



# **Economic Growth Centre Working Paper Series**

## **A Regime Switching Macro-finance Model of the Term Structure**

*by*

**ZHU Xiaoneng and Shahidur RAHMAN**

Economic Growth Centre  
Division of Economics  
School of Humanities and Social Sciences  
Nanyang Technological University  
Nanyang Avenue  
SINGAPORE 639798

Website: <http://www.ntu.edu.sg/HSS/egc/>

**Working Paper No: 2009/01**

Copies of the working papers are available from the World Wide Web at:

Website: <http://www.hss.ntu.edu.sg/egc/>

The author bears sole responsibility for this paper. Views expressed in this paper are those of the author(s) and not necessarily those of the Economic Growth Centre.

# A regime switching macro-finance model of the term structure

Xiaoneng Zhu<sup>1</sup>

Department of Economics, HSS  
Nanyang Technological University

Shahidur Rahman

Department of Economics, HSS  
Nanyang Technological University

January 2009

<sup>1</sup>corresponding to: zhux0002@ntu.edu.sg

## **Abstract**

This paper presents and estimates a regime switching macro-finance model of the term structure with latent and macroeconomic factors. The joint dynamics of the yield and macro factors are examined simultaneously. Both the canonical yields-only model and the macro-finance model capture two regimes in the state equation that relate to a turbulent period and a tranquil period. Statistically, the formal tests indicate significant bidirectional linkages between the yield curve and economic activity. I also examine how the yield factors respond to shocks to the macro factors and the feedback of the macro factors to the yield curve. Finally, I find that the theoretical level implied by the expectations hypothesis is a good approximation of the actual level factor in the regime-shifting macro-finance model framework.

# 1 Introduction

Understanding the joint dynamics of macroeconomic and yield factors is important for monetary policy-making and bond portfolio management. The yield curve contains important information about the future economic activity (e.g., among others, Estrella and Hardouvelis (1991), Estrella and Mishkin (1998)). On the other hand, the conduct of monetary policy, according to the Taylor (1993) rule, transmits the movement in macroeconomic factors into the dynamics of the short end of the yield curve. Through the expectations hypothesis with the addition of a partially predictable time-varying risk premium, it also moves the long end of the yield curve. Since the interactions between the yield and macroeconomic factors are expected to be bidirectional and simultaneous, they should be investigated in one system. The joint system, labeled as the ‘macro-finance’ model, implicitly implies a monetary policy rule.

Recently an extensive literature focuses on examining the linkages between the yield curve and the economic driving force in the term structure models with macro factors. Ang and Piazzesi (2003) imposed no-arbitrage restrictions on a VAR model with latent yield factors derived from an affine term structure model and found that macroeconomic factors explain up to 85% of the variation in the bond yields in addition to improved forecasting performance. Diebold, Rudebusch and Aruoba ((2006), DRA henceforth) provided strong evidence of the dynamic interactions between the yield curve and the macroeconomy<sup>1</sup> in a framework of the Nelson-Siegel (1987) type term structure model.

The model proposed in this paper extends the DRA dynamic Nelson-Siegel model by incorporating regime shift into the joint dynamics. Nowadays the regime shift stands as a stylized fact in the term structure modeling. Some recent studies (see, for example, Bansal and Zhou (2002), Dai, Singleton and Yang (2007)) show that the regime-shifting term structure models can account for some well-documented puzzles, for instance, the violation of the expectation hypothesis and the predictability of excess bond return. Regimes are typically interpreted as low and high volatility states and are intimately related to business cycles. Neglecting the regime shift might lead to an infinite VAR specification instead of a VAR model with short lag length.

The DRA dynamic Nelson-Siegel model with regime shift has several advantages. Firstly, the DRA dynamic Nelson-Siegel model provides more accurate forecasting (Diebold and Li (2006)) of the dynamics of the yield curve over time in contrast to the

---

<sup>1</sup>See, among others, Ang, Piazzesi and Wei (2006), Dewachter and Lyrio (2003), Kozicki and Tinsley (2001), Rudebusch and Wu (2008)

typically no-arbitrage models in finance literature (Duffee (2002)). Secondly, it allows a bidirectional feedback mechanism with which the entire yield curve responds to the macroeconomic information, and vice versa. Thirdly, with regime-shifting VAR we can investigate the impulse responses between macro and yield factors under two different regimes. Since regimes are intimately related with business cycles, using regime shifting model may shed light on the monetary policy transmission and market expectation formation mechanism during economic recessions and booms. Finally, the model is flexible enough to match the changing shape of the yield curve, and it is still parsimonious and easy to estimate.

The selected sample covers a period of January 1980 to March 2008<sup>2</sup>. According to the NBER dating, this period covers four economic recessions, including two relatively consecutive recessions<sup>3</sup>. This period can be identified as a volatile regime in the yields-only and macro-finance models. During the sample period, there were also substantial changes in monetary policy (Walsh (2003), Chapter 9). In particular, the Fed operating procedure has shifted from a non-borrowed-reserves targeting to a borrowed-reserves targeting in 1982. From 1988 onwards, however, the Fed changed to target the federal funds rate. The identification of regimes coincides with the monetary policy shift in 1982.

The disadvantage of the DRA dynamic Nelson-Siegel model is that the model doesn't explicitly impose no-arbitrage restrictions. DRA makes a defense on this theoretically unappealing feature. If the arbitrage opportunities are hedged away immediately in the financial markets, the data should reflect the matter of fact. Therefore the dynamic Nelson-Siegel model approximately does not admit arbitrage opportunities. In addition, the arbitrage-free model might be subject to misspecification if there exist some transitory arbitrage opportunities in the market.

The rest of this paper is organized as follows. In section 2, I present and estimate the benchmark 'yields-only' model subject to regime shifts. Section 3 proposes and estimates the regime switching macro-finance model. The implications and relationship between the yield and macro factors are analyzed. Furthermore, I examine the implication of the expectations hypothesis in the framework of the macro-finance model. Section 4 is concluding remarks.

---

<sup>2</sup>January 1980 is the earliest observation on the EconStat<sup>TM</sup> database.

<sup>3</sup>Respectively, from January 1980 to July 1980, and from July 1981 to November 1982.

## 2 Yields-only model

Principle component analysis shows that a few factors can explain over 97% (Piazzesi (2004)) of the variance of yield changes. These factors are usually labeled as ‘level’, ‘slope’ and ‘curvature’ according their effect on the yield curve. Because it seems to be a stable interpretation of factors across different specifications and sample selections, most term structure models use three factors to capture stylized facts of yields. These term structure models impose cross-section restrictions on the yield equation to achieve parsimony. With term structure models, we can generate forecasting on future path of yields and recover missing bond yields from observed yields.

In finance literature, the cross-section restrictions are typically derived from no-arbitrage conditions. This is consistent with the reasonable assumption that a riskless arbitrage opportunity should be traded away immediately in liquid and deep markets. Unfortunately, the theoretical consistency doesn’t provide a good forecasting performance (Duffee (2002)). Another strand of literature employs empirical appealing models, for example, Nelson and Siegel (1987), Diebold and Li (2006). Although these models are not theoretically well-grounded, they show good predictive power in time-series and fit in the cross-section. The proceeding properties make this type of models widely applied in central banks and investment banks. This article follows DRA and goes along this strand of literature.

### 2.1 model representation

By properly restricting factor loadings in a statistical factor model, Diebold and Li (2006) propose the dynamic Nelson-Siegel model for the yield with maturity  $\tau$ ,

$$y_{t(\tau)} = L_t + S_t \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + C_t \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) + \varepsilon_t \quad (1)$$

where  $L_t$  is the level factor,  $S_t$  denotes the slope factor and  $C_t$  represents the curvature factor. Empirically, the level factor is corresponding to long-term interest rate, the slope factor is associated with the difference between the short-term yield and long-term yield, and the curvature factor corresponds to two times of medium-term yields minus the sum of long- and short-term yields. Therefore, the level factor is a long-term factor, the slope factor is a short-term factor and the curvature is a medium-term factor. These factors contain information of the macroeconomic dynamics and vice versa (Diebold, Rudebusch and Aruoba (2006), Tam and Yu (2008)). The  $\lambda_t$  is the rate of changes of factors loadings along the maturity horizons, it also determines the

maturity at which the curvature loading achieves its maximum.

For the entire yield curve with different maturities, the observation equation can be specified as:

$$\begin{bmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{bmatrix} = \begin{bmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1}e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2}e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N}e^{-\lambda\tau_N} \end{bmatrix} \begin{bmatrix} L_t \\ S_t \\ C_t \end{bmatrix} + \begin{bmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_N) \end{bmatrix} \quad (2)$$

with  $\varepsilon_t \sim N(0, \Omega)$ . The dynamic Nelson-siegel has superior out-of-sample forecasting performance, especially at long horizon. In contrast, some affine term structure models that impose no-arbitrage restrictions give poor forecasting performance. Although the dynamic Nelson-Siegel is neither general equilibrium model nor no-arbitrage model, it provides empirical fit, simplicity and parsimony.

