratio is also known as the *Bayes factor* of Model 1 relative to Model 2. The model with the highest marginal likelihood, or a Bayes factor greater than one, is then the obvious choice. We apply this idea to our setting and show in the next section that Model M is the best model that fits the data.

4. BAYESIAN INFERENCE

It can be seen from the previous section that the posterior distributions are vital for the inference of unknown parameters and the calculation of asset prices that essentially depends on investors' subjective beliefs about unknown parameters and latent states. The commonly used Monte Carlo simulation methods no longer applies here, because it is impossible to directly generate simulations from the posterior distribution, which is a joint distribution of all unknown parameters and latent states. Thus the inference problem of our model is significantly more complexed due to the latent Markovian states.

Fortunately, the advent of a computational technique called *the Markov Chain Monte Carlo* (MCMC) has tremendously eased the inference problems in Bayesian models. The technique is so powerful that virtually any Bayesian statistical models can be estimated. While we do not intend to give a literature review about the MCMC method, we provide some references for interested readers. For an in-depth coverage of the topic, see Gamerman and Lopes (2006). For introductory expositions of the *Gibbs* Sampling and its more general version of the *Metropolis-Hastings* algorithm, which are two important and widely-used algorithms of the MCMC family, see Casella and George (1992) and Chib and Greenberg (1994), respectively.

We use the technique Gibbs sampling to make posterior inference about the parameters and the unobservable states. We illustrate the idea and procedure of the MCMC method through our context. Our model differs from traditional Bayesian models in that both the conditional means and variances follow some regime switching processes that are unobservable. However, these unobservable regimes or states pose no problems for posterior inference, as pointed out by Albert and Chib (1993). These authors devise a technique to generate random samples of the latent states in their Gibbs sampler and average the latent states out to obtain the p.d.f. of the posterior distribution of only unknown parameters. More details about this technique will be provided later in this section.

Formally, there are nine unknown parameters $\Theta = (\lambda, \xi, \rho, \theta, \eta, p, q, p_v, q_v)$[6] and $2T$ unobservable states $Z_T = (S_T, V_T)$, where $S_T = (s_1, \ldots, s_T)$ and $V_T = (v_1, \ldots, v_T)$ in our model. We first specify the prior distributions and then outline the simulation procedure.

The conjugate prior distributions are of our choice. It is further assume that the prior distributions of all parameters and states are independent:

$$
\begin{aligned}
&\pi(\lambda) \sim N(\lambda_0, L^{-1}) \\
&\pi(\xi) \sim N(\xi_0, X^{-1}) I_{\{\xi > 0\}} \\
&\pi(\theta) \sim Gamma(a_0^\theta, b_0^\theta) I_{\{\theta > \delta\}}, \text{ where } \delta > 0. \\
&\pi(\eta) \sim Gamma(a_0^\eta, b_0^\eta) I_{\{\theta > \delta\}} \text{ and } \eta < \theta \text{ a.s.} \\
&\pi(p) \sim Beta(u_{01}, u_{00}) \text{ and } \pi(q) \sim Beta(u_{10}, u_{11}) \\
&\pi(p_v) \sim Beta(v_{01}, v_{00}) \text{ and } \pi(q_v) \sim Beta(v_{10}, v_{11})
\end{aligned}
$$

Note that ξ is restricted to be positive since the state of $s = 1$ is interpreted as the high-mean state. In a similar sense, η is required to be less than θ because the state $v = 1$ is the high-variance state, meaning that $1/\eta$ is higher than $1/\theta$, the variance when $v = 0$. Both precisions are truncated due to the same reason as in the previous chapter. Namely, investors with CRRA preferences would be unrealistically scared of the zero-probability event that precisions of consumption growth are zero and the variance could go to infinity. They would be willing to pay infinity for the riskless bond

[6]In this chapter, we do not use asterisks to distinguish the true parameters from subjectively perceived random variables as we did in the previous chapter. This should not cause confusion because the context always makes it clear which is intended. For example, those that appear after the conditional signs are considered constant, and those that appear before conditional signs are treated as running variables of density functions, and those that show up elsewhere are considered as random variables.

to hedge against such an event. We bound the support of precisions away from zero to eliminate this extreme case a priori.

