Three essays on empirical study of the term structure of interest rates

Zhu, Xiaoneng

2009

Zhu, X. N. (2009). Three essays on empirical study of the term structure of interest rates.
Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/18890>

<https://doi.org/10.32657/10356/18890>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**Three Essays on Empirical Study of the Term
Structure of Interest Rates**

Xiaoneng Zhu

SCHOOL OF HUMANITIES AND SOCIAL SCIENCES

2009

Three Essays on Empirical Study of the Term Structure of Interest Rates

Xiaoneng Zhu

School of Humanities and Social Sciences

**A thesis submitted to the Nanyang Technological University
In partial fulfillment of the requirement for the degree of**

**Doctor of Philosophy
in
Economics**

2009

Acknowledgements

This thesis was written in the period from July 2005 to May 2009 while I was enrolled as a PhD student at the Division of Economics, School of Humanities and Social Science, Nanyang Technological University. I would like to take this opportunity to thank those who have contributed to the making of this thesis.

First of all, I am grateful to the division for providing a good research environment for PhD students. During my studies, I have also benefited tremendously from the excellent seminar series organized by Division of Banking & Finance, Nanyang Business School, my thanks go to them as well.

Second, I would like to thank my academic advisor Professor Shahidur Rahman for his continuing patience and encouragement during my PhD studies. I would also like to thank Professor Choy Keen Meng, Professor Joseph Dennis Alba, Professor Tan Kim Heng, and two anonymous external examiners for helpful comments in an earlier version of my thesis. I have received help from the academic members and the staff at both Division of Economics and Division of Banking & Finance, for which I am grateful.

I would also like to thank my fellow PhD students for academic and practical discussions. Without all of you, the last three years would not have been nearly as pleasant. Some of my good friends went to the United States in between, but I did enjoy those days with them.

Finally, I would like to thank my family in China for their enduring understanding and loving support through the years. Without such kind of support, I could not have completed this thesis.

Xiaoneng Zhu

Singapore, June 2009

Contents

Acknowledgements	1
List of Figures	4
ABSTRACT	5
Chapter 1. Introduction	7
Chapter 2. Information, Expectations Hypothesis and Regime Shift	11
2.1. Introduction	11
2.2. The Expectations Hypothesis Under Regime Shift	15
2.3. The Predictability of Excess Bond Return and the EH	27
2.4. Data Issues and Yield Factors	32
2.5. Empirical Analysis	35
2.6. Concluding Remarks	39
Appendix A	40
Appendix B	41
Chapter 3. A Regime Switching Macro-finance Model of the Term Structure	44
3.1. Introduction	44
3.2. Yields-only model	47
3.3. Macro-finance model	59
3.4. Conclusions	73
Appendix A. The MCMC Algorithm	73
Appendix B. Impulse Responses	77

Appendix C. Variance Decomposition	78
Chapter 4. Global Yield Curves and Sovereign Bond Market Integration	80
4.1. Introduction	80
4.2. Country-specific yield curve factors	83
4.3. Global yield curve factors	89
4.4. Global market interactions and integration	94
4.5. Conclusions	100
Chapter 5. Conclusions and Recommendations for Future Study	120
References	122

List of Figures

2.1 U.S. Yield Curves, 1983.01–2009.05	33
2.2 Yield factors and empirical counterparts.	35
2.3 The Wald statistics	37
2.4 The probabilities of being in a tranquil regime.	37
2.5 Theoretical and actual spreads.	40
3.1 U.S. Yield Curves, 1983.01–2009.05	52
3.2 Level, slope, curvature factors and the empirical counterparts.	57
3.3 Probabilities of being in a tranquil regime.	58
3.4 Probabilities of being in a tranquil regime. The shaded bars indicate the NBER economic recessions.	58
3.5 Macro factors and level, slope, curvature factors from the macro-finance model.	61
3.6 Impulse responses of yield and macro factors on each other under regime L.	69
3.7 Impulse responses of yield and macro factors on each other under regime H.	70
3.8 The level from the macro-finance model and the theoretical level implied by the expectations hypothesis.	72
4.1 Yield Curves Across Countries	109
4.2 AFDNS Model Estimates of Level Factor Across Countries	110
4.3 AFDNS Model Estimates of Slope Factor Across Countries	110
4.4 AFDNS Model Estimates of Curvature Factor Across Countries	111

4.5 Factors from AFDNS Estimates and DNS Estimates: Germany and Japan	111
4.6 Factors from AFDNS Estimates and DNS Estimates: U.K. and U.S.	112
4.7 US level, slope and curvature factors and empirical factors	112
4.8 Error Terms of AFDNS Estimates across Countries	113
4.9 Average and AFDNS-fitted Yield Curves across Countries and Time:	113
4.10 The ACF and PACF of Errors from the AFDNS Estimation: Part 1	114
4.11 The ACF and PACF of Errors from the AFDNS Estimation: Part 2	114
4.12 Global Factors from Kalman Filter Estimates and Principle Component Analysis(with appropriate adjustment)	115
4.13 Dynamic Conditional Correlations: Level Factor	115
4.14 Dynamic Conditional Correlations: Slope Factor	116
4.15 Dynamic Conditional Correlations: Curvature Factor	116
4.16 Global Yield Curves	117
4.17 Filtered and Predicted Probabilities of Integration for Germany, Japan and the UK	118
4.18 Filtered and Predicted Probabilities of Integration: the US and World	119

ABSTRACT

Three Essays on Empirical Study of the Term Structure of Interest Rates

Xiaoneng Zhu

The modeling of the term structure dynamics is important for a variety of reasons. Forecasting is a first reason. The current yield curve contains information about future economic activity. Monetary policy constitutes a second reason for the term structure modeling. The transmission mechanism of monetary policy is related with the movements of the yield curves. The pricing and hedging of interest rate derivatives is a third reason. The price of many securities, such as coupon bonds, swaps, futures and options on interest rate is calculated based on some specification of term structure models. Bond portfolio diversification provides a fourth reason. Bonds and securities are traded in well-organized international market, to diversify bond portfolios in world market, we need the term structure models. This thesis is one more attempt to contribute to this discipline.

Essay I (Chapter 2) reexamines the expectations hypothesis, which is a central view of the term structure of interest rates. The empirical failure of the expectations hypothesis has been well-documented. There are usually three interpretations for the failure of the hypothesis. These are, respectively, the failure of the rational expectations assumption and unlimited arbitrage assumption, the presence of time-varying risk premiums, and poor properties of the statistical tests in finite samples caused by

peso problems. In this essay, I take into account peso problems and time-varying risk premiums by assuming that the data generating process is a Markov-switching vector autoregression model and investor's decision is based on a larger information set. In so doing, the testing of the expectations hypothesis is more robust. In particular, I find that the deviations from the expectations hypothesis are insignificant with my testing framework.

The term structure of interest rates and macroeconomic activity are closely related. Hence, the term structure models should also be able to identify the economic forces behind these movements. On the other hand, the yield curve contains important information for forecasting the future path of the economy. Therefore, we should model the yield curve and macroeconomic variables jointly. Essay II (chapter 3) attempts to contribute to jointly modeling yield and macro factors. The model proposed in this essay extends the dynamic Nelson-Siegel model by incorporating regime shifts. To estimate the proposed model, A MCMC (Markov chain Monte Carlo) method is presented. The main finding is that there are significant bidirectional linkages between the yield curve and economic activity.

Government securities are traded in well-organized international markets. Therefore, to what extent bond markets are integrated is a fundamental question in international finance. Furthermore, the degree of market integration is important for international public policy coordination. Essay III (Chapter 4) tries to answer this question. I propose a dynamic measure of bond-market integration based on the affine arbitrage-free dynamic Nelson-Siegel model. It contributes to the extant literature in several ways. First, my measure is consistent in both cross-section and time series. Second, this measure is theoretically consistent with no-arbitrage theory. Third, it is a dynamic measure. The empirical study demonstrates that global government bond markets are integrated to some extent.

CHAPTER 1

Introduction

Term structure modeling has generated a large literature. Recently, we see that the term structure models are evolving in some clear directions. First, the term structure models should be able to identify bidirectional linkages between yield and macroeconomic factor. Macro factors improve forecasting performance of the term structure models. Meanwhile, the term structure models with macro factors are able to identify the economic driving force behind the yield movements. Second, the term structure models should be able to capture time-varying risk premiums and explain the empirical failure of the expectations hypothesis. Third, the term structure models should have good forecasting performance. My thesis is one more attempt to contribute to the term structure modeling.

This thesis consists of three self-contained essays that conduct empirical analyses on the term structure of interest rates. The first essay (chapter 2) connects the classic expectations hypothesis and the modern term structure models. In essay II (chapter 3), I propose and estimate a macro-finance model of the term structure. Essay III (chapter 4) tries to measure the global government bond market integration.

Though three essays focus on the different aspects of the term structure of interest rates, there are some methodological interconnections among three essays. The first common emphasis is regime shifts. Nowadays 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-switching term structure models can account for some well-documented puzzles, for instance, the violation of the

expectations hypothesis and the predictability of excess bond return. A second common emphasis is the dynamic Nelson-Siegel model (Nelson and Siegel (1987), Diebold and Li (2006)). The affine arbitrage-free dynamic Nelson-Siegle term structure model is a theoretically rigorous yield curve model that simultaneously displays empirical tractability, good fit, and good forecasting performance.

Due to its simplicity and intuitive appeal, the expectations hypothesis remains the benchmark model of determining long-term interest rates. Essay I shows that an appropriate information set and regime switches can account for the empirical failure of the expectations hypothesis of the term structure of interest rates. It indicates that on a prior basis, time-varying risk premiums are second-order effects of yield dynamics. This contrasts with the in-sample predictability of time-varying risk premiums in the literature. My results may suggest that the departures of the expectations hypothesis are not profitable in practice. Furthermore, two regimes are found to be intimately related to business cycles.

Recently, evidence has accumulated that a few financial variables such as interest rates and yield spreads have enduring power of predicting aggregate economic activity. This predictive usefulness of interest rates and spreads thereafter has been well-established across countries. The countercyclical monetary policy, expectations and time-varying risk premium are the main reasons accounting for the predictive power of interest rates. On the other hand, the conduct of monetary policy transmits the movements 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 factors and macroeconomic factors are expected to be bidirectional and simultaneous, understanding the joint dynamics of the macro and yield factors is important for monetary policy-making and bond portfolio management.

Essay II 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.

The main focus of the third chapter is the time-varying degree of government bond market integration. If bond markets are not closely integrated, that implies investment institutions can benefit from further diversifying fixed income portfolios, otherwise there must be costs of diversification. From a monetary policy-making perspective, policy-makers in an open economy need take into account the yield dynamics in another market if there are interactions between markets. Policy coordination is therefore necessary in such case as indicated by the theoretical model (Chang 1997).

Essay III proposes a dynamic measure of bond-market integration based on the affine arbitrage-free dynamic Nelson-Siegel model (Christensen et. al. 2007). It contributes to the existing literature in several ways. First, our measure is consistent in both cross section and time series. So far, the empirical studies in this area have mainly focused on the long-term bond yield or weighted bond yield index that neglects the maturity structure of the term structure. However the cross-section spread of the term structure has important information about the future path of interest rates and economic activity. Further, we need to rule out arbitrage opportunities in the cross-section of the term structure. It is therefore of great interest to investigate bond market

integration in a yield curve model that takes into account the maturity structure of the term structure. Second, the yield curve model employed has good in-sample and out-of-sample fits (Diebold and Li 2006). The generalizations of the Nelson-Siegel model are extensively applied by financial institutions and central banks. This is hence an important advantage. Third, the proposed measure allows time-varying partial segmentation of national and global government bond markets. The measure also circumvents the polar cases of completely segmented or integrated bond markets. In addition, it nests the uncovered interest rate parity as a polar case given that the expected exchange rate difference is a martingale process.

CHAPTER 2

Information, Expectations Hypothesis and Regime Shift

2.1. Introduction

Simplicity and intuitive appeal make the expectations hypothesis (EH) a central view of the term structure of interest rates. According to the EH, the long-term yield is a weighted average of the expected future short-term yields plus a maturity-specific constant risk premium. The EH has been intensively examined using a variety of tests and data since it has long been recognized as a basic workhorse model of the term structure of interest rates. Contrary to prior intuition, most empirical studies reject the expectations hypothesis¹. In addition to the statistical rejection, some recent studies (for example, Campbell and Shiller (1991), Cochrane and Piazzesi (2005), Fama and Bliss (1987)) found that yield spreads or forward rates have predictive power on future excess bond returns. This appears to be the critical evidence on the empirical failure of the expectations hypothesis since the EH implies that excess returns should be unpredictable.

The standard tests of the EH have missing motivations. Bekaert and Hodrick (2001) presented three main potential reasons for the usual rejections of the EH, respectively, the failure of the assumption of rational expectations and unlimited arbitrage, the presence of variable risk premiums, and the poor properties of the statistical tests in finite samples caused by peso problems. This article sticks to the assumption of rational expectations and extends empirical tests along the second and third dimensions. The omission of the variables capturing time-varying risk premiums might distort the

¹There are some exceptions. For example, Longstaff (2004) found that the EH is valid at the very short end of the yield curve. Sola and Driffill (1994) resurrected the EH in a Markov-switching vector autoregression model.

statistical inference. Thus, the conditional information set for testing the EH should include the variables that are expected to capture time-varying risk premiums. Furthermore, peso problems imply that the data generating process for testing the EH should incorporate regime shifts.

Modern term structure models provide some guidance on the extension of the conditional information set. Recently, Bansal and Zhou (2002), Dai and Singleton (2002), Dai, Singleton and Yang (2007), among others, have accounted for the well-documented violations of the expectations hypothesis within sophisticated no-arbitrage models of the term structure. Researchers build these term structure model upon the standard intuition of the current yield curve containing all information relevant to predicting future bond yields. In particular, a few (usually three) factors that sufficiently capture the cross-section of yields can also capture the dynamics of yields. By specifying the evolution of factors under both the physical and risk-neutral measures, these term structure models describe the dynamics of the entire yield curve.

