REGIME SHIFTS IN SHORT TERM RISKLESS INTEREST RATES

Clifford A. Ball[1]

Owen Graduate School of Management
Vanderbilt University
Nashville, TN 37203

Walter N. Torous

Anderson School of Management
University of California at Los Angeles
Los Angeles, CA 90095

August 25, 1995

[1]Ball acknowledges the support of a grant from the Dean's fund and the Financial Markets Research Center at the Owen Graduate School of Management, Vanderbilt University. We thank Peter Boessarts and Ron Masulis for many helpful discussions. An earlier version of this paper was presented at the 1994 Wisconsin Symposium on Estimating and Managing Multiperiod Financial Risks. We thank the discussant, Angelo Melino, for his comments. The authors are solely responsible for any remaining errors.

Abstract

Chan, Karolyi, Longstaff, and Sanders [1992] find no evidence that the October 1979 change in Federal Reserve operating policy resulted in a once-and-for-all deterministic break in the behavior of short term riskless interest rates. In contrast, we provide evidence of such a regime shift even after allowing the volatility of interest rate changes to depend on the level of interest rates. However, rather than modeling this regime shift as a permanent event with no further shifts possible, it is more realistic to model the change in regimes itself as a random variable. Accordingly, we put forward a stochastic volatility interest rate model which generalizes previous specifications of interest rate dynamics and allows testing for stochastic regime shifts. This Markov regime shifting model provides a more accurate description of the behavior of U.S. short term riskless interest rates. We also consider a specification that allows interest rate volatility to follow a diffusion process and we provide a statistically efficient integration-based filtering procedure to estimate its parameters. Given U.S. short term riskless interest rate data, we cannot statistically distinguish between these alternative models. In either case, once the stochastic nature of interest rate volatility is taken into account, we find little or no evidence of a deterministic structural break in corresponding stochastic volatility interest rate dynamics around October 1979.

REGIME SHIFTS IN SHORT TERM RISKLESS INTEREST RATES

1 Introduction

Modeling the stochastic behavior of short term riskless interest rates is of considerable importance in financial economics. As such, numerous specifications of interest rate dynamics have been put forward. Common to all of these alternatives, however, is the assumption that the conditional distribution of interest rate changes is time varying. For example, in the Cox, Ingersoll, and Ross [1985] mean-reverting, square-root specification of interest rate dynamics,

$$dr = (a + br)dt + \sigma\sqrt{r}dz, \tag{1}$$

both the conditional mean and the conditional variance of interest rate changes depend upon the level of the interest rate, r.

It is also typically assumed that the structural form of this conditional distribution remains unchanged. That is, the parameters which characterize interest rate dynamics are assumed to be constant. However, structural breaks in the conditional distribution of interest rate changes may occur in response to, say, changes in monetary policy so that the parameters will not remain constant but rather will shift over time. The Cox, Ingersoll, and Ross specification, for example, implies that high volatility in interest rates is associated with high interest rate levels. While this may have been an accurate description of interest rate behavior in the late 1970s, casual empiricism suggests that it does not adequately characterize their behavior in the late 1980s when interest rates were relatively low but still quite volatile.[1]

[1]For example, the one month Treasury bill yield was approximately 10% in December 1979 as compared to approximately 5% in June 1987. However, the volatility of monthly interest rate changes, measured by the square root of the annualized 7-point smoothed variance estimate centered at these dates, was approximately

This paper investigates the empirical evidence of regime shifts or structural breaks in the post-1964 times series behavior of U.S. short term riskless interest rates. The possibility of regime shifts is important for at least two reasons. First, if there is evidence of regime shifts, estimation procedures which ignore this may systematically misestimate the parameters of the interest rate process. Second, if regime shifts have occurred in the past, they may occur again in the future. If so, hedging and valuation techniques which ignore the possibility of regime shifts may prove to be inaccurate.

Earlier studies by, among others, Huizinga and Mishkin [1984], Campbell [1987], and Sanders and Unal [1988], conclude that the October 1979 change in Federal Reserve operating policy did result in a once-and-for-all deterministic break in interest rate behavior. In contrast, more recently, Chan, Karolyi, Longstaff, and Sanders (CKLS) [1992] find no evidence of such a structural break once the instantaneous volatility of interest rate changes is explicitly allowed to depend on the level of interest rates.

The keys to reconciling these conflicting conclusions lie in the assumed specification of interest rate dynamics, and in the applicability of the statistical techniques used. Given CKLS's specification of interest rate dynamics, we provide clear evidence of a deterministic break in interest rate dynamics around October 1979. *A fortiori*, these tests explicitly acknowledge that the precise timing of this shift is potentially unknown.

The assumption of a deterministic regime shift implies that this structural break in interest rate behavior is a permanent event with no further shifts possible. As an alternative, it may be more realistic to model the change in regimes itself as a random variable. We do so by following Hamilton [1990] and putting forward a stochastic volatility interest rate

the same, at 15%.

model which generalizes CKLS's specification of interest rate dynamics and allows testing for stochastic regime shifts. This Markov regime switching model provides an accurate description of the time series behavior of post-1964 U.S. short term riskless interest rates and, in fact, cannot be rejected in favor of a competing specification in which interest rate volatility follows a diffusion process. However, little or no evidence exists of a once-and-for-all deterministic break in these stochastic volatility interest rate dynamics around October 1979.

The plan of this paper is as follows. Section 2 introduces our specifications of interest rate dynamics and discusses their statistical estimation. Section 3 provides reliable evidence of a deterministic regime shift in post-1964 U.S. short term riskless interest rates around October 1979. Recognizing that it is unrealistic to assume that this change in Federal Reserve operating policy represented a permanent event with no further regime shifts in interest rates possible, Section 4 tests for stochastic regime shifts given a Markov switching model for interest rate dynamics. We also consider a specification that allows interest rate volatility to follow a diffusion process. The stability of these stochastic volatility interest rate specifications around October 1979 is also examined. Section 5 provides a summary and conclusions.