The joint p.d.f. of T data $Y_T = (y_1, \ldots, y_T)$ conditional on the all parameters Θ and unobservable states Z_T is $f(Y_T|\Theta, Z_T)$. The joint posterior distribution of interest is

$$f(\Theta, Z_T|Y_T) \propto \pi(\Theta, Z_T) f(Y_T|\Theta, Z_T) \tag{12}$$

If this joint distribution could be directly simulated from, then posterior inference and asset pricing, which are related to some moments of the posterior distribution, would be straightforward because this would be the pure Monte Carlo method that depends on random simulations from the target distribution of (12). Unfortunately this distribution is only available for direct simulation under very special cases like the one studied in the previous chapter. Nevertheless, simulation from (12) can still be carried out in an indirect way. This is where the MCMC method comes into play. It involves constructing a Markov chain that has the desired distribution (12) as its stationary distribution. The samples generated from this process can be approximately viewed as being generated from the limiting distribution of (12). Posterior moments can be then estimated by the corresponding ergodic averages of the simulated sample—just like what is done in the pure Monte Carlo method. Hence the name "Markov Chain Monte Carlo." Various algorithms of the MCMC family differ in how the Markov chain is constructed.

The Gibbs sampler, in particular, builds a Markov Chain out of a series of full conditional distributions[7] that eventually converge to the joint distribution in (12). These full conditional distributions are usually some well known distributions that can be directly simulated from. We illustrate the idea of MCMC in the context of Model MV. The procedure is as follows:

1. Given $\Theta^{(i)}$ that represents the i-th draws of all parameters $(\xi^{(i)}, \rho^{(i)}, \theta^{(i)}, \eta^{(i)}, p^{(i)}, q^{(i)}, p_v^{(i)}, q_v^{(i)})$, generate the i+1-th draws of $Z_T^{(i+1)} = (S_T^{(i+1)}, V_T^{(i+1)})$ from the conditional distribution of $f(Z_T|Y_T, \Theta^{(i)})$.

2. Generate $\lambda^{(i+1)}$ from the conditional distribution of $f(\lambda|Y_T, Z_T^{(i+1)}, \Theta^{(i)})$, given the i-th draws of $\Theta^{(i)}$ and the $(i+1)$-th draws of $Z_T^{(i+1)}$ obtained from Step 1.

3. Generate $\xi^{(i+1)}$ from the conditional distribution of $f(\xi|Y_T, Z_T^{(i+1)}, \lambda^{(i+1)}, \rho^{(i)}, \theta^{(i)}, \eta^{(i)}, p^{(i)}, q^{(i)}, p_v^{(i)}, q_v^{(i)})$. Notice that the $(i+1)$-th draw of λ from the last step replaces the i-th draw.

4. Generate $\rho^{(i+1)}$ from the conditional distribution of $f(\rho|Y_T, Z_T^{(i+1)}, \lambda^{(i+1)}, \xi^{(i+1)}, \theta^{(i)}, \eta^{(i)}, p^{(i)}, q^{(i)}, p_v^{(i)}, q_v^{(i)})$. Similar to Step 3,

[7] A full conditional distribution of one random variable is defined as the distribution of this random variable conditional on all the other variables in a model.

the $(i+1)$-th draw of ρ is plugged into the conditional density. The procedure continues this way as newly generated samples recursively replace old ones.

5. Generate $(\theta^{(i+1)}, \eta^{(i+1)})$ from the conditional distribution of $f(\theta, \eta | Y_T, Z_T^{(i+1)}, \lambda^{(i+1)}, \xi^{(i+1)}, \rho^{(i+1)}, p^{(i)}, q^{(i)}, p_v^{(i)}, q_v^{(i)})$.
6. Generate $(p^{(i+1)}, q^{(i+1)})$ from the conditional distribution of $f(p, q | Y_T, Z_T^{(i+1)}, \lambda^{(i+1)}, \xi^{(i+1)}, \rho^{(i+1)}, \theta^{(i+1)}, \eta^{(i+1)}, p_v^{(i)}, q_v^{(i)})$.
7. Generate $(p_v^{(i+1)}, q_v^{(i+1)})$ from the conditional distribution of $f(p_v, q_v | Y_T, Z_T^{(i+1)}, \lambda^{(i+1)}, \xi^{(i+1)}, \rho^{(i+1)}, \theta^{(i+1)}, \eta^{(i+1)}, p^{(i+1)}, q^{(i+1)})$.
8. Repeat Step 1.