There are two key findings from the no-arbitrage term structure model analyses. First, predictable time-varying risk premiums that can be generated from the dependence of risk prices on risk factors are important for matching the EH. This finding has notable implications for the testing of the EH. The EH must be rejected if we fail to use the information in yield factors. The non-inclusion of yield factors in the conditional information set means that we miss important information in the market. Even capital markets are efficient in the sense that they fully and correctly reflect all relevant information in determining security prices, our testing procedure is flawed because we have missing information in the testing of the EH. Second, regime shifts play an important role in explaining deviations from the EH. This finding implies that we should take into account regime shifts in the testing of the EH. This is consistent with the conjecture of Bekaert and Hodrick (2001).

Since aforementioned studies do not rule out the possibility that investor's actual historical predictions of risk premiums are not as accurate as those from the in-sample predictions found in today's statistical analysis (Piazzesi and Schneider, 2009)², in other words, the departures of the EH are not profitable based on criteria of economic significance (Corte, Sarno and Thornton, 2008), it makes sense to reexamine the EH with a larger information set and a Markov-switching VAR as a data generating process. The reexamination is also motivated by the theoretical consistency of the EH and the capital asset pricing model as will be shown in section 2.

A larger information set has several advantages. First, the extension of information set may alleviate simultaneity bias in the estimation (Carriero, Favero and Kaminska, 2006). One more reason for a larger information set is the improved power properties of the extended testing procedure (Sarno, Thornton and Valente, 2007). Finally, yield factors may partially capture time-varying risk premiums.

Another contribution of this paper is to show that regime shifts might contribute to the predictability of excess bond returns in addition to risk premiums. The non-recognition of this point might lead to single-regime term structure models overstate or understate the volatility of risk premiums for understanding the predictability of excess bond returns³. It has notable implications for the testing of the EH. It is possible that within each regime variable term premiums are second-order effects of yield dynamics on a prior basis. In particular, the EH holds within each regimes, but the violation of the EH is due to regime switches. This is one more motivation for reexamining the EH after explicitly taking into account regime shifts. To net out risk premiums effect, I show that regime shifts can generate predictability in a discrete Vasicek (1977)

²Using survey data on interest rate forecasts, Piazzesi and Schneider (2009) found that subject premia are less volatile and not very cyclical.

³Dai, Singleton and Yang (2007) found that a single-regime term structure model overstate the volatility of risk premiums during the tranquil times of yields.

model. This is achieved with the introduction of regime-dependent heteroscedasticity. Furthermore, economic agents in this model are assumed to be uncertain about the prevailing regime. This assumption is reasonable for two reasons. In the extant literature, regimes are usually related with economic recessions and booms. However, it is well-known that the NBER's Business Cycle Dating Committee announces an economic contraction with a lag of several quarters to avoid major revisions⁴. This provides a first reason for the assumption. Secondly, if regimes are associated with the conduct of monetary policy, this assumption is still reasonable since the Fed does not announce its targeting in advance. Usually, regime shift is a low frequency phenomena and it is supposed to be persistent and partially predictable. The uncertainty about the prevailing regime and the partial predictability of regimes generate the predictability on excess bond returns.

I find that the EH cannot be rejected consistently. The results highlight the importance of including yield factors in the conditional information set and taking into account regime shifts. The results have two implications for risk premiums. On the one hand, the empirical findings may suggest that within each regime variable term premiums are second-order effects of yield dynamics on a prior basis. On the other hand, yield factors may partially account for time-varying risk premiums.

The remaining paper is organized as follows. Section 2 discusses the expectations hypothesis in an arbitrage-free framework and derives the restrictions implied by the EH. Without loss of generality, section 3 proves that regime shifts can generate the predictability of excess bond returns. The data descriptions and summary statistics of

⁴For example, the November 2001 trough was announced July 17, 2003, and the March 2001 peak was announced November 26, 2001. According to the NBER definition, contractions start at the peak of a business cycle and end at the trough.

yields and yield factors are presented in section 4. The arbitrage-free dynamic Nelson-Siegel model is briefly introduced in this section, too. Section 5 applies the tests to the data. Section 6 make up of concluding remarks.

2.2. The Expectations Hypothesis Under Regime Shift

This section introduces two approaches of testing the EH in an arbitrage-free framework under regime shifts. One approach is the present value model of the expectations hypothesis (Campbell and Shiller, 1987, CS1 henceforth). The employment of the present value model attempts to shed light on the entire yield curve in the testing of the EH. Instead of a pair of long- and short-term yields, we can use the level, the slope and the curvature factors from the dynamic Nelson-Siegel model (DNS) (Diebold and Li, 2007) to test the EH. The level factor from the DNS model is a long-term factor. Empirically, the level factor is highly correlated with long-term yields. Furthermore, the sum of the level and slope factors has a nice interpretation of short-term yield. Using the yield factors, we need not test each pair of long- and short-term yields.

The second approach is a commonly cited statement of the EH, that a long-term n -period yield is an equal weighted average of future short-term one-period yields in addition to a maturity-specific constant risk premium (Campbell and Shiller, 1991, CS2 hereafter). This approach is an empirical benchmark so that my results are comparable with other empirical studies. Both approaches can provide evidence on the economic significance of the EH in addition to a statistical rejection or non-rejection. Since the focus of this article is to examine the EH under regime shifts, restrictions implied by the EH are derived from a Markov-switching VAR framework.

2.2.1. The Present Value Model

The present value model of the term structure of interest rates posits that the n -period, continuously compounded yield, $i_{t,n}$ equals a weighted average of the current and expected short yields plus a maturity-specific constant risk premium.

$$(2.1) \quad i_{t,n} = (1 - \delta) \sum_{j=0}^{\infty} \delta^j E_t i_{t+j,1} + c_n$$

where δ , the discount factor, is a parameter that reflects the impatience of economic agents. c_n denotes a constant maturity-specific premium. E_t is the conditional expectations based on the information set at time t . Due to its theoretical appeal, a class of modern asset pricing models impose no-arbitrage restrictions. It is straightforward to demonstrate that the EH in Eq.(2.1) is consistent with arbitrage-free conditions. Let M_{t+1} denote the pricing kernel, it is well-known that any gross return R_{t+1} in an economy that doesn't admit arbitrage opportunities can be correctly priced by

$$(2.2) \quad E_t(M_{t+1}R_{t+1}) = 1$$

To be statistically tractable, it is assumed that returns and pricing kernels are conditionally log-normal. Following Bekaert and Hordrick (2001), Eq.(2.2) implies that

$$(2.3) \quad E_t(m_{t+1}) + 0.5V_t(m_{t+1}) + E_t(r_{t+1}) + 0.5V_t(r_{t+1}) + Cov_t(m_{t+1}, r_{t+1}) = 0$$

where V_t and Cov_t respectively represent conditional variance and covariance. The lower letters denote the logs of the corresponding uppercase letters so for example, $m_{t+1} = \log(M_{t+1})$. $i_{t,1}$, the return of one-period yield, is observable at time point t , so

the last two items on the left-hand side of Eq.(2.3) disappear. Thus the expression for the one-period yield is

$$(2.4) \quad i_{t,1} = -[E_t(m_{t+1}) + 0.5V_t(m_{t+1})]$$

Now let r_t in Eq.(2.3) represent the excess return of the holding period return ($h_{t+1,n}$) of a long-term bond over an one-period bond. Combined with Eq.(2.4), the expected excess return can be given by the following equation:

$$(2.5) \quad E_t(h_{t+1,n}) - i_{t,1} = -[Cov_t(m_{t+1}, h_{t+1,n}) + 0.5V_t(h_{t+1,n})]$$

The right-hand side of Eq.(2.5) is a constant conditional on the time t information set. Let a_n denote the constant, and Eq.(2.5) can be expressed as:

$$(2.6) \quad E_t(h_{t+1,n}) = i_{t,1} + a_n$$

The one-period holding return on a n-period bond can be approximated (Shiller (1979)) by a linear function in the neighborhood of $i_{t,n} = i_{t+1,n-1}$.

$$(2.7) \quad h_{t+1,n} = \frac{i_{t,n} - \delta i_{t+1,n-1}}{1 - \delta}$$

By taking expectations of Eq.(2.7) and using Eq.(2.4) and Eq.(2.5), after rearrangement, we have

$$(2.8) \quad i_{t,n} = (1 - \delta)i_{t,1} + \delta E_t i_{t+1,n-1} + (1 - \delta)a_n$$

Using recursive substitution and letting $n \rightarrow \infty$, combined with the terminal condition $i_{t,0} = 0$, we have the present value version of the expectations hypothesis in Eq.(2.1) with $c_n = (1 - \delta) \sum_{j=0}^{\infty} \delta^j a_n$.

The nonstationarity of time series may invalid the statistical inference. The realization that yields are usually persistent and integrated of order one motivates CS1 to test the EH using the yield spread $S_{t,n} = i_{t,n} - i_{t,1}$ and the first difference of the short-term yield. The stationarity of the yield spread imposes a restriction on the long-run dynamics. It is clear by subtracting $i_{t,1}$ from both sides of equation (2.1), and rearranging, that we have

$$(2.9) \quad S_{t,n} = \sum_{j=1}^{\infty} \delta^j E_t \Delta i_{t+j,1} + c_n$$

The yield spread is a weighted average of a stationary variable — the first difference of the short-term yield. However, this is a necessary but not sufficient condition of the EH since the validity of the EH also requires restrictions to be imposed on the short-run dynamics.

Eq.(2.9) provides the testable restrictions implied by the present value model of the expectations hypothesis. $y_t = [MS_t, \Delta LS_t, C_t]'$ is a 3×1 vector of the state variables. In particular, ML_t is minus slope factor from the DNS model, it represents $S_{t,n}$ in Eq.(2.9)⁵. Δ is a first-difference operator. LS_t is the sum of the level and slope factor from the DNS model and represents the short-term yield Eq.(2.9)⁶. C_t is the curvature factor from the DNS model. I assume that the state vector y_t follows a vector

⁵The ML is a good proxy of the yield spreads between a long-term yield, say, 10-year yield and a short-term yield, such as, 3-month yield. Panel D of Table 1 shows that the correlation between the slope factor and the spreads between 10-year yield and 3-month yield is 0.9928.

⁶Also refer to subsection "Yield Factors".

autoregressive process of finite order l

$$(2.10) \quad y_t = \mu + \sum_{j=1}^l \phi_j y_{t-j} + u_t$$

For simplicity, the intercepts are removed from Eq.(2.10) since the EH doesn't impose any restriction on the constant risk premium. By expanding the state vector to the companion form $Y_t = [y_t \dots y_{t-l}]$, we can rewrite the state dynamics in a first-order representation:

$$(2.11) \quad Y_t = \Phi Y_{t-1} + U_t$$

The information set $\Theta = [MS_{t-i}, \Delta LS_{t-i}, C_{t-i}, i \geq 0]$ is observed by econometricians at time point t . Let $g' = [1, 0, \dots, 0]$ and $h' = [0, 1, 0, \dots, 0]$ be the selection vectors with $3l$ elements such that $MS_t = g'Y_t$ and $\Delta LS_t = h'Y_t$. It is easy to show that the i -period ahead optimal forecast is $\Delta \hat{LS}_{t+i} = h'\Phi^i Y_t$. Next, by projecting restrictions Eq.(2.9) onto the data generating process, Eq.(2.11), we obtain⁷

$$(2.12) \quad g'Y_t = \delta h'\Phi(I - \delta\Phi)^{-1}Y_t$$

The testable cross-equation restrictions implied by the EH of the term structure of interest rate are

$$(2.13) \quad g' = \delta h'\Phi(I - \delta\Phi)^{-1}$$

If the present value model of the expectations hypothesis is true, Eq.(2.13) hold for sure and the selection of information set is not relevant. The intuitions is that given

⁷The details are given in Campbell and Shiller (1987).

a constant term premium, all the relevant information of investors is embodied in the yield spread, which is included in y_t . However, any economic model is an approximation, in this sense, all economic model are false. The question is to what extent an economic model approximates the truth. In this case, the selection of information set is important. First, selected variables should have enough information about the question being asked. Second, statistical tests based on selected variables should have good power and size properties. These two criteria account for the selection of state variables in y_t .

The CS test is also motivated by the unsatisfactory of pure econometric tests that may over-reject or under-reject the null hypothesis⁸. Furthermore, econometric tests do not tell us the economic significance (or adequacy) of the EH. If the EH can explain most of the variations in actual yield spread $S_{t,n}$, economically, it is a good approximation regardless of the statistical rejection or non-rejection of the EH. It is also possible that the observed sample contains little information about the expectations hypothesis. Hence, many economists are reluctant to see the statistical rejection or non-rejection of restrictions as a definitive answer. The economic significance should also play an important role in evaluating the EH. To shed light on how well the model explains the economic significance, instead of resorting only to statistical significance, we can evaluate the theoretical spread

$$(2.14) \quad S_{t,n}^* = \delta h' \Phi (I - \delta \Phi)^{-1} Y_t$$

⁸Bekaert and Hordick (2001) provide the Monte Carlo simulation results for the frequently used Wald test and LM test. The Wald test is found to grossly overreject the EH in a small sample. In contrast, the LM test slightly underrejects the null hypothesis.

The theoretical yield spread in Eq.(2.14) is the optimal forecast given the information set and the entertained model. If the present value model is a good enough approximation and our econometric model describes the joint dynamics of the yield factors well, the theoretical spread would be observed in financial markets. The difference between theoretical and actual spreads contains information about the adequacy of the EH. We can visually inspect the deviation by plotting both series in a diagram. The good fit of the $S_{t,n}$ and $S_{t,n}^*$ indicates a good economic significance. Another measure is to calculate the correlation between $S_{t,n}$ and $S_{t,n}^*$. The high degree of comovement suggests that economic agents accurately forecast the future changes of spread. They hence incorporate the predictions into present investment decisions, as a result, no profitable arbitrage opportunity is available in bond markets.

