Yield-Factor Volatility Models

Christophe Pérignon* Daniel R. Smith*

First Draft: May 25, 2004
This Daft: October 4, 2005

Abstract

The term structure of interest rates is often summarized using a handful of yield factors that capture shifts in the shape of the yield curve. In this paper, we develop a comprehensive model for volatility dynamics in the level, slope, and curvature of the yield curve that simultaneously includes level and GARCH effects along with regime shifts. We show that the level of the short-rate is useful in modeling the volatility of the three yield factors and that there are significant GARCH effects present even after including a level effect. Furthermore, we find that allowing for regime shifts in the factor volatilities dramatically improves the model's fit and strengthens the level effect. Our empirical results have some important implications for the estimation of term structure models. We argue that the Efficient Method of Moment, which is currently the standard estimation technique in the term structure literature, is inconsistent with the main features of the U.S. yield curve that are identified by our model.

JEL Code: E43, C32, C51

---

[*] Faculty of Business Administration, Simon Fraser University, Canada. We thank Avi Bick, Peter Christoffersen, Jefferson Duarte, Eric Ghysels, Peter Klein, Andrey Pavlov, Elvezio Ronchetti, seminar participants at Simon Fraser University, University of Geneva, University of New South Wales, and University of Technology Sydney, and participants at the 2004 Australasian Meeting of the Econometric Society and Northern Finance Association 2004 Conference for their comments. We would also like to thank Robert Bliss for providing us with data and his programs for constructing yield curves. We are responsible for all errors and omissions. We gratefully acknowledge financial support from the Social Sciences and Research Council of Canada. Emails: cperigno@sfu.ca (Pérignon); drsmith@sfu.ca (Smith). The corresponding author is Daniel Smith (Tel.: +1-604-291 4675).

Yield-Factor Volatility Models

The term structure of interest rates is often summarized using a handful of yield factors that capture shifts in the shape of the yield curve. In this paper, we develop a comprehensive model for volatility dynamics in the level, slope, and curvature of the yield curve that simultaneously includes level and GARCH effects along with regime shifts. We show that the level of the short-rate is useful in modeling the volatility of the three yield factors and that there are significant GARCH effects present even after including a level effect. Furthermore, we find that allowing for regime shifts in the factor volatilities dramatically improves the model's fit and strengthens the level effect. Our empirical results have some important implications for the estimation of term structure models. We argue that the Efficient Method of Moment, which is currently the standard estimation technique in the term structure literature, is inconsistent with the main features of the U.S. yield curve that are identified by our model.

I Introduction

The term structure of interest rates is often summarized using a handful of yield factors that capture shifts in the shape of the yield curve, i.e., changes in the overall level, slope, and curvature of the yield curve (see Litterman and Scheinkman, 1991). This factor decomposition provides a parsimonious representation of the term structure and is extensively used in fixed-income derivative pricing (Driessen, Klaassen and Melenberg, 2003), to model the linkages between interest rates and macroeconomic variables (see Ang and Piazzesi, 2003), or to estimate term structure models (Brandt and Chapman, 2003). Despite this wide application in financial economics, very little is known about the volatility of these factors. In this paper, we study the dynamics of yield-factor conditional volatility.

The key features of our econometric model are inspired by early models of the short-term interest rates. The main conclusions from this literature is that a level effect, in which the volatility is a positive function of the level of interest rates, GARCH effects, and regime shifts are required to adequately model the short-rate volatility.[1] The dependence of the interest rate volatility on the level of the short-rate is systematically studied in Chan et al. (1992, hereafter CKLS). The first model that combines both level and GARCH effects for the short-rate volatility is proposed in Longstaff and Schwartz (1992). Furthermore, Brenner, Harjes and Kroner (1996) show that models that include both GARCH and level effects are better able to predict volatility than models that only include one of these effects. Gray (1996) extends the GARCH-level model to allow for multiple regimes in the short-rate volatility and concludes that all three effects are needed to adequately model interest rate volatility.

Less research has been devoted to understanding the joint-dynamics of the level, slope, and curvature of the yield curve, which we call the yield factors. Pérignon and Villa (2004) show that the volatility of the yield factors is the primary source of time variation in the covariance matrix of interest rates. The role of conditional heteroscedasticity in the dynamics of the volatility of the yield factors has been highlighted by Christiansen and Lund (2002) and Christiansen (2004). The empirical challenge of estimating a level effect in a multi-factor framework has been tackled by Boudoukh, Richardson, Stanton and Whitelaw (1998), Brandt and Chapman (2003), and Christiansen (2005). Another strand of research examines

---

[1] See Chapman and Pearson (2001) for a survey of recent progress in the estimation of term structure models.

the role of regime shifts in the dynamics of yield-factor volatilities. Using international data, Kugler (1996) and Ang and Bekaert (2002a,b) estimate a two-state regime-switching VAR model of the level and the slope factors with a constant covariance matrix in each regime (i.e., without any level nor GARCH effects). Christiansen (2004) extends the latter approach by fitting a two-state regime-switching ARCH model to the U.S. level and the slope factors. A broad conclusion of this research is that regime shifts are a central feature of yield-factor volatilities.

Our primary contribution is the development of a comprehensive model for yield factors that simultaneously includes level and GARCH effects along with regime shifts in the factor volatilities. There are two main motivations for modeling the variance of the three yield factors as a function of the short term interest rate. The first one is the empirical observation that the volatility of all three yield factors tends to be higher when short-term interest rates are higher. The second motivation is theoretical. We demonstrate that within the class of affine term structure models in which the volatility of all state variables is determined by a single state variable, i.e. the $\mathbb{A}_1(3)$ class in the terminology of Dai and Singleton (2000), both the level of the short-term interest rate and the variance of all three yield factors are determined by this very same state variable. As a result, we model the variance of all three yield factors using a level effect in which conditional volatility is a nonlinear function of the level of short-term interest rates. In our level effect, the short-rate serves as a proxy for the unobserved latent volatility state variable. Furthermore, we endeavor to combine both GARCH and level effects, while maintaining the traditional interpretation that a GARCH(1,1) model implies an ARMA(1,1) representation for the squared residual. In addition, our model explicitly includes regime shifts, a feature which has been demonstrated to be important in fitting short-term interest rates. Each model we consider is nested within this encompassing model, so we are able to directly measure the marginal contribution of each component of the model.

The estimation of our econometric model leads to the following conclusions. Using monthly U.S. bond yields over the 1970-2002 period, we show that all three yield factors display a significant level effect. In particular, we find that the level effect for the slope factor is better captured by the overall level of interest rates rather than by the level of the slope factor. A similar conclusion is reached for the curvature factor. Furthermore, the three yield factors exhibit strong GARCH effects even after including a level effect. We also examine regime-switching models that recognize different regimes in the volatility of the yield factors. We find that allowing for regime shifts dramatically improves the model's fit and strengthens the level effect.

Our empirical results have some important implications for the estimation of term structure models. We argue that the Efficient Method of Moment, which is currently the standard estimation technique in the term structure literature, is inconsistent with the main features of the U.S. yield curve that are identified by our model. Consequently, EMM remains silent on which term structure model is (un)able to generate the kind of GARCH and level effects or regime shifts observed in the real data.

The remainder of the paper proceeds as follows. Section II details the model, section III describes the data, section IV presents the empirical results, section V discusses the main implications of our results for the estimation of term structure models, and section VI offers some concluding comments.

II Model Development

A Single-Regime Models

We are interested in modeling the dynamics of the following three yield factors, the level of interest rates (L), the slope of the yield curve (S), and the curvature of the yield curve (C). Denote by F_{it} the i^{th} yield factor, with $i = L, S, C$, whose dynamics is modeled as:

$$dF_{it} = (a_i + b_i \cdot F_{it})dt + \sigma_{it} F_{jt}^{\gamma_i} dW_{it} \tag{1}$$

where W_{it} is a standard Brownian motion and W_{Lt}, W_{St}, and W_{Ct} may be correlated. This model is inspired by early models of the short-rate which include both mean reversion and allow the conditional volatility to be a function of the level of the short-rate (see among others Cox, Ingersoll and Ross, 1985). When allowing for a level effect in a multi-factor framework, one can either model the residual volatility of a given factor as a function of the value of this very factor ($j = i$) or of the level factor ($j = L$). We argue in this paper that the approach based on the level factor should be preferred. First, the own-level approach appears rather inconsistent since the residual standard-deviation becomes negative when the value of the factor (slope or curvature) is negative. Second, given that the level effect is so important in modeling volatility of short-term interest rates, it is likely to also affect the volatilities of the slope and curvature of the yield curve. As will be seen below, this conjecture is born out by the data. Third, the approach based on the level factor is consistent with the popular affine model with N state variables, among which one drives the state variable volatilities. Indeed, if the yield factors are defined as linear combinations of bond yields, as they are in the literature, the conditional volatility of all the yield factors depends on the same single source of risk (see Appendix A for a proof).

In our empirical work, we discretize the process in Equ. (1) as:

$$\Delta F_{it} = \alpha_{0i} + \alpha_{1i} \cdot F_{it-1} + e_{it} \tag{2}$$

for $i = L, S, C$ and $t = 1, ..., T$. We approximate dW_{it}, which is normally distributed with variance dt, by a normally distributed innovation e_{it}. We decompose the conditional volatility of e_{it} into the product of two terms, $E(e_{it}^2|\psi_{t-1}) = \delta_{it}^2 = \sigma_{it}^2 F_{jt-1}^{2\gamma_i}$ with ψ_{t-1} denoting the information set at time $t-1$. This specification allows heteroscedasticity to enter through a time-varying coefficient σ_{it}^2, which depends on past shocks on the residuals factors, and through the level effect.