2 Models of Interest Rate Dynamics and their Estimation

CKLS introduce the following continuous-time specification for the dynamics of the short term riskless rate r:

$$dr = (a + br)dt + \sigma r^{\gamma} dz, \tag{2}$$

where the parameters a and b characterize the linear drift component, σ measures the 'base' instantaneous volatility, while $\gamma > 0$ allows the instantaneous volatility of changes in r to functionally depend, in a power fashion, on its level. While, as noted by CKLS, this specification includes as special cases many models of interest rate dynamics previously put forward in the literature, it does not, however, allow interest rate volatility to evolve stochastically.

To estimate the parameters of their model, $\vartheta = \{a, b, \sigma^2, \gamma\}$, CKLS use the Generalized Method of Moments (GMM) assuming discrete-time observations from the posited continuous-time model:

$$r_{t+1} - r_t = a + br_t + \epsilon_{t+1}, \tag{3}$$

$$E[\epsilon_{t+1}] = 0, \qquad E[\epsilon_{t+1}^2] = \sigma^2 r_t^{2\gamma}. \tag{4}$$

Defining the vector:

$$f_t(\theta) = [\epsilon_{t+1}, \quad \epsilon_{t+1} r_t, \quad \epsilon_{t+1}^2 \sigma^2 r_t^{2\gamma}, \quad (\epsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma}) r_t]^T, \tag{5}$$

GMM requires finding those parameter values which minimize the distance between the population moment conditions, $E[f_t(\theta)]$, and their sample counterparts, $g(\theta)$.

In practice, with this choice of moment conditions, GMM reduces approximately to generalized least squares (GLS). On an annualized basis, monthly observations correspond to a small time increment and renders any consequent temporal aggregation problem negligible. This follows from the fact that increments to diffusion processes are locally Brownian and, hence, for small time increments, are approximately multivariate normally distributed.[2]

[2]See Kearns [1993] for a discussion of the various problems with using GMM to estimate parameters of interest rate processes.

However, particular care must still be taken in this estimation. For example, for $b \sim 0$, we are near a unit root in the interest rate process and therefore can expect systematically biased estimates of b as well as unreliable estimates of σ.[3] For GMM estimators to be well-behaved, regularity conditions must also be imposed on the interest rate process (Hansen [1982]). In the CKLS case, the following parameter restrictions are necessary as well as sufficient for the stationarity and ergodicity required by GMM: $b > 0$ and $0 \leq \gamma \leq 1$ (see Broze, Scaillet, and Zakoian [1993], especially Proposition 3.3, as well as Ait-Sahalia [1995]).[4]

Unfortunately, CKLS often obtain estimates of $\gamma > 1$, calling into question the reliability of their statistical conclusions. The resultant explosive behavior of short term riskless interest rates given CKLS's estimate of $\hat{\gamma} = 1.4999$ can clearly be seen in Figure 1 where we plot a representative sample path of $\{r_t\}$ assuming $r_0 = 0.07$ with $a = 0.004$, $b = 0.06$[5], and $\sigma = 1.721$. Holding all else equal, larger values of σ consistently result in floating point overflows when using a 32-bit version of GAUSS-386. Intuitively, if r_t is sufficiently large then the resultant volatility of interest rate changes increases correspondingly for $\gamma > 1$ so that an even larger r_{t+1} becomes likely.

As noted earlier, the CKLS specification does not permit stochastic volatility. However, it is straightforward to generalize their framework to accommodate this feature. To begin with, we follow Hamilton [1990] and consider the following Markov switching regime specification for interest rate dynamics:

$$dr = (a + br)dt + \sigma(t)r_t^{\gamma}dz, \tag{6}$$

[3]For further details, see Ball and Torous [1995].

[4]If these conditions do not hold then sample moments (means, variances, etc.) will converge to different values depending on the particular sample. Hence, estimates of the parameters (which are transformations of the moments) will be sample-dependent even asymptotically. In other words, if we allow $\gamma > 1$, then estimates of γ would change from one sample to another, regardless of the size of the sample.

[5]These values of the a and b parameters correspond approximately to the least squares estimates obtained using the CKLS interest rate data series.

$$\sigma(t) = \theta + \beta I(t), \tag{7}$$

where $I(t)$ is a zero-one indicator variable following a simple two state Markov chain corresponding to low and high levels of volatility, respectively. The two stage Markov chain is characterized by $p = \text{prob}(I_t = 1 \mid I_{t-1} = 1)$, the probability of being in the high volatility state given we were previously in the high volatility state, and $q = \text{prob}(I_t = 1 \mid I_{t-1} = 0)$, the probability of being in the high volatility state given we were previously in the low volatility state. According to (6) and (7), the regime shift in interest rates, from a state of low volatility to one of high volatility, is itself a random variable. This specification may be estimated using maximum likelihood techniques. As the level of stochastic volatility may take on only two values, the resulting likelihood function may be efficiently evaluated using a nonlinear filter. We initiate the recursion at $\text{prob}(I_0 = 0 \mid r_0) = \text{prob}(I_0 = 1 \mid r_0) = 1/2$, say, and run the filter for a warm-up period. The one-step-ahead forecast is given by

$$\text{prob}(I_t = 1 \mid r_{t-1}^{-}) = p\ \text{prob}(I_{t-1} = 1 \mid r_{t-1}^{-}) + q\ \text{prob}(I_{t-1} = 0 \mid r_{t-1}^{-})$$

where r_{t-1}^{-} corresponds to all interest rate observations through $(t-1)$, while the contribution to the likelihood function is

$$\text{prob}(r_t \mid r_{t-1}^{-}) = \text{prob}(I_t = 1, r_t \mid r_{t-1}^{-}) + \text{prob}(I_t = 0, r_t \mid r_{t-1}^{-}).$$

