

An Empirical Examination of Term Structure Models with Regime Shifts

John Driffill^a Turalay Kenc^b
Martin Sola^{a,c}

^a Birkbeck College, University of London,

^b Imperial College, University of London,

^c Universidad Torcuato Di Tella

February 10, 2003

Abstract

We examine several continuous-time term structure models in which the short rate is subject both to continuous changes and to discrete shifts. Several regime-switching term structure models are developed, with regime-dependence in various combinations of their drift and diffusion parameters. We examine their predictive power. Our empirical analysis suggests that it is important to attempt to specify the switching model correctly: badly parameterized switching models may not be an improvement (in terms of pricing) over models which do not allow for regime switching, even when there are clear breaks in the data.

Keywords: Term structure of interest rates; bond yields; stochastic discount factor/pricing kernel; and regime switching.

JEL Classification numbers: E43, G12

1 Introduction

Modelling the term structure of interest rates plays a significant role in pricing fixed income derivatives, in risk management and in designing macroeconomic policies. A number of models belonging to a particular class, called “affine” models [Duffie and Kan (1996) and Dai and Singleton (2000)], have been widely used in the theoretical and empirical literatures because of their tractability and functional richness. However, their empirical performance is somewhat unsatisfactory [Ghysels and Ng (1998)]. There are new developments in the affine term structure literature aimed at improving their predictive power. These involve richer parameterizations (i) for the price of risk [Duarte (2000), Dai and Singleton (2001) and Duffee (2002)] and (ii) to capture discrete shifts in interest rates [Naik and Lee (1997), Evans (1998) and Bansal and Zhou (2002)].

The second approach is appealing since it adds more realism by capturing discrete movements in the economy. We argue that for two reasons interest rates are subject to discrete shifts as well as continuous changes.¹ These are: (i) there is increasing evidence in the literature that both short and long term interest rates are characterized by stochastic regime switching processes [see, for example, Hamilton (1988), Sola and Driffill (1994), Garcia and Perron (1996), Gray (1996), and Dahlquist and Gray (2000)]; and (ii) economic regimes such as business cycle expansion and recession have regime switching effects on interest rates. Empirical evidence for regime switching interest rates can also be found indirectly in the form of the parameter instability of single-regime interest rate models over the 1979-82 period [see Cai (1994),

¹A possible third reason is that regime switching interest rate models may capture some of the non-linearities in interest rates which may show up in higher order unconditional moments [Ang and Bekaert (2002b)].

Pearson and Sun (1994) and Brenner, Harjes and Kroner (1996)].

The models that currently exist in the literature, which allow for discrete shifts (see for example, Naik and Lee (1997), Evans (1998) and Bansal and Zhou (2002)), are solved under different assumptions about the evolution of the instantaneous interest rate. For example, with the exception of Naik and Lee (1997) (the Vasicek model), the most widely used is the CIR model (since it is rich enough to mimic most of the possible shapes of the term structure). In Naik and Lee (1997), only the volatility term is allowed to switch while in Bansal and Zhou (2002) the full set of parameters is allowed to switch.² This paper provides a framework for deciding which of these assumptions is best supported by the data.

A key element of this paper is the search for a suitable parameterization for the short-term interest rate. When looking for it we have to take into account several criteria. An obvious first criterion is the goodness of fit of the short-term interest rate itself. But we are also interested in how well the prices of longer maturity bonds, that it predicts via the term-structure model, fit the data. The model for the short rate has to be rich enough to be capable of reproducing the properties of the term structure data.

We compare a variety of models for the short term interest rate that allow for regime switching in some or all of the following aspects: the volatility of the short rate; the long-run value of the short rate, and the speed of adjustment to the long-run value. More precisely, we use an extended *mean-reverting square root* process due to Cox, Ingersoll and Ross (1985; CIR) to account for situations in which there is the possibility of changes in parame-

²In estimating parameters including market prices of risks in both regimes Bansal and Zhou (2002) use the entire term structure data. Papers by Landen (2000) and Elliott, Hunter and Jamieson (2000) use a Hidden Markov model in mixing continuous changes in the short rate with discrete changes.

ters. We then use each of these models to price bonds of different maturities. The versions of the CIR short rate process that we use include: (1) a benchmark case with no regime-switching; models with regime-switching in: (2) volatility; (3) volatility and the speed of adjustment; (4) volatility and the long-run value of the short rate; and (5) volatility, the speed of adjustment and the long-run value of the short rate. This allows us to find the model that best matches the data for bond prices of all maturities rather than using model selection criteria that use only the data on the short-term interest rate itself.³

The issue of which univariate parameterization best characterizes short term interest rates has been left largely unanswered in the literature. Choosing among the alternatives is rather difficult since a model that allows all the parameters to switch typically has a very flat likelihood. Nevertheless, different parameterizations can produce very different bond prices. In this paper we attempt to obtain the best model for bond prices (in the sense that its predictions are closest to the observed ones) and as a by-product use the term structure of interest rates as a criterion to assess which parameterization of the short term interest rate is preferred. In this paper we recursively estimate the different parameterizations of the switching CIR process for the short term interest rate described above and use the results to price bonds for different maturities. In this way we generate a series of prices which we then compare with the actual prices in terms of fit and also in terms of the

³Gray (1996) shows that a switching CIR is the preferred model to characterize the US short term data. Since then this model has been widely used to characterize the short term data (see for example Dahlquist and Gray (2000) and Ang and Bekaert (2002a)). None of these papers ask the question of how different simplifications of this model perform in terms of fit, or what is more important in terms of forecasting, which is crucial for pricing bonds.

shape of the term structure at different points in time. Interestingly, the results obtained for the whole sample using standard likelihood ratio tests and goodness of fit criteria do not coincide with those obtained using the pricing strategies described above. The possible interpretation of these results is that the model which provides the best fit does not necessarily provide the best forecast and therefore the best price.

2 Pricing Bonds when the instantaneous rate switches between two Brownian motions

We consider a financial market model in which all activity takes place in the time interval $[0, \mathcal{T}]$ and operates on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(t \in [0, \mathcal{T}]}, \mathcal{P})$. Elements of the stochastic basis are explained as follows: Ω denotes the set of all possible states of the economy. \mathcal{F} is the set of distinguishable events in the economy and is a given σ -field of subsets of Ω . The term $(\mathcal{F}_t)_{(t \in [0, \mathcal{T}]}$ (with $\mathcal{F}_T = \mathcal{F}$) represents a filtration describing the information arrival. Finally, \mathcal{P} denotes the probability beliefs of the agents. Stochastics in the economy are generated by a standard Wiener process W in \mathbb{R}^1 and a Markov process.

It is assumed that short term interest rates are driven by a stochastic process with state dependent drift and diffusion parameters, where the discrete states X are said to be regime 0 and 1. The switch between the two regimes is governed by a Markov chain with intensity (rate) matrix $H = (h_{ij})$.

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} -h_{12} & h_{12} \\ h_{21} & -h_{21} \end{bmatrix} \quad (1)$$

In the above continuous-time Markov chain model, the probability that a transition occurs from a given source state depends not only on the source state itself but also on the length of the interval of observation. Then, the

probability that a transition occurs from state i (say $i = 1$) to state j (say $j = 2$) in the interval $[t, t + \Delta t)$ is equal to $h_{12}\Delta t + o(\Delta t)$. Similarly, $1 - h_{12}\Delta t + o(\Delta t)$ is the probability that the process remains in state i . Moreover, the above Markov chain has a stationary distribution property, and hence the long-run probability and conditional probability given the source state i can be easily calculated. $\pi_\infty(1) = \frac{h_{21}}{h_{12}+h_{21}}$ and $\pi_\infty(2) = \frac{h_{12}}{h_{12}+h_{21}}$ are the long run probabilities that the process is in the regime 1 and 2, respectively. $\frac{h_{21}}{h_{12}+h_{21}} + \frac{h_{12} \exp[-(h_{12}+h_{21})(s)]}{h_{12}+h_{21}}$ is the probability that the process is in regime 1 at time $t + s$ given that it is in regime 1 at time t . For analytical tractability, it is assumed that the discrete states X_t are independent of the Wiener process $W(t)$. It is also assumed that agents in the financial markets know the actual state of the system X_t .⁴

The short rate process r is thus formally represented by the stochastic differential equation for which we specify a “regime-switching mean-reverting square root process” (the RSCIR process)⁵,

$$dr(t) = \kappa(X_t)[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t). \quad (2)$$

where $dW(t)$ is the increment from a standard Wiener process. The drift term, $\kappa(X_t)[\alpha(X_t) - r(t)]dt$, in (2) captures regime-dependent mean-reversion by setting $\kappa(X_t) > 0$. The parameter $\alpha(X_t)$ (> 0) is the regime-dependent implied long-run mean interest rate, and $\kappa(X_t)$ determines the regime-dependent adjustment speed of r toward the long-term mean. $\sigma(X_t)^2 r$ is the regime-dependent variance of unexpected interest rate changes. The term $\sigma(X_t)$ is

⁴However, the econometrician does not, and has to make inferences of it based on the observable history of the system.