2.2.2. The CS2 Approach

In CS2, the expectations hypothesis states that a n -period yield $i_{t,n}$ is an equal weighted average of expected future 1-period yields,

$$(2.15) \quad i_{t,n} = \frac{1}{n} \sum_{j=0}^{n-1} E_t i_{t+j,1} + c_n$$

The term premium c_n may vary with n but is assumed to be constant through time. Bekaert and Hodrick (2001) shows that this form of expectations hypothesis is consistent with the capital asset pricing model. Now the VAR approach for evaluating the present value model can be easily modified to evaluate the expectations hypothesis in Eq.(2.15). Let the yield spread $S_t = i_{t,n} - i_{t,1}$. By subtracting $i_{t,1}$ from both sides of Eq.(2.15), Simple algebra shows that S_t is

$$(2.16) \quad S_{t,n} = \sum_{j=0}^{n-1} \left(1 - \frac{j}{n}\right) E_t \Delta i_{t+j,1} + c_n$$

Eq.(2.16) provides the testable restrictions implied by the expectations hypothesis Eq.(2.15). Let $y_t = [S_{t,n}, \Delta i_{t,1}, C_t]'$ be a 3×1 vector of the state variables. The state vector y_t is assumed to follow a VAR(l) process as in Eq.(2.10). Accordingly, a more convenient VAR(1) form is given in Eq.(2.11). We have $g' = [1, 0, \dots, 0]$ and $h' = [0, 1, 0, \dots, 0]$ be the selection vectors with $3l$ elements such that $S_{t,n} = g'Y_t$ and $\Delta i_{t,1} = h'Y_t$. As a result, we have the i -period ahead optimal forecast $\hat{\Delta i_{t+i}} = h'\Phi^i Y_t$. we can project Eq.(2.16) onto the data generating process Eq.(2.11). With tedious algebra manipulation, we have the following expression⁹

$$(2.17) \quad g'Y_t = h'\Phi \left[I - \frac{1}{n}(I - \Phi^n)(I - \Phi)^{-1} \right] (I - \Phi)^{-1} Y_t$$

Then, instead of directly testing the EH in Eq.(2.15), we can test the following cross-equation restrictions

$$(2.18) \quad g' = h'\Phi \left[I - \frac{1}{n}(I - \Phi^n)(I - \Phi)^{-1} \right] (I - \Phi)^{-1}$$

The theoretical spread implied by the EH in Eq.(2.15) can be computed with

$$(2.19) \quad S_{t,n}^* = h'\Phi \left[I - \frac{1}{n}(I - \Phi^n)(I - \Phi)^{-1} \right] (I - \Phi)^{-1} Y_t$$

2.2.3. The Expectations Hypothesis Under Regime Shift

The evidence has accumulated that regime-switching models describe the historical yields better than single-regime models (e.g., Gray (1996), Hamilton (1988)). Regime shifts usually relate to business cycles or monetary policy shifts. During economic recessions and booms, there are asymmetric dynamics of the yield curve. Furthermore, monetary policy shifts significantly change the behavior of the yield curve. The Markov-switching model captures these important features. Assuming that the dynamics of

⁹The details are given in Campbell and Shiller (1991).

state variables can be described by a Markov-switching (MS) VAR model instead of the VAR in Eq.(2.11):

$$(2.20) \quad Y_t = \Phi_{k_t} Y_{t-1} + U_{k_t}$$

where the subscript $k_t = \{1, 0\}$ denotes the unobservable state variable, which is assumed to be governed by a discrete-time Markov chain¹⁰. The specification of a first-order Markov chain is not as restrictive as it seems. The first order chain offers a good approximation to higher order Markov chain regime shift (e.g. Hamilton (1994), chapter 22). In addition, it also provides an approximation to some continuous regime shifts. The nonlinear process with regime shift may better characterize the yield dynamics of the selected sample than a linear process. Thus, the regime switching specification provides a parsimonious way to express complicated dynamics, which might otherwise require an ARIMA model with long lags. Alternatively, Eq.(2.20) can be rewritten as

$$(2.21) \quad Y_t = k_t \Phi_1 Y_{t-1} + (1 - k_t) \Phi_0 Y_{t-1} + k_t U_{1t} + (1 - k_t) U_{0t}$$

The restrictions in Eq.(2.13) and Eq.(2.18) can be projected onto a Markov-switching VAR model where the data generating process accommodates a regime shift. The Markov chain that governs the state variable is

$$(2.22) \quad P = \begin{bmatrix} p & 1 - q \\ 1 - p & q \end{bmatrix}$$

¹⁰This is a general specification, it allows regime-dependent coefficients and heteroscedasticity. In Krolzig (1997) terminology, this is a MSAH model. When testing the EH against data, the model selection is an empirical issue.

To simplify the projection, I defined a matrix M as

$$(2.23) \quad M = \begin{bmatrix} p\Phi_1 & (1-q)\Phi_1 \\ (1-p)\Phi_0 & q\Phi_0 \end{bmatrix}$$

Now it is straightforward to produce the optimal forecasts given the Markov-switching data generating process. If 0 is the prevailing regime, then

$$(2.24) \quad Y_{t+i}^\wedge = JM^i Q_0 Y_t$$

Alternatively, if we start from the regime 1,

$$(2.25) \quad Y_{t+i}^\wedge = JM^i Q_1 Y_t$$

where

$$(2.26) \quad J = \iota \otimes I_{2l}; \quad \iota = (1, 1)$$

and

$$(2.27) \quad Q_i = e_i \otimes I_{2l}$$

e_i is the i th column of 2×2 identity matrix in association with the state we are in, and I_{2l} is $2l$ -dimension identity matrix.

With the above well-defined notions, given the regime 0, we can project Eq.(2.9) onto the data generating process Eq.(2.21)

$$(2.28) \quad g'Y_t = \delta h'JM(I - \delta M)^{-1}Q_0Y_t$$

Starting from regime 1, the restrictions are

$$(2.29) \quad g'Y_t = \delta h'JM(I - \delta M)^{-1}Q_1Y_t$$

Eqs. (2.28) and (2.29) are counterparts of Eq.(2.12) in a MS-VAR model. Thus, In regime 0 the cross-equation restrictions implied by the present value model of the expectations hypothesis in Eq.(2.1) are:

$$(2.30) \quad g' = \delta h'JM(I - \delta M)^{-1}Q_0$$

Starting from regime 1, the restrictions are

$$(2.31) \quad g' = \delta h'JM(I - \delta M)^{-1}Q_1$$

In CS1, the cross-equation restrictions can be transformed to linear restrictions, but in the framework of Markov-switching VAR model, the restrictions are highly nonlinear. Eqs.(2.30) and (2.31) deliver the testable restrictions implied by the EH in a matrix tractable way and the test of restrictions become operational.

Accordingly, we can project Eq.(2.16) onto the data generating process Eq.(2.21). In regime 0, we have

$$(2.32) \quad g'Y_t = h'JM \left[I - \frac{1}{n}(I - M^n)(I - M)^{-1} \right] (I - M)^{-1}Q_0Y_t$$

In regime 1, we have

$$(2.33) \quad g'Y_t = h'JM \left[I - \frac{1}{n}(I - M^n)(I - M)^{-1} \right] (I - M)^{-1}Q_1Y_t$$

Eqs. (2.32) and (2.33) are counterparts of Eq.(2.17) in a MS-VAR model. Thus, in regime 0 the restrictions implied by Eq.(2.15) are

$$(2.34) \quad g' = h'JM \left[I - \frac{1}{n}(I - M^n)(I - M)^{-1} \right] (I - M)^{-1}Q_0$$

Given that the prevailing regime is 1, the restrictions implied by Eq.(2.15) are

$$(2.35) \quad g' = h'JM \left[I - \frac{1}{n}(I - M^n)(I - M)^{-1} \right] (I - M)^{-1}Q_1$$

The maximum likelihood estimation of MS-VAR with nonlinear restrictions is complicated. Instead of the likelihood ratio test, the Wald test is proposed to serve as an alternative¹¹. Because all restricted parameters are presented in matrix M , in regime k , the first-order derivatives of the restrictions in Eqs.(2.30) and (2.31) with respect to the parameters are

$$(2.36) \quad \frac{\partial C_{(\hat{\theta})}}{\partial \hat{M}_{ij}} = \delta h'JM \frac{dM}{dM_{ij}} (I - \delta M)^{-1}Q_k + \delta^2 h'JM (I - \delta M)^{-1} \frac{dM}{dM_{ij}} (I - \delta M)^{-1}Q_k$$

where $\frac{dM}{dM_{ij}}$ is the derivative of matrix M with respect to parameter M_{ij} . With the above derivatives, the Wald test statistics can be calculated easily. And the first-order

¹¹Since the Monte Carlo simulation results show that the Wald test usually overrejects the EH in a small sample, the statistical non-rejection of the EH in the empirical section is persuasive.

derivatives of the restrictions in Eqs.(2.34) and (2.35) are

$$\begin{aligned}
(2.37) \quad \frac{\partial C_{\hat{\theta}}}{\partial M_{ij}} &= h'JM \frac{dM}{dM_{ij}} (I - M)^{-1} Q_k + \\
&h'JM(I - M)^{-1} \frac{dM}{dM_{ij}} (I - M)^{-1} Q_k + \\
&\frac{1}{n} h'JM \frac{dM}{dM_{ij}} (I - M^n)(I - M)^{-2} Q_k - \\
&\frac{1}{n} h'JM \left(\sum_{\rho=1}^n M^{\rho-1} \frac{dM}{dM_{ij}} M^{n-\rho} \right) (I - M)^{-2} Q_k + \\
&\frac{1}{n} h'JM(I - M^n) \left[\sum_{\rho=1}^2 M^{\rho-1} (I - M)^{-1} \frac{dM}{dM_{ij}} (I - M)^{-1} M^{2-\rho} \right] Q_k
\end{aligned}$$

Given the well-defined restrictions and the data generating process, the Wald test is operational now.

2.3. The Predictability of Excess Bond Return and the EH

Eq.(2.5) and the EH in Eq.(2.1) imply that excess holding period returns cannot be predicted by current yields. Since forward rates can be recovered from the current yields as implied by the no-arbitrage conditions, this is equivalent to the statement that forward rates do not predict excess returns. This section shows that regime shifts might contribute to the predictability of excess bond returns in addition to risk premiums. To see the point, we begin with a Vasicek model discretized by Campbell, Lo and Mackinlay (1997, pp429-431). The Vasicek model is a fundamental term structure model, which, like all other fundamental models, share two properties: a time-homogeneous short rate process and an explicit specification of the market prices of risk. It is a good starting point because the Vasicek model is consistent with the expectations hypothesis. Suppose that a single state variable forecasts the stochastic discount factor,

$$(2.38) \quad -m_{t+1} = x_t + \epsilon_{t+1}$$

It is assumed that ϵ_{t+1} is normally distributed with constant variance.

x_{t+1} is assumed to evolve according to a univariate AR(1) process with mean μ , persistence ϕ and shock ε_{t+1}

$$(2.39) \quad x_{t+1} = (1 - \phi)\mu + \phi x_t + \varepsilon_{t+1}$$

The correlation between the innovations to m_{t+1} and x_{t+1} is captured by a time-series process,

$$(2.40) \quad \epsilon_{t+1} = \beta \varepsilon_{t+1} + \eta_{t+1}$$

where ε_{t+1} is normally distributed with constant variance σ^2 , η_{t+1} is also normally distributed with constant variance and is uncorrelated with ε_{t+1} . Note that the shock η_{t+1} does not affect the slope or dynamics of the term structure although it does affect the average level of the yield curve. To simplify the notations, we accordingly drop it and assume that $\epsilon_{t+1} = \beta \varepsilon_{t+1}$. Eq.(2.38) can then be rewritten as

$$(2.41) \quad -m_{t+1} = x_t + \beta \varepsilon_{t+1}$$

The only innovation that shocks the system now is ε_{t+1} .

From Eq.(2.2), we can solve for the log price of $p_{t,n}$ of an n-period nominal bond at time t by recursively solving the relation

$$(2.42) \quad p_{t,n} = E_t(m_{t+1}p_{t+1,n-1})$$

with the terminal condition $p_{t,0} = 1$. The resulting bond prices are linear functions of the state variable

$$(2.43) \quad -p_{t,n} = A_n + B_n x_t$$

The absence of arbitrage implies that the coefficients can be calculated from the following difference equations (refer to Campbell et. al. (1997))

$$(2.44) \quad \begin{aligned} A_n &= A_{n-1} + (1 - \phi)\mu B_{n-1} - (\beta + B_{n-1})^2 \sigma^2 / 2 \\ B_n &= 1 + \phi B_{n-1} = (1 - \phi^n) / (1 - \phi) \end{aligned}$$

Bond yields are then affine functions of the state variable

$$(2.45) \quad \begin{aligned} i_{t,n} &= -p_{t,n} / n \\ &= A_n / n + B_n / n x_t \end{aligned}$$

The discrete Vasicek model is a homoskedastic bond pricing model. This salient feature has several interesting implications. It is clear from the Eqs. (2.44) and (2.45) that the coefficient B_n measures the sensitivity of the n -period yield to the one-period yield. Given that x_t is a stationary process with $|\phi| < 1$, B_n increases at a decreasing rate, accordingly, the impact of a shock in x_t on bond prices rises but at a decreasing rate. As $n \rightarrow \infty$, B_n approaches the limit $1/(1 - \phi)$.

This feature is in accordance with the expectations hypothesis of the term structure. The EH requires that the right-hand side of Eq.(2.5) is conditionally a constant, thus the expected excess returns are unpredictable. This is exactly what is implied by the model. The expected excess return on an n -period bond over a one-period bond $E_t(h_{t+1,n}) - i_{t,1} = E[p_{t+1,n-1}] - p_{t,n} + p_{t,1}$, is given by

$$(2.46) \quad E_t(h_{t+1,n}) - i_{t,1} = -B_{n-1}\beta\sigma^2 - B_{n-1}^2\sigma^2/2$$

This is a maturity-specific constant, so the EH holds. The derivation of Eq.(2.46) is provided in the appendix A¹².

Recently, strong evidence against the unpredictability of excess returns has accumulated. It ruins the ground of the EH. However, this section shows that the EH can be consistent with the predictive power of forward rates if parameters governing the state variables' dynamics are regime-dependent. The EH is therefore defensible on a theoretical ground. The validity of the EH is still an empirical issue.

A large literature has suggested that the conduct of monetary policy in the U.S. has changed substantially. The yield movements respond to a discrete, but probably long-lasting change in the monetary policy significantly. Another source of discrete regime shifts on the yield curve is business cycles, with asymmetric dynamics of the yield curve in good or bad economic conditions. The term structure models with regime shifts can explain a few puzzles, for example, the failure of the expectations theory. It is therefore an important stylized fact in modeling yield dynamics.