After repeating the above procedure of M times, a simulated sample, denoted by $\{\Theta^{(i)}, Z_n^{(i)}\}_{i=m+1}^{M}$, is obtained by discarding the first m draws.[8] Under some regularity conditions, the random draws of $(Z_T^{(M)}, \Theta^{(M)})$ converges to the joint posterior distribution of $\pi(\Theta, Z_T | Y_T)$ as M approaches infinity. Asset prices, which are some functions $g(\cdot)$ of the posterior moments, can then be calculated with the simulated sample because the Law of Large Numbers holds:

$$\frac{1}{M-m} \sum_{i=m+1}^{M} g(\Theta^{(i)}, Z_n^{(i)}) \overset{a.s.}{\to} E_\pi g(\Theta, Z_T), \text{ as } M \to \infty \tag{13}$$

where E_π indicates that the expectation is taken with respect to the joint posterior distribution in (12).

We can now derive all the full conditional distributions after specifying the conjugate prior distributions. It can be shown that the full conditional distributions are all known distributions that are easy to sample from. The details of the derivation of these full conditional distributions are in the Appendix.

Given all the prior distribution and full conditional distributions derived, the Gibbs sampler can be readily initiated by picking some arbitrary start values for all the parameters and states, which are specified in the Appendix.

We run the above iterations for 10000 times and discard the first 1000 runs. The posterior moments are then estimated by the ergodic average by virtue of the Law of Large Numbers:

$$E_\pi g(\Theta, Z_T) \approx \frac{1}{M-m} \sum_{i=m+1}^{M} g(\Theta^{(i)}, Z_n^{(i)}) \tag{14}$$

TABLE 1.

Estimation results for Model MV

Model MV: Log-marginal likelihood = 832.74				
		Posterior		
Parameters	Prior	Mean	Std.	90% Credible Set
λ	$N(-.002, .2)$	.0050	.0009	[.0029, .0066]
ξ	$N(.002, .2)I_{\{\xi>0\}}$	.0015	.0011	[.0001, .0040]
ρ	$N(0, 1)I_{\{\lvert\rho\rvert<1\}}$	.3084	.0911	[.1311, .4854]
$1/\sqrt{\theta}$	$\theta \sim \Gamma(.01, .0005)I_{\{\theta\geq 10\}}$	.0061	.0004	[.0052, .0069]
$1/\sqrt{\eta}$	$\eta \sim \Gamma(.01, .001)I_{\{\eta\geq 10\}}$	.0068	.0004	[.0061, .0078]
p	$Beta(1, 4)$	.2083	.1085	[.0445, .4563]
q	$Beta(1, 4)$	.1629	.0932	[.0309, .3805]
p_v	$Beta(1, 4)$	.2541	.1037	[.0825, .4849]
q_v	$Beta(1, 4)$	.1322	.0606	[.0416, .2755]

TABLE 2.

Estimation results for Model M

Model M: Log-marginal likelihood =857.93				
		Posterior		
Parameters	Prior	Mean	Std.	90% Credible Set
λ	$N(-.002, .2)$	.0049	.0011	[.0021, .0064]
ξ	$N(.002, .2)I_{\{\xi>0\}}$	.0024	.0020	[.0002, .0072]
ρ	$N(0, 1)I_{\{\lvert\rho\rvert<1\}}$	.2800	.0792	[.1180, .4324]
$1/\sqrt{\theta}$	$\theta \sim \Gamma(.01, .0005)I_{\{\theta\geq 10\}}$	.0052	.0002	[.0048, .0058]
$1/\sqrt{\eta}$	—	—	—	—
p	$Beta(1, 4)$	.1232	.0833	[.0119, .3204]
q	$Beta(1, 4)$	.1495	.0973	[.0188, .3868]
p_v	—	—	—	—
q_v	—	—	—	—

4.1. Estimation Results and Model Selection

We proceed using the numerical procedure outlined in the previous section and estimate each one of the four models postulated. The calculation of the marginal likelihood function follows the method suggested by Chib (1995). All prior and posterior estimates are summarized in the tables labeled from Table 1 to Table 4.