Alternatively, the residual in Equ. (2) can be written as $e_{it} = \sigma_{it} F_{jt-1}^{\gamma_i} z_{it}$ where z_{it} is i.i.d. $N(0, 1)$. In this modeling, σ_{it}^2 is the volatility of the scaled residual $v_{it} = e_{it}/F_{jt-1}^{\gamma_i} = \sigma_{it} z_{it}$ and is modeled as a GARCH process:

$$\sigma_{it}^2 = \beta_{0i} + \beta_{1i} \cdot v_{it-1}^2 + \beta_{2i} \cdot \sigma_{it-1}^2. \tag{3}$$

Our specification for σ_{it}^2 differs from previous processes proposed in the literature. For instance, in a univariate setting, Brenner, Harjes and Kroner (1996) use a standard GARCH model for the residual volatility of the short-rate:

$$\sigma_t^2 = \beta_0 + \beta_1 \cdot e_{t-1}^2 + \beta_2 \cdot \sigma_{t-1}^2. \tag{4}$$

Alternatively, Longstaff and Schwartz (1992), Brenner, Harjes and Kroner (1996), Gray (1996), and Hamilton and Kim (2002) add the level term directly to the GARCH model:

$$\sigma_t^2 = \beta_0 + \beta_1 \cdot e_{t-1}^2 + \beta_2 \cdot \sigma_{t-1}^2 + \beta_3 \cdot F_{Lt-1}^{2\gamma} \tag{5}$$

where γ is either fixed or estimated. It is well known that "*if* u_t *is described by a GARCH*(r, m) *process, then* u_t^2 *follows an ARMA*(p, r) *process, where* p *is the larger of* r *and* m" (Hamilton, 1994, page 666). We maintain this interpretation of the GARCH model when including a level effect. Indeed, we assume that v_{it}^2 evolves as an ARMA(1,1) process yielding a GARCH model for the scaled residual v_{it}. This model can be compared with the stochastic volatility literature (see Andersen and Lund, 1997 and Ball and Torous, 1999) where the conditional volatility of the short-rate is modeled as $\sigma_t^2 F_{t-1}^{2\gamma}$ and the conditional volatility of the scaled residual σ_t^2 follows an autoregressive process. Here, we also model the conditional volatility of the scaled residual but using a GARCH model. Although there is nothing wrong with Equ. (4) and (5) as empirical models, they are somewhat ad hoc extensions of the GARCH model and are inconsistent with this traditional interpretation of the GARCH model.

We assume that ΔF_t is a tri-dimensional vector of the changes in yield factors with conditional mean vector $\mu_t = E(\Delta F_t|\psi_{t-1})$ and conditional covariance matrix $\Sigma_t = H_t^{1/2} \rho H_t^{1/2}$, where ρ is a (3×3) conditional correlation matrix and H_t is a (3×3) diagonal matrix with conditional volatility of the i^{th} factor on the i^{th} element of the principal diagonal (see Bollerslev, 1990). We estimate the parameter vector using quasi-maximum likelihood, where lnL=$\sum_{t=1}^{T}$ln $f(e_t|\psi_{t-1})$ is the quasi-loglikelihood function, and $f(e_t|\psi_{t-1})$ is the probability density function of the multivariate normal density with mean 0 and covariance matrix $\Sigma_{t|t-1}$. The initial observation is assumed to be drawn from the unconditional distribution of ΔF_t.

Many classical models for interest rates are nested in the model derived above. Firstly, a multivariate homoscedastic-AR(1) model, labelled as the NO GARCH-NO LEVEL model, can be derived by assuming that the residual volatility of each factor is constant through time ($\beta_{ki} = 0$, $i = L, S, C$ and $k = 1, 2$, and $\gamma_i = 0$, $i = L, S, C$). Secondly, a multivariate version of the CKLS model, which is called the LEVEL model, is obtained by assuming that σ_{it}^2 is constant. In the latter model, the volatility remains time-varying but depends solely on the level of the factor ($\beta_{ki} = 0$, $i = L, S, C$ and $k = 1, 2$). Thirdly, a multi-factor model, labelled as the GARCH model, allows σ_{it}^2 to follow a GARCH process but does not permit volatility to be a function of the level of the factors ($\gamma_i = 0$, $i = L, S, C$). Finally, the unrestricted version of the model, which is referred to as the GARCH-LEVEL model, permits both the coefficient σ_{it}^2 to vary through time as new information arrives and the residual volatility to depend on the level of the factors.

B Regime-Switching Models

We extend our basic model to allow for different regimes in the volatility of the yield factors. There are economic reasons to believe that the entire yield curve is subject to regime shifts caused by transitions between periods of economic expansion and contraction. This approach is also motivated by the extensive empirical literature suggesting that regime-switching models describe historical interest rates better than single-regime models (see Hamilton, 1988, Gray, 1996 and Smith, 2002). Further, Ang and Bekaert (2002a,b), Bansal and Zhou (2002), and Dai, Singleton and Yang (2003) show that regime shifts are also important in capturing the dynamics of interest rates using multi-factor term structure models.

We denote by S_t the random state of the world at time t which can take two values, $s_t = \{1, 2\}$, where 1 denotes the "high-volatility regime" and 2 the "low-volatility regime". We assume that these regimes are common to the level, slope, and curvature of the yield curve. This assumption is primarily to keep the state space parsimonious[2], but it also seems more reasonable to assume that the state of the economy would affect all characteristics of the yield curve jointly rather than only affecting short-term interest rates without an effect on the slope and curvature. Furthermore, most term structure models assume that the entire yield curve be priced with the same underlying state variables, which would demand a common regime. Finally, a simple perusal of Figure 2 indicates that the interesting high-volatility episodes of one series appear also in the other two series. To also keep the model simple we allow only the unconditional mean and volatility of each series to be state-dependent. The conditional mean is given by:

$$\Delta F_{it} = \alpha_{0is_t} + \alpha_{1i} \cdot F_{it-1} + e_{it} \tag{6}$$

and the conditional volatility of the scaled residual is:

$$\hat{\sigma}^2_{it|s_t,s_{t-1}} = \beta_{0is_t} + \beta_{1i} \cdot \hat{v}^2_{it-1} + \beta_{2i} \cdot \hat{\sigma}^2_{it-1|s_{t-1}}. \tag{7}$$

This specification of the GARCH model follows Dueker (1997) by defining $\hat{\sigma}^2_{it|s_t,s_{t-1}} = E(v_t^2 | S_t = s_t, S_{t-1} = s_{t-1}, \psi_{t-1})$ and:

$$\hat{v}_{it-1} = \sum_{s_{t-1},s_{t-2}=1}^{2} P(S_{t-1} = s_{t-1}, S_{t-2} = s_{t-2} | \psi_{t-1}) v_{it|s_{t-1}}. \tag{8}$$

---

[2] Without this assumption, the state space would enlarge to $2^3 = 8$ regimes. A similar assumption is made by Kugler (1996), Ang and Bekaert (2002a,b), and Christiansen (2004).

Note that for any time point t, the conditional volatility depends only on the regimes in the current period and in the previous period. The dependence of lagged volatility on states in previous periods is integrated out by substituting the entire path dependent $\sigma^2_{it-1|s_{t-1},s_{t-2},\ldots}$ with $\hat{\sigma}^2_{it-1|s_{t-1}} = E(v^2_{t-1}|S_{t-1} = s_{t-1}, \psi_{t-1})$:

$$\hat{\sigma}^2_{it|s_t} = \sum_{s_{t-1}=1}^{2} P(S_{t-1} = s_{t-1}|S_t = s_t, \psi_{t-1})\sigma^2_{it|s_t,s_{t-1}} \tag{9}$$

$$P(S_{t-1} = s_{t-1}|S_t = s_t, \psi_{t-1}) = \frac{P(S_t = s_t, S_{t-1} = s_{t-1}|\psi_{t-1})}{\sum_{s_{t-1}=1}^{2} P(S_t = s_t, S_{t-1} = s_{t-1}|\psi_{t-1})}. \tag{10}$$

The transition between the two latent states is modelled as a first-order Markov process with constant transition probabilities.[3] We define $\xi_{t|t-1}$ as:

$$\xi_{t|\tau} = \begin{bmatrix} P(S_t = 1, S_{t-1} = 1|\psi_\tau) \\ P(S_t = 1, S_{t-1} = 2|\psi_\tau) \\ P(S_t = 2, S_{t-1} = 1|\psi_\tau) \\ P(S_t = 2, S_{t-1} = 2|\psi_\tau) \end{bmatrix} \tag{11}$$

with ψ_τ denoting three possible information sets. For $\tau = t-1$, we get the forecast probabilities, which are used to construct the loglikelihood function; for $\tau = t$, we get the filtered probabilities, which are a product of the updating algorithm; for $\tau = T$, we get the full-sample smoothed probabilities, which use all information and are helpful when making inference regarding states. The transition matrix from one point to another is given by:

$$\xi_{t|t-1} = P\xi_{t-1|t-1}. \tag{12}$$

Imbedded in this formula is that the previous regime is integrated out at each point in time. The transition matrix P is given by:

$$P = \begin{bmatrix} p & p & 0 & 0 \\ 0 & 0 & 1-q & 1-q \\ 1-p & 1-p & 0 & 0 \\ 0 & 0 & q & q \end{bmatrix} \tag{13}$$

where p is $P(S_t = 1, S_{t-1} = 1)$ and q is $P(S_t = 2, S_{t-1} = 2)$. We follow Hamilton (1994) and set the initial probability vector $\xi_{1|0}$ to the ergodic steady state probabilities.

---

[3] Alternatively, the transition probabilities may depend on the level of interest rates (see Gray, 1996). However, as acknowledged by Ang and Bekaert (2002a, p. 172), multi-factor models with time-varying transition probabilities are likely to be overparameterized, which leads to many insignificant coefficients in the probability terms.