Suppose there are two regimes $k_t = \{1, 0\}$. Furthermore, economic agents in this model are uncertain about the prevailing regime. They infer regimes from the observed yields and macro factors through some econometric techniques, such as the Hamilton

¹²The derivation is a detailed version of that presented in Campbell et al. 1997. Because this is a key point of the paper, I provide it in the appendix A. It is also for cross-reference.

filter (1989, 1990, 1994). Suppose the filtered probabilities of being in state 0 $\xi_t^0 := \Pr_t[k = 0]$ conditional on the time t information set, is a truncated normal distribution (or any other distribution) in the interval $[0, 1]$,

$$(2.47) \quad \xi_t^0 \sim TN(v_t, \sigma_t^2)$$

A specific expression for σ_t^2 is given in the Eq.(3.17) of appendix B. In regime k , Eq.(2.46) implies that the expected holding period return is constant. The model also implies that the forward rate $f_{t,n}$ is a constant as given in Campbell et al. (1997, pp432). If there doesn't exist a regime shift, we run such a regression as in Fama and Bliss (1987)

$$(2.48) \quad h_{t+1,n} = a + b * f_{t,n} + w_t$$

then the estimate of b shouldn't be different with 0 significantly because $Cov_t(h_{t+1,n}^k, f_{t,n}^k) = 0$ under rational expectations in which w_t is a white noise process.

However, if regime shift is an important factor in yield movements, $Cov_t(h_{t+1,n}, f_{t,n})$ is not equal to zero except for some pathological cases. Suppose there exists at least one regime-dependent parameter¹³ in Eq.(2.46)) that determines the expected holding period return. And let $h_{t+1,n}^k$ denote the excess return in regime k . Assuming that we are in regime k at time point t , we have

$$(2.49) \quad h_{t+1,n}^{kk} = \sum_{m=0}^1 p_{km} * h_{t+1,n}^k, \quad k = 0, 1$$

and

¹³It can be σ^2 , ϕ , or β .

$$(2.50) \quad f_{t,n}^{kk} = \sum_{m=0}^1 p_{km} * h_{t,n}^k, \quad k = 0, 1$$

where $p_{km} = p, 1 - p, q$ or $1 - q$, are transition probabilities shown in Eq. (2.22). It can be derived that

$$(2.51) \quad \begin{aligned} Cov_t(f_{t,n}, h_{t+1,n}) &= (h_{t+1,n}^{00} f_{t,n}^{00} + h_{t+1,n}^{00} f_{t,n}^{11} + h_{t+1,n}^{11} f_{t,n}^{00} + h_{t+1,n}^{11} f_{t,n}^{11}) \\ &\quad (1 + p - q) \sigma_t^2 \end{aligned}$$

The derivation of the covariance between forward rates and excess returns is presented in appendix B. The non-zero covariance have notable implications for the testing of the EH. It is possible that within each regime variable term premiums are second-order effects of yield dynamics on a prior basis. In particular, the EH holds within each regimes, but the violation of the EH is due to regime switches.

2.4. Data Issues and Yield Factors

2.4.1. The Data

The yield curve consists of the end-of-month observations of 1, 3, 6, 12, 24, 36, 60, 84, 120 months zero-coupon yields on treasury securities. The sample covers a period from January 1983 to May 2009. The data source is econstatsTM. Figure 3.1 plots the U.S. yield curves. One stylized fact of yields is that they tend to exhibit considerable persistence and are thus believed to be nonstationary or better approximated by an integrated process. This feature has profound implications for the estimation and the statistical inference. The upper panel of Table 2.1 provides the evidence of persistence of unit root tests. The classic CS regression requires that yield spreads are stationary.

Figure 2.1. U.S. Yield Curves, 1983.01–2009.05

The middle panel of Table 2.1 is the Johansen cointegration analysis results¹⁴. The cointegrations may explain another important stylized fact of the yield curve: spreads are less persistent than yields.

2.4.2. Yield Factors

Yield factors are obtained by estimating the dynamic Nelson-Siegel model. Most term structure models use three factors to capture stylized facts of yields in cross-section and time series. By properly restricting the factor loadings in the statistical factor model, Diebold and Li (2006) proposed the dynamic Nelson-Siegel model for the τ -period yield

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

where L_t is the level factor, S_t denotes the slope factor and C_t represents the curvature factor. The parameter λ_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. Empirically, the level factor corresponds to the long-term interest rates, 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

¹⁴The pairwise cointegration test results are available upon request.

Table 2.1: Summary Statistics of Yields and Yield Factors

Panel A: Yields						
Maturity (months)	Mean	Std. Dev.	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$	ADF
1	4.7161	1.5176	0.9785	0.6623	0.1708	-1.0878
3	4.9747	1.5576	0.9857	0.6807	0.2130	-1.2846
6	5.7164	1.5752	0.9857	0.6870	0.2234	-1.3507
12	5.3656	1.5877	0.9854	0.7002	0.2622	-1.4357
24	5.7606	1.6064	0.9848	0.7202	0.3340	-1.1759
36	5.9690	1.5950	0.9846	0.7296	0.3745	-1.2097
60	6.2964	1.5682	0.9848	0.7427	0.4268	-1.2274
84	6.5403	1.5522	0.9856	0.7557	0.4595	-1.2260
120	6.6908	1.5277	0.9859	0.7663	0.4812	-1.3084

Panel B: Johansen Cointegration Analysis			
	Rank	Test Statistics	p_value
Trace Test:	0	505.85	[0.000]
	1	319.90	[0.000]
	2	212.12	[0.000]
	3	139.38	[0.000]
	4	86.55	[0.001]
	5	44.476	[0.100]
	6	18.937	[0.508]
	7	8.942	[0.378]
	8	1.309	[0.253]

Panel C: Yield Factors						
	Mean	Std. Dev.	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$	ADF
Level	7.0027	1.4924	0.9854	0.7793	0.5039	-1.4655
Slope	-2.2428	1.2430	0.9760	0.4289	-0.2388	-2.5387
Curvature	-0.2290	1.4623	0.9250	0.4340	-0.0494	-3.2828

Panel D: Correlations	
Level and Empirical Level	0.9817
Slope and Empirical Slope	0.9928
Curvature and Empirical Curvature	0.9743

Note: $\rho(i)$ is autocorrelation with lag length i ; ADF is the augmented Dickey-Fuller test with lag length selected by AIC.

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. The three factors contain information on the macroeconomic dynamics and vice versa (Diebold, Rudebusch and Aruoba (2006), Zhu and Shahidur (2009)).

Figure 2.2. Yield factors and empirical counterparts.

The dynamic Nelson-Siegel model is flexible enough to match the changing shape of the yield curve, but it is still parsimonious and easy to estimate. I stick to the tradition of Diebold and Lin (2006) and estimate the DNS model by OLS with fixed $\lambda = 0.0603$. In so doing, the yield factors at time point t only depend on the observable yields at time t . Thus, adding the yield factors in the information set will not lead to the use of posterior information. The summary statistics for the extracted factors are presented in panel C of Table 2.1. Panel D of Table 2.1 is the pairwise correlations among the yield factors and empirical counterparts. Figure 2.2 plots three yield factors¹⁵ from the DNS model and their empirical counterparts. The proxy for the empirical level factor is 10-year interest rate. The proxy for the empirical slope factor is the yield spreads between 10-year and 3-month yields. The curvature factor is proxied by the average of 10-year, 2-year and 3-month yields.

2.5. Empirical Analysis

The present value model of the expectations hypothesis circumvents the pairwise investigation of yields, but still shed light on the accuracy of the expectations hypothesis. The curvature factor is a missing factor in CS1. It contains information about the

¹⁵Instead of plotting the slope factor, I plot the minus slope factor.

middle section of the yield curve that are not captured by the level and slope factors. The omission of the curvature factor means that we do not use market information efficiently. The inclusion of the curvature factor also help identify two regimes.

The standard Hamilton (1989, 1994) algorithm can be used to estimate the model in Eq.(2.20). The smoothed probabilities that are usually parameters of interest can be calculated from the Kim (1996) filter. The entertained model is a MSH-VAR(1) model. MSH means a Markov-switching heteroscedasticity model with a VAR lag length 1. The model is selected by specification analysis on the error terms. The recursive tests are conducted since January 2001 because this is the point after which the transition probabilities p and q are usually significant. The recursive Wald statistics of testing the present value model are plotted in Figure 2.3 (marked as CS1 in the figure). The asymptotic distribution of the test statistics is a $\chi^2(6)$ distribution. The bold straight bold line is ten percent critical value. The validity of the expectations hypothesis cannot be rejected consistently¹⁶.

A large body of empirical evidence suggests that regime shifts are important for explaining stylized facts of the yield curve. These studies usually relate regimes with business cycles. The upper panel of Figure 2.4 graphs the regime classifications¹⁷ from the present value model of the term structure. They are smoothed probabilities of being in regime 1 that are extracted by the Kim filter. Because the only regime-dependent parameters vector is the variance-covariance matrix, regimes here have clear interpretation of high and low volatility states. For identification, regime 1 is labeled as the high volatility regime. As indicated in the plot, regime shifts relate to business cycles. The sample covers a monetary policy shift. Before 1988, the Fed takes a borrowed-reserves operating procedure. Since 1988, It is well-known that the Fed has been targeting the funds rate directly. The monetary policy shift may contaminate the regime identifications. Thus, two regimes do not have a clear interpretation of being economic booms and recession.

¹⁶In the conduction of empirical tests, the discount rate δ is set to equal to $1/(1+L/12)$, L is the mean of level factor abstracted from the AFDNS model. It is equal to 0.9926.

¹⁷Both the regime classifications and below the computation of theoretical spreads are based on the full-sample estimation.

Figure 2.3. Recursive tests for the validity of the cross-equation restrictions implied by the expectations hypothesis. The bold straight line is 10 percent critical values. CS1 is the Wald statistics of testing the present value model of the EH. CS2_1 is the Wald statistics of testing CS2 model of the EH with 10-year and 3-month yields. CS2_2 is The Wald statistics of testing CS2 model of the EH with 5-year and 3-month yields.

Figure 2.4. The probabilities of being in a tranquil regime. The shaded bars indicate NBER recessions.

The consistent non-rejection of the EH implies that both the variables relating to time-varying risk premiums and peso problems play a role in the failure of the expectations hypothesis. Bekaert and Hodrick (2001) summarized three potential reasons for the rejection of the EH. The second interpretation is omitted variables that might capture time-varying risk premiums. The extension of the information set for testing the EH might provide an insight on this issue if time-varying risk premiums can be captured by the yield factors. This testing framework explicitly takes regime shifts into account, thus it also sheds light on the peso problems or learning. Although the rational expectations assumption is made in the testing, the extension along the second and third lines in Bekaert and Hodrick (2001) can resurrect the EH. In contrast, Bekaert, Hodrick and Marshall (2001) considered peso problems, but they did not find evidence in favor of the expectations hypothesis. They stick to Campbell and Shiller (1991) single-equation regression that may lead to simultaneous bias. Furthermore, the regression also neglects the information in the yield factors.

The CS2 form of the EH in Eq.(2.15) is a commonly cited form of the expectations hypothesis. The evaluation of the EH based on Eq.(2.15) makes my results comparable with other related empirical works. Furthermore, it can serve as a robustness check on the present value model of the expectations hypothesis. $y_t = [S_{t,n}, \Delta i_{t,1}, C_t]$ is the conditional information set at time point t for testing Eq.(2.15). Since C_t , the curvature factor, is a medium-term factor, $S_{t,n}$ and $\Delta i_{t,1}$ are supposed to contain information about the short- and long-end of the yield curve. In the empirical tests of the EH, 3-month yield is commonly used short-term interest rate, and 10-year yield is frequently used long-term interest rate. Thus, I choose 3-month yield as the short-term interest rate. For the long-term interest rate, two candidate yields are 10-year yield and 5-year yield. So I test Eq.(2.15) based on two pair of interest rates. One pair is 10-year yield and 3-month yield, and the other pair is 5-year yield and 3-month yield.

The line CS2_1 and CS2_2 in Figure 2.3 are the Wald statistics of testing the Eq.(2.15) based on a pair of 10-year and 3-month yields and a pair of 5-year and 3-month yields. Likewise, the EH are usually not rejected. The results provide new evidence on the non-rejection of the EH. All results together imply that an appropriate information set and regime switches can account for the empirical failure of the EH.

The regime classifications for the pair of 10-year and 3-month yields are plotted in the middle panel of Figure 2.4. The lower panel of Figure 2.4 is a plot of the regime classifications for the pair of 5-year and 3-month yields. Since both the present value model and the CS2 form of the EH use information in the level, slope and curvature factors, it is reasonable that a similar regime identification is generated.

The empirical results indicate that within each regime time-varying risk premiums are second-order effects of the term structure of interest rates. This contradicts to a large literature of the term structure modeling that concludes that time-varying risk premiums are predictable. My interpretation for the discrepancy is that the predictability of variable risk premiums are based on the posterior analysis. However, it is possible that on a prior basis, risk premiums are less predictable. For example, using survey data on interest rate forecasts, Piazzesi and Schneider (2009) found that subject premiums are less volatile and not very cyclical. Furthermore, Corte, Sarno and Thornton (2008) found that though the EH is statistically rejected, the departures of the EH are not profitable based on criteria of economic significance in the context of a simple trading strategy.

Figure 2.5 is a plot of the theoretical and actual spreads. The upper panel plots the MS factor and the theoretical MS factor computed from Eqs. (2.28) and (2.29). The MS and theoretical MS factors are highly correlated with a correlation coefficient 0.90. The middle panel draws the actual yield spreads between 10-year and 3-month yields and the theoretical spreads calculated from Eqs. (2.32) and (2.33). The correlation between the actual and theoretical spreads is 0.89. The lower panel is a plot of the yield spreads between 5-year and 3-month yields and the corresponding theoretical spreads¹⁸. The correlation coefficient is 0.71.

2.6. Concluding Remarks

This paper has reexamined the expectations hypothesis of the term structure of interest rates. The empirical results indicate that the expectations hypothesis cannot be rejected. The non-rejection of the expectations hypothesis is achieved through using the yield factors to capture time-varying risk premiums and taking into account regime

¹⁸The theoretical spreads are smoother than the actual spreads. I plot two times of the theoretical spreads with the actual spreads.

Figure 2.5. Theoretical and actual spreads.

switches. Furthermore, the regimes relate to business cycles. My interpretation for the resurrection of the expectations hypothesis is the use of an appropriate information set and a Markov-switching VAR model as the data generating process for testing the EH.

The above results indicate a promising direction for future research. The interpretation of the empirical results are based on the argument that time-varying risk premiums are not very predictable on a prior basis. Using survey data on interest rate forecasts, Piazzesi and Schneider (2009) found that subject premiums are less volatile and not very cyclical. However, a large literature finds that time-varying risk premiums are predictable on a posterior basis. It seems interesting to further investigate the prior predictability of variable risk premiums.