As can be seen from these results, the estimates of the AR(1) coefficient differ little across various models. They are all very close to the the first

[8]This is because the chain has not yet converged to the target distribution during the early iterations. The sample drawn during this period should be therefore thrown away.

TABLE 3.

Estimation results for Model V

Model V: Log-marginal likelihood =838.46				
		Posterior		
Parameters	Prior	Mean	Std.	90% Credible Set
λ	$N(-.002, .2)$	.0060	.0006	[.0047, .0072]
ξ	—	—	—	—
ρ	$N(0,1)I_{\{\lvert\rho\rvert<1\}}$	.3062	.0864	[.1377, .4750]
$1/\sqrt{\theta}$	$\theta \sim \Gamma(.01, .0005)I_{\{\theta\geq 10\}}$	.0058	.0004	[.0049, .0065]
$1/\sqrt{\eta}$	$\eta \sim \Gamma(.01, .001)I_{\{\eta\geq 10\}}$	.0065	.0004	[.0058, .0075]
p	—	—	—	—
q	—	—	—	—
p_v	$Beta(1,4)$	.1559	.0798	[.0321, .3366]
q_v	$Beta(1,4)$	.0899	.0527	[.0189, .2223]

TABLE 4.

Estimation results for Model MVS

Model MVS: Log-marginal likelihood = 828.53				
		Posterior		
Parameters	Prior	Mean	Std.	90% Credible Set
λ	$N(-.002, .2)$	.0053	.0008	[.0037, .0068]
ξ	$N(.002, .2)I_{\{\xi>0\}}$	.0015	.0010	[.0001, .0039]
ρ	$N(0,1)I_{\{\lvert\rho\rvert<1\}}$	.3066	.0961	[.1178, .4920]
$1/\sqrt{\theta}$	$\theta \sim \Gamma(.01, .0005)I_{\{\theta\geq 10\}}$	.0062	.0005	[.0053, .0080]
$1/\sqrt{\eta}$	$\eta \sim \Gamma(.01, .001)I_{\{\eta\geq 10\}}$	.0070	.0005	[.0063, .0081]
p	$Beta(1,4)$	.1390	.0625	[.0435, .2868]
q	$Beta(1,4)$	.2502	.1014	[.0839, .4861]
p_v	$Beta(1,4)$	.2502	.1014	[.0839, .4861]
q_v	$Beta(1,4)$	.1390	.0625	[.0435, .2868]

order sample autocorrelation of .31. The jump size of the mean ξ from the low-mean state to the high-mean state is larger for Model M than for Model MV or Model MVS. This is because Model M gives a better fit to the data and the estimation tends to be more efficient.

Table 1 shows that the posterior estimates of the standard deviations in the low-variance and high-variance state, $1/\sqrt{\theta}$ and $1/\sqrt{\eta}$, are respectively .0061 and .0068, which are very close to each other. For Model V in which only the variance but not the mean switches between two states, Table 3 reports a similar result that the standard deviations in two different states are .0058 and .0065, whose difference is only .0007. This is mainly because the consumption growth series is well known to be smooth and less volatile than GNP or GDP. Considering the variance of consumption growth as a two-state Markov switching process perhaps makes little difference from modeling the variance as a constant. Weitzman (2007) assumes that the variance of consumption growth follows a stochastic process defined on a continuous state space. In light of our result that even a two-state Markov switching process of variance does not yield distinguishing estimates, one is wondering whether modeling the low variance of consumption growth as a stochastic process is necessary at all.

Our estimation results show that Model M in which only the mean follows a two-state Markov switching process has the highest marginal likelihood and therefore is the best model by the Bayesian model selection criterion discussed above. In the next subsection, we use Model M to calculate the return of the riskless bond.