⁵Evans (1998) and Bansal and Zhou (2002) use a discrete version of the RSCIR process. For the connection between discrete-time and continuous-time single regime models see Sun (1992).

the regime-dependent volatility factor and serves as a scale factor. This implies that the volatility of the interest rate is parameterized as a function of interest rate levels and produces conditional heterokedasticity which is the cause of the leptokurtosis in the unconditional distribution of changes in the short rate.

In order to price bonds based on the short rate specified in (2) we use the stochastic discount factor process of the form (due to Cochrane (2001))

$$\frac{d\Lambda(t)}{\Lambda(t)} = -r(t)dt - \sigma_{\Lambda}\sqrt{r(t)}dW(t) \quad (3)$$

where $\Lambda(t)$ is the stochastic discount factor, equivalent to the pricing kernel or state-price deflator [see for example Fisher and Gilles (2000)], which is determined by marginal utility

$$\Lambda(t) = e^{-\delta t}u'(c(t))$$

where c is the consumption and δ is the time preference rate. Note that for simplicity it is assumed that uncertainty in the instantaneous interest rate and the stochastic discount rate is governed by the same Brownian motion $Z(t)$.⁶

We assume that there is a market for every bond for every choice of maturity T and the market is arbitrage free. We assume furthermore that, for very T , the price of a maturity T -bond has the form

$$P(t, T) = F(t, r(t), X, T), \quad (4)$$

where $F(t, r(t), X, T)$ is a function of four variables and more precisely takes a semi-affine form

$$\ln F(t, r(t), X, T) = A(r, X, T) - B(r, X, T)r \quad (5)$$

⁶In the literature it is sometimes assumed that they are governed by different sources of uncertainty. But at later stages of the modelling it is either explicitly or often implicitly assumed that the discount factor and shocks are perfectly correlated, e.i., $\rho = 1$.

where A and B are deterministic functions. This leads to a **semi-affine form term structure(SATS)**.⁷

Our aim is now to price bonds using (i) the above semi-affine term structure and (ii) no-arbitrage condition. To do this we make two further assumptions: one about the hedging behavior of investors and the other about the market price of risk. The first assumption is that investors do not hedge the regime switching risk while they hedge the risk of the continuous changes in short rates. The second is that the market price of risk is the same in different regimes as in Naik and Lee (1997) and Landen (2000). To satisfy the no-arbitrage condition we will follow the fundamental pricing equation

$$\mathcal{E}(t) \left(\frac{dF}{F} \right) - \left(\frac{1}{P} \frac{\partial F}{\partial T} \right) dt = rdt - \mathcal{E}(t) \left(\frac{dF}{F} \frac{d\Lambda}{\Lambda} \right) \quad (6)$$

where \mathcal{E} is the expectation operator. The left hand side is the expected (holding period) rate of return on the T-Bond which should be equal to the sum of the risk-free rate and the covariance of the return with the discount factor or marginal utility (the last term in (6)). This is the obvious continuous-time analogue to the CAPM expression. From the Itô formula we get the following price dynamics for the T-bond (dF)

$$dF = aFdt + bFdW \quad (7)$$

where

$$a = \frac{\kappa(X_t)[\alpha(X_t) - r]F_r + \frac{1}{2}\sigma(X_t)^2 r F_{rr} + \sum_{i=0}^1 \sum_{j=0}^1 h_{ij} \Delta F}{F} \quad (8)$$

$$b = \frac{\sigma(X_t)\sqrt{r}F_r}{F} \quad (9)$$

Plugging this in (6) and cancelling dt , we obtain the differential equation for

⁷Duarte (2000) also uses a similar terminology “semi-affine square-root” model in which he develops a different (flexible) parameterization for the price of risk.

bonds,

$$\kappa(X_t)[\alpha(X_t) - r]F_r + \frac{1}{2}\sigma(X_t)^2 r F_{rr} - F_T + \sum_{i=0}^1 \sum_{j=0}^1 h_{ij} \Delta F - rF = F_r \sigma \sigma_\Lambda r. \quad (10)$$

The right hand-side of (10) is the risk premium. Setting $\sigma_\Lambda = \lambda$ the market price of risk and rearranging (10) we obtain the following “term structure equation”

$$(\kappa(X_t)[\alpha(X_t) - r] - \sigma(X_t)\lambda r)F_r + \frac{1}{2}\sigma(X_t)^2 r F_{rr} - F_T + \sum_{i=0}^1 \sum_{j=0}^1 h_{ij} \Delta F - rF = 0, \quad (11)$$

with the boundary condition

$$F(T, X, r, T) = 1.$$

Using (5) we now easily compute the various partial derivatives of F , and since F must solve the term structure (11), after collecting terms we thus obtain

$$\left\{ \begin{array}{l} B_{0t}(t, T) - \kappa_0 B_0(t, T) - \frac{1}{2}\sigma_0 B_0(t, T)^2 + 1 \\ B_{1t}(t, T) - \kappa_1 B_1(t, T) - \frac{1}{2}\sigma_1 B_1(t, T)^2 + 1 \end{array} \right\} r + \quad (12)$$

$$\left\{ \begin{array}{l} A_{0t}(t, T) - (\kappa_0 \alpha_0 - \lambda_0 \sigma_0) B_0(t, T) + \sum_{j=0}^1 h_{0j} \Delta F_0 \\ A_{1t}(t, T) - (\kappa_1 \alpha_1 - \lambda_1 \sigma_1) B_1(t, T) + \sum_{j=0}^1 h_{1j} \Delta F_1 \end{array} \right\} = 0$$

where

$$\Delta F_0 = e^{A_j(t, T) - r B_j(t, T) - [A_0(t, T) - r B_0(t, T)]},$$

$$\Delta F_1 = e^{A_j(t, T) - r B_j(t, T) - [A_1(t, T) - r B_1(t, T)]}.$$

The boundary value $F(T, r, X, T) \equiv 1$ implies

$$\left\{ \begin{array}{l} A_i(T, T) = 0, \quad i = 0, 1, \\ B_i(T, T) = 0, \quad i = 0, 1. \end{array} \right.$$

Equation (12) may be solved by applying the commonly adopted log-linear approximation $e^y - 1 \approx y$ [see, for example, Bansal and Zhou (2002)].

$$\left\{ \begin{array}{l} B_{0t}(t, T) - \kappa_0 B_0(t, T) - \frac{1}{2} \sigma_0 B_0(t, T)^2 + h_{01}[B_0(t, T) - B_1(t, T)] + 1 \\ B_{1t}(t, T) - \kappa_1 B_1(t, T) - \frac{1}{2} \sigma_1 B_1(t, T)^2 + h_{10}[B_1(t, T) - B_0(t, T)] + 1 \end{array} \right\} r +$$

$$\left\{ \begin{array}{l} A_{0t}(t, T) - (\kappa_0 \alpha_0 - \lambda_0 \sigma_0) B_0(t, T) + h_{01}[A_1(t, T) - A_0(t, T)] \\ A_{1t}(t, T) - (\kappa_1 \alpha_1 - \lambda_1 \sigma_1) B_1(t, T) + h_{10}[A_0(t, T) - A_1(t, T)] \end{array} \right\} = 0$$

3 Data Description

There are two widely used data sets of US bond yields⁸: (i) McCulloch and Kwon (1993) which use a cubic spline to construct the yield curve and (ii) Fama and Bliss (1987) which use bootstrap methods to construct the yield curve. We use the data (monthly estimates of annualized continuously-compounded zero-coupon US government bonds yields) constructed by Duffee (2002) which are based on the Bliss (1997) extension of McCulloch and Kwon (1993). The data set ranges from January 1952 to December 1998.⁹ There are 564 monthly observations with 6 maturities 3, 6 month and 1, 2, 5, 10 year.

To find out how the shapes of the yield curves evolve over the sample period we plot the surfaces of yield curves in Figure 1. Likewise, Table 2 reports the summary statistics for the yields. The yield curve is typically upward sloping, but inverted around 1979 to 1981. The yield volatilities, measured by standard deviations, decrease with the maturities. Figure 1 reveals the widely reported empirical fact that there is a significant increase

⁸We assume that measurement errors such as not using the actual trade data, nonsynchronous quotations, and quotation errors are negligible.