Appendix A

As implied by Eq. (2.5),

$$(A1) \quad E_t(h_{t+1,n}) - i_{t,1} = -Cov_t(m_{t+1}, h_{t+1,n}) - 0.5V_t(h_{t+1,n})$$

By combining Eq. (2.44) and (2.45), the holding period return is

$$\begin{aligned}
h_{t+1,n} &= p_{t+1,n-1} - p_{t,n} \\
&= A_{n-1} + B_{n-1}x_{t+1} - A_n - B_n x_t \\
&= B_{n-1}x_{t+1} + (A_{n-1} - A_n - B_n x_t)
\end{aligned} \tag{A2}$$

Since the item in the parenthesis is in the time t information set,

$$\begin{aligned}
Cov_t(m_{t+1}, h_{t+1,n}) &= B_{n-1}Cov_t(m_{t+1}, x_{t+1}) \\
V_t(h_{t+1,n}) &= B_{n-1}V_t(x_{t+1})
\end{aligned} \tag{A3}$$

Eqs. (2.39) and (2.41) implies a ARMA process for m_{t+1} ,

$$(A4) \quad m_{t+1} = (\phi - 1)\mu + \phi m_t - \beta \varepsilon_{t+1} - (\beta \phi + 1)\varepsilon_t$$

Obviously, the only shock for m_{t+1} and x_{t+1} is ε_{t+1} , then $Cov_t(m_{t+1}, x_{t+1}) = -\beta\sigma^2$ and $V_t(x_{t+1}) = \sigma^2$. It comes naturally

$$(A5) \quad E_t(h_{t+1,n}) - i_{t,1} = -B_{n-1}\beta\sigma^2 - B_{n-1}^2\sigma^2/2$$

Appendix B

Assume the regimes follow a discrete Markov chain process. As in Hamilton (1994, chapter 22), a useful representation is obtained by letting ξ_t denote a random 2×1 vector defined as

$$(B1) \quad \xi_t = \begin{cases} (1, 0)' & \text{if } k_t = 0 \\ (0, 1)' & \text{if } k_t = 1 \end{cases}$$

Thus, the Markov chain is a vector AR(1) process

$$(B2) \quad \xi_{t+1} = P\xi_t + v_t$$

with P given in Eq.(2.22). For some specific element i of the vector ξ_t , we have

$$(B3) \quad Var(\xi_{it}|Y_t) = \hat{\xi}_{it}^\wedge(1 - \hat{\xi}_{it}^\wedge)$$

Where $\hat{\xi}_{it}^\wedge$ is from the Hamilton filter. Because regimes are latent variables, Eqs. (2.47) and (2.49) gives

$$\begin{aligned} h_{t+1,n} &= \xi_{t+1}^0 * p * h_{t+1,n}^0 + \xi_{t+1}^0 * (1 - p) * h_{t+1,n}^0 + \\ &\quad (1 - \xi_{t+1}^0) * q * h_{t+1,n}^1 + (1 - \xi_{t+1}^0) * (1 - q) * h_{t+1,n}^1 \\ &= \xi_{t+1}^0 h_{t+1,n}^{00} + (1 - \xi_{t+1}^0) h_{t+1,n}^{11} \end{aligned} \quad (B4)$$

Moreover, Eqs.(2.47) and (2.50) gives

$$\begin{aligned} f_{t,n} &= \xi_t^0 * p * f_{t,n}^0 + \xi_t^0 * (1 - p) * f_{t,n}^0 + \\ &\quad (1 - \xi_t^0) * q * f_{t,n}^1 + (1 - \xi_t^0) * (1 - q) * f_{t,n}^1 \\ &= \xi_t^0 f_{t,n}^{00} + (1 - \xi_t^0) f_{t,n}^{11} \end{aligned} \quad (B5)$$

By combining Eqs.(3.18) and (B5), we find the covariance between forward rates and excess returns implied by the discrete Vasicek model is given by

$$\begin{aligned} Cov_t(f_{t,n}, h_{t+1,n}) &= (h_{t+1,n}^{00} f_{t,n}^{00} + h_{t+1,n}^{00} f_{t,n}^{11} + h_{t+1,n}^{11} f_{t,n}^{00} + h_{t+1,n}^{11} f_{t,n}^{11}) \\ &\quad * Cov(\xi_{t+1}^0, \xi_t^0) \end{aligned} \quad (B6)$$

While from Eq. (3.16), we know

$$(B7) \quad \xi_{t+1}^0 = p * \xi_t^0 + (1 - q) * (1 - \xi_t^0) + v_{1t}$$

Because $cov(R_t, v_{t1}) = 0$, then,

$$(B8) \quad Cov_t(\xi_{t+1}^0, \xi_t^0) = p * Var(\xi_t^0) + (1 - q) * Var(\xi_t^0)$$

Eqs. (B5), (B7) and (2.47) imply

$$(B9) \quad Cov_t(f_{t,n}, h_{t+1,n}) = D(1 + p - q)\sigma_t^2$$

with

$$D = (h_{t+1,n}^{00}f_{t,n}^{00} + h_{t+1,n}^{00}f_{t,n}^{11} + h_{t+1,n}^{11}f_{t,n}^{00} + h_{t+1,n}^{11}f_{t,n}^{11})$$

Except some pathological case where $D = 0$, the covariance is not equal to zero. This has clear implication: the predictive power of forward rates is consistent with the expectations hypothesis of the term structure. In Eq. (B8), $(1 + p - q)\sigma_t^2$ is positive-definitive. In Eq. (B4), clearly $f_{t,n}$ can't be negative (nominal interest rate). Only if the excess return in Eq. (B3) is positive, as usually it is, then Eq. (B8) provides a positive correlations. While it provides no indication on the forecasting structure as maturity n changes. Yet this can be done by simulation.

CHAPTER 3

A Regime Switching Macro-finance Model of the Term Structure

3.1. 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 future economic activity (e.g., among others, Estrella and Hardouvelis (1991), Estrella and Mishkin (1998)). For example, investors require higher risk premia on long-term bonds in economic recessions (bad times). This implies that premia on long bonds are countercyclical. Meanwhile, it has been well-documented that yield spread is a leading indicator of economic recessions. 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 interactions between 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 forces in the term structure models with macro factors¹. 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 bond yields in addition to improved forecasting performance. Diebold, Rudebusch and Aruoba (2006) (DRA henceforth) provided some strong evidence of dynamic interactions between the yield

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

curve and economic activity in a framework of the Nelson-Siegel (1987) type term structure model.

The macro-finance model proposed in this paper extends the DRA dynamic Nelson-Siegel model by incorporating regime shift into the joint dynamics. Nowadays regime shifts stand as a stylized fact in the term structure modeling. Statistically, the term structure models with regime switches describe the historical yield curves better than single-regime models². From an asset pricing perspective, the term structure models with regime switches can account for some well-documented puzzles of the term structure (see, for example, Bansal and Zhou (2002), Dai, Singleton and Yang (2007)), for instance, the violation of the expectations hypothesis and the predictability of excess bond returns. From the central bank's perspective, it is important to understand the role of the yield curve in the monetary transmission mechanism in different regimes. In addition, 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 a short lag length. In this sense, the regime-switching term structure models represent a parsimonious way to capture non-linear interactions between the term structure dynamics and business cycle effects. The proposed macro-finance model includes latent yield factors and unobserved regimes, to estimate the macro-finance model, I introduce a Markov Chain Monte Carlo (MCMC) algorithm that allows us to simultaneously fit the yield curves and filter out regimes.

The DRA dynamic Nelson-Siegel model with regime shifts has several advantages. First, 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 typically no-arbitrage models in finance literature (Duffee (2002)). Second, it allows a bidirectional feedback mechanism with which the entire yield curve responds to the macroeconomic information, and vice versa. The bidirectional feedback mechanism allow us to answer some important questions, such as, how empirically relevant is the feedback from the yield curve? In contrast, to be statistically tractable, many affine term structure models with macroeconomic variables disregard the feedback from the yield curve to economic activity. Third, the macro-finance model with regime switches

²See, for example, Cecchetti, Lam and Mark (1993), Gray (1996), Garcia and Perron (1996), and Ang and Bekaert (2002).

allow us to investigate bidirectional feedbacks in each regime. 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 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 financial markets, the data should reflect this matter of fact. Empirical evidence (Duffee, 2008) indicates that the no-arbitrage restrictions have no practical effect on forecast accuracy. 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.

In my empirical analysis of the macro-finance model, I find that both regime switches and bidirectional interactions between macroeconomic and yield factors are important components of the term structure model. In the turbulent regime, the proportion of the forecast variance of yields attributable to macro factors is stable along the maturity spectrum of yields. Around 40% of the forecast variance is attributable to macro factors at a 1-month forecast horizon³. At a 60-month forecast horizon, around 50% of the forecast variance is due to macro factors. In the tranquil regime, the proportion of the forecast variance of yields attributable to macro factors increases along the forecast horizon. Furthermore, the short-end of the yield curve is less subject to the shocks of macro factors. At a 60-month forecast horizon, 20% of the forecast variance is attributable to macro factor for 3-month yields. In contrast, 45% of the forecast variance is attributable to macro factors for 5-year yield. In the tranquil regime, the level factor dominates the forecast variance at a long forecast horizon. At a 60-month forecast horizon, the level factor can account for 66%, 55%, 64% of the forecast variance for CU, inflation and FFR. In contrast, at a 1-month forecast horizon, only 5%, 0% and 12% of the forecast variance are explained by the level factor. Interestingly, in the

³The macroeconomic factors include the capacity utilization, inflation, and the federal fund rate.

high volatility regime, the feedback from yield factors to macro fact seems break down in the sense that the forecast variance of macro variables is not attributable to yield factors. However, the impulse responses and statistical tests confirm the interactions between yield and macro factors in both regimes.

The rest of this paper is organized as follows. In section 2, I present and estimate the benchmark ‘yields-only’ model⁴ subject to regime shifts. Section 3 proposes and estimates the regime switching macro-finance model. The implications and relationship between the yield curve 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.

3.2. Yields-only model

Principal component analysis shows that a few factors can explain over 97% (Piazzesi (2004)) of the variability of yields in the cross-section and time series. These factors are usually labeled as ‘level’, ‘slope’ and ‘curvature’ according to their effect on the yield curve. Because the interpretation of yield factors seems to be stable across different specifications and sample selections, most term structure models use three factors to capture stylized facts of yields. In order to achieve parsimony, cross-section restrictions are usually imposed on these term structure models.

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 hedged 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 empirical fit makes this type of models widely applied in central banks and investment banks. This article follows DRA and goes along this strand of literature.

⁴A model of the yield curve without macroeconomic variables.

3.2.1. 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 τ ,

$$(3.1) \quad i_{t(\tau)} = L_t + S_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + C_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) + \varepsilon_t$$

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. λ 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 is

$$(3.2) \quad \begin{bmatrix} i_{t(\tau_1)} \\ i_{t(\tau_2)} \\ \vdots \\ i_{t(\tau_N)} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{bmatrix} \begin{bmatrix} L_t \\ S_t \\ C_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t(\tau_1)} \\ \varepsilon_{t(\tau_2)} \\ \vdots \\ \varepsilon_{t(\tau_N)} \end{bmatrix}$$

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 process⁵

⁵This specification allows regime-dependent heteroscedasticity, but autoregression coefficients are not regime-dependent.

$$(3.3) \quad \begin{bmatrix} L_t - \mu_L \\ S_t - \mu_S \\ C_t - \mu_C \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \begin{bmatrix} L_{t-1} - \mu_L \\ S_{t-1} - \mu_S \\ C_{t-1} - \mu_C \end{bmatrix} + \begin{bmatrix} \eta_{\xi_t, 1t} \\ \eta_{\xi_t, 2t} \\ \eta_{\xi_t, 3t} \end{bmatrix}$$

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,

$$(3.4) \quad \begin{aligned} \eta_H &\sim N(0, \Sigma_H) \\ \eta_L &\sim N(0, \Sigma_L) \end{aligned}$$

For optimality of the Kalman filter, I assume the disturbances η_t and ε_t are uncorrelated with each other, and initial state X_0 is orthogonal to the realization of η_{ξ_t} and ε_t

$$(3.5) \quad \begin{aligned} E(\varepsilon_t \eta_{\xi_t}) &= 0 \text{ for } t = 1, 2, \dots, T; \xi_t = H, L \\ E(\varepsilon_t X_0) &= 0 \text{ for } t = 1, 2, \dots, T \\ E(\varepsilon_t X_0) &= 0 \text{ for } t = 1, 2, \dots, T \end{aligned}$$

If we stack the state variables in a 3×1 vector $X_t = (L_t, S_t, C_t)$ and stack the yields in $\mathbf{y}_t = (i_{t(\tau_1)}, \dots, i_{t(\tau_n)})$. The state space model can be succinctly written in matrix notation as

$$(3.6) \quad \begin{aligned} \mathbf{y}_t &= \Lambda X_t + \varepsilon_t \\ X_t &= \Phi X_{t-1} + \eta_{\xi_t}; \xi_t = H \text{ or } L \end{aligned}$$

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

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

Now the standard Hamilton (1989, 1994) and Krolzig (1997) algorithms can be used to extract the probabilities of 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)). Nowadays regime shifts stand as a stylized fact in the term structure modeling. From an asset pricing perspective, the term structure models with regime switches can account for some well-documented puzzles of the term structure (see, for example, Bansal and Zhou (2002), Dai, Singleton and Yang (2007)), for instance, the violation of the expectations hypothesis and the predictability of excess bond returns. From the central bank's perspective, it is important to understand the role of the yield curve in the monetary transmission mechanism in different regimes. In addition, 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 a short lag length. In this sense, the regime-switching term structure models represent a parsimonious way to capture nonlinear interactions between the term structure dynamics and business cycle effects.

The coefficient matrix Λ in Eq.(3.6) plays three roles in our analysis. Three yield factors and two regimes are unobserved components in the system (3.6). As usual, there are some identification conditions that must be imposed to estimate a model with latent factors. The matrix Λ provides such identification restrictions. Since three yield factors explain most of variations of yield dynamics, they are supposed to be highly correlated with three principal components (Zhu (2008)). The matrix Λ 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 Λ . Admissibility (Dai and Singleton (2002)) constitutes a third role of the matrix Λ . As discussed in DRA, the Nelson-Siegel form avoids a negative forward rate at all horizons.