To identify possibly turbulent and tranquil periods in the term structure of interest rates, the latent yield factors are assumed to follow a Markov-switching vector autoregression process<sup>4</sup>

$$\begin{bmatrix} L_t - \mu_L \\ S_t - \mu_S \\ C_t - \mu_C \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \begin{bmatrix} L_{t-1} - \mu_L \\ S_{t-1} - \mu_S \\ C_{t-1} - \mu_C \end{bmatrix} + \begin{bmatrix} \eta_{\xi_t, 1t} \\ \eta_{\xi_t, 2t} \\ \eta_{\xi_t, 3t} \end{bmatrix} \quad (3)$$

where  $\xi_t = H, L$  indicates a high or low volatility regime prevailing at time  $t$  and  $\eta_{\xi_t} = (\eta_{\xi_t, 1t}, \eta_{\xi_t, 2t}, \eta_{\xi_t, 3t})'$  allows regime-dependent heteroscedasticity,

$$\begin{aligned} \eta_H &\sim N(0, \Sigma_H) \\ \eta_L &\sim N(0, \Sigma_L) \end{aligned} \quad (4)$$

For optimality of the Kalman filter, I assume the disturbances  $\eta_t$  and  $\varepsilon_t$  are uncorrelated with each other, and initial state  $X_0$  is orthogonal to the realization of  $\eta_{\xi_t}$  and  $\varepsilon_t$

$$\begin{aligned} E(\varepsilon_t \eta_{\xi_t}) &= 0 \text{ for } t = 1, 2, \dots, T; \xi_t = H, L \\ E(\varepsilon_t X_0) &= 0 \text{ for } t = 1, 2, \dots, T \\ E(\varepsilon_t X_0) &= 0 \text{ for } t = 1, 2, \dots, T \end{aligned} \quad (5)$$

---

<sup>4</sup>This specification allows regime-dependent heteroscedasticity, but autoregression coefficients are not regime-dependent.

If we stack the state variables in a  $3 \times 1$  vector  $X_t = (L_t, S_t, C_t)$ , The state space model can be succinctly written in matrix notations as

$$\begin{aligned} \mathbf{y}_t &= \Lambda X_t + \varepsilon_t \\ X_t &= \Phi X_{t-1} + \eta_{\xi_t}; \xi_t = H \text{ or } L \end{aligned} \tag{6}$$

A discrete Markov chain governs switches between two regimes, the transition matrix is given by

$$P = \begin{bmatrix} p & 1 - q \\ 1 - p & q \end{bmatrix} \tag{7}$$

Now the standard Hamilton (1989, 1994) and Krolzig (1997) algorithms can be used to extract probabilities staying in each regime.

I build this model upon a growing literature suggesting that regime-shifting models describe yields dynamics better than single regime model (Ang and Bekaert (2002), Garcia and Perron (1996), Gray (1996)). Two regimes are usually characterized by low and high volatility. The transitions between regimes are governed by a discrete-time Markov process. From the perspective of macroeconomics, these regimes are frequently related to business cycles. From the perspective of finance, they are connected with bond risk premium (e.g., Bansal and Zhou (2002), Dai, Singleton and Yang (2007)). Because the Nelson-Siegel models are heavily applied in central banks and investment institutions, my model tries to investigate the capability of the DRA dynamic Nelson-Siegel model on capturing regimes shifts.

The matrix  $\Lambda$  plays three roles in my analysis. Three yield factors and regimes are unobserved components in the system (6). As usual, there are some identification conditions that must be imposed to estimate a model with latent factors. The matrix  $\Lambda$  provides such identification restrictions. Since three yield factors explain most of variations in yields dynamics, they are supposed to be highly correlated with three principle components (Zhu (2008)). It also gives three latent yield factors a nice interpretation, respectively, the level, the slope and the curvature factor. These factors have empirical counterparts and are related to economic activities. In contrast, an unrestricted vector autoregression doesn't provide us such a clear interpretation. In addition, the restricted DRA dynamic Nelson-Siegel model seems to be stable over sample selection and set of yields chosen. This is the second role played by the matrix  $\Lambda$ . Admissibility (Dai and Singleton (2002)) constitutes a third role of the matrix  $\Lambda$ . As discussed in DRA, the Nelson-Siegel form avoids a negative forward rate at all horizons.

The entertained model achieves parsimony by diagonal  $\Omega$  assumption. Since three underlying latent factors explain a large fraction of yield variation<sup>5</sup>, the diagonal  $\Omega$  is expected to be a good approximation. Thus the efficiency loss from diagonal restrictions shouldn't be significant. This is an usual strategy, for example, Christensen, Diebold and Rudenbush (2008) show that this diagonal covariance model has good forecasting performance, it offers a more accurate prediction than non-diagonal model in many cases. For affine term structure model with no-arbitrage conditions, Ang and Piazzesi (2003) assume some yields are measured with errors. Computational tractability is a second reason for the diagonal covariance matrix assumption.

## 2.2 Yields

The yields are at a monthly frequency and the sample covers a period from January 1980 to March 2008. I examine U.S. treasury yields with maturities 1, 3, 6, 12, 24, 36, 60, 84, 120 months. The yields are retrieved from Econstats<sup>TM</sup>. One stylized fact of yields is that they tend to exhibit considerable persistence and are believed to be nonstationary or better approximated by an integrated process. This feature has profound implications for the macro-finance model estimation. Figure 1 plots the yields with different maturities. It is clear from the figure that yields are volatile over time. Furthermore, the yield curve shows a variety of shapes.

## 2.3 The Gibbs sampling

The state-space system Eqs.(6) is estimated by the Markov chain Monte Carlo (MCMC) method, specifically, a Gibbs sampling algorithm (see Appendix for details). Three main reasons account for our choice of a Bayesian method instead of the classical maximum likelihood estimation. First, in classical estimation, inference on the latent factors is conditional on the estimated parameters. In contrast, the Bayesian method describes the joint distribution of the latent yield factors, unobserved regimes and other parameters. It thus incorporates the parameters' variability.

Second, the reliability of the Bayesian inference is less dependent on the sample size of the real data. Even in a single equation regime-shifting regression, Monte Carlo experiment indicates (Psaradakis and Sola (1998)) that the conventional asymptotic approximations to the distribution of the maximum likelihood estimator are not good until the sample size approaches 800. For regime-shifting vector autoregression with

---

<sup>5</sup>The model can explain over 98% of the variance of yield changes (Diebold and Li (2006), Diebold, Rudebusch and Aruoba (2006)).

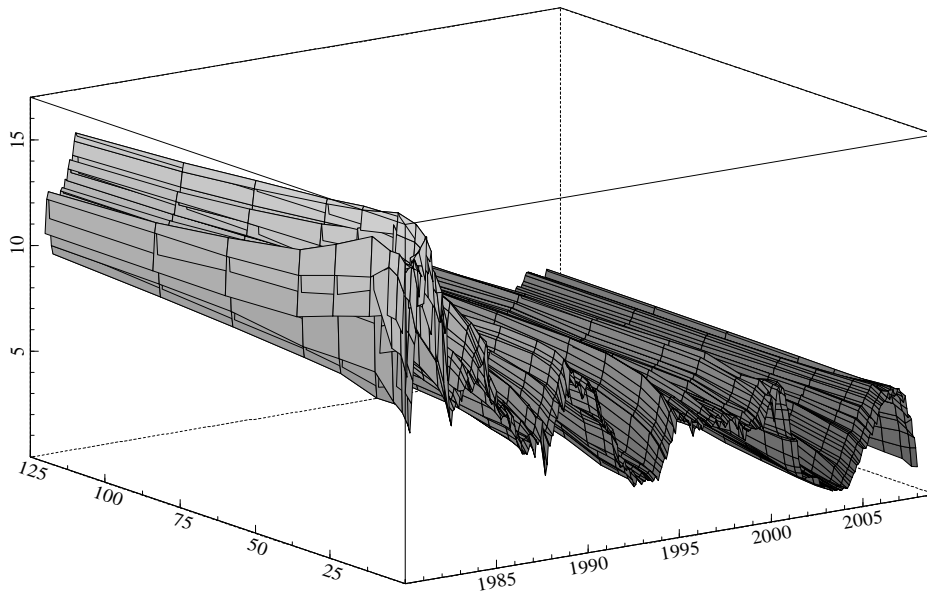


Figure 1: U.S. Yield Curves, 1980.01–2008.03

large number of parameters, the reliability of asymptotic theory is problematic with our sample size. With the Bayesian method, however, the size of the sample is under control of researcher.

Third, one shortcoming of the maximum likelihood estimation inspires the use of the MCMC. Kim (1994) and Kim and Nelson (1999) provide an approximation method and make the maximum likelihood estimation of the state-space models with regime switching feasible. However, the properties of approximation method is unknown. In some cases, the accuracy provided by the approximation method is probably not good enough. Furthermore, for high dimensional model like macro-finance model in the next section, the likelihood function may be subject to multiple local optima.