4.2. Asset Pricing Implications

Figure 3 plots investors' subjective unconditional probabilities of high-mean states of expansions. These are the probabilities such that each state equals one given the data Y_t ($\Pr(s_t = 1|Y_t)$) regardless of the previous state. For example, the representative agent's perceived unconditional probability of an expansion in the first quarter of 1980 when she only has data up to that time is .25. As can be seen from the graph, the agent's estimates of the probabilities of high-mean states are low at the beginnings of each recession and then becoming rising near the ends of each recession. This implies that the representative agent in our model does a pretty good job in learning about which state of the economy she lives in. Her successful learning about the alternations of business cycles has a profound impact on her consumption and saving behavior that determines bond returns in our model. At the beginning of an economic downturn, her view about the future economy, in terms of her estimate of next period's consumption growth, is not optimistic. Consequently, she would have an incentive to increase her demand of the safe bond and drives down the returns. When the recession is ending, the agent becomes optimistic and tend to reduce

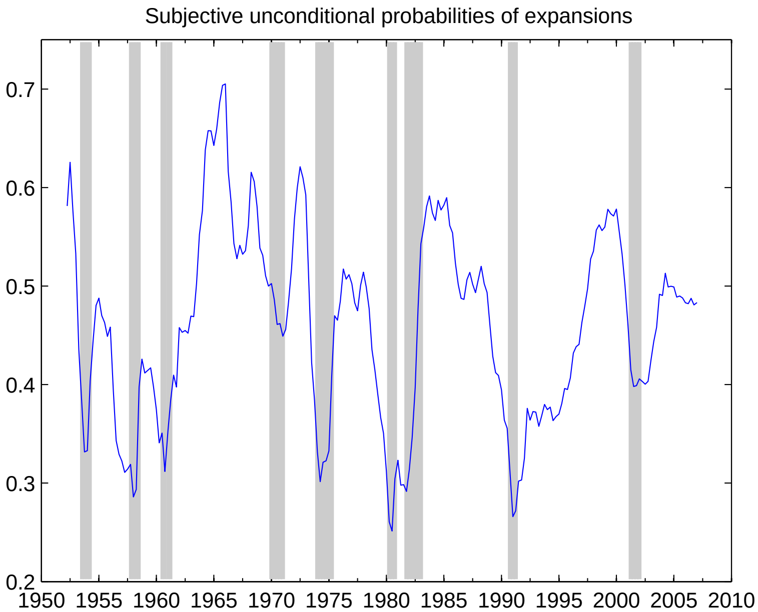

FIG. 3. The subjective unconditional probabilities of high-mean states of expansion ($\Pr(s_t = 1)$) across time, together with the shaded areas that denote the historical recessions.

her bond holdings. The real interest rates are higher to prevent the agent from trading. As will be seen from the upcoming results, our model's predictions of bond returns fall at the beginnings and then rise at the ends of each recession.

Before we present the results, we first derive the expression of bond returns. From the asset pricing section in the previous chapter, we know that the return of the riskless bond is

$$R_t^f = E_t^S[\exp(-\gamma y_{t+1})] \tag{15}$$

where the operator E_t^S denotes the subjective expectation of next period consumption growth conditioned on this period's information. By the law of iterative expectations, the above equation can be expressed as

$$R_t^f = E_t^{\Theta,Z} E[\exp(-\gamma y_{t+1})|\Theta, Z_t] \tag{16}$$

Identifying the inner expectation as the moment generating function of a normal distribution with mean $\overline{\mu}_{t+1} = \mu_{s_{t+1}} + \rho(y_t - \mu_{s_t})$ and variance $1/\theta$, we have that

$$E[\exp(-\gamma y_{t+1})|\Theta, Z_t] = E[\exp(-\gamma\overline{\mu}_{t+1} + \frac{\gamma^2\theta}{2})|s_t] \tag{17}$$

Depending on the state at time t, the expectation in the above equation equals

$$C(\Theta, s_t)E[\exp(-\gamma\xi s_{t+1})|s_t = 0]^{1-s_t}E[\exp(-\gamma\xi s_{t+1})|s_t = 1]^{s_t} \tag{18}$$

where

$$C(\Theta, s_t) = \exp\{-\gamma[(1-\rho)\lambda + \rho(y_t - \xi s_t) + \frac{\gamma^2\theta}{2}]\} \tag{19}$$

and

$$E[\exp(-\gamma\xi s_{t+1})|s_t = 0] = 1 - p + p\exp\{-\gamma\xi\} \tag{20}$$

and

$$E[\exp(-\gamma\xi s_{t+1})|s_t = 1] = q + (1-q)\exp\{-\gamma\xi\} \tag{21}$$

Therefor the riskfree rate of return can be expressed as

$$R_t^f = E_t^{\Theta,Z}g(\Theta, Z) \tag{22}$$

where the function $g(\Theta, Z)$ is just the term in (18). The random samples $(\Theta^{(i)}, Z^{(i)})_{i=m+1}^{M}$ generated from the above numerical procedure can be used to calculate the riskfree rate of return:

$$R_t^f \approx \frac{1}{M-m}\sum_{i=m+1}^{M} g(\Theta^{(i)}, Z^{(i)}) \tag{23}$$