⁹This data set is available on the web page associated with Duffee (2002). The address is <http://www.haas.berkeley.edu/~duffee/affine.htm>.

in interest rate volatility in the last two decades.

[Figure 1 approximately here]

[Table 2 approximately here]

4 Estimation

We use the 3 month T-Bill yield as the proxy for the instantaneous interest rate¹⁰ as in Andersen and Lund (1997) and Duffee (2002) to estimate the parameters of the regime-switching Cox, Ingersoll and Ross model (RSCIR). To avoid the serial correlation induced by overlapping expectations we converted the monthly data set into a quarterly one and our sample is 1964:1–1998:4. Five models that are specified in Table 1 are estimated. Note that a model without regime switching in the volatility scale factor σ is not included as the data reject such models.

Estimation and testing in the context of the Markov-switching models presented in Table 1—where the Markov chain $\{X_t\}$ is unobserved—can be carried out by using the recursive algorithm discussed in Hamilton (1994, ch. 22). This gives as a by-product the sample likelihood function

¹⁰Since our model is in continuous time, its state variable is the instantaneous interest rate which is unobservable. Chapman, John B. Long and Pearson (1997) explores this proxy problem and show that it is not economically significant for single-factor affine models. But it can be economically significant when applied to a two-factor affine model and a nonlinear single-factor model. In the literature, the one-month rate is also used a popular proxy for the instantaneous rate [see for example Chan, Karolyi, Longstaff and Sanders (1992) and Nowman (1997)]. But it is well documented that estimating the model with the one-month rate is relatively difficult. Another argument against using 1-month yields is that they are more likely to be influenced by liquidity needs. For the same reason Bansal and Zhou (2002) use the 6 month T-Bill yield.

which can be maximized numerically with respect to $(\kappa_0, \kappa_1, \alpha_0, \alpha_1, \sigma_0, \sigma_1, p, q)$, subject to the constraint that p and q lie in the open unit interval.

Table 1: Models to be estimated

- Model 1: No regime switching.

$$- dr(t) = \kappa[\alpha - r(t)]dt + \sigma\sqrt{r(t)}dW(t)$$

- Model 2: Regime switching in volatility.

$$- dr(t) = \kappa[\alpha - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$$

- Model 3: Regime switching in volatility and adjustment speed.

$$- dr(t) = \kappa(X_t)[\alpha - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$$

- Model 4: Regime switching in volatility and long-run rate.

$$- dr(t) = \kappa[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$$

- Model 5: Regime switching in all parameters.

$$- dr(t) = \kappa(X_t)[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$$

To estimate the RSCIR process (2) we discretize it using the exact discretization method as follows:

$$r_{t+\Delta} = e^{-\kappa(X_t)\Delta t}r_t + (1 - e^{-\kappa(X_t)\Delta t})\alpha(X_t) + \sigma(X_t)\sqrt{r_t}\sqrt{\frac{1 - e^{-2\kappa(X_t)\Delta t}}{2\kappa(X_t)}}\epsilon_{t+\Delta}. \quad (13)$$

$\epsilon_{t+\Delta}$ is assumed to be i.i.d. standard normal in the model. Notice that as Δt becomes small equation (13) reduces to the Euler discretization equation

$$r_{t+\Delta} = \varphi(X_t)r_t - (1 - \varphi(X_t))\alpha(X_t) + \sigma(X_t)\sqrt{r_t\Delta t}\epsilon_{t+\Delta}, \quad (14)$$

where $\varphi \approx 1 - \kappa\Delta t$ captures the autocorrelation of the interest rate.

In Table 3, we report Gaussian S-PML estimates of the parameters of the models presented in Table 1, along with corresponding asymptotic standard errors.¹¹ Estimates of the parameters vary (i) model to model and (ii) regime to regime. For example, estimates of the long-run rate range from 0.0161 (Model 1) to 0.0337 (Model 4) in regime 0. The estimate of the same parameter for the model 4 is 0.0149 in regime 1. There is an interesting result to report: a comparison between the versions of the RSCIR process, i.e. Models 2 through 4, reveals that the volatility parameter, σ does not vary much, ranging from 0.0310 to 0.0399 in regime 0 and from 0.0128 to 0.0130 in regime 1.

The allocation of time periods to the two states for the four switching models under consideration is shown in Figure 2. The period between 1965 and 1980 is assigned with high probability to state 1, with a brief departure from it around 1971. The period from 1980 to 1982 is assigned to the high interest rates/low-variance state (state 0). The remaining observations fall into state 1.

Table 3 shows that the hypothesis that model 4 and model 2 are valid simplifications of model 5 are rejected (the likelihood ratio test statistics are 5.9938, distributed $\chi^2(1)$ and 6.5392, distributed $\chi^2(2)$ respectively), while the null hypothesis that model 3 is a valid reduction is not rejected (the likelihood ratio test statistic is 2.299, distributed $\chi^2(1)$). The Akaike, Schwarz, and Hannan-Quinn specification criteria, give conflicting results. While model 5 is favored by the AIC, model 2 is favored by the SIC and model 3 by the HQ criteria.

¹¹The likelihood function was maximized by using the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm with numerically computed derivatives.

Neither the likelihood ratio test nor the selection criteria give us a clear cut indication of which should be the preferred model. Therefore we use the information contained in the term structure to decide which is the model that produces the best bond prices. To do that we recursively estimate the five models described above, price them with the pricing equation derived in Section 2 and compare the generated prices (returns) with the actual data. We therefore use the observations from 1964:1-1980:4 to start the pricing exercise and sequentially enlarge the sample up to 1998:4 obtaining as a result generated prices for 6 month, 1, 2 and 5 year bonds from 1980:4 to 1998:4. More formally at time τ a yield curve can be constructed simply by using the pricing model derived in Section 2 and instantaneous interest rates for the alternative models using information up to time $\tau = t_1, \dots, T - 1, T$. Therefore we recursively estimate the alternative models of the three months U.S. Treasury Bill rates and compute the different yield curves using $F(\tau, r(\tau), X_\tau, T)$ prices where information from 1 to τ is used to estimate the short run rate. This will produce a series of $T - t$ long term interest rates for each maturity and estimation model. Then we compare the generated data with the actual.

Notice that in order to obtain these prices we need to calculate the market price of risk λ in order to generate the yield curve. The market price of risk λ is calculated such that the observed yield on the 10 year bond equals the model generated one [for further information, see Backus, Foresi and Telmer (2000)]. Using the estimates of the alternative models for each period of time from 1980:4 to 1998:4 and the calculated market prices of risk we generate yield curves for our 5 models. Figures 4-6 plot the model generated yield curves and actual yield curve as of 1988:1, 1989:3 and 1998:4, respectively. A simple visual inspection shows that only model 4 is able to generate a

yield curve which resembles the actual curve for the three chosen dates. All the figures reveal that one can get differently shaped yield curves from the 5 models. Nevertheless it seems that even though the fact that model 4 (and to some extent model 2) manages to reproduce the shape of the actual yield curve is a desirable feature of the model, visual inspection of the shapes of the yield curve for every data point does not seem a feasible or formal enough strategy for use as a selection criterion. Therefore we attempt to use all the information contained in our generated prices to assess which of the models has best predictive power. Figure 7 plots the model generated and actual yields for the four maturities 6 months, 1 year, 2 years and 5 years. Table 4 reports some goodness of fit statistics which are used to compare the empirical performance of each of the 5 models. We present performance indicators such as mean squared error (MSE), relative mean square error (RMSE), mean absolute error (MAE) and relative mean absolute error (RMAE) of the difference between the generated yields and the actual data for each maturity and also an aggregate measure (column 5) which capture both the time series and the cross section dimension. A general feature of all the pricing models is that they perform better in predicting the lower maturities than in predicting higher maturities (the reduction of the error for the 5 year bond has to be attributed to the way the price of the risk is computed). More importantly, we find that Models 2 and 4 significantly outperform the other models in terms of producing prices closer to the actual data. Interestingly a straight comparison between Model 1 and Model 5 (advocated by Evans (1998) and Bansal and Zhou (2002)) shows that using the very general Markov switching parameterization may not produce better prices than the standard CIR model even when there are apparent structural breaks in the sample under scrutiny, implying that any

improvement of Model 5 over Model 1 in terms of fit is undone by the poor forecasting performance of Model 5. Figure 7 complements the information presented in Table 4 mostly by showing the deterioration in the fit of the simple CIR model after first half of the 1980s.