## 2.4 Convergence checks

Some diagnostics on the reliability of the MCMC are available. The basic idea of most convergence statistics is to compare moments of the sampled parameters. A visual check on the plot of sampled parameters can provide information about the convergence. For a converged MCMC implementation, the drawings shouldn't deviate from some mean for a long period. Although this is subjective in the sense that there is no clear measure of deviation and duration. To further access the convergence of

the Gibbs sampling, I implement another two practical statistics. The first statistic is Yu and Mykland (1998) plot of CUSUM path, for a specific parameter  $\theta$  with sample variance  $\sigma_\theta^2$  and mean  $\mu_\theta$  for the sample up to iteration  $t$ ,

$$CUSUM_t = \frac{1}{t\sigma_\theta} \sum_{i=1}^t (\theta^i - \mu_\theta), \quad t = 1, 2, \dots, T$$

This method is intuitive, if CUSUM diverges from zero for a prolonged period, it is an indication of non-convergence. Therefore a visual check on the CUSUM plot provides us information about the convergence of Gibbs sampling.

The second criterion of the convergence is the relative numerical efficiency (RNE) proposed by Geweke (1992). As the drawings for the latents factors and unobserved regimes are from a serially correlated distribution, the RNE shed light on the efficiency of the Gibbs sampling since the RNE measures the quality of a correlated sample. The rationale of the RNE is to compare the empirical variance with the Newey-West (Newey and West 1987) heteroscedasticity and autocorrelation consistent variance  $\sigma_{NW,q}^2$ ,

$$RNE = \frac{\sigma_\theta^2}{\sigma_{NW,q}^2}$$

where  $q$  is the length of the Barlett window for the Newey-West estimator.

## 2.5 Empirical results

Diebold and Li (2006, 2008) fix the  $\lambda$  and set it equal to a value that maximize the loading on the curvature factor at 30 months. As three yield factors are time-varying, the dynamic Nelson-Siegel model can generate a variety of yield curve shapes, such as, upward-sloping, inverted, hump, and S shapes. The estimated model explains the main stylized facts regarding the yield curve. Yield forecasts based on the entertained model produce encouraging results, especially at long horizon. The dynamic Nelson-Siegel model beats various benchmark models in terms of predictive power. I stick to this tradition by fixing  $\lambda$  at a value of 0.0598<sup>6</sup> and expect little loss of generality from fixing  $\lambda$  at a constant. The simplified version makes the MCMC simulation feasible, otherwise it is a challenge to draw this parameter because the conditional posterior distribution function of  $\lambda$  doesn't correspond to a well-know distribution.

---

<sup>6</sup>My calculation is based on maximization of the curvature factor at 30 months. In Diebold and Li (2006), it is set equal to 0.0608, but they define a month as 30.4375 days, this day-counting scheme might account for the difference.

Table 1 About Here

For a large-scale dynamic factor model, the Bayesian method is preferred to the classical maximum likelihood due to the aforementioned reasons. There is a large number of parameters from the Bayesian approach perspective because the latent yield factors and unobserved regimes are all seen as parameters in a Bayesian estimation. Given the yield factors and regimes, there are twenty-six parameters to estimate: 3 parameters in diagonal variance-covariance matrix  $\Omega$ ; for each regime, 6 parameters in non-diagonal covariance matrix  $\Sigma_{\xi_t}$ ; 9 parameters in autoregressive coefficient matrix  $\Phi$ ; and 2 regime transition parameters.

The details of the Gibbs sampling are presented in the Appendix. Three yield factors are drawn based on the multi-move Gibbs sampling algorithm (Carter and Kohn 1994) where the entire conditional posterior distributions are from other parameters and the Kalman filter. This method simplifies the MCMC simulation because we can draw yield factors jointly by a recursive method. Specifically, we use the Kalman filter to process yields forward, then we take random draws of the posterior distributions backward. This forward filtering and backward sampling (FFBS) method make the simulation more efficient because this scheme draws serially correlated yield factors jointly. Using the FFBS scheme combined with the Hamilton (1989, 1994) filter, we can also generate the unobserved regimes prevailing at each time point  $t$ . The smoothed regimes are usually parameters of interest, The FFBS scheme combined with Kim (1994) filter produces drawings of the smoothed regimes.

To facilitate the convergence of the Gibbs sampling iterations, I initialize the MCMC by a two-step estimation. The first step runs the OLS to estimate yield factors by fixing  $\lambda$  at 0.0598. With estimated yield factors, the state equation (3) can be estimated by the Gaussian maximum likelihood method where we get the autoregressive parameters, regime probabilities and transition probabilities. These parameters from the two-step estimation are catered to the MCMC scheme. This initialization makes the Gibbs sampling converged quickly. I also try other initials for yields-only model, they produce similar results.

The two-step estimation indicates that the state equation is stationary since all roots of the autoregressive coefficient matrix are smaller than one. However, the Gibbs sampling drawings are usually nonstationary after dozens of iterations. I don't drop nonstationary iterations and the entertained state equation is nonstationary. In contrast, drawings of the macro-finance model are usually stationary. Thus, for the macro-finance model, I control the nonstationary iterations by dropping them to facilitate the calculation of impulse response functions. I simulate 10000 iterations with an initial

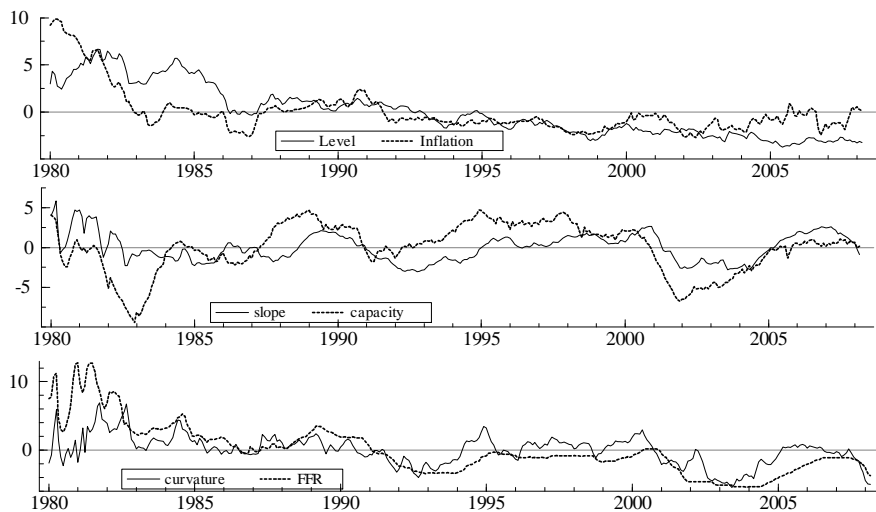


Figure 2: Level, slope, curvature factors and the empirical counterparts.

burn-in period of 7000 observations. All three measures of convergence, respectively, visual plot of parameters, CUSUM and NW, indicate that the sampled parameters are converged.

Table 1 presents the parameter estimates of the yields-only model. The estimates of the autoregressive coefficient matrix  $\Phi$  contain much information. First, three latent yield factors are highly persistent. This is consistent with typical results found in the term structure modeling. Second, there is some difference in the time-series properties of the yield factors. It seems that  $C_t$  is the most persistent factor, and  $S_t$  is the least persistent factor. This contrasts to the evidence typically found in empirical studies where  $L_t$  is most persistent and  $C_t$  is least persistent. Third, cross-correlations across yield factors are small but still significant.

The extracted level, slope and curvature are plotted in figure 2. For the purpose of comparison, the empirical counterparts of three yield factors are depicted in the same figures. The empirical level factor is defined as the 10-year yield. The proxy for the empirical slope factor is the difference between the 10-year yield and 3-month yield. The empirical curvature is the twice the 2-year yield minus the sum of the 10-year and 3-month yields. The correlation among the extracted factors and the empirical factors are respectively 0.99 for the level, 0.88 for the slope and 0.78 for the curvature. The correlation analysis indicates why the latent factors are labeled as ‘level’, ‘slope’, and ‘curvature’.

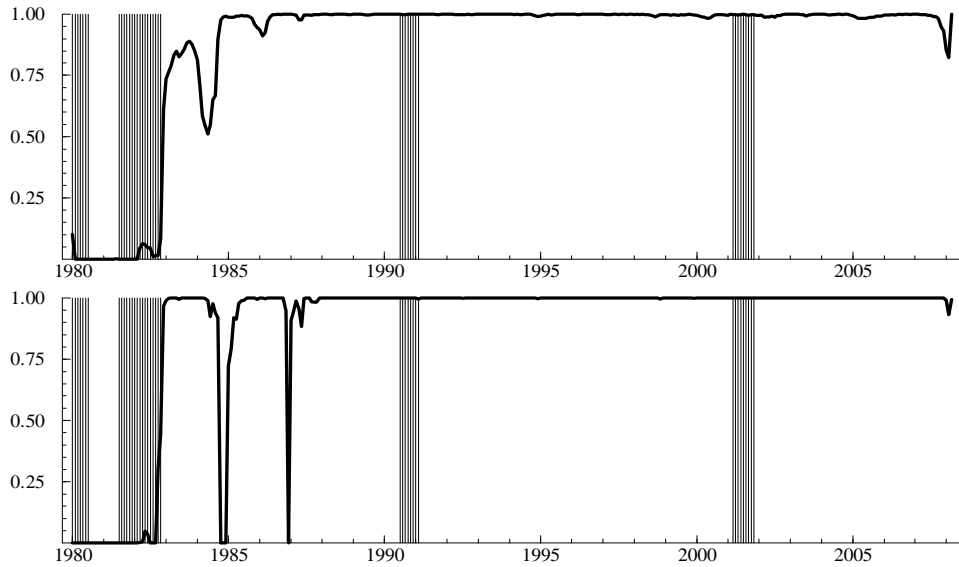


Figure 3: Probabilities of being in a tranquil regimes, the shaded bars indicate the NBER recessions.