The riskfree rate of returns are calculated this way at each time point t in our sample. Correspondingly, the MCMC procedure is performed by assuming that the sample is only available up to time t. This is to mimic the situation in which when making decisions at time t, investors do not have information beyond that time to rationally calculate the return of the riskless bond. Therefore the MCMC simulations need to be run in T sets in order to produce a times series of bond returns, which are plotted as a dotted line and compared with the historical real interest rates (solid line) in Figure 4.

The model's results largely mismatch the dynamics of real interest rates. Although it seems to successfully predict the drops of interest rates in the first three recessions prior to 1965 and in the recession of 1980, the model fails to catch a few spikes in the 1960s, underestimates the trough in the

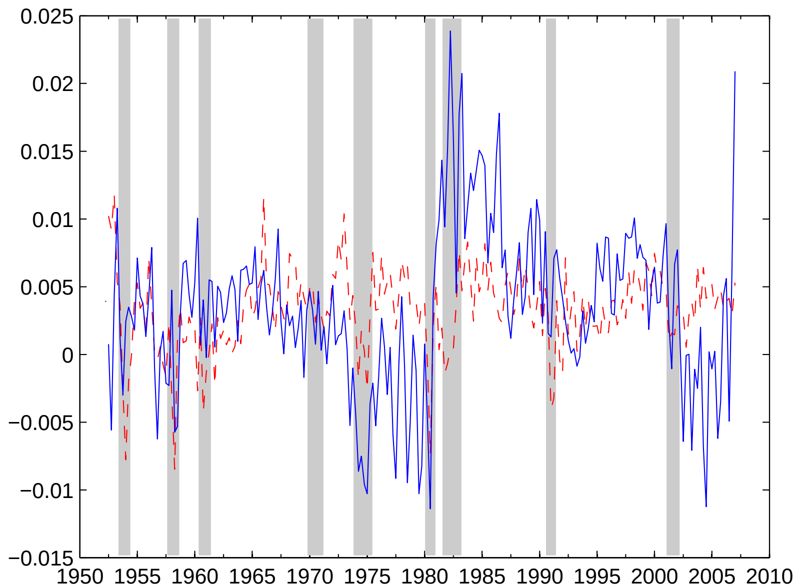

FIG. 4. Model M's predictions of returns of the riskless bond across time. $\gamma = 2$ and $\beta = .99$. The broken line represents the model's predictions and the solid line is the series of historical real interest rates.

recession of 1974 and overestimates the rate drop in the 1990 recession. Most importantly, the spectacular failure of the model is that it fails to explain the hype of the real interest rate and provides a completely opposite prediction of the interest rate during the 1981-1982 recession. The huge jump of the interest rate in 1981 is known to be associated with the shift of monetary policy from an accommodating to a restrictive one. Therefore, if one is to use the Lucas-type asset pricing models to explain the interest rate fluctuations seen in the data, especially the dramatic increase in the early 1980s, one cannot hope to do a good job without incorporating a monetary perspective.

5. CONCLUSIONS

I investigate whether business cycles can explain the dynamics of real interest rates in the US economy. In particular, I use a Lucas-type consumption-based asset pricing model with an exchange economy and a representative agent, with an added feature that the representative investor does not have perfect information and has to learn about business cycles. To incorporate learning about economic expansions and recessions, I model the exogenous consumption growth process as an autoregressive process of order one with two two-state Markov-switching conditional means and variances. This model is a generalized version of many models that are previously studied in the literature. A numerical technique of the MCMC is used to estimate parameters and compare competing econometric models. Our estimation results indicate that the model with only Markov switching conditional means best describe the data of consumption growth. We further argue that modeling the low-volatility of consumption growth as a continuous stochastic process is not necessary because our model's estimates of the high variance and the low variance are virtually indifferent. Finally, we show that our model's predictions largely duplicates the fluctuations in the real interest rates before 1980, but miss the drastic increase in the early 1980s. We conclude that, despite the recent success of using structural uncertainty to explain the *average* level of real interest rates in the data, the consumption-based asset pricing models without a monetary perspective are difficult to fully capture the *dynamics* of real interest rates in the US data.