We have also reported in Table 5 the descriptive statistics of yield differences. The results show that Model 1 overestimates (underprices) yields (bonds) while other models underestimate (overprice) yields (bonds). In terms of the first two moments (mean and variance) Models 2 and 4 perform better than the other 3 models. When the third (skewness) and fourth (kurtosis) moments are considered, Model 4 gains some support.

A clear message can be drawn from our analysis if we consider an individual who intends to price bonds using a Markov switching model in 1998:4 (the end of the sample). If she attempts to use standard statistical criteria, she would find that model 2 and model 4 are not valid simplifications of model 5, while model 3 is preferred to model 5. On the other hand by carrying out a more extensive analysis, that is, by obtaining the bond prices by means of estimating recursively a Markov switching model, we find that model 4 (and to some extent model 2) not only are the only models that manage to reproduce the actual shape of the yield curve, but also seem to have produced on average much better prices than the other parameterizations. Most importantly the poor performance of model 5 compared with model 1 highlights how important it is (especially when the model is used to forecast) to attempt to specify the switching model correctly. This exercise seems to suggest that badly parameterized switching models may not be an improvement (in terms of pricing) over models which do not allow for regime switching, even when there are clear breaks in the data.

[Table 3 approximately here]

[Figures 2 – 6 approximately here]

[Table 4 approximately here]

[Figure 7 approximately here]

[Table 5 approximately here]

5 Conclusions

In this paper, we have provided an analysis of several regime-switching characterizations of the CIR term structure process. In order to make our results comparable with the existing literature, we keep intact the standard continuous-time framework and affine setup when modelling the term structure. We investigate how the pricing performance of the model is affected by different assumptions about which parameters (drift and diffusion) are specified as regime-dependent to capture. We employ corresponding/equivalent models to examine the empirical significance of alternative regime-dependent characterization of drift and diffusion parameters. Using these models we have estimated recursively Markov switching models for short term interest rates and generated bond yields which are then compared with actual yields. Our results reveal that simpler specifications, such as a term structure model with only a regime-dependent volatility parameter and a term structure model with both a regime-dependent volatility parameter and a regime-dependent long-run rate perform better than models with no regime switching, models where all the parameters switch, and models with both regime-dependent volatility and regime-dependent speed of adjustment parameters. These results have the interesting feature that the preferred pa-

parameterizations differ from those that one would have chosen on the basis of goodness of fit. This can only be interpreted as a sign that the models which better fit the data for the sample under scrutiny are not necessarily those that better forecast and this is crucial for pricing bonds.

Figure 1: The Historical U.S. Term Structure of Interest Rates

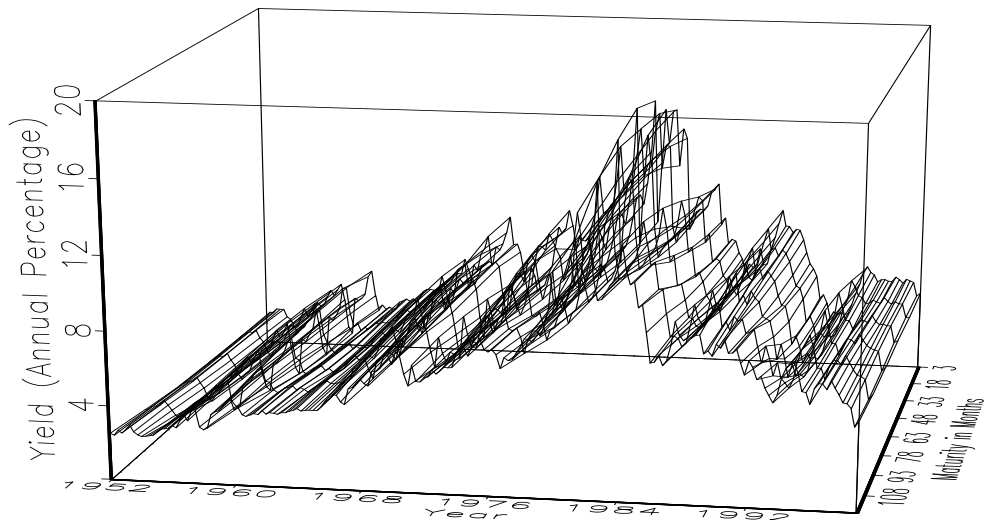


Table 2: Summary Statistics for the U.S. Term Structure of Interest Rates

Maturity	Period 1964-1978		Period 1979-1981		Period 1982-1998	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
3	4.15023	1.98836	12.0278	2.47187	6.50504	2.26995
6	4.38941	2.04902	12.2170	2.43729	6.74247	2.36220
12	4.56006	2.03908	12.1098	2.30643	7.07753	2.43233
24	4.72129	1.96039	11.8012	2.24064	7.55160	2.40562
60	4.96496	1.87292	11.4083	2.15478	8.15998	2.28925
120	5.10668	1.84290	11.2681	1.92634	8.54856	2.12881

Table 3: Parameter Estimates of Models

The table reports the recursive Markov switching maximum likelihood estimates as of 1998:4 for the 5 models which are

Model 1: $dr(t) = \kappa[\alpha - r(t)]dt + \sigma\sqrt{r(t)}dW(t)$

Model 2: $dr(t) = \kappa[\alpha - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$

Model 3: $dr(t) = \kappa(X_t)[\alpha - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$

Model 4: $dr(t) = \kappa[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$

Model 5: $dr(t) = \kappa(X_t)[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$

(The numbers in parentheses are the corresponding asymptotic standard errors).

	Model 1	Model 2	Model 3	Model 4	Model 5
Regime 0					
α_0	0.0161 (0.0028)	0.0154 (0.0024)		0.0337 (0.0239)	0.0296 (0.0016)
σ_0	0.0177 (0.0011)	0.0399 (0.0082)	0.0310 (0.0064)	0.0388 (0.0080)	0.0311 (0.0064)
κ_0	0.0684 (0.0344)	0.0635 (0.0332)	1.0234 (0.3387)	0.0725 (0.0360)	0.9905 (0.3417)
Regime 1					
α_1			0.0296 (0.0016)	0.0149 (0.0021)	0.0154 (0.0025)
σ_1		0.0128 (0.0009)	0.0130 (0.0008)	0.0128 (0.0008)	0.0129 (0.0008)
κ_1			0.0118 (0.0080)		0.0614 (0.0340)
Transition Probability $Pr(X_{t+1} X_t)$					
$Pr(X_{t+1} = 0 X_t = 0)$		0.9899 (0.0108)	0.9925 (0.0075)	0.9906 (0.0098)	0.9926 (0.0075)
$Pr(X_{t+1} = 1 X_t = 1)$		0.9039 (0.0849)	0.9115 (0.0796)	0.9107 (0.0806)	0.9136 (0.0777)
log likelihood	363.1274	384.9890	387.1091	385.2617	388.2586
AIC	-720.2548	-757.9780	-760.2182	-756.5230	-760.5172
SIC	-711.4298	-740.3281	-739.6267	-735.9319	-736.9840
HQ	-716.6681	-750.8056	-751.8504	-748.1556	-750.9540

Figure 2: Probability of Regime 1 - univariate filter. (These are filtered probabilities conditional on information available at time t .)