The conditional density $f(\Delta F_t|S_t = s_t, S_{t-1} = s_{t-1}, \psi_{t-1})$ is a multivariate normal density with conditional mean $\mu_{t|s_t,s_{t-1},t-1} = \left\{\Delta\hat{F}_{it}\right\}_{i=L,S,C}$ and conditional covariance matrix $\hat{\Sigma}_{t|s_t,s_{t-1},t-1} = \hat{H}^{1/2}_{t|s_t,s_{t-1}}\rho\hat{H}^{1/2}_{t|s_t,s_{t-1}}$.

Although the states are latent, the forecast probabilities can be used to calculate the joint density of ΔF_t and the states as:

$$f(\Delta F_t, S_t = s_t, S_{t-1} = s_{t-1}|\psi_{t-1}) =$$

$$f(\Delta F_t|S_t = s_t, S_{t-1} = s_{t-1}, \psi_{t-1}) \times P(S_t = s_t, S_{t-1} = s_{t-1}|\psi_{t-1}). \tag{14}$$

The marginal density of ΔF_t is found by integrating the joint density of ΔF_t over all possible states and is given by:

$$f(\Delta F_t|\psi_{t-1}) = \sum_{s_t,s_{t-1}=1}^{2} f(\Delta F_t, S_t = s_t, S_{t-1} = s_{t-1}|\psi_{t-1}). \tag{15}$$

The loglikelihood function is calculated as lnL= $\sum_{t=1}^{T} \log f(\Delta F_t|\psi_{t-1})$ and is maximized to estimate the parameters. Finally the updated filter probabilities of the latent states (the appropriate elements of $\xi_{t|t}$) can be obtained using the definition of the conditional probability:

$$P(S_t = s_t, S_{t-1} = s_{t-1}|\psi_t) = \frac{f(\Delta F_t, S_t = s_t, S_{t-1} = s_{t-1}|\psi_{t-1})}{f(\Delta F_t|\psi_{t-1})}. \tag{16}$$

The various models fitted in this section recognize diverse sources of conditional heteroscedasticity in the yield factors:

- In the RS-NO GARCH-NO LEVEL model, conditional heteroscedasticity can only be driven by switches between regimes.
- In the RS-LEVEL model, conditional heteroscedasticity comes from either time-variation in the level of interest rates or from switches between regimes.
- In the RS-GARCH model, conditional heteroscedasticity is driven by serial correlation in volatility or by switches between regimes.
- In the RS-GARCH-LEVEL model, conditional heteroscedasticity comes from the three different sources of time-variation.

III Data and Specification Tests

We use the Fama-Bliss (1987) monthly data on U.S. Treasury zero-coupon bond yields over the 1970:01 - 2002:12 period. We denote by $y_t^{(\tau)}$ the bond yield with a τ-month maturity observed at time t. Following a prevalent practice, we build the three yield-factor series from a short-term, medium-term, and long-term yields. Specifically, we associate the level factor with the 3-month yield ($F_{Lt} = y_t^{(3)}$), the slope factor with the difference between the 120-month yield and the 3-month yield ($F_{St} = y_t^{(120)} - y_t^{(3)}$), and the curvature factor with a linear transformation of the short, medium, and long-term yields ($F_{Ct} = y_t^{(3)} - 2y_t^{(24)} + y_t^{(120)}$).[4]

Table 1 presents some descriptive statistics for the yield factors and the yield-factor residuals extracted from a first-order autoregressive model. For each series, we provide the first four central moments, the Bera-Jarque normality test, the correlation with other factors (or factor residuals), the first-order (cross-)autocorrelation, and the Box-Pierce statistics to test for the k^{th}-order autocorrelation of the series.[5] We observe that the factor series are strongly autocorrelated and depart from normality. The level factor turns out to be negatively correlated with the other two factors, while the slope and curvature factors exhibit a positive correlation. The factor residuals are negatively correlated among each other and far from being normal since their distributions are clearly leptokurtic. The Box-Pierce test suggests that the factor residuals are much less persistent than the factor levels.

The time series of each factor is plotted in Figure 1. While the level factor is always positive, both the slope and curvature factors take negative and positive values. There are several episodes when the yield curve is downward-sloping: the 1973 OPEC oil crisis, during a significant portion of the 1979-1982 monetary experiment, 1989, and towards the end of 2000. Note also that the monetary experiment had a great impact on the curvature of the term structure. Figure 2 displays the absolute value of the factor residuals and shows that all three series exhibit volatility clustering. This suggests that an appropriate model for the yield-factor volatility should include ARCH effects. Moreover, factor volatilities appear to

---

[4] We use the 3-month yield to proxy the level factor because the evolution of the 1-month yield is known to be idiosyncratic (Duffee, 1996). The 2-year maturity is an important intermediate maturity since the term structure of volatility peaks at this maturity (see Dai and Singleton, 2003). We proxy the long-term rate by the 10-year maturity since this yield is less subject to liquidity problems than longer-maturity yields. The transformation used to derive the curvature factor is a discrete measure of the second derivative of the yield curve, which captures its curvature.

[5] The Box-Pierce statistics are distributed as a chi-squared random variable with k degrees of freedom.

depend on the level of interest rates as attested by the superimposed level series. Indeed, the volatility of both slope and curvature factors tends to be high when interest rates are high. This suggests that a diffusion model, in which volatility is a positive function of the level of interest rates, may be able to account for this effect.

< Insert Table 1 >

< Insert Figures 1 and 2 >

To initially assess the relative importance of the level and ARCH effects in yield-factor volatilities, we implement the robust, regression-based specification tests of Wooldridge (1990). The null hypothesis for these specification tests is homoscedasticity. If the data are homoscedastic, then $e_t^2 - \hat{\sigma}^2$ will be uncorrelated with any function of lagged information variables $\lambda(V_{t-1})$. Wooldridge's test is a conditional moment test that determines whether $E[(e_t^2 - \hat{\sigma}^2)\lambda(V_{t-1})] = 0_{K\times 1}$ for some K-dimensional vector $\lambda(V_{t-1})$. The alternative hypothesis is that the expectation is non-zero, which implies that at least one of the variables in $\lambda(V_{t-1})$ is useful in explaining conditional volatility. The size of each test statistic provides a crude metric of the relative ability of each component of $\lambda(V_{t-1})$ in explaining the time-varying volatility. Unlike Engle's (1982) test for ARCH, which proceeds by regressing e_t^2 on a number of lagged squared residuals, Wooldridge's test is robust to non-normality. A major advantage of this test is that it can be constructed using nothing more sophisticated than OLS. In this framework, a robust test for p^{th} order ARCH effects is obtained using $\lambda(V_{t-1})^{\intercal} = (e_{t-1}^2 \; e_{t-2}^2 \; ... \; e_{t-p}^2)$ and a robust test for level effects is derived with $\lambda(V_{t-1}) = |F_{j,t-1}|^{\gamma}$ for some suitably defined γ, such as 0.5 or 1.

We report the Wooldridge's test statistics with the associated p-values in Table 2. Since the largest p-value is 0.0228, there is some clear indication that factor changes are conditionally heteroscedastic. Further, the conditional volatility is strongly related to the overall level of interest rates, but not to the level of the slope or curvature factors. Indeed, when modeling the volatility of the slope (respectively curvature) as a function of the value of the slope (curvature) factor, no level effect can be detected. It appears in Table 2 that setting the elasticity parameter γ equal to one is optimal for the three yield factors, though the significance remains even when $\gamma = 0.5$ as implied by the square root process of Cox, Ingersoll and Ross (1985). This preliminary analysis highlights the following features that a correctly specified volatility model should possess. First, the residuals exhibit volatility clustering, which suggests using a GARCH process to model the conditional residual volatility. Second, the volatility of the slope and curvature factors depends strongly on the level of interest rates, but much less on the levels of the slope and curvature factors. As a result, in the following empirical analysis, we primarily model the level effect using the overall level of interest rates.

< Insert Table 2 >

IV Empirical Results

A Single-Regime Models

In this section, we report the results of fitting the competing models, i.e., the NO LEVEL-NO GARCH, LEVEL, GARCH, and GARCH-LEVEL models, to the U.S. term structure of interest rates over the 1970-2002 period.

We begin by estimating the univariate version of the four models. Table 3 reports for each yield factor the parameter estimates and the Bollerslev-Wooldridge (1992) robust standard-errors. The first column of Table 3 reports the estimates of the homoscedastic model. There appears to be mean reversion in all three yield factors ($\hat{\alpha}_{1i} < 0$, $i = L, S, C$), though not statistically significant for the level. When a level effect is introduced, the loglikelihood function increases dramatically ($\Delta \ln \mathrm{L}_L = 136.70$, $\Delta \ln \mathrm{L}_S = 53.68$, and $\Delta \ln \mathrm{L}_C = 11.40$). The elasticity parameter $\hat{\gamma}_i$ is significant for all three yield factors. Interestingly, if the level effect is modeled using the level of each factor, instead of the overall level of interest rates, the fit of the LEVEL model is significantly reduced ($\ln \mathrm{L}_S = -297.58$ and $\ln \mathrm{L}_C = -234.91$, not reported in the tables). Results of the GARCH model suggest that explicitly modeling the serial correlation in volatility leads to a superior fit. Further, the variance processes exhibit high persistence, though the persistence is lower for the slope and curvature factors ($\hat{\beta}_1 + \hat{\beta}_2 = 0.9897$ for the level, 0.9316 for the slope, and 0.9070 for the curvature). Finally, the GARCH-LEVEL model gives rise to lower estimates of the elasticity parameters than in the LEVEL model. The marginal contribution of the GARCH effect turns out to be stronger than the marginal contribution of the LEVEL effect for the three yield factors.