The entertained model achieves parsimony by a diagonal Ω assumption. Since three underlying latent factors explain a large fraction of yield variation⁶, the diagonal

⁶The model can explain over 98% of the variance of yield changes (Diebold and Li (2006), Diebold, Rudebusch and Aruoba (2006)).

Ω 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.

3.2.2. Yields

The yield curve consists of the end-of-month observations of 1, 3, 6, 12, 24, 36, 60, 84, 120 months zero-coupon yields on treasury securities. The sample covers a period from January 1983 to May 2009. The data source is econstatsTM. Figure 3.1 plots the U.S. yield curves. It is clear from the figure that the yield curves have many different shapes. One stylized fact of yields is that they tend to exhibit considerable persistence and are thus believed to be nonstationary or better approximated by an integrated process. This feature has profound implications for the macro-finance model estimation. Table 3.1 provides the evidence of persistence of yields and the summary statistics of yields.

A long sample makes statistical inference more reliable. However, the relationships between the yield curve and macroeconomic factors have changed in last decades. In 1982, the Fed operating procedure has shifted from a non-borrowed-reserves targeting to a borrowed-reserves targeting (Walsh (2003), Chapter 9). This may contaminate the relationships between the yield curve and macroeconomic factors. Thus, I choose January 1983 as the starting point of our sample.

3.2.3. The MCMC algorithm

The state-space system Eq.(3.6) is estimated by the Markov chain Monte Carlo (MCMC) method, specifically, a Gibbs sampling algorithm combined a random walk Metropolis step (see Appendix A 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

Figure 3.1. U.S. Yield Curves, 1983.01–2009.05

Table 3.1: Summary Statistics of Yields and Yield Factors

Maturity (months)	Mean	Std. Dev.	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$	ADF
1	4.7161	1.5176	0.9785	0.6623	0.1708	-1.0878
3	4.9747	1.5576	0.9857	0.6807	0.2130	-1.2846
6	5.7164	1.5752	0.9857	0.6870	0.2234	-1.3507
12	5.3656	1.5877	0.9854	0.7002	0.2622	-1.4357
24	5.7606	1.6064	0.9848	0.7202	0.3340	-1.1759
36	5.9690	1.5950	0.9846	0.7296	0.3745	-1.2097
60	6.2964	1.5682	0.9848	0.7427	0.4268	-1.2274
84	6.5403	1.5522	0.9856	0.7557	0.4595	-1.2260
120	6.6908	1.5277	0.9859	0.7663	0.4812	-1.3084

Note: $\rho(i)$ is autocorrelation with lag length i ; ADF is augmented Dickey-Fuller test with lag length selected by AIC.

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 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

a 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 the 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 the approximation method is unknown. In some cases, the accuracy provided by the approximation method is probably not good enough. Furthermore, for a high dimensional model like the macro-finance model in next section, the likelihood function may be subject to multiple local optima.

3.2.4. Convergence checks

Some diagnostics are available on the reliability of our estimation method. 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 assess the convergence of the MCMC-Gibbs algorithm, I implement another three practical statistics. First, the MCMC-Gibbs drawings allow us to compute the Raftery and Lewis (1992) minimum burns and minimum runs of required to estimate the 0.025 quantile to within ± 0.025 with probability 0.95. For all the parameters, the minimum required burn-in is only several hundred and the minimum number of runs is several thousand.

The second statistic is the Yu and Mykland (1998) plot of CUSUM path, for a specific parameter θ with sample variance σ_θ^2 and mean μ_θ for the sample up to iteration t ,

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

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 third criterion for the convergence is the relative numerical efficiency (RNE) proposed by Geweke (1992). As the drawings for the latent 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.

3.2.5. Empirical results

Diebold and Li (2006, 2008) fix the λ 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. Instead of fixing λ at a constant, in this paper I use a random walk Metropolis step to draw λ as guided by Johannes and Polson (2004). The relaxation allows the dynamic Nelson-Siegel model fit the yield curves more flexible and increases the efficiency of the estimation.

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 thirty-three parameters to estimate: a parameter reflecting the change rate of factor loadings λ , 9 parameters in diagonal variance-covariance matrix Ω ; for each regime, 6 parameters in non-diagonal covariance matrix Σ_{ξ_t} ; 9 parameters in autoregressive coefficient matrix Φ ; and 2 regime transition parameters.

Table 3.2: Yields-only Model^{a,b}

The rate of factor loading changes λ :			0.0516 (0.00002)	
Φ	L_{t-1}	S_{t-1}	C_{t-1}	μ
L_t	0.900 (0.0001)	-0.022 (0.0001)	0.067 (0.0003)	7.780 (1.634)
S_t	-0.053 (0.0001)	0.938 (0.0001)	0.088 (0.0004)	-1.958 (1.344)
C_t	0.001 (0.0001)	-0.008 (0.0001)	0.946 (0.0004)	-0.321 (1.452)
Estimated Covariance matrix Σ_L				
	L_t	S_t	C_t	
L_t	0.089 (0.013)			
S_t	0.029 (0.013)	0.221 (0.036)	0.034 (0.018)	
C_t	0.034 (0.018)	-0.044 0.027	0.381 0.042	
Estimated Covariance matrix Σ_H				
	L	S_t	C	
L_t	0.463 (0.129)	0.641 (0.235)		
S_t	0.641 (0.235)	2.192 (0.614)		
C_t	0.073 (0.241)	-0.014 (0.511)	3.621 (1.166)	
Transition Probabilities p and q				
	p		q	
	0.63 (0.0053)		0.76 (0.0055)	
Test for Diagonality of Σ Matrix				
	Wald statistic ^c		P-value	
Regime L	16.155		0.001	
Regime H	7.847		0.049	
Test for no Regime-dependent Heteroscadasticity ^d				
	Wald statistic		P-value	
	19.514		0.0034	

^aBold entries indicate 5% significance. Standard errors are in the parentheses.^b L denotes low volatility regime and H is high volatility regime.^cWald statistics are asymptotically Chi-square with 3 degrees of freedom.^dWald statistic has a Chi-square distribution with 6 degrees of freedom.

The details of the MCMC algorithm are presented in the Appendix A. 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, I use the Kalman filter to process yields forward, then I 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 the Kim (1994) filter produces drawings of the smoothed regimes.

To facilitate the convergence of the MCMC iterations, I initialize the MCMC by a two-step estimation. The first step runs the OLS to estimate yield factors by fixing λ at 0.0598. With estimated yield factors, the state equation (3.3) 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 MCMC 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 MCMC drawings are usually nonstationary after dozens of iterations. I don't drop nonstationary iterations. On average, the entertained state equation is stationary⁷. 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 15000 iterations with an initial burn-in period of 5000 observations. All three measures of convergence, respectively, visual plot of parameters, CUSUM and NW, indicate that the sampled parameters are converged.

⁷The largest eigenvalue of the coefficient matrix A is 0.9464.

Figure 3.2. Level, slope, curvature factors and the empirical counterparts.

Table 3.2 presents the parameter estimates of the yields-only model. The estimates of the autoregressive coefficient matrix Φ contain much information. First, the value of λ is consistent with our prior belief that it maximizes the curvature factor in the medium-term⁸. Second, three latent yield factors are highly persistent. This is consistent with typical results found in the term structure modeling. Third, 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. Fourth, cross-correlations across yield factors are small but still significant.

The extracted level, slope and curvature are plotted in figure 3.2. 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 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 0.98 for the level, 0.98 for the slope and 0.97 for the curvature.

⁸The scaling factor ς in the Random Walk Metropolis steps (see Appendix A) for generating λ is set to be 0.000295. The acceptance rate of the Random Walk Metropolis steps is 0.2504.

Figure 3.3. Probabilities of being in a tranquil regime.

Figure 3.4. Probabilities of being in a tranquil regime. The shaded bars indicate the NBER economic recessions.

The correlation analysis indicates why the latent factors are labeled as ‘level’, ‘slope’, and ‘curvature’.

Table 3.3 reports the fits of the yields-only model and macro-finance models. For each model, We 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 salient 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 significant.

The upper panel of Figure 3.4 plots the smoothed probabilities 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 Σ_L and Σ_H in Table 3.2. The regime classifications confirm the well-documented observation that regime H tends to be associated with economic recessions in the sense that the probabilities of being in regime L is close to 0. Meanwhile, it seems that the regime classifications have plentiful information content beyond only signaling economic recessions.

Table 3.3: Summary Statistics for Measurement Errors of Yields

Maturity	Yields-only model		Macro-finance model	
	Mean	Std. Dev.	Mean	Std. Dev.
1-month	-23.84	10.03	-22.48	6.86
3-month	-7.08	1.39	-4.41	0.64
6-month	0.35	0.08	4.27	0.18
12-month	-3.28	0.58	1.53	0.56
24-month	0.71	0.15	4.71	0.15
36-month	-4.41	0.12	-1.43	0.09
60-month	-5.40	0.17	-2.47	0.21
84-month	-0.72	0.17	3.55	0.08
120-month	2.47	0.38	3.84	0.07

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

3.3. 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 investors 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.

3.3.1. 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)⁹ 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, we try to characterize the relationship between the yield curve and economic activity at a higher frequency. Inflation is included in the extended macro-finance model because it is a key variable in shaping the nominal yield curve through the level of inflation and inflation risk premia (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. We consider several other variables, such as average weekly hours at a monthly frequency, they have similar implications for regime identification purpose.

3.3.2. 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 6×1 vector $X_t^{MF} = (L_t, S_t, C_t, CU_t, \pi_t, FFR_t)'$ be factors in the macro-finance model, then the state equation is

⁹On 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 3.5. Macro factors and level, slope, curvature factors from the macro-finance model.

$$(3.7) \quad \begin{bmatrix} L_t - \mu_L \\ S_t - \mu_S \\ C_t - \mu_C \\ CU_t - \mu_{CU} \\ \pi_t - \mu_\pi \\ FFR_t - \mu_{FFR} \end{bmatrix} = \Phi^{MF} \begin{bmatrix} L_{t-1} - \mu_L \\ S_{t-1} - \mu_S \\ C_{t-1} - \mu_C \\ CU_{t-1} - \mu_{CU} \\ \pi_{t-1} - \mu_\pi \\ FFR_{t-1} - \mu_{FFR} \end{bmatrix} + \begin{bmatrix} \eta_{\xi_t, 1t} \\ \eta_{\xi_t, 2t} \\ \eta_{\xi_t, 3t} \\ \eta_{\xi_t, 4t} \\ \eta_{\xi_t, 5t} \\ \eta_{\xi_t, 6t} \end{bmatrix}$$

where subscript $\xi_t = H$ or L , denote high and low volatility regimes¹⁰. And disturbances in regime L and H follow the Gaussian distribution.

$$(3.8) \quad \begin{aligned} \eta_L &\sim N(0, \Sigma_L^{MF}) \\ \eta_H &\sim N(0, \Sigma_H^{MF}) \end{aligned}$$

The observation equation is the same as in the state-space system Eq.(3.2)¹¹. The macro-finance model maintains the assumption of a diagonal covariance matrix Ω .

¹⁰This intuitive label is a little abuse in notation. However our empirical results justify two regimes with high and low volatility.

¹¹In DRA's representation [Eq.(6') in DRA], the state vector is a 6×1 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.

The state equation Eq.(3.7) is subject to regime switches, 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

$$(3.9) \quad \begin{aligned} \mathbf{y}_t &= \Lambda X_t^{MF} + \varepsilon_t \\ X_t^{MF} &= \Phi^{MF} X_{t-1}^{MF} + \eta_{\xi_t}; \quad \xi_t = L \text{ or } H \end{aligned}$$

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 90 parameters: the rate of change of factor loadings λ , the autoregressive coefficient matrix includes 36 parameters; the covariance matrix Σ_L and Σ_H respectively have 21 parameters; 2 parameters in transition matrix P ; and 9 parameters in the diagonal matrix Ω . The MCMC method is a preferred method for estimating the state-space system Eqs.(3.9). 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, we drop all non-stationary drawings to ensure that the estimated system is stationary. In particular, our implementation consists of 15000 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 1.089 for the most persistent level factor¹², while for the least persistent slope factor it is 0.839. 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. The value of λ is 0.0752, it is still consistent with our prior belief that it maximizes the curvature factor in the medium-term¹³. There are significant cross-correlations among the yield factors and the macro factors. Three yield factors play an important role in accounting for the macroeconomic dynamics, and vice versa. It is also important to note that inflation is not significant for explaining the dynamics of other factors.

¹²The state equation is stationary because the largest eigenvalue of the coefficient matrix A is 0.9804.

¹³The scaling factor ς in the Random Walk Metropolis steps (see Appendix A) for generating λ is set to be 0.0003. The acceptance rate of the Random Walk Metropolis steps is 0.3319.