Table 2 reports the fits of the yields-only model and macro-finance models. For each model, I present the estimated means and standard deviations of the measurement equation residuals. It seems that both models fits the yield curve well. Meanwhile, it is important to note one sailent feature of the model that it fits the middle region of the yield curve best. In particular, the fitting errors at the short end of the yield curve is singificant.

#### Table 2 About Here

The upper panel of Figure 3 plots the smoothed regimes of being in low volatility regime. In affine term structure models, regimes are typically labeled as high and low volatility. In this analysis the filtered regimes still have such a clear interpretation as is clear from  $\Sigma_L$  and  $\Sigma_H$  in Table 1. Clearly form the plot, the regime identification conincides the Fed operating procedure shift in 1982. This is no surprise since the entire yield curve responds to the conduct of the yield curve (see, for example, Ang, Boivin and Dong 2007). Furthermore, The regime identification captures the two relatively consecutive economic recessions in early 1980s. Unfortunately, it neglects two economic recessions in 1990 and 2001. This possibly implies that a two-regime model may be not enough to capture business cycles and monetary policy shifts in one system. It necessitates a more regimes model (Garcia and Perron (1996)) to describe the term structure of interest rates. This conjecture constitutes an interesting future research

agenda.

### 3 Macro-finance model

This section tries to shed light on the joint dynamics of the yield curve and economic activity that incorporates an implicit monetary policy rule. For modeling interest rates, the yields-only model provides a good description of the yield curve on the cross-section and time series. For other purposes, for example, monetary policy modeling and economic activity forecasting, we need relate yield factors to macroeconomic variables. The conduct of monetary policy shifts the short-end of the yield curve, through risk-adjusted expectations, it further shift the long-end of the yield curve. According to the Taylor (1993) rule, the Fed sets short interest rates by responding to the output gap and inflation. The key intersection of macroeconomic dynamics and the yield dynamics is short-term interest rate. The yields-only model has a missing motivation that the Fed ignores the information from economic activity or the bond market ignores the information from the Fed. This section extends the yields-only model by including macroeconomic variables. The extended macro-finance model is estimated using the MCMC and result analysis is reported.

#### 3.1 Macroeconomic factors

Three proxies for economic activity are the capacity utilization (CU), the federal fund rate (FFR) and inflation. The CU, FFR and consumer price index (CPI)<sup>7</sup> are retrieved from the economic database, Federal Reserve Bank of St. Louis. The year-over-year inflation rate is defined by taking the yearly percentage change in the CPI index,

$$\pi_t = 100 * (\ln CPI_t - \ln CPI_{t-4})$$

The capacity utilization is a measure of the deviation of economic activity from its natural level. For modeling business cycles and monetary policy, quarter is a typical frequency. At a quarterly frequency, the GDP is an obvious proxy for economic activity. Alternatively, this study exploits the availability of monthly data. In so doing, I try to characterize the relationship between the yield curve and economic activity at a higher frequency. The inflation is included in the extended macro-finance model because it is

---

<sup>7</sup>On the database, three variables are labeled, respectively, as "the total industry capacity utilization", "the effective federal fund rate" and "consumer price index for all urban consumers: all items"

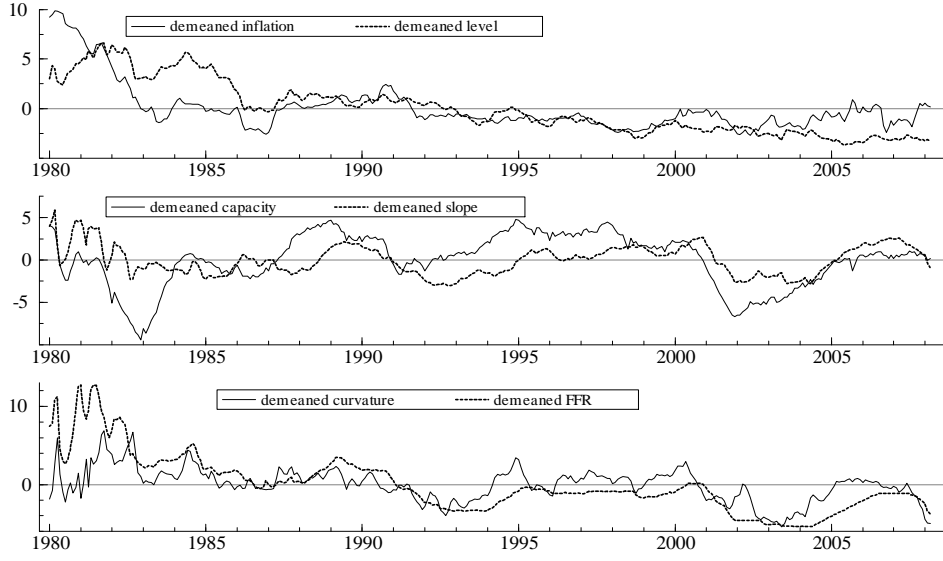


Figure 4: Macro factors and level, slope, curvature factors from the macro-finance model.

a key variable in shaping nominal yield curve through the inflation risk premium (Ang, Bekaert and Wei (2008)) and in making monetary policy, such as the Taylor principle (Taylor (1993)). The federal fund rate is a monetary policy instrument that moves the yield curve. The selection of macroeconomic variables is consistent with the DRA model that is the foundation of regime-shifting macro-finance model. I consider several other variables, such as average weekly hours at a monthly frequency, they have similar implications for regime identification purpose.

### 3.2 The macro-finance model and estimation

It is straightforward to extend the yields-only model by adding macroeconomic variables to the information set. Let the  $6 \times 1$  vector  $X_t^{MF} = (L_t, S_t, C_t, FFR_t, \pi_t, CU_t)'$  be factors in the macro-finance model, then the state equation is

$$\begin{bmatrix} L_t - \mu_L \\ S_t - \mu_S \\ C_t - \mu_C \\ FFR_t - \mu_{FFR} \\ \pi_t - \mu_\pi \\ CU_t - \mu_{CU} \end{bmatrix} = \Phi^{MF} \begin{bmatrix} L_{t-1} - \mu_L \\ S_{t-1} - \mu_S \\ C_{t-1} - \mu_C \\ FFR_{t-1} - \mu_{FFR} \\ \pi_{t-1} - \mu_\pi \\ CU_{t-1} - \mu_{CU} \end{bmatrix} + \begin{bmatrix} \eta_{\xi_t,1t} \\ \eta_{\xi_t,2t} \\ \eta_{\xi_t,3t} \\ \eta_{\xi_t,4t} \\ \eta_{\xi_t,5t} \\ \eta_{\xi_t,6t} \end{bmatrix} \quad (8)$$

where subscript  $\xi_t = H$  or  $L$ , denote high and low volatility regimes<sup>8</sup>. With disturbances in regime  $L$  and  $H$  follow the Gaussian distribution.

$$\begin{aligned}\eta_L &\sim N(0, \Sigma_L^{MF}) \\ \eta_H &\sim N(0, \Sigma_H^{MF})\end{aligned}\tag{9}$$

The observation equation is all the same as in the state-space system Eq. (2)<sup>9</sup>. The macro-finance model maintains the assumption of a diagonal covariance matrix  $\Omega$ . The state equation Eq. (8) is subject to regime shifts, and the standard Hamilton filter can be used to extract regimes. The state-space model for the macro-finance model can be succinctly represented by

$$\begin{aligned}y_t &= \Lambda X_t^{MF} + \varepsilon_t \\ X_t^{MF} &= \Phi^{MF} X_{t-1}^{MF} + \eta_{\xi_t}; \xi_t = L \text{ or } H\end{aligned}\tag{10}$$

This is a large-scale dynamic model. Even if we don't take into account the latent yield factors and unobserved regimes, there are still 89 parameters: the autoregressive coefficient matrix includes 36 parameters; the covariance matrix  $\Sigma_L$  and  $\Sigma_H$  respectively have 21 parameters; 2 parameters in transition matrix  $P$ ; and 9 parameters in the diagonal matrix  $\Omega$ . The Gibbs sampling method is a preferred method for estimating the state-space system Eqs.(10). The initial values are provided by the two-step estimation: the OLS regression of the observation equation and the Markov-switching regression of the state equation. In MCMC, I drop all non-stationary drawings to ensure that the estimated system is stationary. In particular, my implementation consists of 8000 stationary iterations, the number of burn-in iteration is 5000. The CUSUM and RNE indicate the convergence of the estimation.

Three latent factors are all persistent although the degree of persistence differs. The autoregressive coefficient is 1.002<sup>10</sup> for the most persistent level factor, while for the least persistent curvature factor it is 0.894. These results are consistent with those typically found in the empirical term structure models. This finding contrasts to the yields-only model where the curvature factor is most persistent. There are significant

---

<sup>8</sup>This intuitive label is a little abuse in notation. However our empirical results justify two regimes with high and low volatility.