We update our assessment of the state variable according to

$$\begin{aligned} \text{prob}(I_t = 1 \mid r_t, r_{t-1}^{-}) &= \text{prob}(I_t = 1, r_t \mid r_{t-1}^{-})/\text{prob}(r_t \mid r_{t-1}^{-}) \\ &= \text{prob}(r_t \mid I_t = 1, r_{t-1}^{-})\ \text{prob}(I_t = 1 \mid r_{t-1}^{-})/\text{prob}(r_t \mid r_{t-1}^{-}). \end{aligned}$$

Proceeding in this fashion, the log-likelihood function can be constructed recursively and then maximized using numerical techniques.

3 The CKLS Evidence Revisited

In October 1979 the Federal Reserve announced a change in its operating policy from targeting interest rates to targeting money supply growth. In this section, we investigate whether this change gave rise to a once-and-for-all structural break in the stochastic behavior of U.S. short term riskless interest rates. Following CKLS, we restrict, for the time being, our attention to (2) where the volatility of interest rate changes is not allowed to evolve stochastically.

As before, the parameter vector is denoted by $\vartheta = \{a, b, \sigma^2, \gamma\}$. Assume we have a sample size of T observations, t = 1,...,T.

The null hypothesis is no structural change in the parameters:

$$H_0 : \vartheta_t = \vartheta_0 \ \text{ for } \ t = 1, \ldots, T. \tag{8}$$

Alternatively, assume that a deterministic structural change has occurred and that the change point is known with certainty. If this known change point is designated by π, $\pi \epsilon (0, 1)$, then the assumed time of change is $T\pi$ (or more precisely $[T\pi]$, where [.] is the integer part operator). In this case, we can write the alternative hypothesis as

$$H_1 : \vartheta_t = \begin{cases} \vartheta_1(\pi) & \text{for } \ t = 1, \ldots, T\pi \\ \vartheta_2(\pi) & \text{for } \ t = T\pi + 1, \ldots, T. \end{cases} \tag{9}$$

For the case where π is known with certainty, we can use likelihood ratio or Wald tests to test H_0 versus H_1. For example, the likelihood ratio test statistic is

$$LR_T(\pi) = 2\ [L(\hat{\vartheta}_1) + L(\hat{\vartheta}_2) - L(\hat{\vartheta}_0)]$$

where $L(\hat{\vartheta}_0)$ is the value of the log-likelihood function under H_0, while $L(\hat{\vartheta}_1) + L(\hat{\vartheta}_2)$ is the value of the log-likelihood function under H_1. Alternatively, the Wald test statistic is

$$W_T(\pi) = T(\hat{\vartheta}_1 - \hat{\vartheta}_2)' \, [\pi^{-1}\hat{V}_1 + (1-\pi)^{-1}\hat{V}_2]^{-1}(\hat{\vartheta}_1 - \hat{\vartheta}_2)$$

where $\hat{V}_1$ is the asymptotic variance of $\hat{\vartheta}_1$ while $\hat{V}_2$ is the asymptotic variance of $\hat{\vartheta}_2$.

However, it is more realistic to assume that the change point π is not known with certainty. In the case of the October 1979 change in Federal Reserve operating policy this implies that interest rates anticipated this policy change or, alternatively, that the resultant change in interest rates was effected with a lag (see Antoncic [1986] for further details). Bliss and Smith [1994] argue that the misspecification of this change point reduces the power of CKLS's test to reject the null hypothesis of no regime shift. If π is taken to be unknown, the resultant statistical inference is complicated by the fact that this parameter is no longer identified under the null hypothesis. As a result, likelihood ratio or Wald tests which treat π as a parameter do not possess their standard asymptotic distributions.

Because of these difficulties, we follow Andrews [1993] and test for a structural break when the change point is unknown by considering test statistics of the form

$$\sup_{\pi \in \Pi} LR_T(\pi) \quad \text{and} \quad \sup_{\pi \in \Pi} W_T(\pi) \tag{10}$$

where Π is some pre-specified subset of [0,1]. That is, if the change point is known to lie in some restricted interval, we calculate the Wald and likelihood ratio test statistics at each potential change point within this test interval and determine corresponding maxima. Andrews derives the nonstandard asymptotic null distribution of these maximal test statistics and tabulates their critical values for a variety of test intervals.

3.1 Data

Following CKLS, we use the one-month Treasury yield[6] to proxy for the short term riskless rate, r. We also consider CKLS's sample period, June 1964 through December 1989, and given annualized monthly data, this provides a sample size of $T = 307$ observations. However, unlike CKLS who use CRSP's Fama 12-month Treasury bill term structure file, we rely on one-month risk-free rates (the average of bid and ask) reported in CRSP's risk-free rate file since this data provides a more accurate measure of prevailing one-month riskless rates.[7]

3.2 Empirical Results

Without further restriction, estimation of (2) can result in inadmissible γ values larger than one; for example, CKLS report $\hat{\gamma} = 1.4999$. By definition, a structural break gives rise to non-stationarity in the underlying data, and so it is not surprising that CKLS cannot detect the October 1979 regime shift once such non-stationarity has been captured by their estimated γ value.[8]

The asymptotic null distribution of Andrew's maximal test statistics requires stationarity

[6]Duffee [1994] argues that the instantaneous riskless rate is better proxied by the one-month Eurodollar yield. However, to ensure the comparability of our results to those of CKLS, we rely on Treasury bill yields throughout.