< Insert Table 3 >

Table 4 reports the parameter estimates and robust standard-errors for the multivariate models. The main difference between the specifications presented in Table 3 and the present specifications is that the residual factors can now be correlated. The point estimates for the correlation coefficients ($\rho_{L,S}$, $\rho_{L,C}$, and $\rho_{S,C}$) are negative in all models, while their magnitude varies across models. As pointed by Dai and Singleton (2000, p. 1945), negative correlation among risk factors is an important feature of the U.S. term structure of interest rates. Allowing the residuals to be correlated improves the fit of each model. For instance, the sum of the loglikelihoods for the three univariate GARCH models is equal to -602.24 and the loglikelihood for the trivariate GARCH model is as high as -477.18. However, correlation matters not only for fitting purposes but it also strongly impacts the point estimate for the elasticity parameters. Indeed, for the LEVEL and GARCH-LEVEL models, the point estimate of γ_i drops considerably for the level and slope factors after accounting for correlation. On the other hand, the point estimate of γ_i increases for the curvature factor. While the level of persistence is comparable to the univariate case, the response of volatility to lagged information shocks ($\hat{\beta}_1$) drops significantly for the level and slope factors.

Comparing the various multivariate volatility models yields some interesting conclusions. Consistent with the univariate results, the GARCH effect seems to dominate the level effect. When only a GARCH effect is introduced, the value of the loglikelihood function increases by 193, which exceeds the rise (126) observed when a level effect is introduced. However, the level effect only requires three extra parameters while the GARCH model requires six extra parameters. Another interesting observation is that the value of the elasticity parameter is weakened when a GARCH effect is introduced, though it remains significant for the level and slope factors. In conclusion, it seems that one needs both level and GARCH effects to adequately model yield-factor volatilities.

< Insert Table 4 >

B Regime-Switching Models

Parameter estimates and robust-standard errors for the regime-switching models are reported in Table 5. Regime-switching models outperform the models estimated in Table 4, which is not overly surprising given that the latter models have been estimated under the assumption that there is only one regime. Allowing for multiple regimes dramatically improves the fit of all four models. Interestingly, when the volatility is allowed to switch from low to high-volatility regimes, the level effect is strengthened and the volatility persistence drops significantly. Furthermore, the performance of the RS-LEVEL model is higher than the performance of the RS-GARCH model, whereas the single-regime GARCH model outperformed the single-regime LEVEL model. Because of its lack of parsimony (it requires the estimation of 26 parameters) and its loglikelihood value, the RS-GARCH level is dominated by the RS-LEVEL model according to the Bayesian information criterion ($\text{BIC}_{RS-LEVEL} = 972.48$ vs. $\text{BIC}_{RS-GARCH} = 1046.59$). In the same way, the RS-LEVEL model is also preferred to the general RS-GARCH-LEVEL model ($\text{BIC}_{RS-GARCH-LEVEL} = 991.07$).

The four panels of Figure 3 contain plots of the smoothed probabilities of high-volatility state for the four considered regime-switching models. The probabilities have been computed using the smoothing algorithm of Kim (1994). Because of their multi-factor nature, our models exploit complementary information on the slope and curvature of the term structure. Our models identify all the major well-known episodes of extreme volatility: the 1973 OPEC oil crisis and its aftermath, the 1979-1982 monetary experiment, the October 1987 stock market crash, and the Russian Ruble devaluation in August 1998. Furthermore, we also identify a period in 1984, which is also identified in Gray (1996), with no clear economic interpretation. Interestingly, the two models that include a level effect identify two additional high-volatility episodes. The first high volatility episode is intermittent and occurs during the 1991-1994 period in which the Federal Reserve held short rates down to combat weakness in the economy and then raised its federal funds rate by 125 basis points during the first four months of 1994 (see Campbell, 1995). The second high volatility regime starts in 2001 and coincides with the massive mortgage refinancing wave that occurred in the US (see Duarte, 2005). These missed episodes illustrate the importance of the level effect. Indeed, during these periods the volatility of all three yield factors was only trivially elevated above previous levels, yet the short-rate was at historically low levels. This coincidence of low interest rates and lightly elevated volatility is explained as a high-volatility episode. This demonstrates that we need to be cautious when interpreting these regimes. A more precise interpretation is that the scaled residuals $e_{it}/Y_{Lt-1}^{\gamma_L}$ have high volatility.

< Insert Table 5 >

< Insert Figure 3 >

V Implications for the Estimation of Term Structure Models

The standard approach to estimate, test, and compare non-nested term structure models is to rely on the Efficient Method of Moments (hereafter EMM) of Bansal et al. (1995) and Gallant and Tauchen (1996). This estimation method can be applied to any class of term structure models: affine models (see Duffie and Kan, 1996 and Dai and Singleton, 2000), quadratic models (Ahn, Dittmar and Gallant, 2002 and Leippold and Wu, 2002), and regime-switching models (see Bansal and Zhou, 2002 and Dai, Singleton and Yang, 2003). Recent applications of EMM to term structure models can be found in Dai and Singleton (2000), Ahn, Dittmar and Gallant (2002), Bansal and Zhou (2002), Ahn et al. (2003), Brandt and Chapman (2003), Bansal, Tauchen and Zhou, (2004), and Duffee and Stanton (2004).

EMM requires that an auxiliary model be used to estimate the conditional density of observed interest rates. The standard approach is to use a VAR-based semi-nonparametric ARCH model as the auxiliary model. This auxiliary model offers a high level of flexibility since it is based on a high-order Hermite series expansion. The popularity of EMM is mainly due to the fact that it does not require the likelihood function of each model to be known in closed-form. Furthermore, it attains the same asymptotic efficiency as maximum likelihood. The moments used in the estimation are the scores of the likelihood function from the auxiliary models, which are known in closed form.

However, the semi-nonparametric auxiliary model cannot capture the main features of the Treasury yields, which are summarized in Table 6. For each considered multi-factor model, we report the value of the log-likelihood function and the number of estimated parameters. We see that the single biggest improvement in statistical fit that arises from adding either a level effect, some GARCH effects, or a second regime, is due to the second regime. The second most important feature of the data is the addition of a level effect. The third most important feature of the data is the GARCH effect. However, the auxiliary model invariably used in EMM estimation only captures low-order ARCH effects and by neglecting the level, GARCH and regime shifts EMM ignores these central features of the data.[6] Instead, EMM relies on a large set of scores that do not always have a clear economic interpretation (e.g., high-order Hermite polynomial parameters).

< Insert Table 6 >

However, extending the standard EMM setting to include those three features is not a straightforward exercise. One could proceed by replacing the semi non-parametric auxiliary model used in EMM with the econometric model developed in this paper. Unfortunately, the scores of this model are not known analytically and must be computed numerically. An alternative is to rely on Indirect Inference of Gouriéroux, Monfort and Renault (1993). While this technique does not use the scores of the auxiliary model, it remains prohibitively computationally intensive since we fit the auxiliary parameters to the simulated data each time we evaluate the objective function in the minimization process. Another alternative is the Simulated Method of Moments (see Duffie and Singleton, 1993) which allows one to specify the important features of the yields and focus only on those moments. Regrettably, there are many features of the data, such as the level effects and regime shifts, that are difficult to specify without explicitly modeling conditional volatility, which cannot be accommodated in the Simulated Method of Moments setting. Consequently, we argue that more research has to be carried out to come up with an estimation technique for term structure models that captures the main features of the yield data in a parsimonious way.

---

[6] Dai and Singleton (2000, p. 1960) argue that one may only need ARCH effects when modeling LIBOR and swap rates as opposed to Treasury yields. In their analysis of U.S. Treasury yields, Ahn, Dittmar and Gallant (2002), Bansal and Zhou (2002), and Bansal, Tauchen and Zhou (2004) only allow for ARCH effects and consider only up to 5 lags.

VI Conclusion

In this paper, we develop a comprehensive model for volatility dynamics in the level, slope, and curvature of the yield curve that simultaneously includes level and GARCH effects along with regime shifts. Our analysis leads to the following conclusions. First, we show that the level of the short-rate is useful in modeling the volatility of the level, slope, and curvature factors of the yield curve. Second, there are significant GARCH effects present even after including a level effect. Third, when the volatility is allowed to switch from low to high-volatility regimes, the model's fit improves dramatically, the level effect is strengthened, and the volatility persistence drops significantly.

Our empirical results have some important implications for the estimation of term structure models. We argue that the Efficient Method of Moment, which is currently the standard estimation technique in the term structure literature, is inconsistent with the main features of the U.S. yield curve that are identified by our model: level and GARCH effects along with regime shifts. Furthermore, the encouraging results obtained with our regime-switching models strengthen the need for including regime-shifts in theoretical term structure models. From this respect, the recent contributions of Dai, Singleton and Yang (2003) and Monfort and Pegoraro (2005) that propose flexible regime-switching term structure models are promising endeavors on this challenging avenue of research.

References

[1] Ahn, D. H., R. F. Dittmar, and A. R. Gallant. "Quadratic Term Structure Models: Theory and Evidence." *Review of Financial Studies*, 16 (2002), 459-485.

[2] Ahn, D. H., R. F. Dittmar, A. R. Gallant, and B. Gao. "Purebred or Hybrid: Reproducing the Volatility in Term Structure Dynamics." *Journal of Econometrics*, 116 (2003), 147-180.

[3] Andersen, T. G., and J. Lund. "Estimating Continuous Time Stochastic Volatility Models of the Short-Term Interest Rates." *Journal of Econometrics*, 77 (1997), 343-377.

[4] Ang, A., and G. Bekaert. "Regime Switches in Interest Rates." *Journal of Business and Economic Statistics*, 20 (2002a), 163-182.

[5] Ang, A., and G. Bekaert. "Short Rate Nonlinearities and Regime Switches." *Journal of Economic Dynamics and Control*, 26 (2002b), 1243-1274.