Table 3.4: Macro-finance Model^{a,b}

The rate of factor loading changes λ : $\begin{matrix} 0.0752 \\ (0.00005) \end{matrix}$						
Autoregressive coefficient matrix Φ^{MF}						
	L_{t-1}	S_{t-1}	C_{t-1}	CU_{t-1}	π_{t-1}	FFR_{t-1}
L_t	1.089 (0.023)	-0.102 (0.022)	0.436 (0.041)	0.239 (0.020)	-0.018 (0.046)	0.276 (0.016)
S_t	0.102 (0.023)	0.839 (0.024)	0.384 (0.043)	0.126 (0.024)	-0.022 (0.045)	0.221 (0.016)
C_t	0.007 (0.009)	0.078 (0.010)	0.937 (0.022)	0.030 (0.015)	-0.001 (0.013)	0.084 (0.008)
CU_t	-0.009 (0.005)	0.009 (0.006)	0.030 (0.011)	0.999 (0.008)	0.012 (0.008)	0.002 (0.004)
π_t	-0.004 (0.011)	0.006 (0.012)	-0.030 (0.288)	-0.061 (0.020)	0.939 (0.017)	-0.013 (0.009)
FFR_t	-0.089 (0.021)	0.057 (0.020)	-0.373 (0.038)	-0.181 (0.019)	0.030 (0.043)	0.714 (0.015)
Estimated Covariance Matrix Σ_L^{MF}						
	L_t	S_t	C_t	CU_t	π_t	FFR_t
L_t	0.051 (0.007)					
S_t	-0.048 (0.008)	0.071 (0.001)				
C_t	0.037 (0.011)	-0.053 (0.014)	0.312 (0.046)			
CU_t	0.009 (0.007)	-0.003 (0.009)	0.036 (0.019)	0.154 (0.010)		
π_t	0.003 (0.005)	0.002 (0.006)	0.001 (0.014)	0.010 (0.010)	0.073 (0.011)	
FFR_t	0.001 (0.002)	0.008 (0.003)	0.011 (0.005)	0.004 (0.004)	0.005 (0.002)	0.011 (0.002)
Estimated Covariance Matrix Σ_H^{MF}						
	L_t	S_t	C_t	CU_t	π_t	FFR_t
L_t	0.109 (0.018)					

Table 3.4 (*continued*)

	L_t	S_t	C_t	CU_t	π_t	FFR_t
S_t	-0.064 (0.016)	0.120 (0.010)				
C_t	0.054 (0.027)	0.012 (0.015)	0.455 (0.088)			
CU_t	0.058 (0.022)	-0.017 (0.009)	0.020 (0.048)	0.361 (0.079)		
π_t	0.044 (0.019)	-0.026 (0.006)	-0.004 (0.039)	0.054 (0.032)	0.244 (0.041)	
FFR_t	0.041 (0.013)	0.036 (0.003)	0.101 (0.029)	0.044 (0.021)	0.031 (0.017)	0.115 (0.019)
Transition Probabilities p and q						
	p			q		
	0.693 (0.054)			0.820 (0.040)		
Test for Diagonality of Σ^{MF}						
	Wald Statistic ^c				P-value	
Regime L	75.13				0.000	
Regime H	53.27				0.000	
Test for no Regime-dependent Heteroscedasticity ^d						
	Wald Statistic				P-value	
	51.605				0.0002	

^aBold entries indicate significant at 5 percent level, standard deviations are in parentheses.

^b L denotes low volatility regime and H is high volatility regime.

^cWald statistics are asymptotically Chi-square with 3 degrees of freedom.

^dWald statistic has a Chi-square with 21 degrees of freedom.

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 4.4 rejects the null hypothesis of equal variance in two regimes. The Wald statistics also indicate that neither the covariance matrix Σ_L^{MF} nor the matrix Σ_H^{MF} are diagonal. The lower panel of Figure 3.4 plots the smoothed probabilities of being in regime L . We see that the macro-finance model have advantage in terms of regime identification.

It is not surprising because the macroeconomic variables contain information about economic recessions. The regime classifications indicate that the high volatility regime is associated with economic recessions. This is consistent with the widely documented observation in the literature.

Although the autoregressive coefficient matrixes are different¹⁴, 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 3.5 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 FFR with the highest correlation 0.72 between the curvature and FFR. The level factor is correlated with the inflation with a correlation 0.46. To a less extent, the correlation between the slope and the capacity utilization is 0.42.

3.3.3. 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 (6×6) matrix Φ^{MF} into four (3×3) blocks

$$(3.10) \quad \Phi^{MF} = \begin{bmatrix} \Phi_1^{MF} & \Phi_2^{MF} \\ \Phi_3^{MF} & \Phi_4^{MF} \end{bmatrix}$$

and similarly partitioning the covariance matrixes Σ_H^{MF} and Σ_L^{MF}

¹⁴The matrix Φ in Table 1 and the upper-left (3×3) sub-matrix of Φ^{MF} in Table 3.

Table 3.5: Tests of Interactions among Yield and Macro Factors^{a,b,c}

	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 L	3300 (0.000)	131.1 (0.000)	691 (0.000)
Regime H	3293 (0.000)	131.1 (0.000)	703 (0.000)

^a L denotes low volatility regime and H is high volatility regime.

^bReported statistics are based on the Wald test that is asymptotically χ^2 distribution.

^cP-values appear in parentheses.

$$(3.11) \quad \Sigma_{\xi_t}^{MF} = \begin{bmatrix} \Sigma_1^{MF} & \Sigma_2^{MF} \\ \Sigma_3^{MF} & \Sigma_4^{MF} \end{bmatrix}; \quad \xi_t = L \text{ or } H$$

where Σ_3^{MF} is the transpose of the Σ_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 3.5 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.

3.3.4. Variance decompositions

Variance decompositions can tell us the relative contributions of the macroeconomic and yield factors to forecast errors. The detail of variance decomposition computation is in Appendix C. Table 3.6 summarizes the variance decomposition results for the high volatility and low volatility regimes. The variance decompositions also tell us something about the interactions between the yield curve and economic activity. In the turbulent regime, the proportion of the forecast variance of yields attributable to macro factors is stable along the maturity spectrum of yields, but slightly increases along the forecast

horizon. Around 40% of the forecast variance is attributable to macro factors at a 1-month forecast horizon. At a 60-month forecast horizon, around 50% of the forecast variance is due to macro factors. In contrast, the forecast variance of macro factors are not attributable to yield factors in volatile ages, though the statistical tests and the impulse responses support interactions between yield and macro factors.

In the tranquil regime, the proportion of the forecast variance of yields attributable to macro factors increases along the forecast horizon. Furthermore, the forecast of the short-end of the yield curve is less subject to the shocks of macro factors. At a 60-month forecast horizon, 20% of the forecast variance is attributable to macro factor for 3-month yields. In contrast, 45% of the forecast variance is attributable to macro factors for 5-year yield.

In the tranquil regime, however, the forecast of macro factors is subject to the shocks of yield factors. Actually, the level factor dominates the variance decomposition of forecast error at a long horizon. Only 7% of the capacity utilization forecast variance is explained by the level factor at a 1-month forecast horizon, but 55% of the CU forecast variance is attributable to the level at a 60-month forecast horizon. Likewise, Less than 1% of the inflation forecast variance is attributable to the level factor at a 1-month forecast horizon, but 57% of the inflation forecast variance can be explained by the level factor. For the Federal Fund Rate, the proportion of the forecast variance attributable to the level factor increases from 12% to 64% at a 1-month or 60-month forecast horizon. Combined with the impulse responses and statistical tests, we conclude that there is a strong feedback from yield factors to macro factors. In conclusion, We find that both regime switches and bidirectional interactions between macroeconomic and yield factors are important components of the term structure model.

3.3.5. Impulse responses

It is intuitive and interesting to examine the factor impulse responses. Appendix B derives the computation of the impulse responses. Figure 3.6 and 3.7 report the impulse responses of the yield and macro factors on each other in regime L and H¹⁵. The Cholesky decomposition of the non-diagonal covariance matrixes Σ_L^{MF} and Σ_H^{MF} is

¹⁵The bootstrapping-based error bands are available upon request.

Table 3.6: Variance Decompositions for the Macro-Finance Model^{a,b}

Horizon		Level	Slope	Curvature	CU	Inflation	FFR
3-month yield							
1	L	0.3325	0.3286	0.3245	0.0080	0.0047	0.0018
	H	0.2070	0.2018	0.1975	0.1405	0.1324	0.1208
12	L	0.3408	0.3290	0.3197	0.0037	0.0035	0.0032
	H	0.1747	0.1661	0.1595	0.1745	0.1676	0.1578
60	L	0.2775	0.2684	0.2613	0.0668	0.0645	0.0616
	H	0.1629	0.1561	0.1509	0.1835	0.1775	0.1691
12-month yield							
1	L	0.1566	0.1674	0.1781	0.1848	0.1689	0.1443
	H	0.1557	0.1669	0.1781	0.1854	0.1695	0.1445
12	L	0.1945	0.1896	0.1821	0.1674	0.1423	0.1241
	H	0.1948	0.1901	0.1826	0.1677	0.1418	0.1230
60	L	0.1920	0.1866	0.1791	0.1656	0.1452	0.1315
	H	0.1933	0.1879	0.1802	0.1660	0.1439	0.1288
60-month yield							
1	L	0.2567	0.2285	0.1944	0.1472	0.0979	0.0753
	H	0.4091	0.2257	0.0733	0.0004	0.1005	0.1904
12	L	0.1880	0.1838	0.1777	0.1664	0.1483	0.1357
	H	0.1903	0.1856	0.1788	0.1662	0.1464	0.1327
60	L	0.1900	0.1844	0.1769	0.1645	0.1474	0.1368
	H	0.1903	0.1846	0.1771	0.1645	0.1472	0.1363
CU							
1	L	0.0517	0.0150	0.0008	0.0940	0	0.0284
	H	0	0	0	0.9999	0	0
12	L	0.4160	0.0716	0.0029	0.2905	0.0269	0.1921
	H	0	0	0	0.9981	0.0009	0.0010
60	L	0.6591	0.0377	0.0034	0.1023	0.0435	0.1540
	H	0	0	0	0.9943	0.0036	0.0021
Inflation							
1	L	0	0.0002	0	0.0002	0.9957	0.0039
	H	0	0	0	0	1	0
12	L	0.0384	0.0070	0.0020	0.0100	0.9371	0.0053
	H	0	0	0	0	1	0
60	L	0.5648	0.0715	0.0049	0.0231	0.2236	0.1121
	H	0	0	0	0	0.9979	0.0021

Table 6 (continued)

FFR							
1	L	0.1185	0.0762	0.0111	0	0.0003	0.7940
	H	0	0	0	0	0	1
12	L	0.5600	0.2842	0.0515	0.0007	0.0012	0.1023
	H	0	0	0	0	0	1
60	L	0.6365	0.1759	0.0268	0.0024	0.0023	0.1562
	H	0	0	0	0	0	1

a. Each entry is the proportion of the forecast variance (at specified forecast horizon) for level, slope, curvature, FFR, inflation and CU that is explained by the particular factors.

b. L in column two means a low volatility regime, H means a high volatility regime.

Figure 3.6. Impulse responses of yield and macro factors on each other under regime L.

based on the ordering $(L_t, S_t, C_t, CU_t, \pi_t, FFR_t)$. Each response is measured in terms of one percentage point shock to residuals. We 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, we 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.

Figure 3.7. Impulse responses of yield and macro factors on each other under regime H.

Figure 3.6 and Figure 3.7 plot the impulse responses of yield and macro factors on each other in regime L and H . From the top left to bottom right in Figure 3.6 and Figure 3.7, the sub-figures respectively plot the impulse response of the level, slope, curvature, CU, inflation and FFR to shock on all factors. 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 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 our 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.

3.3.6. 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 yields¹⁶. 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 influential framework is a bivariate model (Campbell and Shiller (1987)) based on the present value model that links the τ -period yield $y_{t(\tau)}$ with the expected one-period yield $y_{t(1)}$,

$$(3.12) \quad y_{t(\tau)} = (1 - \delta) \sum_{i=0}^{\tau-1} \delta^i E_t y_{t+i(1)} + c_\tau$$

where c_τ is a maturity-dependent constant, δ 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

¹⁶For survey, refer to Piazzesi 2004.

Figure 3.8. The level from the macro-finance model and the theoretical level implied by the expectations hypothesis.

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, we do not use 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

$$(3.13) \quad L_t = (1 - \delta) \sum_{i=0}^{\tau-1} \delta^i E_t(L_{t+i} + S_{t+i}) + c_\tau$$

Suppose that the data generating process is given by the state-space system Eqs.(3.9), the implication of the Eq.(3.13) is

$$(3.14) \quad L_t = (1 - \delta) h' (I - \delta \Phi^{MF})^{-1} X_t^{MF}$$

where $h' = [1, 1, 0, \dots, 0]$ is a (6×1) selecting vector such that $L_t + S_t = h' X_t^{MF}$. Eq.(3.14) is the theoretical level factor implied by the expectation hypothesis. This section concentrates on the comparison of the actual¹⁷ and theoretical level factors. This tells us how well the expectations hypothesis approximates the observed yield

¹⁷The actual level is the level factor from the macro-finance estimation.

curve, or in terminology of Campbell and Shiller (1987), the economic significance¹⁸. Figure 3.8 plots two series of the theoretical and actual levels¹⁹. The theoretical level frequently predicts the directions of the actual level factor with a correlation of 0.92, but it is less volatile than the actual level with a volatility ratio of 0.45.

3.4. Conclusions

We 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, the existence of 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 yield factors and macro factors. The formal tests provide strong evidence in favor of interactions among yield and macro factors. This conclusion is robust across both high volatility and low volatility regimes.

In future research, we plan to extend our macro-finance model to allow for more than two regimes. The Fed changed to target the federal funds rates in 1988. It seems that a two-regime model may not be enough to capture business cycles and monetary policy shifts. This necessitates a three-regime model (Garcia and Perron (1996)) to describe the term structure of interest rates. Thus, it constitutes an interesting future research agenda.

Appendix A. The MCMC Algorithm

(1) *Generation of coefficient matrix Φ* ; Assume the prior distribution of $\text{vec}(\Phi)$ 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

¹⁸I 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, Facero and Kaminska (2006) conducted the recursive and rolling-window tests of the expectation hypothesis based on a simulation method. Zhu (2009) took into account the regime shift and provided some supporting evidence on the expectation hypothesis.

¹⁹The discount rate is set to equal to $1/(1 + L_t/12)$. We plot $TL_t + 1.8$ in the figure (6), TL_t is the theoretical level.

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 ι_k is a column vector of 1s, ξ 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

$$W = V_0 vec(Y_T) \odot Z_0 + V_1 vec(Y_T) \odot Z_1$$

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

(2) *Generation of regimes ξ* ; We 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 ξ 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 ξ_T is generated with a uniform distribution generator. The second step is to generate ξ_t conditional on ξ_{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 multimove Gibbs sampling method and provide the partition of joint distribution. There are also two steps like in the generation of regimes ξ_t , but we run the Kalman filter instead of the Hamilton filter. Given the measurement equation (3.2) and the state equation (3.3), 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

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

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

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

as shown in Kim and Nelson (1999, pp. 193). In this case, Σ_t is a weighted average of Σ_0 and Σ_1 . Specifically

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

(4) *Generation of the rate of factor loading changes λ* ; We use a Random Walk Metropolis step to draw λ .

$$\lambda_m = \lambda_{m+1} + \varsigma v$$

where $v \sim N(0, 1)$, ς is the scaling factor used to adjust the acceptance rate. The acceptance probability α for λ is given by

$$\begin{aligned} \alpha &= \min\left\{\frac{g(\lambda_{m+1}|Y_T, \Psi_{-\lambda})q(\lambda_m|\lambda_{m+1})}{g(\lambda_m|Y_T, \Psi_{-\lambda})q(\lambda_{m+1}|\lambda_m)}, 1\right\} \\ &= \min\left\{\frac{g(\lambda_{m+1}|Y_T, \Psi_{-\lambda})}{g(\lambda_m|Y_T, \Psi_{-\lambda})}, 1\right\} \end{aligned}$$

where $q()$ is a symmetric proposal distribution in the Metropolis step. Furthermore, the posterior $g(\lambda_m|Y_T, \Psi_{-\lambda_{m+1}})$ is given by

$$g(\lambda_m|Y_T, \Psi_{-\lambda}) \propto g(Y_T|\Psi)g(\lambda_m)$$

Thus, in case of the draw of λ , the acceptance rate is the posterior ratio of the new and old draws of λ .