[7]See Duffee [1994] for further details. The Fama 12-month Treasury bill term structure file is based on the longest bill with at least 11 months and 10 days to maturity on a given date. The yield on this bill when there is approximately one month until its maturity is taken to be the one-month Treasury yield. In contrast, the CRSP risk-free rate file's one-month series is constructed by selecting that Treasury bill closest to 30 days to maturity, regardless of the bill's original term to maturity; if more than one bill is closest to the targeted 30 day maturity, that bill with shortest original term maturity is chosen. Therefore, unlike the 12-month file, there is little variation between the target and actual maturities in the risk-free file. For example, over our sample period, the days-to-maturity underlying the one-month yield series obtained from the Fama 12-month Treasury bill file range between 10 and 41 days. In addition, given that with one month to maturity the 12 month bill is furthest off-the-run, liquidity effects may further jeopardize the accuracy of these quotes.

[8]It is interesting to note from CKLS's Table V that only when γ is assumed or estimated to be larger than one does their χ^2 test not reject the null hypothesis of no deterministic regime shift.

to be applicable. To ensure this, we fix γ and restrict our attention to $0 \leq \gamma \leq 1$. By varying γ we investigate whether allowing the volatility of interest rate changes to depend on the level of interest rates affects the likelihood of detecting a deterministic regime shift.

Using the CRSP risk-free rate file's one month series, Figure 2 displays the values of the likelihood ratio and Wald test statistics for each $0 \leq \gamma \leq 1$, in increments of 0.01. We use the test interval $\Pi = [0.45, 0.55]$; that is, assuming the CKLS specification, we investigate for each given γ value whether there is evidence of a deterministic structural break in interest rates anytime between September 1978 and November 1980. Since for a given γ the model of interest rate dynamics now involves $p = 3$ parameters, $\{a, b, \sigma^2\}$, the 5% asymptotic critical value under the null hypothesis of no deterministic regime shift is 10.15 (Table I, page 840, Andrews [1993]).

For any admissible γ value, we see reliable evidence of a once-and-for-all deterministic break in short term riskless interest rates between September 1978 and November 1980. The likelihood ratio test appears to provide more significant evidence against the null hypothesis of no structural break than does the Wald test.[9] In either case, however, the reported test statistics are well in excess of their 5% asymptotic critical value.[10] As γ increases, both test statistics tend to decline in value, consistent with CKLS's intuition that the evidence of a structural break depends upon whether the volatility of interest rate changes is modeled as

[9]We also investigated the properties of these test statistics under the null hypothesis of no regime shift. To do so, we simulated $T = 307$ observations according to $r_{t+1} - r_t = a + br_t + c\sqrt{r_t}\epsilon_{t+1}$, where $\{\epsilon_{t+1}\}$ are simulated iid standard normals, assuming $r_0 = 0.07$ with $a = 0.004$, $b = 0.06$, and $c = 0.01$. The maximal Wald and likelihood ratio test statistics were then calculated assuming $\Pi = [0.45, 0.55]$. We repeated this experiment 500 times and tabulated the resultant empirical distributions. Under the simulated null hypothesis, the test statistics had very similar sampling characteristics. For example, the 5% (10%) empirical cutoff value for the Wald statistic was found to be 11.1621 (9.3003) as compared to 11.6174 (9.8630) for the likelihood ratio statistic. Different γ values, $0 \leq \gamma \leq 1$, did not significantly alter these results.

[10]We repeated this analysis for one-month Treasury yields obtained from CRSP's 12-month Treasury bill file used by CKLS and found similar results.

being dependent upon the level of interest rates. However, even for $\gamma = 1$, we still have significant evidence of a deterministic structural break.

We now turn our attention to whether a more accurate modeling of volatility dynamics, beyond simply assuming that the volatility of interest rate changes depends on the level of interest rates, is consistent with this evidence. Such a model would provide the basis for an improved description of short term riskless interest rate dynamics.

4 Stochastic Regime Shifts in Short Term Riskless Interest Rates

The previous section provided evidence of a deterministic regime shift in U.S. short term riskless interest rates. Even though the test procedures allowed for a potentially unknown change point, we explicitly assumed that the regime shift was a permanent event with no further shifts possible. However, this may not be a realistic assumption. For example, the October 1979 change in Federal Reserve operating policy was short-lived, ending in September 1982 and this suggests the possibility of other regime shifts.[11] It may, therefore, be more appropriate to model the change in regimes itself as a random variable. One way to capture this behavior is by Hamilton's [1990] Markov switching model which assumes that regime shifts follow a Markov chain.[12]

[11] Romer and Romer [1990] relying on the *Minutes* of the Federal Open Market Committee identified a number of dates in addition to October 1979 when Federal Reserve operating policy changed, including December 1968, April 1974, and August 1978.

[12] Variations on this model for interest rate dynamics have been suggested by Cai [1994] and Naik and Lee [1994]. Unfortunately, Cai does not allow the volatility of changes in r to explicitly depend on the level of r and so cannot address CKLS's claim that doing so eliminates evidence of a deterministic structural break. Unlike Naik and Lee, we later examine the case where volatility follows a diffusion process.