<sup>9</sup>In DRA's representation [Eq.(6') in DRA], the state vector is a  $6 \times 1$  vector for their yields-macro model, but they set the three rightmost columns all equal zeros. This setting implies that the yields are still priced only by three yield factors.

<sup>10</sup>Although the coefficient is 1.002, the estimated system is stationary with the biggest eigenvalue of the matrix  $\Phi^{MF}$  is 0.98.

cross-correlations among the yield factors and the macro factors. The parameters in the lower-right ( $3 \times 3$ ) sub-matrix of  $\Phi^{MF}$  are significant except one parameter. It is important to note that the yield factors play an important role in accounting for the macroeconomic dynamics. In contrast, macro factors are less important in explaining the yield curve. It is not surprising since three-factor term structure models is a good approximation for the joint dynamics of yields.

Table 3 About Here

Two regimes continue to be labeled as L and H regimes according to the estimation results. It is clear that the residual variances of the yield and macro factors in regime L are significantly less than those in regime H. Statistically, the Wald test in the table 3 rejects the null hypothesis of equal variance in two regimes. The Wald statistics also indicate that neither the covariance matrix  $\Sigma_L^{MF}$  nor the matrix  $\Sigma_H^{MF}$  are diagonal. The lower panel of Figure 3 plots the smoothed probabilities of being in regime L. This is a similar result as regimes from the yields-only model. The 1980-1982 period is turbulent, thereafter all regimes are identified as tranquil age. This may correspond to the frequent switch between the economic recession and booms. It is also possible that the Fed operating procedure change in 1982 plays a role. The plotted regime identification misses three components: the Fed operating procedure change in 1988, the NBER economic recessions in 1990 and 2001. To recover the missing information, a macro-finance model allowing more than 2 regimes might be a better choice.

Although the autoregressive coefficient matrixes are different<sup>11</sup>, similar filtered time series of the level, slope and curvature factors are obtained from the yields-only model and the macro-finance model. The correlations round to 1 for pairs of levels and slopes, for curvatures, it is 0.99. Figure 4 depicts the yield factors from the macro-finance model with three macroeconomic factors. It is clear from the graph that the yield factors are closely linked to the macroeconomic factors with the highest correlation 0.66 between the curvature and FFR. The level factor is correlated with the inflation with a correlation 0.65. To a less extent, the correlation between the slope and the capacity utilization is 0.39.

### 3.3 Impulse responses

It is intuitive and interesting to examine the factor impulse responses. Figure 5 and 6 report the impulse responses of the yield and macro factors on each other in regime L and H. The Cholesky decomposition of the non-diagonal covariance matrixes  $\Sigma_L^{MF}$  and

---

<sup>11</sup>The matrix  $\Phi$  in Table 1 and the upper-left ( $3 \times 3$ ) sub-matrix of  $\Phi^{MF}$  in Table 3.

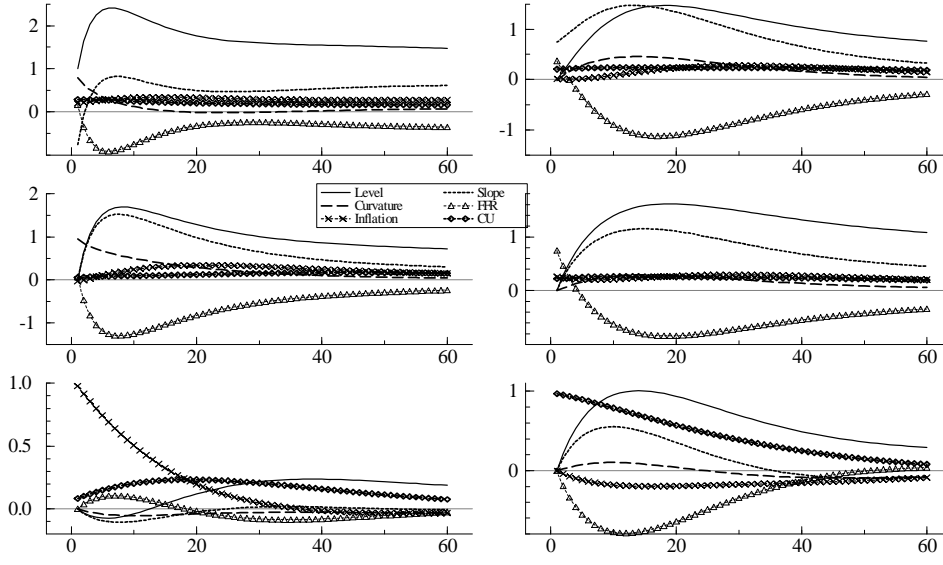


Figure 5: Impulse responses of yield and macro factors on each other under regime L.

$\Sigma_H^{MF}$  is based on the ordering  $(L_t, S_t, C_t, FFR_t, \pi_t, CU_t)$ . Each response is measured in terms of one percentage point shock to residuals. I consider two classifications of the impulse responses. One focus of the macro-finance model is regime shift, thus the first classification is to compare the impulse responses in regime L and H. According to another focus of the macro-finance model, that is, linkages among the yield and macro factors, I split the impulse responses into four groups as in DRA: macro-to-macro responses, macro-to-yield responses, yield-to-yield responses and yield-to-macro responses.

As is clear in figures, the IRs only have marginal difference in two regimes in terms of the shape of impulse responses curves, but they are different in terms of magnitude. There are volatile and stable periods in financial markets, and the economy goes through booms and recessions. Yet these fluctuations don't significantly change the relationship of factors on each other. The macro factors usually respond to shocks to the yield factors in both regimes, to a less extent, the yield factors respond to shocks to the macro factors. In turbulent periods, the IRs are more significant as indicated by the scale of y-axis.

In the group of the yield factors, there is no initial response of the level factor to the slope factor. On subsequent periods, this response rises and keeps persistently at that level. We can see that the level factor is the most persistent factor with respect to the IR. The slope and curvature factors respond to shocks to the yield factors, but

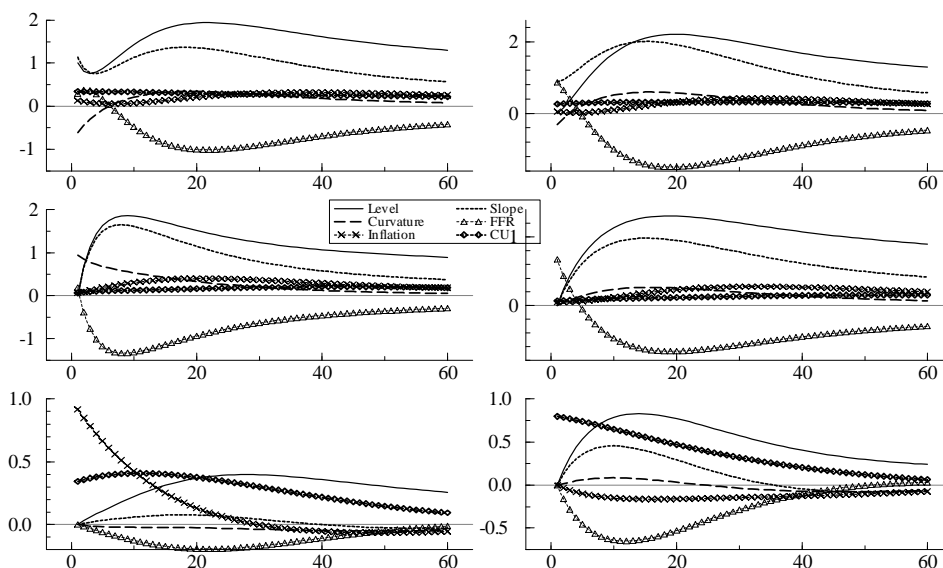


Figure 6: Impulse responses of yield and macro factors on each other under regime H.

they usually decays to near zero quickly.

At the initial stage, the FFR and the capacity utilization strongly respond to shocks to the level. As time lapses, the capacity utilization responses fall down rapidly to a lower level. It implies that the yield level mainly affect the capacity utilization in the short-run. Similarly, the effect of the slope on the capacity utilization disappears quickly. Contrary to my intuition, the inflation doesn't respond significantly to three yield factors.

On the other hand, there are consistently weak responses of the yield factors to shocks to inflation and capacity utilization. This is not surprising since three factors can explain most of variation of yields. As a Fed instrumental variable, the FFR has effect on the yield factors on the medium-run. In the group of the macro factors, all macro factors respond to shocks to the FFR. The monetary policy hence has effects on the inflation and economic activity. Inflation shocks affect two factors: the inflation own and the capacity utilization. To a least extent, all other factors don't significantly respond to shocks to the capacity utilization.

### 3.4 Testing interactions across the yield and macro factors

There are three interesting null hypothesis about interactions across the yield and macro factors. The first hypothesis is totally no interaction among the yield factors

and the macroeconomic variables. A less strong assumption is the dynamics of the yield factors do affect the dynamics of the macro factors, but not vice versa. Opposed to the second assumption, the last hypothesis postulates that the unidirectional linkage is from the macro factors to the yield factors.