APPENDIX A

We now derive the full conditional distributions. We randomly draw T mean states $S_n^{(0)} = \{s_1^{(0)}, s_2^{(0)}, \dots, s_n^{(0)}\}$ and T variance states $V_n^{(0)} = \{v_1^{(0)}, v_2^{(0)}, \dots, v_n^{(0)}\}$ according to a Markov chain in which the transition probabilities are 1/2. The initial values of the parameters to start the iterations are given as follows:

$\lambda^{(0)}$	$\xi^{(0)}$	$\rho^{(0)}$	$\theta^{(0)}$	$\eta^{(0)}$	$p^{(0)}$	$q^{(0)}$	$p_v^{(0)}$	$q_v^{(0)}$
0	0	0	200	100	1/2	1/2	1/2	1/2

The generation of all latent states follows the procedure suggested by Chib (1996).

The conditional distribution for λ given all other parameters $\Theta_{-\lambda}$ and latent states Z_T is a normal distribution:

$$\pi(\lambda | Y_T, \Theta_{-\lambda}, Z_T) \sim N(\overline{\lambda}_T, \overline{\sigma}^2_{\lambda T})$$

where

$$\overline{\lambda}_T = \frac{(1-\rho)\sum_{t=2}^{T} \frac{(y_t - \xi s_t) - \rho(y_{t-1} - \xi s_{t-1})}{\sigma^2(v_t)} + \lambda_0 L}{(1-\rho)^2 \sum_{t=2}^{T} 1/\sigma^2(v_t) + L}$$

and

$$\overline{\sigma}^2_{\lambda T} = \frac{1}{(1-\rho)^2 \sum_{t=2}^{T} 1/\sigma^2(v_t) + L}$$

Similarly for ξ, it is a normal distribution:

$$\pi(\xi|Y_T, \Theta_{-\xi}, Z_T) \sim N(\overline{\xi}_T, \overline{\sigma}^2_T) I_{\{\xi>0\}}$$

where

$$\overline{\xi}_T = \frac{\sum_{t=2}^{T} \frac{(s_t - \rho s_{t-1})[y_t - \rho y_{t-1} - (1-\rho)\lambda]}{\sigma^2(v_t)} + \xi_0 X}{\sum_{t=2}^{T} (s_t - \rho s_{t-1})^2/\sigma^2(v_t) + X}$$

and

$$\overline{\sigma}^2_{\xi T} = \frac{1}{\sum_{t=2}^{T} (s_t - \rho s_{t-1})^2/\sigma^2(v_t) + X}$$

.

For the precision of θ, the conditional distribution is truncated gamma:

$$\pi(\theta|Y_T, \Theta_{-\theta}, Z_T) \sim Gamma(a_t^\theta, b_t^\theta) I_{\{\theta>\delta>0\}}$$

where $a_t^\theta = a_0^\theta + \frac{n_0}{2}$ and

$$b_t^\theta = b_0^\theta + \frac{\Sigma_{J_0}[y_j - \mu(s_j) - \rho(y_{j-1} - \mu(s_{j-1}))]^2}{2}$$

where J_0 is the set of v's that equal zero $\{t|v_t = 0, t = 1, \ldots, n\}$ and n_0 the number of elements in J_0 $(n_0 = n - n_1)$. In other words, n_0 is the number of states in which the variance is low $(v = 0)$ in the sample.