4.1 Testing for Markov Regime Shifts

The Markov regime shifting specification for interest rate dynamics is summarized by (6) and (7). We estimate the drift parameters a and b by OLS; these estimators are consistent even when the variance is itself stochastic. As a result, we consider the following discrete-time specification

$$res_t = \sigma(t) r_t^{\gamma} \epsilon_t,$$

and

$$\sigma(t) = \theta + \beta I(t),$$

where $I(t)$ summarizes the state of the system and is governed by a two state Markov chain parameterized by (p, q), res_t is the residual from the OLS estimation at time t, while the $\{\epsilon_t\}$ are assumed iid standard normal.[13]

The parameters to be estimated are summarized by $\vartheta = (\beta, \Gamma, \theta, \gamma)$, with $\Gamma = (p, q)$. To test for the presence of Markov regime shifts, we consider the following null and alternative hypotheses:

$$H_0 : \beta = 0 \qquad H_1 : \beta \neq 0. \tag{11}$$

However, when $\beta = 0$, the variance $\sigma^2(t)$ is constant and the corresponding likelihood function is flat with respect to $\Gamma = (p, q)$. In other words, the transition probabilities p and q are not identified under the null hypothesis. Furthermore, the scores with respect to β, p, and q are all identically zero. Consequently, standard asymptotic theory does not apply. For example, the usual $t - statistic$ on β will no longer have a standard t distribution.[14]

[13] We also considered Markov switching models which allowed shifts in mean as well as variance. However, likelihood ratio tests provided no evidence of mean shifts once volatility shifts had been taken into account.

[14] One might attempt to resolve these problems by simulating under the null hypothesis and then estimating

Fortunately, Hansen [1992] has developed a theory of statistical testing precisely under these nonstandard conditions. Whereas standard statistical theory requires that the mean of the likelihood ratio be well-behaved, Hansen relies on much weaker regularity conditions on the deviation of the likelihood ratio from its mean. While the resultant statistical procedure is not optimal when standard asymptotic theory applies, Hansen's method has reasonable power and is not overly conservative for the nonstandard conditions underlying the Markov regime switching model. See Appendix A for a summary of this statistical method as well as further details.

4.1.1 Empirical Results

Table 1 presents the maximum likelihood parameter estimates of the Markov switching model of interest rate dynamics. As before, short term riskless interest rates are proxied by one-month Treasury yields obtained from the CRSP risk-free rate file.

Our empirical results are consistent with the presence of Markov regime shifts, $\beta \neq 0$.[15] There is persistence in these regimes as the transition probability p of remaining in the high volatility state is quite large. The persistence of the low volatility state, as measured by $1 - q$, is also correspondingly high.

---

the switching model to obtain the sampling distribution of the corresponding likelihood ratio test statistic. Unfortunately, there are severe difficulties with this approach. In particular, under the null hypothesis, the likelihood function for the nonlinear switching model is ill-behaved, usually with numerous local maxima, rendering global optimization extremely difficult.

[15]Hansen's nonstandard test procedure requires a grid search to obtain the appropriate test statistic. We varied β from -0.1 to -2.0 in steps of 0.1, while the transition probabilities p and q were varied from 0.1 to 0.9 in steps of 0.1. This grid search required $20 \times 9 \times 9 = 1620$ evaluations of the test statistic. The test statistic's maximum value was determined to be 5.6821 at $p = 0.9$, $q = 0.8$ and $\beta = -0.2$ with $\hat{\alpha} = -5.869$ and $\hat{\gamma} = 0.674$. We then simulated, with 1,000 replications, the sampling distribution of this test statistic and obtained the following empirical critical values: 10% = 1.28, 5% = 1.63, 1% = 2.30, and 0.5% = 2.63. These results provide reliable evidence of Markov regime shifts in our short term riskless interest rate data.

The maximum likelihood estimate of γ is 0.6792 with an asymptotic standard error of 0.1679. That is, with the presence of Markov regime shifts, we cannot reject Cox, Ingersoll, and Ross's hypothesis that $\gamma = 0.5$. This result is also consistent with CKLS's claim that much of the variability in short term riskless interest rates is attributable to their level. However, even after accounting for this heteroscedasticity, our results indicate that a significant Markov switching component remains.

To see this more clearly, Figure 3 plots the time series behavior of $\text{prob}(I_t = 1 \mid \bar{r}_t)$ at the Markov switching model's maximum likelihood parameter estimates. For example, the $1979 - 1983$ time interval represents a regime of highly volatile interest rates as evidenced by the corresponding likelihood of being in the high volatility state, $I_t = 1$. Figure 4 plots the resultant estimated volatility of monthly changes in one month Treasury yields. As expected, estimated volatilities are highest over the $1979 - 1983$ time period. While the $1974 - 1975$ and post-1985 time periods also appear to correspond to the high volatility state, estimated volatilities are lower since prevailing interest rates were much lower.

We also test for stability in the Markov switching model of interest rate dynamics. In other words, we consider whether this switching model is itself subject to a once-and-for-all deterministic regime shift around October 1979. To do so in the presence of an unknown change point, as before, we select the test interval $\Pi = [0.45, 0.55]$ and, restricting our attention to $\gamma = 0.5$, compute likelihood ratio test statistics for each possible monthly change point between September 1978 and November 1980.[16] The resultant maximum likelihood ratio test statistic (10.84) actually obtains at October 1979. To reflect the uncertainty in the change point, the critical values for $p = 4$ parameters, from Andrews [1993], are 10.35 and

[16]Recall that we cannot reject $\gamma = 0.5$ for the full sample. We assume $\gamma = 0.5$ to make the required computations feasible.

12.27 at the 10% and 5% significance levels, respectively. Consequently, when uncertainty in the change point is acknowledged, we see but marginal evidence against the null hypothesis of no deterministic regime switch in the Markov switching model.

4.2 A Stochastic Volatility Model for Short Term Riskless Interest Rates

In the previous section we statistically rejected the assumption of deterministic interest rate volatility in favor of stochastic volatility characterized by Markov regime shifting. While appealing, this specification is limited since the 'base' instantaneous volatility, $\sigma(t)$, can attain only two values.