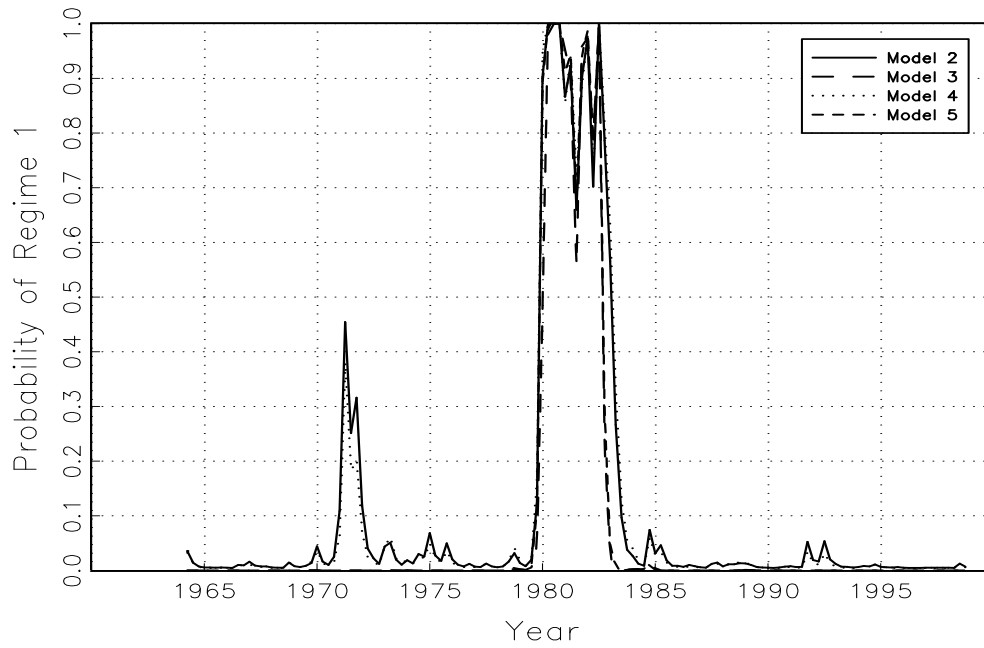


Figure 3: Actual Yields on 3-Month Bond



Table 4: Performance Results of Models

Maturity					
	6 Month	1 Year	2 Year	5 Year	Total
Mean Squared Error (MSE)					
Model 1	0.5306	1.0997	0.8339	0.1728	2.6370
Model 2	0.0614	0.2042	0.2986	0.1181	0.6823
Model 3	0.3592	1.2558	1.7106	0.4848	3.8104
Model 4	0.0704	0.2711	0.4625	0.2185	1.0225
Model 5	0.3635	1.2511	1.6432	0.4449	3.7027
Relative Mean Square Error (RMSE)					
Model 1	0.0237	0.0382	0.0238	0.0039	0.0896
Model 2	0.0011	0.0035	0.0047	0.0016	0.0109
Model 3	0.0167	0.0591	0.0667	0.0124	0.1549
Model 4	0.0014	0.0055	0.0086	0.0032	0.0187
Model 5	0.0166	0.0561	0.0613	0.0112	0.1452
Mean Absolute Error (MAE)					
Model 1	0.5960	0.8395	0.7163	0.3301	2.4819
Model 2	0.1498	0.2951	0.3858	0.2529	1.0836
Model 3	0.4736	0.9414	1.1535	0.6303	3.1988
Model 4	0.1828	0.4129	0.5902	0.4098	1.5957
Model 5	0.4836	0.9439	1.1258	0.6013	3.1546
Relative Mean Absolute Error (RMAE)					
Model 1	0.1083	0.1375	0.1082	0.0462	0.4002
Model 2	0.0221	0.0420	0.0515	0.0314	0.1470
Model 3	0.0915	0.1846	0.2106	0.0960	0.5827
Model 4	0.0281	0.0616	0.0837	0.0523	0.2257
Model 5	0.0926	0.1818	0.2025	0.0909	0.5678

Model 1: $dr(t) = \kappa[\alpha - r(t)]dt + \sigma\sqrt{r(t)}dW(t)$

Model 2: $dr(t) = \kappa[\alpha - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$

Model 3: $dr(t) = \kappa(X_t)[\alpha - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$

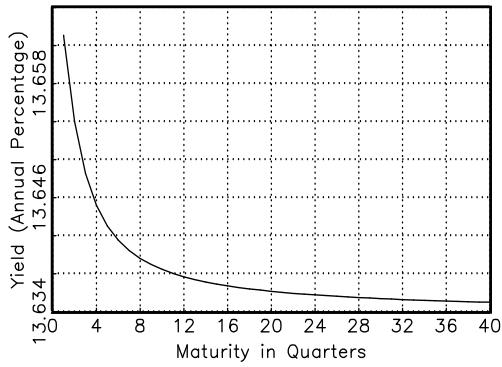
Model 4: $dr(t) = \kappa[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$

Model 5: $dr(t) = \kappa(X_t)[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$

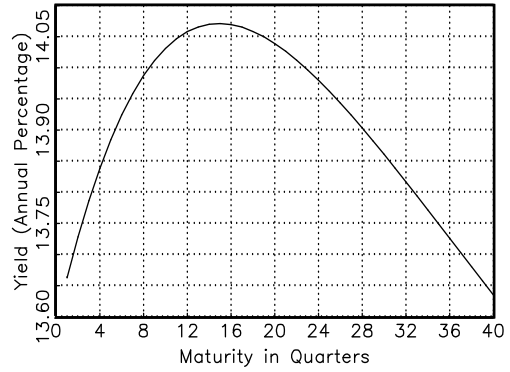
Figure 4: Model Generated Yield Curves

Date: 1981:1.

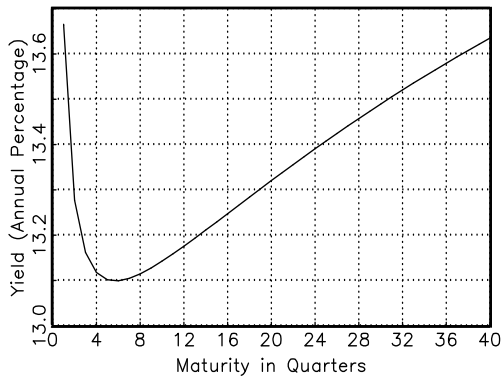
Model 1



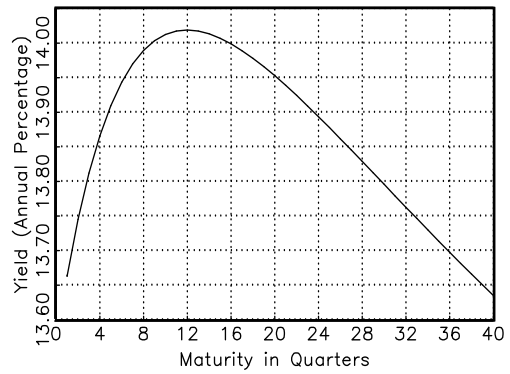
Model 2



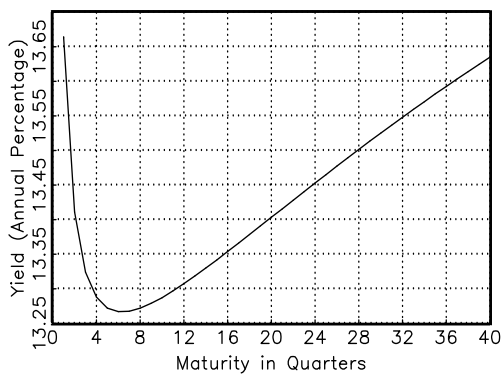
Model 3



Model 4



Model 5



Actual

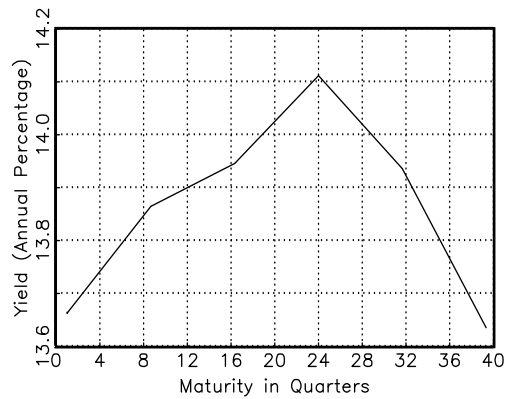
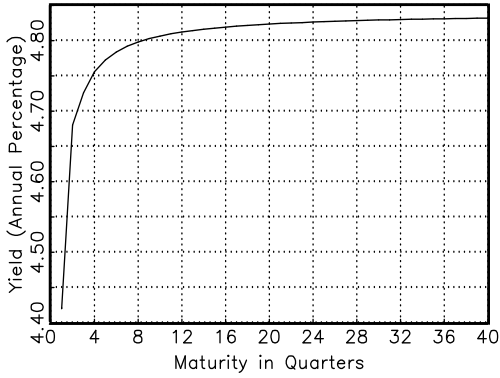


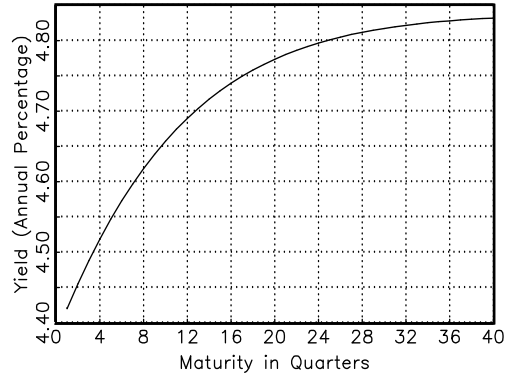
Figure 5: Model Generated Yield Curves

Date: 1989:3.