[6] Ang, A., and M. Piazzesi. "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables." *Journal of Monetary Economics*, 50 (2003), 745-787.

[7] Ball, C., and W. N. Torous. "The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence." *Journal of Finance*, 56 (1999), 2339-2359.

[8] Bansal, R., A. R. Gallant, R. Hussey, and G. Tauchen. "Nonparametric Estimation of Structural Models for High-Frequency Currency Market Data." *Journal of Econometrics*, 66 (1995), 147-180.

[9] Bansal, R., G. Tauchen, and H. Zhou. "Regime Shifts, Risk Premiums in the Term Structure, and the Business Cycle." *Journal of Business and Economic Statistics*, 22 (2004), 396-409.

[10] Bansal, R., and H. Zhou. "Term Structure of Interest Rates with Regime Shifts." *Journal of Finance*, 57 (2002), 1997-2043.

[11] Bollerslev, T. "Modeling the Coherence in Short Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model." *Review of Economics and Statistics*, 72 (1990), 498-505.

[12] Bollerslev, T., and J. M. Wooldridge. "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances." *Econometric Reviews*, 11 (1992), 143-172.

[13] Boudoukh, J., M. Richardson, R. Stanton, and R. F. Whitelaw. "The Stochastic Behavior of Interest Rates: Implications from a Multifactor, Nonlinear Continuous-Time Model." Working Paper, New York University and U. C. Berkeley (1998).

[14] Brandt, M. W., and D. Chapman. "Comparing Multifactor Models of the Term Structure." Working Paper, Boston College and Duke University (2003).

[15] Brenner, R. J., R. Harjes, and R. Kroner. "Another Look at Models of Short Term Interest Rates." *Journal of Financial and Quantitative Analysis*, 31 (1996), 85-107.

[16] Campbell, J. Y. "Some Lessons from the Yield Curve." *Journal of Economic Perspectives*, 9 (1995), 129-152.

[17] 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*, 52 (1992), 1209-1227.

[18] Chapman, D. A., and N. D. Pearson. "Recent Advances in Estimating Term-Structure Models." *Financial Analysts Journal*, 57 (2001), 77-95.

[19] Christiansen, C. "Regime Switching in the Yield Curve." *Journal of Futures Markets*, 24 (2004), 315-336.

[20] Christiansen, C. "Multivariate Term Structure Models with Level and Heteroscedasticity Effects." *Journal of Banking and Finance*, 29 (2005), 1037-1057.

[21] Christiansen, C., and J. Lund. "Revisiting the Shape of the Yield Curve: The Effect of Interest Rate Volatility." Working Paper, Aarhus School of Business (2002).

[22] Cox, J. C., J. E. Ingersoll, and S. A. Ross. "A Theory of the Term Structure of Interest Rates." *Econometrica*, 53 (1985), 385-408.

[23] Dai, Q., and K. J. Singleton. "Specification Analysis of Affine Term Structure Models." *Journal of Finance*, 55 (2000), 1943-1978.

[24] Dai, Q., and K. J. Singleton. "Term Structure Modeling in Theory and Reality." *Review of Financial Studies,* 16 (2003), 631-678.

[25] Dai, Q., K. J. Singleton, and W. Yang. "Regime Shifts in a Dynamic Term Structure Model of the U.S. Treasury Yields." Working Paper, Stanford University (2003).

[26] Driessen, J., P. Klaassen, and B. Melenberg. "The Performance of Multi-Factor Term Structure Models for Pricing and Hedging Caps and Swaptions." *Journal of Financial and Quantitative Analysis*, 38 (2003), 635-672.

[27] Duarte, J. "The Causal Effect of Mortgage Refinancing on Interest-Rate Volatility: Empirical Evidence and Theoretical Implications." Working Paper, University of Washington (2005).

[28] Dueker, M. J. "Markov Switching in GARCH Processes and Mean-reverting Stock Market Volatility." *Journal of Business and Economic Statistics*, 15 (1997), 26-35.

[29] Duffee, G. "Idiosyncratic Variation of Treasury Bill Yields." *Journal of Finance*, 51 (1996), 527-552.

[30] Duffee, G., and R. Stanton. "Estimation of dynamic term structure models." Working Paper, Berkeley (2004).

[31] Duffie, D., and K. J. Singleton. "Simulated Moments Estimation of Markov Models of Asset Prices." *Econometrica*, 61 (1993), 929-952.

[32] Duffie, D., and R. Kan. "A Yield-Factor Model of Interest Rates." *Mathematical Finance*, 6 (1996), 379-406.

[33] Engle, R. F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation." *Econometrica*, 50 (1982), 987-1008.

[34] Fama, E. F., and R. R. Bliss. "The Information in Long-Maturity Forward Rates." *American Economic Review*, 77 (1987), 680-692.

[35] Gallant, A. R., and G. Tauchen. "Which Moments to Match?" *Econometric Theory*, 12 (1996), 657-681.

[36] Gouriéroux, C., A. Monfort, and E. Renault. "Indirect Inference." *Journal of Applied Econometrics*, 8 (1993), 85-118.

[37] Gray, S. "Modeling the Conditional Distribution of Interest Rates as a Regime Switching Process." *Journal of Financial Economics*, 42 (1996), 27-62.

[38] Hamilton, J. D. "Rational Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates." *Journal of Economic Dynamics and Control*, 12 (1988), 385-423.

[39] Hamilton, J. D. *Time Series Analysis*, Princeton NJ: Princeton University Press (1994).

[40] Hamilton, J. D., and D. H. Kim. "A Re-Examination of the Predictability of the Yield Spread for Real Economic Activity." *Journal of Money, Credit, and Banking*, 34 (2002), 340-360.

[41] Kim, C. J. "Dynamic Linear Models with Markow-Switching." *Journal of Econometrics*, 60 (1994), 1-22.

[42] Kugler, P. "The Term Structure of Interest Rates and Regime Shifts: Some Empirical Results." *Economics Letters*, 50 (1996), 121-126.

[43] Leippold, M., and L. Wu. "Asset Pricing under the Quadratic Class." *Journal of Financial and Quantitative Analysis*, 37 (2002), 271-295.

[44] Litterman, R., and J. Scheinkman. "Common Factors Affecting Bond Returns." *Journal of Fixed Income*, 1 (1991), 54-61.

[45] Longstaff, F. A. and E. S. Schwartz. "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model." *Journal of Finance*, 47 (1992), 1259-1282.

[46] Monfort, A. and F. Pegoraro. "Switching VARMA Term Structure Models." Working Paper, CREST (2005).

[47] Pérignon, C. and C. Villa. "Sources of Time Variation in the Covariance Matrix of Interest Rates." *Journal of Business*, forthcoming (2004).

[48] Smith, D. R. "Markov-Switching and Stochastic Volatility Diffusion Models of Short-Term Interest Rates." *Journal of Business and Economic Statistics*, 20 (2002), 183-197.

[49] Wooldridge, J. "A Unified Approach to Robust, Regression-Based Specification Tests." *Econometric Theory*, 6 (1990), 17-43.

Appendix A

In affine models, the bond yield with a τ maturity is an affine function of N state variables:

$$y_t^{(\tau)} = \frac{1}{\tau}[-A(\tau) + B(\tau)^\top X_t]$$

where $A(\tau)$ is a scalar, and $B(\tau)$ and X_t are N-vectors (see Duffie and Kan, 1996). When the volatility of all N state variables depends on a single state variable, i.e., a $\mathbb{A}_1(N)$ model in the terminology of Dai and Singleton (2000), the dynamics of the state variables is:

$$dX_t = \mu(X_t)dt + \sigma(X_{1t})dW_t^Q$$

where W_t^Q is a N-vector of independent standard Brownian motions. Assume that there are M bond yields with different maturities. Let F_{it} be a given yield factor defined as a linear combination of L bond yields, with $L \leqslant M$. We denote by ω_l the weight associated with the bond yield with a τ_l maturity, $l = 1, ..., L$ and $\omega_l \in \mathbb{R}$. In this case, the yield factor is also affine in the state variables:

$$\begin{aligned} F_{it} &= \frac{\omega_1}{\tau_1}[-A(\tau_1) + B(\tau_1)^\top X_t] + ... + \frac{\omega_L}{\tau_L}[-A(\tau_L) + B(\tau_L)^\top X_t] \\ &= -\sum_{l=1}^{L} \frac{\omega_l}{\tau_l} A(\tau_l) + \sum_{i=1}^{N}\sum_{l=1}^{L} \frac{\omega_l}{\tau_l} B_i(\tau_L) X_{it}. \end{aligned}$$

The yield factor F_{it} is a function of the state variables and time, which is twice differentiable with respect to X_t. Applying Ito's lemma leads to:

$$\begin{aligned} dF_{it}(X_{1t}, ..., X_{Nt}, t) &= \frac{\partial F_{it}}{\partial X_{1t}} dX_{1t} + ... + \frac{\partial F_{it}}{\partial X_{Nt}} dX_{Nt} + \frac{\partial F_{it}}{\partial t} dt \\ &= [\sum_{i=1}^{N}\sum_{l=1}^{L} \frac{\omega_l}{\tau_l} B_i(\tau_l)\mu_i(X_t) + \frac{\partial y_t}{\partial t}]dt + \sum_{i=1}^{N}\sum_{l=1}^{L} \frac{\omega_l}{\tau_l} B_i(\tau_l)\sigma_i(X_{1t})dW_{it}^Q \end{aligned}$$

since the second-order terms are zero. As a result, the conditional volatility of all yield factors depends on the same single source of risk, X_{1t}.