This section addresses this deficiency within the CKLS framework by allowing $\sigma(t)$ itself to follow a diffusion process[17]. We then develop and implement statistical techniques to efficiently estimate the parameters of this stochastic volatility model for short term riskless interest rates. In particular, we assess the extent to which interest rate volatility is dependent upon the level of interest rates within a more realistic stochastic framework where $\sigma(t)$ can attain a continuum of values (or regimes).

As before, the drift parameters a and b are estimated by OLS with the resultant residuals denoted by $\{res_t\}$. The logarithm of $\sigma^2(t)$ is denoted by x_t. We now posit the following

[17] A number of stochastic volatility models have previously been put forward in the literature. Hull and White [1987] and Wiggins [1989], for example, introduce stochastic volatility models for equity return dynamics with a view to pricing options on these assets. Melino and Turnbull [1990] consider stochastic volatility models for exchange rates, while Harvey, Ruiz, and Shephard [1993] apply these models to a multiple exchange rate setting. Longstaff and Schwartz [1992] provide a closed-form two factor model for pricing interest rate derivatives using the short term riskless interest rate and its volatility as state variables. Most recently, Jacquier, Polson, and Rossi [1994] investigate a Bayesian approach to stochastic volatility estimation by using Monte Carlo Markov chain simulation methods. In all of these cases, the estimation of the underlying stochastic variables is subject to substantial econometric difficulties.

dynamics for short term riskless interest rates

$$res_t = exp(\frac{1}{2}x_t)r_{t-1}^{\gamma}\epsilon_1 \quad (12)$$

$$x_t = \alpha + \beta x_{t-1} + \xi\epsilon_2, \quad (13)$$

where $\{\epsilon_1\}$ and $\{\epsilon_2\}$ are iid standard normals. Notice that we are assuming that $\ln\sigma^2(t)$ is itself stochastic and follows an $AR(1)$ process.

Taking the logarithm of the squared observations gives the following equivalent specification:

$$y_t = ln(res_t^2) = x_t + 2\gamma ln(r_{t-1}) + ln(\epsilon_1^2) \quad (14)$$

$$x_t = \alpha + \beta x_{t-1} + \xi\epsilon_2. \quad (15)$$

If we assume that $ln(\epsilon_1^2)$ is normally distributed[18], we may follow Harvey, Ruiz, and Shephard [1993] and rely on the Kalman filter to estimate the parameters of this model using quasi-maximum likelihood. That is, we maximize the likelihood function obtained via the Kalman filter by treating the conditional distribution of the observable (y_t) given the state (x_t) as if it were normal.

While the quasi-maximum likelihood procedure is straightforward to implement, its statistical efficiency hinges on the adequacy of the normal approximation to the *log* χ^2 distribution with 1 degree of freedom. To assess the adequacy of the approximation, Figure 5 graphs this distribution and its associated normal approximation. Clearly the approximate Kalman filter will be adversely affected by outlying observations in the *log* χ^2 distribution's skewed left tail which are highly irregular under the normal approximation.

[18]The mean and standard deviation of the normal distribution are chosen to match the corresponding mean and standard deviation of the log χ^2 distribution with one degree of freedom: $-(ln2 + euler)$ and $euler + \pi^2/6$, respectively, where *euler* is Euler's constant ≈ 0.57721.

To circumvent this problem and minimize any consequent loss of statistical efficiency, we use an alternative procedure (Fruhwirth-Schnatter [1994]) which *explicitly* incorporates the non-normality of the observable's conditional distribution. This integration-based Kalman filter is detailed in Appendix B. Rather than assuming the measurement errors to be normally distributed, we make the less stringent assumption that the prior on the state is normally distributed. As a result, this technique is likely to be more accurate than Harvey, Ruiz, and Shephard's quasi-maximum likelihood procedure[19].

4.2.1 Empirical Results

The results of implementing the integration-based Kalman filter are reported in Table 2. These estimates are consistent with mean reversion in the volatility of short term riskless interest rates ($\beta \neq 0$). The volatility also appears to be dependent upon the prevailing level of interest rates, $\hat{\gamma} = 0.6871$, and this relationship is estimated relatively precisely with a standard error of 0.2576.[20] As in the Markov switching model, we again cannot reject Cox, Ingersoll, and Ross's hypothesis that $\gamma = 0.5$ and, in fact, we obtain remarkably similar γ estimates in the Markov switching and diffusion models (0.6792 and 0.6871, respectively).

Given this similarity of γ estimates, it is interesting to statistically compare the fit of the Markov switching and diffusion models. Unfortunately, we cannot rely on a standard likelihood ratio test in making this comparison since the competing models are not nested. Rather, we use Vuong's [1989] modified likelihood ratio test procedure to formally compare

[19] As compared to the corresponding full non-linear filter (which takes approximately 100 times longer to computationally implement), we find the integration-based filter to be extremely accurate. Unreported experiments indicate that the likelihood obtained under the full non-linear filter differs by less than 0.002% from the integration-based Kalman filter's result.

[20] This can be compared with the result obtained using the quasi-maximum likelihood procedure where $\hat{\gamma} = 0.9281$ with a standard error of 0.3404.

these non-nested hypotheses. Vuong's approach uses the standard deviation of the difference in the logarithm of the competing models' densities to standardize the likelihood ratio statistic and can also be modified to account for differences in their degrees of freedom. See Appendix C for an overview of this method and further details.