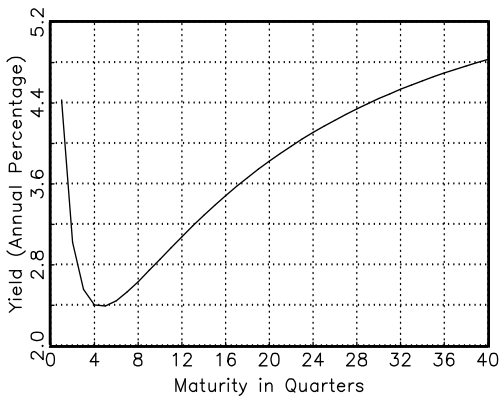
Model 1



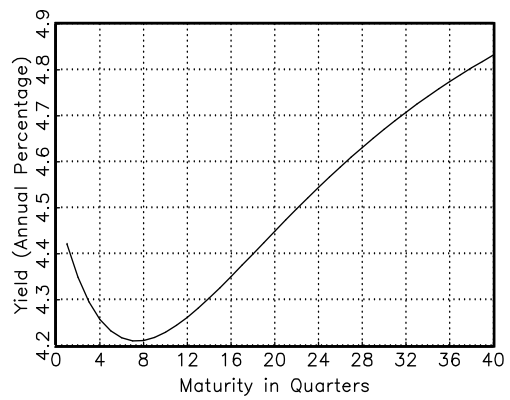
Model 2



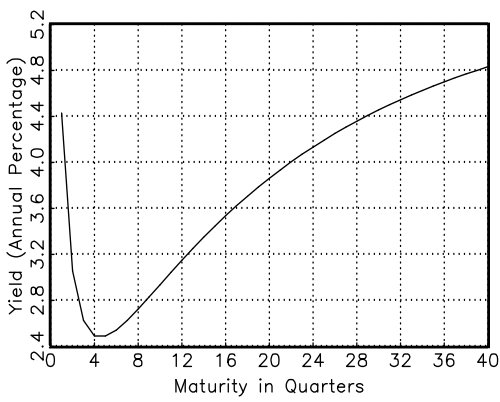
Model 3



Model 4



Model 5



Actual

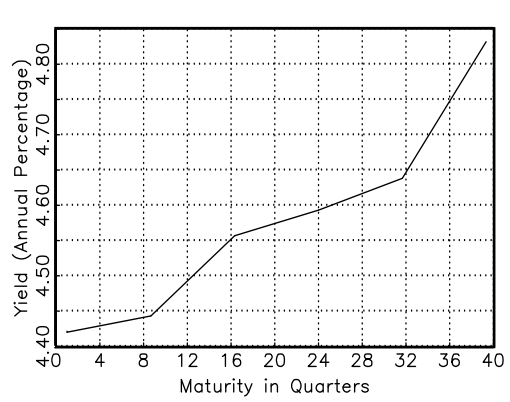
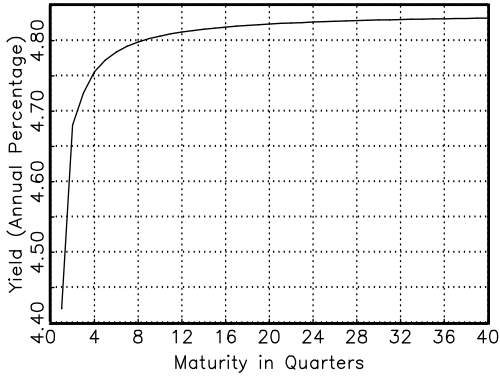


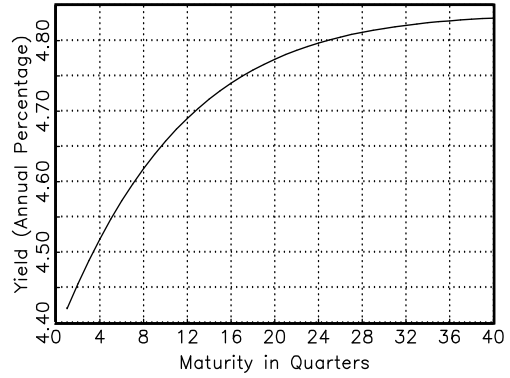
Figure 6: Model Generated Yield Curves

Date: 1998:4.