As an illustration, we consider the case of three state variables and three yield factors defined as $F_{Lt} = y_t^{(\tau_1)}$, $F_{St} = y_t^{(\tau_3)} - y_t^{(\tau_1)}$, $F_{Ct} = y_t^{(\tau_1)} - 2y_t^{(\tau_2)} + y_t^{(\tau_3)}$, where τ_1, τ_2, and τ_3 denote a short, medium, and long maturity, respectively. The dynamics of the yield factors is:

$$
\begin{aligned}
dF_{Lt} &= [\sum_{i=1}^{3} \frac{B_i(\tau_1)}{\tau_1}\mu_i(X_t) + \frac{\partial y_t}{\partial t}]dt + \sum_{i=1}^{3} \frac{B_i(\tau_1)}{\tau_1}\sigma_i(X_{1t})dW_{it}^{Q} \\
dF_{St} &= [\sum_{i=1}^{3}(\frac{B_i(\tau_3)}{\tau_3} - \frac{B_i(\tau_1)}{\tau_1})\mu_i(X_t) + \frac{\partial y_t}{\partial t}]dt + \sum_{i=1}^{3}(\frac{B_i(\tau_3)}{\tau_3} - \frac{B_i(\tau_1)}{\tau_1})\sigma_i(X_{1t})dW_{it}^{Q} \\
dF_{Ct} &= [\sum_{i=1}^{3}(\frac{B_i(\tau_1)}{\tau_1} - \frac{2B_i(\tau_2)}{\tau_2} + \frac{B_i(\tau_3)}{\tau_3})\mu_i(X_t) + \frac{\partial y_t}{\partial t}]dt \\
&\quad + \sum_{i=1}^{3}(\frac{B_i(\tau_1)}{\tau_1} - \frac{2B_i(\tau_2)}{\tau_2} + \frac{B_i(\tau_3)}{\tau_3})\sigma_i(X_{1t})dW_{it}^{Q}.
\end{aligned}
$$

VII Tables and Figures

Table 1: Descriptive Statistics

	Level	*Slope*	*Curvature*	e_L	e_S	e_C
Mean	6.4947	1.3590	-0.0428	-	-	-
Variance	7.7562	2.0399	0.5712	0.3681	0.2633	0.1938
Skewness	1.0471	-0.5966	0.0435	-1.1751	0.7211	-0.6965
Kurtosis	4.4361	3.1934	4.0297	14.6185	9.3072	7.8707
BJ	106.39	24.11	17.62	2318.47	690.70	423.45
Corr(F_L, i)	-	-0.6180	-0.4070	-	-0.8134	-0.0533
Corr(F_S, F_C)	-	-	0.1404	-	-	-0.1448
CACorr($F_{L,t-1}, i$)	0.9711	-0.5928	-0.4043	0.1219	-0.0336	-0.2118
CACorr($F_{S,t-1}, i$)	-0.5721	0.9299	0.1794	-0.1074	0.0463	0.1779
CACorr($F_{C,t-1}, i$)	-0.3930	0.1339	0.8044	-0.0385	-0.0624	-0.0920
BP_1	376.24	345.01	258.20	5.93	0.86	3.38
BP_{12}	3325.19	2063.31	1051.63	42.29	22.12	35.48

Note: This table presents the mean, variance, skewness, and kurtosis of the three yield factors (Level F_L, Slope F_S, Curvature F_C) and yield-factor residuals (e_L, e_S, e_C). BJ stands for the Bera-Jarque normality test, Corr for correlation, CACorr for first-order cross-autocorrelation, and BP_1 and BP_{12} for the Box-Pierce test with one and twelve lags respectively. The latter two statistics are distributed as chi-squared with 1 and 12 degrees of freedom and then the 5 percent critical values are 3.84 and 21.03 respectively.

Table 2: Tests for Level Effect and Heteroscedasticity

$\lambda(V_{t-1})$	*Level*	*Slope*	*Curvature*
$e_{t-1}^2\ e_{t-2}^2\ ...\ e_{t-6}^2$	14.6891 [0.0228]	15.9148 [0.0142]	21.8449 [0.0013]
${F_{L,t-1}}^{0.5}$	7.2969 [0.0069]	6.8220 [0.0090]	6.5559 [0.0105]
$F_{L,t-1}$	7.8533 [0.0051]	8.8714 [0.0029]	7.1608 [0.0075]
$\lvert F_{j,t-1}\rvert^{0.5}$	-	0.6067 [0.4360]	2.8213 [0.0930]
$\lvert F_{j,t-1}\rvert$	-	0.2095 [0.6472]	3.2472 [0.0715]

Note: This table reports the Wooldridge (1990) robust specification tests for level and ARCH effects in yield-factor volatilities. This conditional moment test determines whether $E[(e_t^2 - \hat{\sigma}^2)\lambda(V_{t-1})] = 0_{K\times 1}$ for some K-dimensional vector $\lambda(V_{t-1})$. The robust test for p^{th} order ARCH effects is obtained using $\lambda(V_{t-1})^\intercal = (e_{t-1}^2\ e_{t-2}^2\ ...\ e_{t-p}^2)$ and the robust test for level effects is derived with $\lambda(V_{t-1}) = |F_{j,t-1}|^\gamma$ with $j = S, L$ and $\gamma = 0.5, 1$. We report the Wooldridge's statistics with the associated p-values into brackets.

Table 3: Parameter Estimates for Univariate Models

	NO LEVEL-NO GARCH			LEVEL			GARCH			GARCH-LEVEL		
	Level	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*
α_0	0.1349 (0.1094)	0.0889 (0.0556)	-0.0045 (0.0214)	-0.0001 (0.0118)	0.0744 (0.0387)	0.0126 (0.0203)	0.1135 (0.0549)	0.0424 (0.0426)	-0.0150 (0.0177)	0.0326 (0.0161)	0.0330 (0.0386)	-0.0157 (0.0182)
α_1	-0.0220 (0.0196)	-0.0661 (0.0281)	-0.1820 (0.0417)	-0.0020 (0.0012)	-0.0448 (0.0168)	-0.1487 (0.0384)	-0.0230 (0.0128)	-0.0420 (0.0208)	-0.1121 (0.0254)	-0.0046 (0.0023)	-0.0356 (0.0197)	-0.1146 (0.0260)
β_0	0.3681 (0.0699)	0.2633 (0.0379)	0.1938 (0.0253)	2.5210 (1.0076)	20.404 (6.4583)	63.638 (23.870)	5.8937 (3.4347)	15.947 (9.8622)	23.613 (7.4734)	0.1443 (0.1915)	1.7201 (1.8911)	14.066 (15.773)
β_1	-	-	-	-	-	-	0.2935 (0.0790)	0.2318 (0.1024)	0.3516 (0.0995)	0.2263 (0.1087)	0.1604 (0.0728)	0.3353 (0.0905)
β_2	-	-	-	-	-	-	0.6962 (0.0795)	0.6998 (0.1184)	0.5554 (0.0779)	0.7516 (0.1147)	0.7737 (0.1023)	0.5617 (0.0799)
γ	-	-	-	1.2022 (0.1069)	0.6410 (0.0872)	0.2961 (0.1046)	-	-	-	1.1062 (0.3342)	0.5754 (0.2296)	0.1563 (0.3401)
$\ln L_i$	-365.59	-298.68	-237.60	-228.70	-245.00	-226.20	-199.81	-216.15	-186.28	-185.27	-210.32	-185.94

Notes: The general model used is the GARCH-LEVEL model:

$$\Delta F_{it} = \alpha_{0i} + \alpha_{1i} \cdot F_{it-1} + e_{it}$$

for $i = L, S, C$ and $t = 1, ..., T$. The conditional volatility of e_{it} is $E(e_{it}^2 | \psi_{t-1}) = \sigma_{it}^2 F_{Lt-1}^{2\gamma_i}$. We model σ_{it}^2, the volatility of the scaled residual $v_{it} = e_{it} / Y_{Lt-1}^{\gamma_i}$, as:

$$\sigma_{it}^2 = \beta_{0i} + \beta_{1i} \cdot v_{it-1}^2 + \beta_{2i} \cdot \sigma_{it-1}^2.$$

The restricted versions of the model impose the following constraints: NO LEVEL-NO GARCH model: $\beta_{ki} = 0$, $i = L, S, C$ and $k = 1, 2$, and $\gamma_i = 0$, $i = L, S, C$; LEVEL model: $\beta_{ki} = 0$, $i = L, S, C$ and $k = 1, 2$; GARCH model: $\gamma_i = 0$, $i = L, S, C$. For each univariate model, we report the maximum likelihood parameter estimates, the value of the loglikelihood function, and Bollerslev-Wooldridge robust standard errors into parentheses. β_0 parameters are multiplied by 1000 and so are they in Tables 4 and 5.