Following DRA, three hypotheses can be formalized by zero restrictions on the autoregressive matrix and the variance-covariance matrix of the state equation. Specifically, we partition the  $(6 \times 6)$  matrix  $\Phi^{MF}$  into four  $(3 \times 3)$  blocks

$$\Phi^{MF} = \begin{bmatrix} \Phi_1^{MF} & \Phi_2^{MF} \\ \Phi_3^{MF} & \Phi_4^{MF} \end{bmatrix} \quad (11)$$

and similarly partitioning the covariance matrixes  $\Sigma_H^{MF}$  and  $\Sigma_L^{MF}$

$$\Sigma_{\xi_t}^{MF} = \begin{bmatrix} \Sigma_1^{MF} & \Sigma_2^{MF} \\ \Sigma_3^{MF} & \Sigma_4^{MF} \end{bmatrix}; \quad \xi_t = L \text{ or } H \quad (12)$$

where  $\Sigma_3^{MF}$  is the transpose of the  $\Sigma_2^{MF}$ . Given the prevailing regime,  $\Phi_2^{MF} = \Phi_3^{MF} = \Sigma_2^{MF} = 0$  is the equivalence of the first null hypothesis. The second hypothesis can be rewritten as  $\Phi_2^{MF} = 0$ . The restrictions for the third hypothesis are  $\Phi_3^{MF} = \Sigma_2^{MF} = 0$ . The Wald test is easily implemented for testing these hypotheses. Table 4 displays the Wald statistics for three hypotheses in regime L and H. All hypotheses are overwhelmingly rejected, thus we can't exclude the bidirectional linkages. Overall, this finding is consistent with a growing literature that relates the term structure of interest rates with economic activity.

Table 4 About Here

### 3.5 Implication of the expectations hypothesis

The expectations hypothesis states that the long-term yield equals to a weighted average of future expected short-term yields plus a constant term premium. It is a benchmark model of determining long-term yields. For example, in modern term structure models long-term yields are usually a risk-adjusted average of future short-term yields<sup>12</sup>. It is interesting to relate the regime-shifting macro-finance model to the expectations hypothesis and see what is implication for the expectations hypothesis.

In last decades, a lot econometric methods and techniques for evaluating the expectations hypothesis have been developed and applied. Among these methods, one

---

<sup>12</sup>For survey, refer to Piazzesi 2004.

influential framework is a bivariate model (Campbell and Shiller (1987)) based on the present value model that links the  $\tau$ -period yield  $y_{t(\tau)}$  with the expected one-period yield  $y_{t(1)}$ ,

$$y_{t(\tau)} = (1 - \delta) \sum_{i=0}^{\tau-1} \delta^i E_t y_{t+i(1)} + c_\tau \quad (13)$$

where  $c_\tau$  is a maturity-dependent constant,  $\delta$  is the discount factor that reflects the impatience of economic agents and  $E_t$  is the conditional expectation based on the information set at time  $t$ . From section 2.2, the level factor has the interpretation of the long-term yield, on the other hand, minus slope factor represents the yield spread between 3-month and 10-year yields. To shed light on the expectations hypothesis as an approximation on the entire yield curve, I circumvent the use of a pair of long- and short-term yields. Alternatively, the short rate is defined equal to  $L_t + S_t$  as in Carriero, Favero and Kaminska (2006), and the long yield is  $L_t$ . Then, we have

$$L_t = (1 - \delta) \sum_{i=0}^{\tau-1} \delta^i E_t (L_{t+i} + S_{t+i}) + c_\tau \quad (14)$$

Suppose that the data generating process is given by the state-space system Eqs.(10), the implication of the Eq.(14) is

$$L_t = (1 - \delta) h' (I - \delta \Phi^{MF})^{-1} X_t^{MF} \quad (15)$$

where  $h' = [1, 1, 0, \dots, 0]$  is a  $(6 \times 1)$  selecting vector such that  $L_t + S_t = h' X_t^{MF}$ . Eq.(15) is the theoretical level factor implied by the expectation hypothesis. This section concentrates on the comparison of the actual and theoretical level factors. This tells us how well the expectations hypothesis approximates the observed yield curve, or in terminology of Campbell and Shiller (1987), the economic significance<sup>13</sup>. Figure 7 plots two series of the theoretical and actual levels<sup>14</sup>. The theoretical level frequently predicts the directions of the actual level factor with a correlation of 0.92, but it is more volatile than the actual level with a volatility ratio of 1.96.

---

<sup>13</sup>I do not test the expectations hypothesis in this study since the focus in this study is the adequacy of the expectations hypothesis, not the statistical rejection or non-rejection. Carriero, Favero and Kaminska (2006) conducted the recursive and rolling-window tests of the expectation hypothesis based on a simulation method. Zhu (2008) took into account the regime shift and provided some supporting evidence on the expectation hypothesis.

<sup>14</sup>The discount rate is set to equal to  $1/(1 + L_t/12)$ . I plot  $TL_t/1.2$  in the figure (6),  $TL_t$  is the theoretical level.

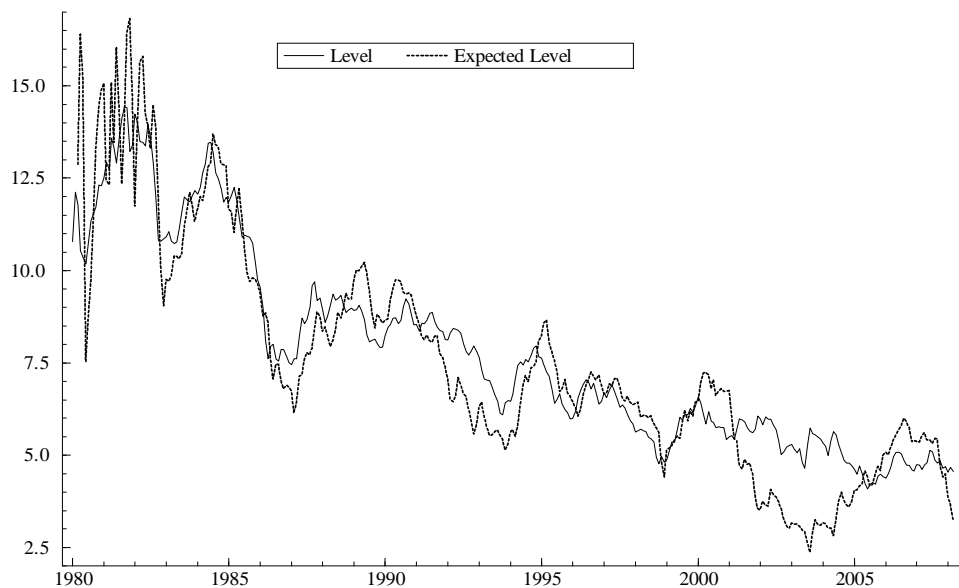


Figure 7: Theoretical and actual levels.

## 4 Conclusions

I have presented and estimated a macro-finance model subject to regime shifts. This approach is inspired by a stylized fact of the term structure of interest rates, that is, existence of the turbulent and tranquil periods in fixed-income securities markets. The DRA state-space representation of the model facilitates the estimation and extraction of the latent yield factors. The proposed macro-finance model allows bidirectional feedback across the yield factors and the macro factors. The formal tests provide strong evidence in favor of the interactions among the yield and macro factors. This conclusion is robust across the high volatility and low volatility regimes.

The empirical results indicate that the existence of two regimes is stable across the yields-only model and the macro-finance model. In early 1980s, the economy keeps switching between the booms and recessions, which is identified as a turbulent period. It coincides with the Fed operating procedure change in 1982. Not as usual, two regimes are not clearly related to the business cycles. My conjecture is that to capture all effects, including both the economic activity and the Fed operating procedure changes, a macro-finance model with more regimes is necessary. This hence constitutes an interesting future research agenda.

## Appendix: Gibbs Sampling

**(1) Generation of coefficient matrix  $\Phi$ ;** Assume the prior distribution of  $vec(\Phi)$  is a normal distribution  $N(a_0, \Omega_0)$ , conditional on all observed yields (and macro factors for the macro-finance model)  $Y_T$  and other parameters  $\Psi_{-\Phi}$ , the posterior distribution of  $vec(\Phi)$  is also a normal distribution  $N(a_1, \Omega_1)$ , with

$$\begin{aligned} a_1|Y_T, \Psi_{-\Phi} &= a_0[\Omega_0^{-1}a_0 + U'W] \\ \Omega_1|Y_T, \Psi_{-\Phi} &= \Omega_0^{-1} + U'U \end{aligned}$$

For simplifying the expressions of  $U$  and  $W$ , we define

$$V_i = (I_T \otimes \Sigma_{\xi_t}^{-1/2})$$

with  $\xi_t = H, L$  represents a high or low volatility regime and  $I$  is identity matrix with dimension  $T$ . Furthermore, let  $Z$  be

$$Z_i = \xi \otimes \iota_k$$

where  $\iota_k$  is a column vector of 1s,  $\xi$  is a matrix from the Hamilton filter consisting of the probabilities in each regime. Moreover,

$$U = V_0(Y_T \otimes I_k) \odot Z_0 + V_1(Y_T \otimes I_k) \odot Z_1$$

and

$$W = V_0vec(Y_T) \odot Z_0 + V_1vec(Y_T) \odot Z_1$$

As usual,  $\otimes$  is Kroneck product and  $\odot$  is element-by-element multiplication. This derivation is based on the multivariate least squares.