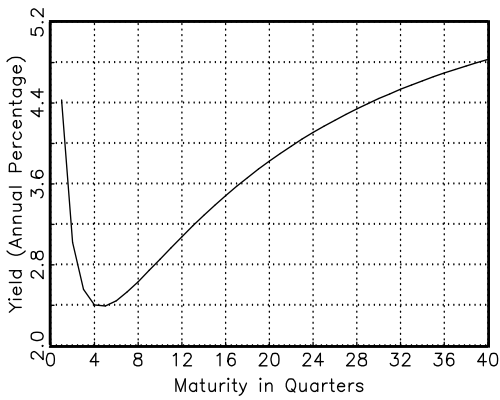
Model 1



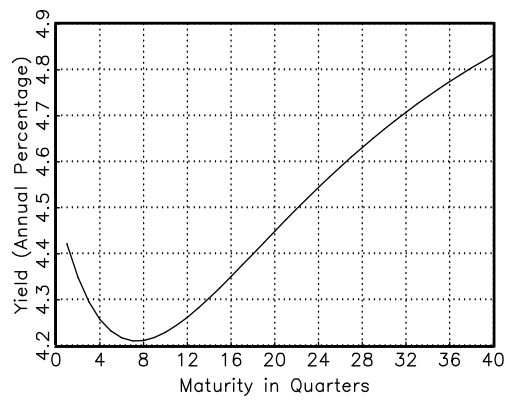
Model 2



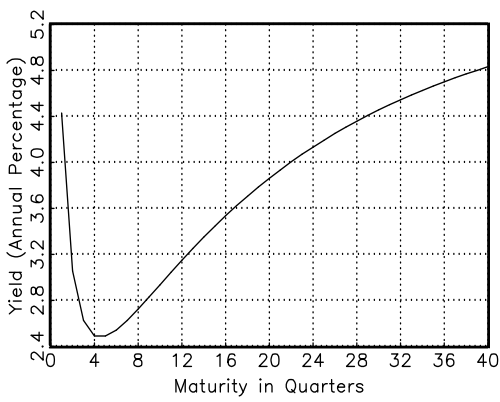
Model 3



Model 4



Model 5



Actual

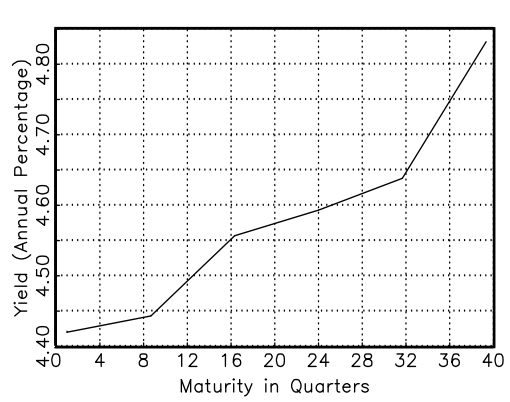


Figure 7: Actual and Model Generated Yields

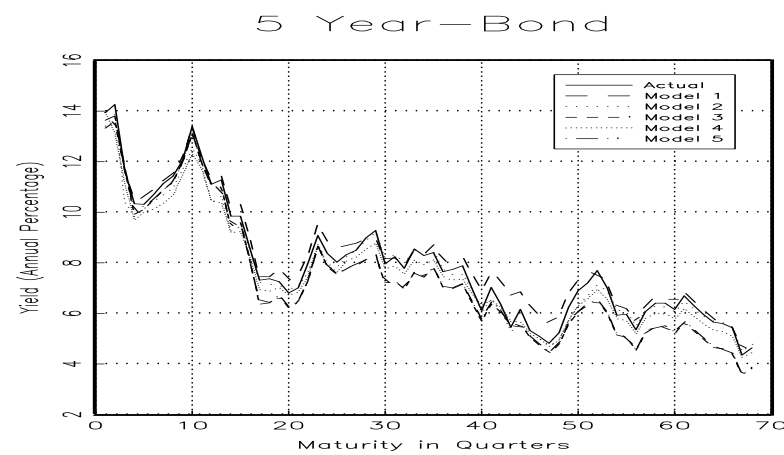
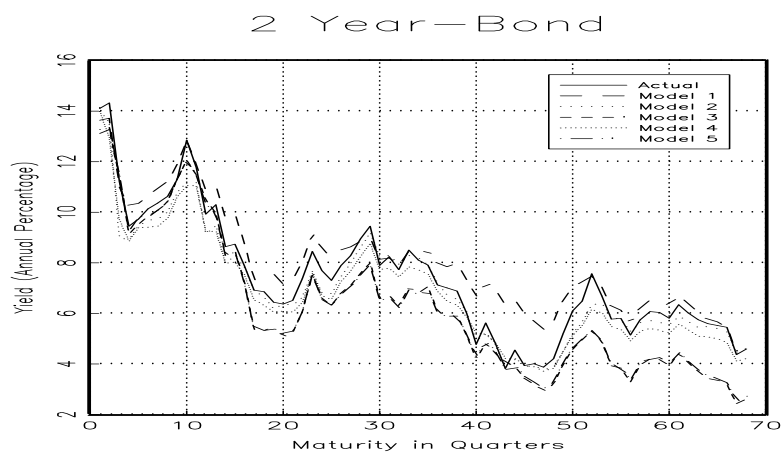
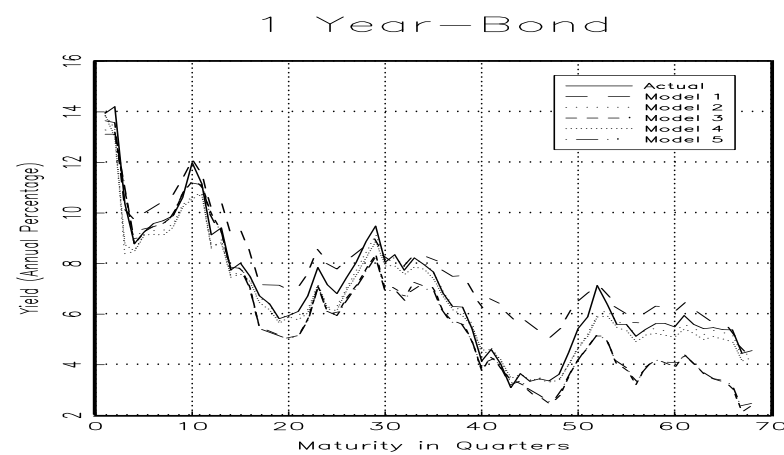
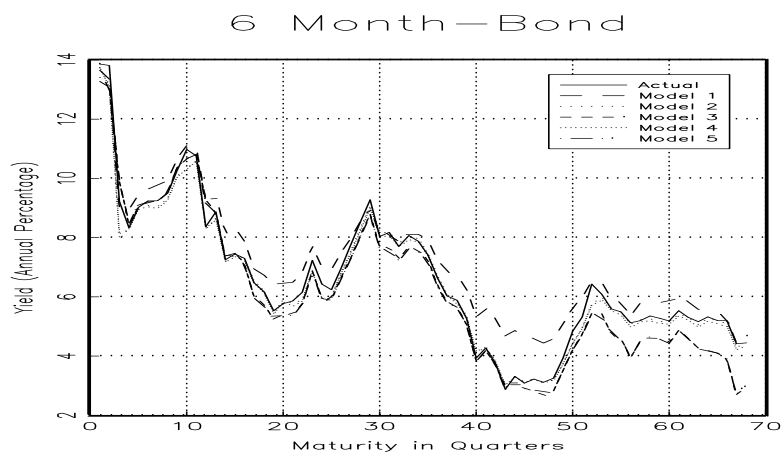


Table 5: Descriptive Statistics of Yield Differences

Maturity	Mean	Variance	Skewness	Kurtosis
Model 1: $dr(t) = \kappa[\alpha - r(t)]dt + \sigma\sqrt{r(t)}dW(t)$				
6 Month	0.5601	0.5306	1.5278	2.7184
1 Year	0.7827	1.0997	1.5763	2.9075
2 Year	0.6570	0.8339	1.6436	3.2214
5 Year	0.2871	0.1728	1.5960	3.2651
Model 2: $dr(t) = \kappa[\alpha - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$				
6 Month	-0.1135	0.0614	-0.0462	0.0466
1 Year	-0.2495	0.2042	-0.2381	0.3687
2 Year	-0.3492	0.2986	-0.3605	0.5860
5 Year	-0.1984	0.1181	-0.0776	0.0668
Model 3: $dr(t) = \kappa(X_t)[\alpha - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$				
6 Month	-0.4118	0.3592	-0.3205	0.4094
1 Year	-0.8881	1.2558	-1.9432	3.4012
2 Year	-1.1255	1.7106	-2.8431	5.0770
5 Year	-0.6273	0.4848	-0.4073	0.3610
Model 4: $dr(t) = \kappa[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$				
6 Month	-0.1659	0.0704	-0.0434	0.0354
1 Year	-0.3812	0.2711	-0.2459	0.2891
2 Year	-0.5654	0.4625	-0.4464	0.5269
5 Year	-0.3965	0.2185	-0.1404	0.1061
Model 5: $dr(t) = \kappa(X_t)[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t)$				
6 Month	-0.4264	0.3635	-0.3260	0.3926
1 Year	-0.8930	1.2511	-1.9041	3.2248
2 Year	-1.1031	1.6432	-2.6718	4.6273
5 Year	-0.5979	0.4449	-0.3614	0.3102

References

- Andersen, T. G. and J. Lund**, “Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate,” *Journal of Econometrics*, 1997, *77*, 343–377.
- Ang, Andrew and Geert Bekaert**, “Regime Switches in Interest Rates,” *Journal of Business and Economic Statistics*, 2002, *20* (2), 163–182.
- and —, “Short Rate Nonlinearities and Regime Switches,” *Journal of Economic Dynamics and Control*, 2002, *26* (7-8), 1243–1274.
- Backus, D., S. Foresi, and C. Telmer**, “Discrete-Time Models of Bond Pricing,” in Narasimhan Jagadeesh and Bruce Tuckman, eds., *Advanced Fixed-Income Valuation Tools*, Wiley & Sons, Inc, 2000, chapter 4, pp. 87–127.
- Bansal, Ravi and Hao Zhou**, “Term Structure of Interest Rates with Regime Shifts,” *Journal of Finance*, October 2002, *forthcoming*.
- Bliss, Robert R.**, “Testing Term Structure Estimation Methods,” *Advances in Futures and Options Research*, 1997, *9*, 197–231.
- Brenner, R. J., R. H. Harjes, and K. F. Kroner**, “Another Look at Models of the Short-Term Interest Rate,” *Journal of Financial and Quantitative Analysis*, 1996, *31* (1), 85–107.
- Cai, J.**, “A Markov Model of Switching-Regime ARCH,” *Journal of Business and Economic Statistics*, 1994, *12* (3), 309–316.
- Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders**, “An Empirical Comparison of Alternative Models of the Short-Term Interest Rate,” *Journal of Finance*, 1992, *47*, 1209–1227.
- Chapman, David A., Jr. John B. Long, and Neil D. Pearson**, “Using Proxies for the Short Rate: When are Three Months Like an Instant?,” *Review of Financial Studies*, 1997, *12* (4), 763–806.
- Cochrane, John H.**, *Asset Pricing*, Princeton University Press, 2001.
- Cox, J. C., Jonathan E. Ingersoll, and S. A. Ross**, “A Theory of the Term Structure of Interest Rates,” *Econometrica*, 1985, *53*, 385–407.

- Dahlquist, Magnus and Stephen F. Gray**, “Regime-Switching and Interest Rates in the European Monetary System,” *Journal of International Economics*, 2000, 50, 399–419.
- Dai, Q. and K. Singleton**, “Specification Analysis of Affine Term Structure Models,” *Journal of Finance*, 2000, 50 (5), 1943–1978.
- and —, “Expectation Puzzles, Time-varying Risk Premia, and Dynamic Models of the Term Structure,” *NBER Working Paper*, 2001, 8167.
- Duarte, Jefferson**, “The Relevance of the Price of Risk in Affine Term Structure Models,” *University of Chicago*, 2000, *Mimeo*.
- Duffee, Gregory R.**, “Term Premia and Interest Rate Forecasts in Affine Models,” *Journal of Finance*, 2002, 57, 405–443.
- Duffie, D. and R. Kan**, “A Yield-Factor Model of Interest Rates,” *Mathematical Finance*, 1996, 6 (4), 379–406.
- Elliott, Robert. J., William C. Hunter, and Barbara M. Jamieson**, “Financial Signal Processing: A Self Calibration Model,” *Federal Reserve Bank of Chicago*, 2000, *Working Paper* (21).
- Evans, Martin**, “Regime Shifts, Risk and the Term Structure,” *Department of Economics, Georgetown University*, 1998, *Working Paper*.
- Fama, E. and R. Bliss**, “The Information in Long-Maturity Forward Rates,” *American Economic Review*, September 1987, 77, 680–692.
- Fisher, Mark and Christian Gilles**, “Modeling the State-Price Deflator and the Term Structure of Interest Rates,” *Federal Reserve Bank of Atlanta*, 2000, *Mimeo*.
- Garcia, R. and P. Perron**, “An Analysis of the Real Interest Rate Under Regime Shifts,” *Review of Economic Studies*, 1996, 78, 111–125.
- Ghysels, Eric and Serena Ng**, “A Semi-Parametric Factor Model of Interest Rates and Tests of the Affine Term Structure,” *Review of Economics and Statistics*, 1998, 80, 535–548.
- Gray, Stephen F.**, “Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process,” *Journal of Financial Economics*, 1996, 42, 27–62.