Similarly, for the precision of η, the posterior conditional distribution is also truncated gamma:

$$\pi(\eta|Y_T, \Theta_{-\eta}, Z_T) \sim Gamma(a_t^\eta, b_t^\eta) I_{\{\eta>\delta>0\}}, \text{ provided that } \eta < \theta \ a.s.$$

where $a_t^\eta = a_0^\eta + \frac{n_1}{2}$,

$$b_t^\eta = b_0^\eta + \frac{\Sigma_{J_1}[y_j - \mu(s_j) - \rho(y_{j-1} - \mu(s_{j-1}))]^2}{2}$$

and where J_1 is the set of v's that equal one $\{t|v_t = 1, t = 1, \ldots, n\}$ and n_1 is the number of elements in J_1 $(n_1 = \Sigma_{t=1}^n v_t)$. That is, n_1 is the number of states in which the variance is high $(v = 1)$ in the sample.

For the transition probabilities in the Markov processes that govern the latent states, the conditional distributions are beta:

$$\pi(p|Y_T, \Theta_{-p}, Z_T) \sim Beta(n_{01} + u_{01}, n_{00} + u_{00})$$

where $n_{01} = \Sigma_{t=1}^{T}(1-s_{t-1})s_t$ and $n_{00} = \Sigma_{t=1}^{T}(1-s_{t-1})(1-s_t)$. That is, n_{10} is the number of transitions that occur from $s_{t-1} = 0$ to $s_t = 1$, and n_{00} is the number of transitions that occur from $s_{t-1} = 0$ to $s_t = 0$. Similarly,

$$\pi(q|Y_T, \Theta_{-q}, Z_T) \sim Beta(n_{10} + u_{10}, n_{11} + u_{11})$$

where $n_{10} = \Sigma_{t=1}^{T} s_{t-1}(1 - s_t)$ and $n_{11} = \Sigma_{t=1}^{n} s_{t-1}s_t$.

Finally, the transition probabilities for the variance given all data, states and parameters are:

$$\pi(p_v|Y_T, \Theta_{-p_v}, Z_T) \sim Beta(n_{v01} + v_{01}, n_{v00} + v_{00})$$

where $n_{v01} = \Sigma_{t=1}^{T}(1 - v_{t-1})v_t$ and $n_{v00} = \Sigma_{t=1}^{n}(1 - v_{t-1})(1 - v_t)$. And

$$\pi(q_v|Y_T, \Theta_{-q_v}, Z_T) \sim Beta(n_{v10} + v_{10}, n_{v11} + v_{11})$$

where $n_{v10} = \Sigma_{t=1}^{T} v_{t-1}(1 - v_t)$ and $n_{v11} = \Sigma_{t=1}^{n} v_{t-1}v_t$.

REFERENCES

Albert, James and Siddhartha Chib, 1993. Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. *Journal of Business & Economic Statistics* **11**, 1-15.

Casella, George and Edward George, 1992. Explaining the Gibbs sampler. *The American Statistician* **46**, 167-174.

Chib, Siddhartha, 1995. Marginal likelihood from Gibbs output. *Journal of American Statistical Association* **90**, 1313-1321.

Chib, Siddhartha and Edward Greenberg, 1995. Understanding the Metropolis-Hastings Algorithm. *The American Statistician* **49**, 327-335.

Chib, Siddhartha, 1996. Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics* **75**, 79-97.

Donaldson, John and Rajnish Mehra, 2008. Risk-Based Explanations of the Equity Premium. Chapter 2 in Rajnish Mehra ed. *Handbook of the Equity Risk Premium*, North-Holland.

Gamerman, Dani and Hedibert F. Lopes, 2006. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference (2nd ed.) *Chapman & Hall*.

Hamilton, James, 1989. A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica* **57**, 357-384.

Lucas, Robert, 1978. Asset prices in an exchange economy. *Econometrica* **46**, 1429-1445.

Mehra, Rajnish and Prescott, Edward C, 1985. The Equity Premium: A Puzzle. *Journal of Monetary Economics* **15**, 145-161.

Kim, Chang-Jin and Charles R. Nelson, 1999. Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle. *The Review of Economics and Statistics* **81**, 608-616.

Wachter, Jessica, 2006. A Consumption-Based Model of the Term Structure of Interest Rates. *Journal of Financial Economics* **79**, 365-399.

Weitzman, Martin, 2007. Subjective Expectations and Asset-Pricing Puzzles. *American Economic Review* **97**, 1102-1130.

Zhang, Yongli, 2008. Three Essays on Asset Pricing—A Bayesian Approach. Chapter 2, *VDM*.