Table 4: Parameter Estimates for Multivariate Models

	NO GARCH-NO LEVEL			LEVEL			GARCH			GARCH-LEVEL		
	Level	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*
α_0	0.1256 (0.0575)	0.0744 (0.0347)	-0.0023 (0.0212)	0.0266 (0.0274)	0.0547 (0.0275)	0.0153 (0.0204)	0.1200 (0.0559)	0.0427 (0.0385)	-0.0109 (0.0174)	0.0461 (0.0462)	0.0245 (0.0208)	-0.0130 (0.0172)
α_1	-0.0212 (0.0106)	-0.0541 (0.0147)	-0.1586 (0.0416)	-0.0069 (0.0051)	-0.0359 (0.0114)	-0.1435 (0.0373)	-0.0215 (0.0127)	-0.0391 (0.0164)	-0.1169 (0.0260)	-0.0073 (0.0076)	-0.0305 (0.0130)	-0.1189 (0.0271)
β_0	0.3682 (0.0697)	0.2633 (0.0383)	0.1941 (80.0254)	6.0018 (1.9586)	38.621 (10.609)	54.799 (20.112)	7.9016 (7.0820)	9.1544 (11.1589)	21.153 (6.8654)	0.5619 (0.4937)	4.7968 (5.2924)	9.3270 (10.768)
β_1	-	-	-	-	-	-	0.1406 (0.1138)	0.0960 (0.0659)	0.3550 (0.0977)	0.0905 (0.0546)	0.0901 (0.0673)	0.3375 (0.0918)
β_2	-	-	-	-	-	-	0.8117 (0.1350)	0.8522 (0.1206)	0.5700 (0.0783)	0.8463 (0.0897)	0.8257 (0.1516)	0.5718 (0.0862)
γ	-	-	-	0.9688 (0.0800)	0.4684 (0.0729)	0.3384 (0.1034)	-	-	-	0.7997 (0.1865)	0.3144 (0.1666)	0.2536 (0.3422)
$\rho_{L,S}$	-0.8076 (0.0324)	-	-	-0.7123 (0.0346)	-	-	-0.6611 (0.0375)	-	-	-0.6508 (0.0364)	-	-
$\rho_{i,C}$	-0.0465 (0.0999)	-0.1550 (0.0959)	-	-0.1068 (0.0825)	-0.1612 (0.0788)	-	-0.1616 (0.0645)	-0.1323 (0.0729)	-	-0.1754 (0.0623)	-0.1330 (0.0697)	-
lnL	-670.637			-544.633			-477.183			-458.124		

Notes: The general model used is the GARCH-LEVEL model:

$$\Delta F_{it} = \alpha_{0i} + \alpha_{1i} \cdot F_{it-1} + e_{it}$$

for $i = L, S, C$ and $t = 1, ..., T$. The conditional volatility of e_{it} is $E(e_{it}^2|\psi_{t-1}) = \sigma_{it}^2 F_{Lt-1}^{2\gamma_i}$. We model σ_{it}^2, the volatility of the scaled residual $v_{it} = e_{it}/F_{Lt-1}^{\gamma_i}$, as:

$$\sigma_{it}^2 = \beta_{0i} + \beta_{1i} \cdot v_{it-1}^2 + \beta_{2i} \cdot \sigma_{it-1}^2.$$

The restricted versions of the model impose the following constraints: NO LEVEL-NO GARCH model: $\beta_{ki} = 0$, $i = L, S, C$ and $k = 1, 2$, and $\gamma_i = 0$, $i = L, S, C$; LEVEL model: $\beta_{ki} = 0$, $i = L, S, C$ and $k = 1, 2$; GARCH model: $\gamma_i = 0$, $i = L, S, C$. For each multivariate model, we report the maximum likelihood parameter estimates, the value of the loglikelihood function, and Bollerslev-Wooldridge robust standard errors into parentheses.

Table 5: Parameter Estimates for Regime-Switching Models

	RS-NO GARCH-NO LEVEL			RS-LEVEL			RS-GARCH			RS-GARCH-LEVEL		
	Level	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*	*Level*	*Slope*	*Curv.*
$\alpha_{0,S1}$	-0.0041 (0.1577)	0.0866 (0.1199)	-0.0970 (0.0905)	-0.0133 (0.0312)	0.1077 (0.0473)	0.0517 (0.0599)	-0.0172 (0.0966)	0.1247 (0.0785)	-0.0650 (0.0525)	-0.0186 (0.0713)	0.1423 (0.0793)	0.0140 (0.0643)
$\alpha_{0,S2}$	0.0589 (0.0403)	0.0238 (0.0221)	0.0197 (0.0160)	0.0988 (0.0669)	0.0058 (0.0199)	-0.0021 (0.0237)	0.0707 (0.0066)	0.0109 (0.0113)	0.0013 (0.0146)	0.1180 (0.2338)	0.0143 (0.0148)	-0.0083 (0.0207)
α_1	-0.0091 (0.0065)	-0.0255 (0.0107)	-0.1317 (0.0287)	-0.0123 (0.0086)	-0.0302 (0.0120)	-0.1435 (0.0401)	-0.0101 (0.0010)	-0.0240 (0.0094)	-0.1165 (0.0244)	-0.0166 (0.0322)	-0.0328 (0.0204)	-0.1281 (0.0386)
$\beta_{0,S1}$	1171.5597 (139.3382)	690.2253 (103.8130)	665.8439 (122.0522)	5.7746 (2.2116)	45.3573 (14.2848)	87.7479 (37.1780)	413.0212 (81.3628)	313.8057 (61.2716)	268.0241 (97.5286)	5.1311 (2.2832)	44.4600 (18.5021)	62.0588 (34.3799)
$\beta_{0,S2}$	100.8820 (15.9410)	104.8961 (9.7203)	93.3185 (11.1286)	1.1366 (0.5002)	10.8774 (3.7007)	12.4080 (6.1394)	5.2555 (3.4850)	18.7982 (14.0511)	11.2662 (5.8319)	0.1904 (0.2545)	5.7374 (5.6103)	2.9259 (2.4358)
β_1	-	-	-	-	-	-	0.2403 (0.0978)	0.0944 (0.0466)	0.3148 (0.0952)	0.1372 (0.0770)	0.0457 (0.0459)	0.2506 (0.0820)
β_2	-	-	-	-	-	-	0.1161 (0.0218)	0.1863 (0.0477)	0.1381 (0.0462)	0.1940 (0.0533)	0.1440 (0.1087)	0.1282 (0.0440)
γ	-	-	-	1.1248 (0.1034)	0.5604 (0.0882)	0.4141 (0.1177)	-	-	-	1.0380 (0.1151)	0.5033 (0.0969)	0.3837 (0.1391)
$\rho_{L,S}$	-0.6356 (0.0358)	-	-	-0.6069 (0.0379)	-	-	-0.6648 (0.0403)	-	-	-0.6102 (0.0386)	-	-
$\rho_{i,C}$	-0.2162 (0.0558)	-0.0959 (0.0591)	-	-0.1630 (0.0616)	-0.1794 (0.0628)	-	-0.2053 (0.0602)	-0.1247 (0.0648)	-	-0.1981 (0.0604)	-0.1493 (0.0593)	-
p_{S1}		2.1466 (0.5189)			2.0022 (0.3964)			2.4406 (0.6495)			2.2527 (0.5224)	
p_{S2}		3.7348 (0.6152)			2.4183 (0.3518)			3.3285 (0.5144)			2.7390 (0.5095)	
lnL		-463.894			-417.453			-445.539			-408.805	

Notes: The general model used is the Regime-Switching GARCH-LEVEL model:

$$\Delta F_{it} = \alpha_{0is_t} + \alpha_{1i} \cdot F_{it-1} + e_{it}$$

for $i = L, S, C$ and $t = 1, ..., T$, where $s_t = 1$ denotes the "high-volatility regime" and $s_t = 2$ the "low-volatility regime". The conditional volatility of e_{it} is $E(e_{it}^2|\psi_{t-1}) = \hat{\sigma}^2_{it|s_t,s_{t-1}} F_{Lt-1}^{2\gamma_i}$ where $\hat{\sigma}^2_{it|s_t,s_{t-1}} = E(v_t^2|S_t = s_t, S_{t-1} = s_{t-1}, \psi_{t-1})$ is modeled as:

$$\hat{\sigma}^2_{it|s_t,s_{t-1}} = \beta_{0is_t} + \beta_{1i} \cdot v^2_{it-1} + \beta_{2i} \cdot \hat{\sigma}^2_{it-1|s_{t-1}}.$$

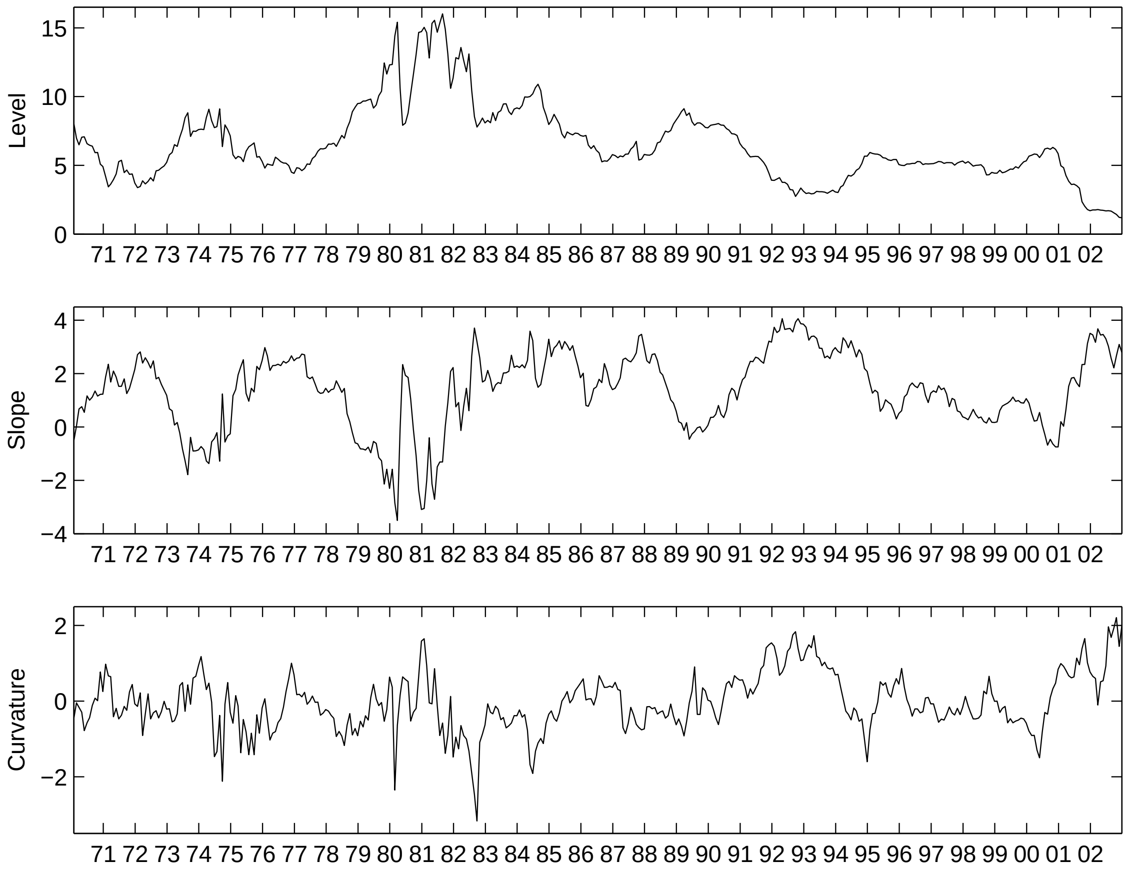
The probability of staying in state j is modeled as $P(S_t = j|S_{t-1} = j) = \exp(p_{Sj})/(1{+}\exp(p_{Sj}))$. The restricted versions of the model impose the following constraints: RS-NO LEVEL-NO GARCH model: $\beta_{ki} = 0$, $i = L, S, C$ and $k = 1, 2$, and $\gamma_i = 0$, $i = L, S, C$; RS-LEVEL model: $\beta_{ki} = 0$, $i = L, S, C$ and $k = 1, 2$; RS-GARCH model: $\gamma_i = 0$, $i = L, S, C$. For each model, we report the maximum likelihood parameter estimates, the value of the loglikelihood function, and Bollerslev-Wooldridge robust standard errors into parentheses.