Using the Akaike [1973] and Schwarz [1978] criteria and relying on Newey-West [1987] standard errors, we compute corresponding Z statistics of 0.1146 and 0.9566, respectively, when comparing the Markov regime switching and diffusion models. By either criteria, we cannot statistically distinguish between these competing models of interest rate dynamics. Intuitively, both models capture the stochastic structure of interest rate volatility. However, since volatility is a second moment property of interest rate changes, even twenty five years of monthly data does not provide sufficient information to decide which model provides the better fit.

Finally, we also test whether the diffusion model is itself subject to a once-and-for-all deterministic regime shift around October 1979. For the test interval $\Pi = [0.45, 0.55]$, we once again restrict our attention to $\gamma = 0.5$ and compute likelihood ratio test statistics for each possible monthly change point between September 1978 and November 1980. This likelihood ratio test statistic is maximized in October 1979 at 9.425, and, from Andrews' Table I for $p = 5$ parameters, we see but marginal evidence against the null hypothesis of no deterministic regime switch in this stochastic volatility model of riskless interest rate dynamics.

5 Summary and Conclusions

Is there a regime shift or structural break in the behavior of short term riskless interest rates surrounding the October 1979 change in Federal Reserve operating policy? Using a generalized specification of interest rate dynamics, this paper provides reliable evidence of a once-and-for-all deterministic regime shift in the behavior of one-month Treasury yields even after acknowledging that the precise timing of this change point is potentially unknown.

A deterministic regime shift, however, assumes such an event is permanent with no further shifts possible. To relax this assumption, we also model the change in regimes as a random variable itself by implementing a Markov switching model of interest rate dynamics. We also generalize this specification by allowing interest rate volatility to follow a diffusion process but cannot statistically distinguish between these competing stochastic volatility specifications.

Our results confirm the importance of modeling interest rate volatility in accurately describing the dynamics of short term riskless interest rates. It is not enough to simply assume that the volatility of interest changes depends solely on the prevailing level of interest rates. While such interest rate level effects are clearly present in the data, an important stochastic volatility component remains.

6 Appendix A: Hansen's Nonstandard Test of a Markov Switching Model

Denote the Markov switching model's log-likelihood function by

$$L_T(\beta,\Gamma,\theta,\gamma) = \sum_{i=1}^{T} li(\beta,\Gamma,\theta,\gamma).$$

Set $\alpha = (\beta,\Gamma)$ and $\delta = (\theta,\gamma)$ so that $L_T(\beta,\Gamma,\theta,\gamma) \equiv L_T(\alpha,\delta)$ and $li(\beta,\Gamma,\theta,\gamma) \equiv li(\alpha,\delta)$. Since δ is identified, we can eliminate δ by concentration

$$\hat{\delta}(\alpha) = max_\delta L_T(\alpha,\delta),$$

giving the concentrated log-likelihood function

$$\hat{L}_T(\alpha) = L_T(\alpha,\hat{\delta}(\alpha)).$$

It is more convenient to work with the corresponding likelihood ratio function defined by

$$\begin{aligned} \hat{LR}_T(\alpha)) &= L_T(\alpha,\hat{\delta}(\alpha)) - L_T(\alpha_0,\hat{\delta}(\alpha_0)) \\ &= \sum_{i=1}^{T}[\hat{li}(\alpha) - \hat{li(\alpha_0}]. \end{aligned}$$

where α_0 is the value of α under the null hypothesis. The likelihood ratio test statistic for testing $H_0 = \alpha_0$ against the composite alternative $H_1 \neq \alpha_0$ is then

$$\hat{LR}_T = sup_\alpha \hat{LR}_T(\alpha).$$

The likelihood ratio function can be decomposed into its mean, $R_T(\alpha)$, and deviation from the mean, $Q_T(\alpha)$:

$$\hat{LR}_T(\alpha) = R_T(\alpha) + Q_T(\alpha).$$

Standard asymptotic theory requires that $R_T(\alpha) = E[\hat{LR}_T(\alpha)]$ be well-behaved. Unfortunately, as argued earlier, this is not the case with the Markov switching model. However, we can avoid this requirement by noting that under H_0 we have $R_T(\alpha) \leq 0$, yielding the following upper bound on the likelihood ratio function

$$\hat{LR}_T(\alpha) \leq Q_T(\alpha).$$

Hansen uses this maximum to test H_0. Since it is a bound, the resultant test will be conservative (under-reject when the null is true) and suffer a consequent loss in power (ability to reject the null when it is false). However, Hansen provides simulation evidence in the context of a Markov switching model for GNP which indicates that the proposed test is not overly conservative and appears to have reasonable power.

Hansen's test procedure is implemented on the basis of the variance standardized $Q_T(\alpha)$ statistic:

$$LR_T^* = sup_\alpha \frac{\hat{LR}_T(\alpha)}{\hat{V}_T^{\frac{1}{2}}(\alpha)},$$

where

$$\hat{q}_i(\alpha) = \hat{l}i(\alpha) - li(\hat{\alpha}_0) - \frac{1}{T}\hat{LR}_T(\alpha)$$

with variance measure

$$\hat{V}_T(\alpha) = \sum_{i=1}^{T} \hat{q}_i^2(\alpha).$$

We may simulate the variance standardized statistic, LR_T^*, for a particular value of α by

$$\sum_{i=1}^{T} \frac{\hat{q}_i(\alpha)u_i}{\hat{V}_T^{\frac{1}{2}}(\alpha)},$$

where $\{u_i\}$ are simulated iid standard normal random variables. The supremum of these test statistics across different values of α allows us to test $H_0 = \alpha_0$ against the composite alternative $H_1 \neq \alpha_0$. This latter computation may, in general, be burdensome as the optimization of this potentially ill-behaved function may necessitate a grid search over values of α.