**(2) Generation of regimes  $\xi$ ;** I use the multimove Gibbs sampling method to generate regimes. Based on Carter and Kohn (1994), Kim and Nelson (1999) partition the joint distribution of regimes  $\xi$  conditional on  $Y_T$  and other generated parameters  $\Psi_{-\xi}$ ,

$$g(\xi|Y_T, \Psi_{-\xi}) = g(\xi_T|Y_T, \Psi_{-\xi}) \prod_{t=1}^T g(\xi_t|\xi_{t+1}, Y_t, \Psi_{-\xi})$$

The forward filtering and backward sampling (FFBS) approach therefore can be applied in two steps. The first step is to run Hamilton's (1989) filter to get filtered probabilities

$g(\xi_t|Y_t, \Psi_{-\xi})$ . The last iteration of the filter is exactly  $g(\xi|Y_T, \Psi_{-\xi})$ , from which  $\xi_T$  is generated with a uniform distribution generator. The second step is to generate  $\xi_t$  conditional on  $\xi_{t+1}$  and  $Y_t$ . We can make use of the following result:

$$g(\xi_t|\xi_{t+1}, Y_t, \Psi_{-\xi}) \propto g(\xi_{t+1}|\xi_t)g(\xi_t|Y_t, \Psi_{-\xi})$$

combined with the matter of fact that  $g(\xi_{t+1}|\xi_t)$  is the transition probability and  $g(\xi_t|Y_t, \Psi_{-\xi})$  has been provided by the Hamilton filter, we have

$$g(\xi_t = 1|Y_t) = \frac{g(\xi_{t+1}|\xi_t = 1)g(\xi_t = 1|Y_t, \Psi_{-\xi})}{\sum_{j=0}^1 g(\xi_{t+1}|\xi_t = j)g(\xi_t = j|Y_t, \Psi_{-\xi})}$$

Then we can generate all regimes recursively.

**(3) Generation of state variables  $X_t$ ;** For generating the state vector, we still employ the FFBS approach. Kim and Nelson (1999) employ Carter and Kohn's multi-move Gibbs sampling method and provide the partition of joint distribution. There are also two steps like in the generation of regimes  $\xi_t$ , but we run the Kalman filter instead of the Hamilton filter. Given the measurement equation (2) and the state equation (3), the  $X_T$  have a conditional normal posterior distribution:

$$X_T|Y_T \sim N(\mathbf{x}_{T|T}, P_{T|T})$$

where  $X_{T|T}$  is the conditional expectation of  $X_T$  from the last step of the Kalman filter.  $P_{T|T}$  is the covariance matrix of  $X_{T|T}$ . For simplification, I suppress the  $\Psi_{-\mathbf{x}}$  in the conditional information set. Consequently, we have

$$X_{t-1}|X_t, Y_{t-1} \sim N(X_{t|t, X_{t+1}}, P_{t|t, X_{t+1}})$$

with

$$X_{t|t, X_{t+1}} = X_{t|t} + P_{t|t}\Phi'(\Phi P_{t|t}\Phi' + \Sigma_t)^{-1}(X_{t+1} - \mu - \Phi X_{t|t})$$

and

$$P_{t|t, X_{t+1}} = P_{t|t} - P_{t|t}\Phi'(\Phi P_{t|t}\Phi' + \Sigma_t)^{-1}\Phi P_{t|t}$$

as shown in Kim and Nelson (1999, pp. 193). In this case,  $\Sigma_t$  is a weighted average of  $\Sigma_0$  and  $\Sigma_1$ . Specifically

$$\Sigma = \Pr(\xi_t = 0)\Sigma_0 + \Pr(\xi_t = 1)\Sigma_1$$

**(4) Generation of diagonal covariance matrix  $\Omega$ ;** Since  $R$  is diagonal, it can be generated element-by-element. Assume  $\sigma_i^2$ , the  $i$ -th element in the diagonal of  $R$ , has an inverted Gamma prior distribution,  $\sigma_i^2 \sim IG(v_0/2, \delta_0/2)$ , the posterior distribution of  $\sigma_i^2$  is still an inverted Gamma distribution,  $\sigma_i^2 \sim IG(v_1/2, \delta_1/2)$ , with

$$v_1 = v_0 + T$$

and

$$\delta_1 = \delta_0 + (\mathbf{y}_i - \mathbf{x}_i\Phi_i)$$

where  $\mathbf{y}_i$ ,  $\mathbf{x}_i$  and  $\Phi_i$  are respectively appropriate column of  $Y_T$ ,  $X_t$  and  $\Phi$ .

**(5) Generation of non-diagonal covariance matrix  $\Sigma_0$  and  $\Sigma_1$ ;** The covariance matrix is sampled from the inverted Wishart distribution. With an informative prior, the posterior distribution of covariance matrix (Chib and Greenberg 1996) follows

$$\Sigma_{\xi_i} | Y_T, \Psi_{-\Sigma} = IW(T_i, \sum_{t=0}^{t_i} \eta_t' \eta_t)$$

where  $i$  denotes two regimes.

**(6) Generation of transition probabilities  $p$  and  $q$ ;** The conjugate prior distribution for  $p$  and  $q$  is a beta distribution.

$$\begin{aligned} p &\sim \text{beta}(u_{11}, u_{10}) \\ q &\sim \text{beta}(u_{00}, u_{01}) \end{aligned}$$

as discussed in Kim and Nelson (1999, pp. 214-215), the posterior distributions are

$$\begin{aligned} p &\sim \text{beta}(u_{11} + n_{11}, u_{10} + n_{10}) \\ q &\sim \text{beta}(u_{00} + n_{00}, u_{01} + n_{01}) \end{aligned}$$

with  $n_{ij}$  refer to the transitions from stat  $i$  to  $j$ , which can be calculated by counting the generated regimes  $\xi_T$ .

## References

- [1] Ang, A., Bekaert, G., 2002. Regime switches in interest rates. *Journal of Business and Economic Statistics* 20, 163-182.
- [2] Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* 50, 745-787.
- [3] Ang, A., Piazzesi, M., Wei, M., 2006. What does the yield curve tell us about GDP growth? *Journal of Econometrics* 131, 359-403.
- [4] Bansal, R., Tauchen, G., Zhou, H., 2004. Regime shifts, risk premium in the term structure, and business cycles. *Journal of Business and Economic Statistics* 20, 396-409.
- [5] Bansal, R., Zhou, H., Term structure of interest rates with regime shifts. *Journal of Finance* 57, 1997-2043.
- [6] Campbell, J., Shiller, R., 1987. Cointegration and tests of present value models. *Journal of Political Economy* 95, 1062-1088.
- [7] Carriero, A., Favero, C.A., Kaminska, I., Financial factors, macroeconomic information and the expectations theory of the term structure of interest rates. *Journal of Econometrics* 131, 339-358.
- [8] Cater, C.K., Kohn, R., 1994. On Gibbs sampling for state space models. *Biometrika* 81, 541-553.
- [9] Chib, S., Greenberg, E., 1994. Bayes inference in regression models with ARMA(p,q) errors. *Journal of Econometrics* 64, 183-206.
- [10] Christensen, J.H.E., Diebold, F.X., Rudebusch, G.D., 2008. An arbitrage-free generalized Nelson-Siegel term structure model. *The Econometrics Journal*, forthcoming.
- [11] Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 1943-1978.
- [12] Dai, Q., Singleton, K.J., Yang, W., 2007. Regime shifts in a dynamic term structure model of U.S. treasury bond yields. *Review of Financial Studies* 20, 1669-1706.

- [13] Dewachter, H., Lyrio, M., 2003. Macro factors and the term structure of interest rates. *Journal of Money, Credit, and Banking* 38, 119-140.
- [14] Diebold, F.X., Li, C., 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* 130, 337-364.
- [15] Diebold, F.X., Li, C., Yue, V., 2008. Global yield curve dynamics and interactions: a generalized Nelson-Siegel Approach. *Journal of Econometrics*, 146, 351-363.
- [16] Diebold, F.X., Rudebusch, G.D., Aruoba, S.B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics* 131, 309-338.
- [17] Estrella, A., Hardouvelis, G.A., 1991. The term structure as a predictor of real economy activity. *Journal of Finance* 46, 555-576.
- [18] Estrella, A., Mishkin, F.S., 1998. Predicting U.S. recessions: financial variables as leading indicators. *Review of Economics and Statistics* 80, 45-61.
- [19] Garcia, R., Perron, P., 1996. An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics* 78, 111-125.
- [20] Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Bayesian statistics 4: Proceedings of the fourth Valencia international meeting*, 169-193.
- [21] Gray, S., 1996. Modeling the conditional distribution of interest rates as a regime switching process. *Journal of Financial Economics* 42, 27-62.
- [22] Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357-384.
- [23] Hamilton, J.D., 1994. *Time series analysis*. Princeton University Press, Princeton, New Jersey.
- [24] Kim, C.-J., 1994. Dynamic linear models with Markov-switching. *Journal of Econometrics* 60, 1-22.
- [25] Kim, C.-J., Nelson, C.R., 1998. *State-space models with regime-switching: classical and Gibbs-sampling approaches with applications*. Cambridge: MIT press.
- [26] Kozicki, S., Tinsley, P.A., 2001. Shifting endpoints in the term structure of interest rates. *Journal of Monetary Economics* 47, 613-652.