Table 6: Which Features of Volatility Are Most Important?

	Single Regime Models	Regime-Switching Models
NO GARCH-NO LEVEL	-670.637	-463.894
	[12]	[20]
LEVEL	-544.633	-417.453
	[15]	[23]
GARCH	-477.183	-445.539
	[18]	[26]
GARCH-LEVEL	-458.124	-408.805
	[21]	[29]

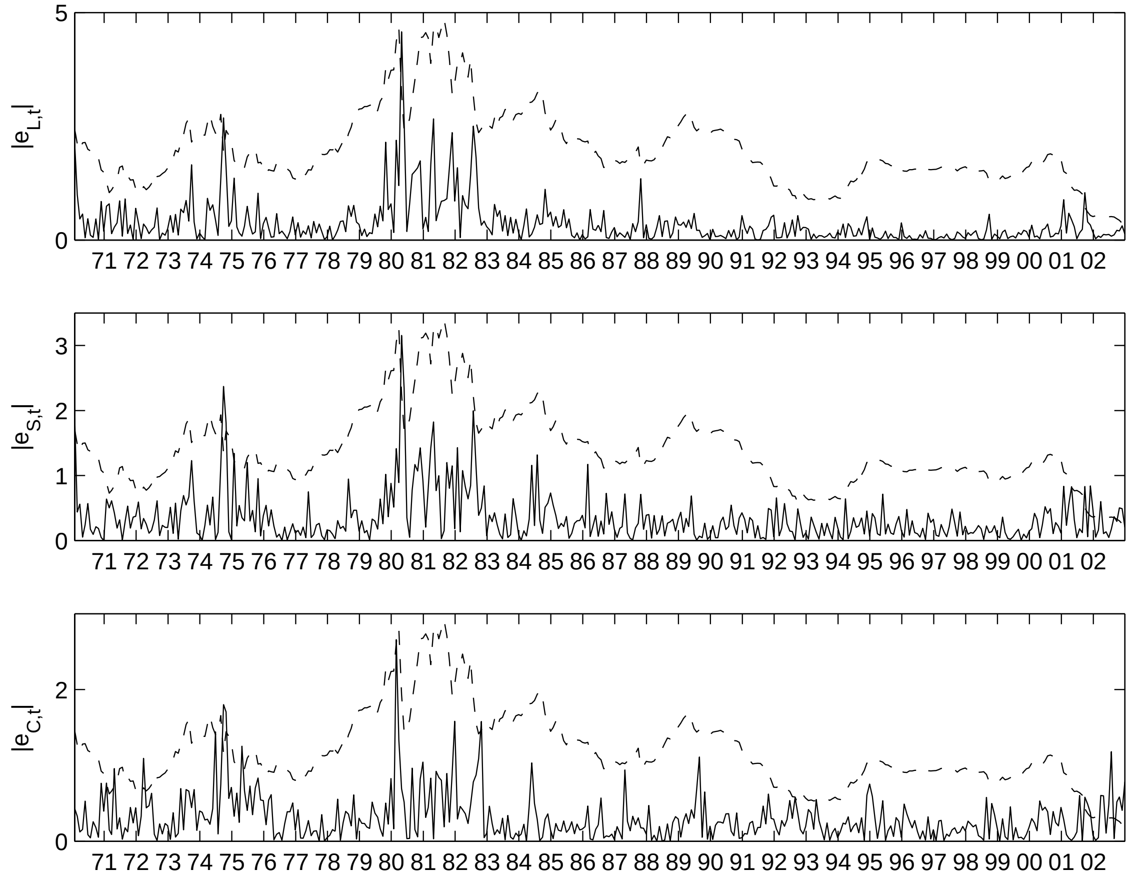
Note: This table presents the value of the log-likelihood function and the number of estimated parameters (in square brackets) for all considered models. We can measure the improvement in statistical fit that arises from adding either a level effect, some GARCH effects, or a second volatility regime.

Figure 1: Value of the Level, Slope, and Curvature Factors



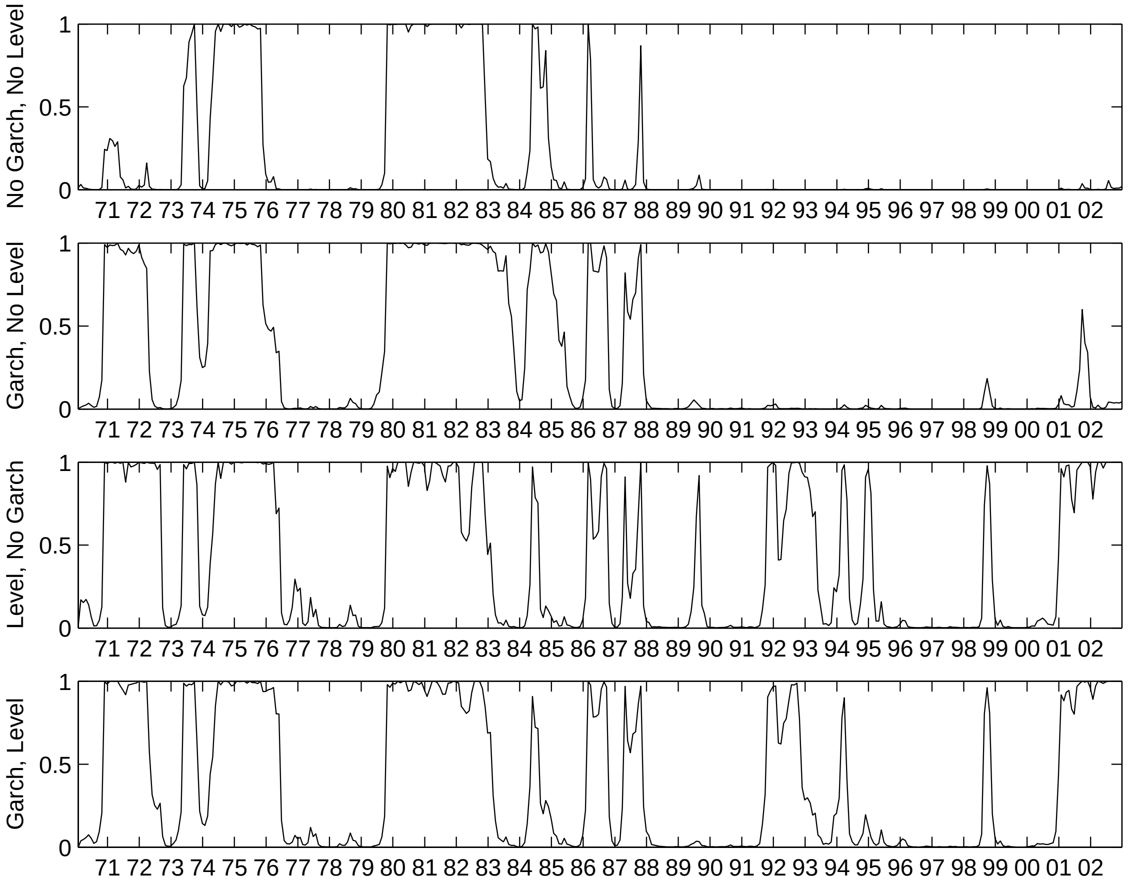
Notes: This figure displays the series for the three yield factors. The level factor is associated with the 3-month yield, the slope factor with the difference between the 120-month yield and the 3-month yield, and the curvature factor with a linear transformation of the short (1-month), medium (24-month), and long-term (120-month) yields.

Figure 2: Factor Volatilities and the Level Factor



Notes: This figure displays the absolute value of the factor residuals (left Y-axis, solid line) with the level factor (right Y-axis, dashed line). $|e_{L,t}|$ denotes the absolute level residual, $|e_{S,t}|$ the absolute slope residual, and $|e_{C,t}|$ the absolute curvature residual.

Figure 3: Smoothed Probability of High-Volatility State $P(S_t = 1|\psi_T)$



Notes: The four panels contain the time series of the smoothed probabilities that the level factor is in the high-volatility regime at time t according to the Regime Switching (RS) NO GARCH-NO LEVEL model, the RS-LEVEL model, the RS-GARCH model, and the RS-GARCH-LEVEL model. The smoothed probability is based on the entire sample: $P(S_t = 1|\psi_T)$.