7 Appendix B: Fruhwirth-Schnatter's Integration-Based Kalman Filter for Non-Gaussian Time Series

Recall that for the general filtering problem the likelihood function is obtained by Bayesian updating. Let Y_t denote the history of the observable through time t, $Y_t \equiv \{y_t, y_{t-1}, \ldots y_0\}$. Given a prior on the state $p(x_{t-1} \mid Y_{t-1})$, there are three stages to the iterative filtering procedure: a projection to obtain $p(x_t \mid Y_{t-1})$; followed by an integration to calculate the conditional likelihood $p(y_t \mid Y_{t-1})$; and finally an updating to obtain $p(x_t \mid Y_t)$. More specifically, given the prior $p(x_{t-1} \mid Y_{t-1})$, the projection is given by:

$$\begin{aligned} p(x_t \quad \mid \quad Y_{t-1}) &= \int_{x_{t-1}} p(x_t, x_{t-1} \mid Y_{t-1}) dx_{t-1} \\ &= \int_{x_{t-1}} p(x_t \mid x_{t-1}, Y_{t-1}) p(x_{t-1} \mid Y_{t-1}) dx_{t-1} \\ &= \int_{x_{t-1}} p(x_t \mid x_{t-1}) p(x_{t-1} \mid Y_{t-1}) dx_{t-1}. \end{aligned}$$

The conditional likelihood is obtained by integrating with respect to x_t:

$$\begin{aligned} p(y_t \mid Y_{t-1}) &= \int_{x_t} p(y_t, x_t \mid Y_{t-1}) dx_t \\ &= \int_{x_t} p(y_t \mid x_t) p(x_t \mid Y_{t-1}) dx_t, \end{aligned}$$

while the updating is given by:

$$\begin{aligned} p(x_t \quad \mid \quad Y_t) &= p(x_t \mid y_t, Y_{t-1}) \\ &= p(x_t, y_t \mid Y_{t-1}) / p(y_t \mid Y_{t-1}) \\ &= p(y_t \mid x_t) p(x_t \mid Y_{t-1}) / p(y_t \mid Y_{t-1}). \end{aligned}$$

For a linear state space model with multivariate normal measurement errors, the Kalman filter gives the projection, conditional likelihood, and updating as simple matrix calculations (see Harvey [1989]). However, for our problem we have a linear state space model but non-normal measurement errors.

Rather than assuming the measurement errors are normally distributed, we follow Fruhwirth-Schnatter [1994] and assume that the prior distribution on the state is normal. Since the projection based on a normal prior preserves normality, it can be implemented analytically. The conditional likelihood, however, requires numerical integration as the measurement error is assumed $log\ \chi^2$ distributed. Taking this into account, we integrate the posterior distribution $p(x_t \mid Y_t)$ to obtain its first and second moments needed in our updating scheme. We then match moments and assume that $p(x_t \mid Y_t)$ is approximately normally distributed. By maximizing the product of the resultant conditional likelihoods across time, we obtain the model's maximum likelihood parameter estimates.

Fruhwirth-Schnatter recommends Gauss-Hermite integration to implement this integration-based Kalman filter. We experimented with Gauss-Hermite as well as the extended Simpson's rule for this problem. The Gauss-Hermite method is a variable-grid numerical quadrature method that is extremely efficient for calculating integrals whose kernels are products of polynomials and normal densities. However, Simpson's rule is extremely accurate as well as being simpler to implement. For example, we obtained 12-digit accuracy using a 141 point extended Simpson's rule in which we updated x_t's range at each time point t by its predicted mean $\pm$ 7 estimated standard deviations.

8 Appendix C: Vuong's Test for Non-Nested Alternative Hypotheses

This appendix briefly describes our implementation of Vuong's test for comparing the Markov switching stochastic volatility model, denoted H_1, with the diffusion stochastic volatility model, denoted H_2. Except for the degenerate case when volatility is constant, the two models do not overlap. We have already rejected the hypothesis of constant volatility using Hansen's nonstandard test, therefore, as noted by Vuong, we may proceed as though the two hypotheses H_1 and H_2 are strictly non-overlapping.

Vuong's test statistic is a likelihood ratio statistic adjusted by the standard deviation of the average log-likelihood functions under the competing models. In particular, let lnf_t denote the loglikelihood of $r_t \mid \tilde{r}_{t-1}$ at the maximized likelihood under H_1 and let lng_t be the corresponding maximum loglikelihood under H_2. For convenience, set $q_t = lnf_t - lng_t$. With T observations, define $LR_T = \sum_{t=1}^{T} q_t$, the maximum loglikelihood ratio and denote by w_T the estimated standard deviation of q based on the T observations. Under mild regularity assumptions, outlined in Vuong, the test statistic $T^{-0.5}LR_T/w_T$ has approximately the standard normal distribution when the competing hypotheses are indistinguishable. It is also possible to adjust this statistic to account for different degrees of freedom in the competing hypotheses and thus give more weight to a particular hypothesis. To do this, we replace LR_T by $LR_T^* = LR_T - K_T$ where, for example, using the Akaike [1973] criterion $K_T = p_1 - p_2$ with p_1 being the number of estimated parameters under H_1 and p_2 the number of parameters under H_2. Alternatively, one could use the Schwarz [1978] criterion with $K_T = ln(T) * (p_1 - p_2)/2$.

Vuong's test is based on the assumption of independent drawings from a common parent distribution. However, for our time series application we may expect autocorrelation or heteroscedasticity in the data which could potentially distort the estimation of the standard deviation of the average loglikelihoods. As a precaution, we use the Newey-West [1987] method to estimate this standard deviation. We experimented with various lag length adjustments ranging from 3 to 12 periods. The results were quite insensitive to the choice of lag and, in addition, were also indicative of no significant deviation from the assumption of independence.