- [27] Nelson, C.R., Siegel, A.F., Parsimonious modeling of yield curves. *Journal of Business* 60, 473-489.
- [28] Piazzesi, M., 2004. Affine term structure models. In Ait-Sahalia, Y., Hansen, L.P. (editors), *Handbook of Financial Econometrics*, forthcoming.
- [29] Psaradakis, Z., Sola, M., 1998. Finite-sample properties of the maximum likelihood estimator in autoregressive models with Markov switching. *Journal of Econometrics* 86, 369-386.
- [30] Rudebusch, G.D., Wu, T., 2003. A macro-finance model of the term structure, monetary policy, and the economy. *Economic Journal* 118, 906-926.
- [31] Tam, C.S., Yu, I.W., 2008. Modeling sovereign bond yield curves of the US., Japan and Germany. *International Journal of Finance and Economics* 13, 82-91.
- [32] Taylor, J.B., 1993. Discretion versus policy rule in practice. *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.
- [33] Walsh, 2003. *Monetary theory and policy*. 2<sup>nd</sup> Edition, Cambridge: the MIT Press.
- [34] Yu, B., Mykland, P., 1998. Looking at Markov samplers through cusum path plots: a simple diagnostic idea. *Statistical and computing* 8, 275-286.
- [35] Zhu, X., 2008. The term structure with macro factors, expectations hypothesis and regime shifts. Working Paper: Nanyang Technological University.

Table 1: Yields-only Model<sup>a,b</sup>

Autoregressive coefficient matrix $\Phi$				
	$L_{t-1}$	$S_{t-1}$	$C_{t-1}$	$\mu$
$L_t$	<b>0.892</b> (0.012)	-0.001 (0.010)	<b>0.067</b> (0.013)	<b>7.791</b> (1.6297)
$S_t$	<b>-0.257</b> (0.022)	<b>0.872</b> (0.017)	<b>0.230</b> (0.023)	-1.957 (1.322)
$C_t$	0.047 (0.028)	<b>0.052</b> (0.022)	<b>0.905</b> (0.029)	-0.316 (1.441)
Estimated Covariance matrix $\Sigma_L$				
	$L_{t-1}$	$S_{t-1}$	$C_{t-1}$	
$L_t$	<b>0.094</b> (0.012)			
$S_t$	0.028 (0.016)	<b>0.218</b> (0.034)		
$C_t$	<b>0.033</b> (0.016)	-0.043 0.025	<b>0.379</b> 0.043	
Estimated Covariance matrix $\Sigma_H$				
	$L$	$S_t$	$C$	
$L_t$	<b>0.458</b> (0.125)			
$S_t$	<b>0.634</b> (0.230)	<b>2.192</b> (0.614)		
$C_t$	0.079 (0.234)	-0.013 (0.502)	<b>3.600</b> (1.066)	
Transition Probabilites $p$ and $q$				
	$p$		$q$	
	<b>0.932</b> (0.050)		<b>0.990</b> (0.008)	
Test for Diagonality of $\Sigma$ Matrix				
	Wald statistic <sup>c</sup>		P-value	
Regime $L$	16.055	27	0.001	
Regime $H$	7.846		0.049	
Test for no Regime-dependent Heteroscedasticity <sup>d</sup>				

Table 2: Summary Statistics for Measurement Errors of Yields

Maturity	Yields-only model		Macro-finance model	
	Mean	Std. Dev.	Mean	Std. Dev.
1-month	-26.84	37.27	-24.99	36.39
3-month	-4.45	8.84	-2.88	6.84
6-month	4.13	9.89	5.35	11.28
12-month	2.46	15.96	3.16	16.56
24-month	4.61	4.81	4.77	5.04
36-month	-1.44	2.80	-1.47	2.30
60-month	-2.72	4.14	-2.77	4.25
84-month	1.75	3.51	1.78	3.40
120-month	0.04	4.39	0.20	4.39

Note: as usual, all means and stand deviations of the yield measurement errors are expressed in basis points.

Table 3: Macro-finance Model<sup>a,b</sup>

Autoregressive coefficient matrix $\Phi^{MF}$						
	$L_{t-1}$	$S_{t-1}$	$C_{t-1}$	$FFR_{t-1}$	$\pi_{t-1}$	$CU_{t-1}$
$L_t$	<b>1.002</b> (0.016)	<b>0.056</b> (0.017)	<b>-0.016</b> (0.008)	-0.027 (0.015)	0.014 (0.008)	-0.004 (0.005)
$S_t$	0.008 (0.025)	<b>0.949</b> (0.026)	<b>0.089</b> (0.011)	-0.045 (0.024)	0.002 (0.011)	0.007 (0.006)
$C_t$	<b>0.639</b> (0.029)	<b>0.611</b> (0.032)	<b>0.894</b> (0.021)	<b>-0.528</b> (0.029)	0.034 (0.024)	0.014 (0.012)
$FFR_t$	<b>0.337</b> (0.022)	<b>0.297</b> (0.022)	<b>0.060</b> (0.007)	<b>0.664</b> (0.020)	0.005 (0.009)	<b>0.011</b> (0.004)
$\pi_t$	<b>-0.052</b> (0.023)	<b>0.050</b> (0.024)	-0.016 (0.010)	<b>0.056</b> (0.022)	<b>0.940</b> (0.010)	<b>0.021</b> (0.006)
$CU_t$	<b>0.225</b> (0.016)	<b>0.164</b> (0.018)	<b>-0.196</b> (0.013)	<b>-0.196</b> (0.016)	<b>-0.039</b> (0.012)	<b>0.984</b> (0.007)
Estimated Covariance Matrix $\Sigma_L^{MF}$						
	$L_t$	$S_t$	$C_t$	$FFR_t$	$\pi_t$	$CU_t$
$L_t$	<b>0.060</b> (0.006)					
$S_t$	<b>-0.045</b> (0.006)	<b>0.077</b> (0.008)				
$C_t$	<b>0.048</b> (0.011)	<b>-0.037</b> (0.013)	<b>0.393</b> (0.041)			
$FFR_t$	<b>0.010</b> (0.003)	<b>0.014</b> (0.032)	<b>0.032</b> (0.008)	<b>0.032</b> (0.003)		
$\pi_t$	<b>0.012</b> (0.005)	-0.008 (0.006)	-0.005 (0.013)	<b>0.007</b> (0.003)	<b>0.106</b> (0.009)	
$CU_t$	<b>0.017</b> (0.007)	-0.001 (0.007)	0.032 (0.017)	<b>0.015</b> (0.004)	0.013 (0.008)	<b>0.171</b> (0.014)

Table 3 (*continued*)

	$L_t$	$S_t$	$C_t$	$FFR_t$	$\pi_t$	$CU_t$
$S_t$	<b>0.353</b> (0.136)	<b>1.708</b> (0.419)				
$C_t$	<b>-0.191</b> (0.164)	-0.675 (0.364)	<b>2.530</b> (0.615)			
$FFR_t$	0.090 (0.149)	<b>1.396</b> (0.424)	-0.065 (0.426)	<b>2.543</b> (0.598)		
$\pi_t$	0.041 (0.038)	0.120 (0.089)	0.094 (0.109)	0.198 (0.113)	<b>0.157</b> (0.040)	
$CU_t$	0.106 (0.071)	<b>0.510</b> (0.177)	-0.050 (0.189)	<b>0.553</b> (0.209)	0.103 (0.054)	<b>0.524</b> (0.123)
Transition Probabilities $p$ and $q$						
	$p$			$q$		
	<b>0.932</b> (0.032)			<b>0.988</b> (0.007)		
Test for Diagonality of $\Sigma^{MF}$ Matrix						
	Wald Statistic <sup>c</sup>			P-value		
Regime $L$	178.31			0.000		
Regime $H$	24.58			0.056		
Test for no Regime-dependent Heteroscedasticity <sup>d</sup>						
	Wald Statistic			P-value		
	51.615			0.0002		

<sup>a</sup>Bold entries indicate significant at 5 percent level, standare deviations are in parentheses.

<sup>b</sup> $L$  denotes low volatility regime and  $H$  is high volatility regime.

<sup>c</sup>Wald statistics are asymptotically Chi-square with 3 degrees of freedom.

<sup>d</sup>Wald statistic has a Chi-square with 21 degrees of freedom.

Table 4: Tests of Interactions among Yield and Macro Factors<sup>a,b,c</sup>

	No Interaction $\Phi_2^{MF} = 0,$ $\Phi_3^{MF} = 0$ $\Sigma_2^{MF} = 0$	No Macro to Yields $\Phi_2^{MF} = 0$	No Yields to Macro $\Phi_3 = 0,$ $\Sigma_2^{MF} = 0$
Restrictions No.	27	9	18
Regime <i>L</i>	2178.9 (0.000)	513.3 (0.000)	699.03 (0.000)
Regime <i>H</i>	2150.1 (0.000)		646.0 (0.000)

<sup>a</sup>*L* denotes low volatility regime and *H* is high volatility regime.

<sup>b</sup>Reported statistics are based on the Wald test that is asymptotically  $\chi^2$  distribution.

<sup>c</sup>P-values appear in parentheses.