- Hamilton, James D.**, “Rational Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates,” *Journal of Economic Dynamics and Control*, 1988, 12, 385–423.
- Hull, John**, *Options, Futures and Other Derivative Securities*, 2 ed., Prentice Hall, 2000.
- Landen, C.**, “Bond Pricing in a Hidden Markov Model of the Short Rate,” *Finance and Stochastics*, 2000, 4 (4), 371–390.
- McCulloch, J. Huston and Heon-Chul Kwon**, “US Term Structure Data, 1947-1991,” *Ohio State University*, March 1993, *Working Paper*, 93-6.
- Naik, Vasant and Moon Hoe Lee**, “Yield Curve Dynamics with Discrete Shifts in Economic Regimes,” *Faculty of Commerce, University of British Columbia*, 1997, *Working Paper*.
- Nowman, Ben**, “Gaussian Estimation of Single-Factor Continuous Time Models of the Term Structure of Interest Rates,” *Journal of Finance*, 1997, 52 (4), 1695–1706.
- Pearson, N. and T. Sun**, “Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model,” *Journal of Finance*, 1994, 49 (4), 1279–1304.
- Sola, Martin and John Driffill**, “Testing the Term Structure of Interest Rates from a Stationary Switching Regime VAR,” *Journal of Economic Dynamics and Control*, 1994, 18, 601–628.
- Sun, Tong-Seng**, “Real and Nominal Interest Rates: A Discrete-Time Model and Its Continuous-Time Limit,” *Review of Financial Studies*, 1992, 5 (4), 581–611.

A Appendix A: Solution

A.1 Model 1: The Single Regime Cox-Ingersoll-Ross Model

Stochastic processes for two state variables (stochastic discount factor and short rate) are given

$$\frac{d\Lambda(t)}{\Lambda(t)} = -r(t)dt - \sigma_\Lambda \sqrt{r(t)}dW(t) \quad (\text{A. 1})$$

$$dr(t) = \kappa[\alpha - r(t)]dt + \sigma \sqrt{r(t)}dW(t). \quad (\text{A. 2})$$

The fundamental bond pricing equation (10) becomes

$$\kappa[\alpha - r]F_r + \frac{1}{2}\sigma^2 r F_{rr} - F_T - rF = F_r \sigma \sigma_\Lambda r. \quad (\text{A. 3})$$

Using the following affine functional form for bond prices

$$F(t, r(t), T) = e^{A(t, T) - B(t, T)r},$$

we obtain the partial derivatives required in (A. 3). Substituting them into (A. 3) and separating the coefficients on the constant and on the terms in r result in a set of *ordinary* differential equations for $A(t, T)$ and $B(t, T)$

$$\begin{aligned} B'(t, T) &= 1 - \frac{1}{2}\sigma^2 B(t, T)^2 - (\sigma\sigma_\Lambda + \kappa)B(t, T), \\ A'(t, T) &= -B(t, T)\kappa\alpha. \end{aligned} \quad (\text{A. 4})$$

One can solve them, subject to the boundary condition imposed by $F(T, r(t), T) = 1$, which implies $A(T, T) = 0$ and $B(T, T) = 0$,

$$\begin{aligned} B(t, T) &= \frac{2(1 - e^{\gamma\tau})}{(\gamma + \kappa + \sigma\sigma_\Lambda)(e^{\gamma\tau} - 1) + 2\gamma}, \\ A(t, T) &= \frac{\kappa\alpha}{\sigma^2} \left(2 \ln \left(\frac{2\gamma}{\psi(e^{\gamma\tau} - 1) + 2\gamma} \right) + \psi\tau \right), \end{aligned}$$

where

$$\gamma = \sqrt{(\kappa + \sigma\sigma_\Lambda)^2 + 2\sigma^2},$$

$$\psi = \kappa + \sigma\sigma_\Lambda + \gamma,$$

$$\tau = T - t.$$

A.2 Models 2-5: The Cox-Ingersoll-Ross Model with Regime Switching

Stochastic processes for two state variables (stochastic discount factor and short rate) are given

$$\frac{d\Lambda(t)}{\Lambda(t)} = -r(t)dt - \sigma_\Lambda \sqrt{r(t)}dW(t) \quad (\text{A. 5})$$

$$dr(t) = \kappa(X_t)[\alpha(X_t) - r(t)]dt + \sigma(X_t)\sqrt{r(t)}dW(t) \quad X = 0, 1. \quad (\text{A. 6})$$

The ‘‘term structure equation’’ is now given by

$$(\kappa(X_t)[\alpha(X_t) - r] - \sigma(X_t)\lambda r)F_r + \frac{1}{2}\sigma(X_t)^2 r F_{rr} - F_T + \sum_{i=0}^1 \sum_{j=0}^1 h_{ij} \Delta F - rF = 0, \quad (\text{A. 7})$$

with the boundary condition

$$F(T, X, r, T) = 1.$$

Using the following affine functional form for bond prices

$$F(t, X, r(t), T) = e^{A(t, X, T) - B(t, X, T)r},$$

we obtain the partial derivatives required in (A. 7). Substituting them into (A. 7) and separating the coefficients on the constant and on the terms in r once again result in a set of *ordinary* differential equations for $A(t, T)$ and $B(t, T)$,

$$\left\{ \begin{array}{l} B_{0t}(t, T) - \kappa_0 B_0(t, T) - \frac{1}{2}\sigma_0 B_0(t, T)^2 + 1 \\ B_{1t}(t, T) - \kappa_1 B_1(t, T) - \frac{1}{2}\sigma_1 B_1(t, T)^2 + 1 \end{array} \right\} r + \left\{ \begin{array}{l} A_{0t}(t, T) - (\kappa_0 \alpha_0 - \lambda_0 \sigma_0) B_0(t, T) + \sum_{j=0}^1 h_{0j} \Delta F_0 \\ A_{1t}(t, T) - (\kappa_1 \alpha_1 - \lambda_1 \sigma_1) B_1(t, T) + \sum_{j=0}^1 h_{1j} \Delta F_1 \end{array} \right\} = 0 \quad (\text{A. 8})$$

where

$$\Delta F_0 = e^{A_j(t, T) - r B_j(t, T) - [A_0(t, T) - r B_0(t, T)]},$$

$$\Delta F_1 = e^{A_j(t, T) - r B_j(t, T) - [A_1(t, T) - r B_1(t, T)]}.$$

Equations system (A. 8) may be solved by applying the commonly adopted log-linear approximation $e^y - 1 \approx y$ [see, for example, Bansal and Zhou (2002)].

$$\left\{ \begin{array}{l} B_{0t}(t, T) - \kappa_0 B_0(t, T) - \frac{1}{2} \sigma_0 B_0(t, T)^2 + h_{01}[B_0(t, T) - B_1(t, T)] + 1 \\ B_{1t}(t, T) - \kappa_1 B_1(t, T) - \frac{1}{2} \sigma_1 B_1(t, T)^2 + h_{10}[B_1(t, T) - B_0(t, T)] + 1 \end{array} \right\} r +$$

$$\left\{ \begin{array}{l} A_{0t}(t, T) - (\kappa_0 \alpha_0 - \lambda_0 \sigma_0) B_0(t, T) + h_{01}[A_1(t, T) - A_0(t, T)] \\ A_{1t}(t, T) - (\kappa_1 \alpha_1 - \lambda_1 \sigma_1) B_1(t, T) + h_{10}[A_0(t, T) - A_1(t, T)] \end{array} \right\} = 0.$$

(A. 9)

Unfortunately, the above equations system (A. 9) has only an approximate numerical solution. One way to improve the approximation is to use a technique similar to the *control variate technique* used as a variance reduction procedures in the option pricing literature [see Hull (2000)]. This involves calculating the bond pricing equation for the single regime CIR model (A. 3) using the approximation adopted in obtaining (A. 9). The difference between two gives the approximation error.