(5) *Generation of diagonal covariance matrix Ω* ; Since R is diagonal, it can be generated element-by-element. Assume σ_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 σ_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 Φ_i are the appropriate columns of Y_T , X_t and Φ .

(6) *Generation of non-diagonal covariance matrix Σ_0 and Σ_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

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

where i denotes two regimes.

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

$$p \sim \text{beta}(u_{11}, u_{10})$$

$$q \sim \text{beta}(u_{00}, u_{01})$$

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

$$p \sim \text{beta}(u_{11} + n_{11}, u_{10} + n_{10})$$

$$q \sim \text{beta}(u_{00} + n_{00}, u_{01} + n_{01})$$

with n_{ij} referring to the transitions from state i to j , which can be calculated by counting the generated regimes ξ_T .

Appendix B. Impulse Responses

To derive the regime-dependent impulse responses (IRs) of the yield and macro-economic factors on each other, consider the state equation in Eqs.(3.9), simplified here,

$$(3.15) \quad X_t = \Phi X_{t-1} + \eta_{\xi_t}; \quad \xi_t = L \text{ or } H$$

where X_t is a 6×1 vector of endogenous variables (factors). The innovations η_{ξ_t} follow

$$(3.16) \quad \begin{aligned} \eta_L &\sim N(0, \Sigma_L) \\ \eta_H &\sim N(0, \Sigma_H) \end{aligned}$$

In order to find IRs, we can use the Choleski decomposition to transform the innovations so that the resulting components are uncorrelated. From now on, we show the computation of the IRs for the high volatility regime, for IRs in the low volatility regime, we just need replace Σ_H with Σ_L in the following derivation. Specifically, there exist a lower-triangular matrix P_H that satisfies $\Sigma_H = P_H P_H'$, Let $v_t = P_H^{-1} \eta_{H,t}$, then

the variance-covariance matrix of v_t is an identity matrix. The MA(∞) representation for Eq.(3.15) is

$$(3.17) \quad X_t = \Psi_0 + \Psi_1 v_{t-1} + \Psi_2 v_{t-2} + \dots = \sum_{i=1}^{\infty} \Psi_i v_{t-i}$$

where $\Psi_i = \Phi^i P_H$. Then, the response of the i th element of X_t to a one unit shock on j -th element v_{t-k} is

$$(3.18) \quad \partial X_t^i / \partial v_{t-k}^j = (\Psi_k)_{ij}$$

where $(\Psi_k)_{ij}$ is the (i, j) element of Ψ_k . Standard bootstrapping techniques can be employed to estimate the standard errors of the IRs.

Appendix C. Variance Decomposition

Following the notation²⁰ in Appendix B, We can compute the regime-dependent variance decompositions. As in Appendix B, the derivation here are for the high volatility regime, by simply replacing $v_t = P_H^{-1} \eta_{H,t}$ with $v_t = P_L^{-1} \eta_{L,t}$, it is easy to calculate the variance decompositions for the low volatility regime. The error of the optimal h -step ahead forecast at time t , $\hat{X}_{t+h|t}$ is

$$\hat{X}_{t+h|t} - X_{t+h} = \sum_{i=0}^{h-1} \Psi_i v_{t+h-i}$$

the i th element of forecast error is,

$$\hat{X}_{t+h|t}^i - X_{t+h}^i = \sum_{j=1}^6 \left[(\Psi_0)_{ij} v_{t+h}^j + \dots + (\Psi_{h-1})_{ij} v_{t+1}^j \right]$$

So the mean squared error of $\hat{X}_{t+h|t}$ is

$$MSE(\hat{X}_{t+h|t}^i) = \sum_{j=1}^6 \left[(\Psi_0)_{ij}^2 + \dots + (\Psi_{h-1})_{ij}^2 \right]$$

Therefore, the contribution of the j th variable to the mean squared error of the h -period-ahead forecast is,

²⁰For computing the variance decompositions of yields, we have the following MA(∞) representation for the observation equation: $\mathbf{y}_t = \sum_{i=1}^{\infty} \Lambda \Psi_i v_{t-i} + \varepsilon_t$

$$VD_{ij,h} = \frac{\sum_{i=0}^{h-1} (\Psi_i)_{ij}^2}{MSE(\hat{X}_{t+h|t}^i)}$$

CHAPTER 4

Global Yield Curves and Sovereign Bond Market Integration

4.1. Introduction

Is the global government bond market integration time-varying? What are the impediments to government bond market integration after financial deregulation? Understanding the dynamic evolution of bond market integration is important for investors and policy makers for a variety of reasons, such as diversifying the risk in the world market, reducing the cost of capital, forecasting the future macroeconomic dynamics and interest rates, pricing the interest rate derivatives, making monetary policy and fiscal policy. The endeavor of understanding the yield curve dynamics has produced a large literature and a lot of models. However little attention has been directed to the integration dynamics of world government bond market. The difficulty associated with defining and measuring the integration is one reason. Another possible reason is bond markets are expected to be closely integrated. A few studies have focused on looking for the driving forces of the comovements (Engsted and Tanggaard (2007), Ilmanen (1995), Sutton(2000)).

There are many reasons in favor of a closely integrated global bond market, for instance, big institution age, financial deregulation, financial networks , free capital flow, among many others. Nevertheless, given the impediments in the global bond markets, we can't argue a priori that the global bond market is completely integrated. Home bias might exist in the bond market because of the information asymmetry about the real activity (Barr and Priestley 2004, hereafter BP). The hedging strategy of institutions with liabilities denominated in domestic currency is usually to manage domestic bond portfolio instead of global portfolio. Tax treatment difference provides one more reason for global bond market segmentation. Local currency denominated government bonds also constitutes a reason. In addition, liquidity and exchange rate risk play a significant role in accounting for the failure of complete integration.

In the literature, the investigation of international linkages of yield curves can be broadly classified in three categories. The first line is the testing of uncovered interest rate parity (UIRP). The market efficiency hypothesis of UIRP is too restrictive and the testing just tells the failure or success of the null hypotheses, but the dichotomy doesn't show how well the model fits as an approximation and the probable time-varying dynamics of bond market integration. Furthermore the focus of the test is on the long-term interest rates parity differentials, so these studies neglect the information contained in the maturity structure of yield curves. As is well-known, the maturity spread helps forecast the real economic activity and interest rates (Ang, Piazzesi, and Wei (2006), Campbell and Shiller (1991), Diebold, Rudebusch, and Aruoba(2006), Estrella and Hardouvelis (1991), Hamilton and Kim (2002)).

The second line uses the one-factor asset pricing framework. The model is theoretically consistent and has firm micro-foundations. However, the one-factor model is difficult to interpret and the omitted factors might change the degree and trend of the market integration. Empirical econometric analysis constitutes the third line, but these models are not theoretically rigorous. Just like the first line literature, the second and third line studies don't pay much attention to the cross section of yield curves that are important for bond portfolio management and economic forecasting. Although the maturity structure is not considered, Sutton (2000) tries to relate the comovement of long-term and short-term yields by the expectation hypothesis of the term structure.

In this paper we apply the affine arbitrage-free dynamic Nelson-Siegel (1987) model (AFDNS) (Christensen, Diebold and Rudebusch(2007)), hereafter CDR) to modeling yield curves. The Nelson-Siegel model has good performance in fitting maturity structure of yield curves, and it is extensively employed by financial institutions and central banks. Diebold and Li (2006) generalized the model to a dynamic specification that is consistent with the main stylized facts of yield curves. Forecasts based on the dynamic model are satisfactory, and three factors of the model, respectively level, slope and curvature, have close interactions with macroeconomic fundamentals (Diebold, Rudebusch and Aruoba (2006), Tam and Yu (2008)). CDR developed the no-arbitrage Nelson-Siegel model, the affine specification make it also a general equilibrium model (Duffie (2001), Chapter 10, Piazzesi (2003)). The price of risk is determined by the

marginal utility process that links the risk-neutral measure and data-generating measure. The AFDNS model shows theoretical consistency and inherits the empirical fit of the Nelson-Siegel model.

There is strong evidence of cross-country bond market interactions (Barr and Priestley (2004), Diebold, Li and Yue (2007), DLY hereafter). In addition to the idiosyncratic factors, there are global yield curve factors driving the bond market in each individual country. We use latent factors for Germany, Japan, the U.K. and the U.S. to extract the global latent factors with the Kalman filter. As expected, the world factors are highly correlated with the country-specific factors by DCC-GARCH (Engle 2001) analysis. The idiosyncratic factors are given by the difference of the country-specific factors and the global factors.

Defining market integration is clearly challenging. There is no consensus. The bond market integration is defined here as movements in world factors determining movements in interest rates. This measure of integration is consistent with the uncovered interest rate parity if the expected exchange rate change is a martingale process, therefore uncovered interest rate parity is a polar case of our model. Following the idea of Bekaert and Harvey (1995), our specification, the Markov-switching Nelson-Siegel model, allows time-varying segmentation of world bond markets, it hence circumvents the polar cases of complete segmented or integrated market and fixed integration. The interactions of global bond markets in this framework are more complicated because of the asymmetry and heterogeneity of three factors. The macroeconomic interpretation of factors hints at the potential impediments in the economic fundamentals. Our conjecture is the segmentation comes from the dynamics of the real economy rather than the nominal dynamics.

The integration is very volatile with our measure. It switches frequently between perfect integration and complete segmentation. This is no surprise since our measure requires markets to fluctuate together, and the first moment is employed for measurement purpose. In contrast, in the international CAPM model the second moment (volatility) is used to interpret the difference in expected return. From the perspective of making investment decisions and policy *ex ante*, the predicted integration is more of our interest. Therefore we suggest applying the expected integration as the measure

instead of the filtered or smoothed integration of the Markov-switching Nelson-Siegel model. It also avoids the volatile situation. Our finding is interesting based on the expected integration measure. The relatively stable integration is consistent with results in BP. But our measure shows lower integration than in BP due to the asymmetry and heterogeneity of factors that are not captured by models not taking into account the maturity structure.

The article proceeds as follows. In section 2, the AFDNS model is presented and the empirical results of country local factors are reported. The global yield curve model is specified in section 3 and the global factors are extracted, then the nature of global factors and country factors are analyzed. In section 4, we present the Markov-switching Nelson-Siegel model for measuring the integration degree of the bond markets, results are interpreted, and the potential impediments are considered. The final section offers some concluding remarks and some conjectures deserving further exploration.

4.2. Country-specific yield curve factors

4.2.1. Affine arbitrage-free dynamic Nelson-Siegel model

Yield factors are obtained by estimating the affine arbitrage-free dynamic Nelson-Siegel model (AFDNS). Most term structure models use three factors to capture stylized facts of yields in cross-section and time series. By properly restricting the factor loadings in the statistical factor model, Diebold and Li (2006) proposed the dynamic Nelson-Siegel model for the τ -period yield

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

where L_t is the level factor, S_t denotes the slope factor and C_t represents the curvature factor. The parameter λ_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. Empirically, the level factor corresponds to the long-term interest rates, 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. The three factors contain information on the macroeconomic dynamics and vice versa (Diebold, Rudebusch and Aruoba(2006), Tam and Yu(2008)).

For the entire yield curve with different maturities (τ) at time t , the model can be specified as:

$$\begin{pmatrix} i_{t(\tau_1)} \\ i_{t(\tau_2)} \\ \vdots \\ i_{t(\tau_N)} \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda_t \tau_1}}{\lambda_t \tau_1} & \frac{1-e^{-\lambda_t \tau_1}}{\lambda_t \tau_1} e^{-\lambda_t \tau_1} \\ 1 & \frac{1-e^{-\lambda_t \tau_2}}{\lambda_t \tau_2} & \frac{1-e^{-\lambda_t \tau_2}}{\lambda_t \tau_2} e^{-\lambda_t \tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda_t \tau_N}}{\lambda_t \tau_N} & \frac{1-e^{-\lambda_t \tau_N}}{\lambda_t \tau_N} e^{-\lambda_t \tau_N} \end{pmatrix} \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{t(\tau_1)} \\ \varepsilon_{t(\tau_2)} \\ \vdots \\ \varepsilon_{t(\tau_N)} \end{pmatrix}$$

The dynamic Nelson-Siegel model has superior out-of-sample forecasting performance, especially at long horizons. In contrast, the affine term structure models, which is important for pricing interest rate derivatives, provide poor forecasts (Duffee 2002). Although the dynamic Nelson-Siegel has the advantage in forecasting, it is neither a general equilibrium model nor a no-arbitrage model, hence it is theoretically inconsistent.

To achieve the theoretical rigor and keep the fit of the model, CDR develops the affine arbitrage-free class of Nelson-Siegel models. The three-dimensional state variable X_t is assumed to be given by a stochastic differential equation (SDE)

$$dX_t = K^Q(t)[\theta^Q(t) - X_t]dt + \Sigma_t \begin{pmatrix} \sqrt{\gamma^1(t) + \delta^1(t)}X_t & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\gamma^3(t) + \delta^3(t)}X_t \end{pmatrix} dW_t^Q$$

where Q is risk neutral-measure, W_t^Q is standard Brownian motions in R^3 under measure Q . The relationship between the data-generating measure P and the risk-neutral measure Q is given by the following equation:

$$dW_t^Q = dW_t^P + \Gamma_t dt$$

with

$$\Gamma_t = \begin{pmatrix} \gamma_1^0 \\ \gamma_2^0 \\ \gamma_3^0 \end{pmatrix} + \begin{pmatrix} \gamma_{11}^1 & \gamma_{12}^1 & \gamma_{13}^1 \\ \gamma_{21}^1 & \gamma_{22}^1 & \gamma_{23}^1 \\ \gamma_{31}^1 & \gamma_{32}^1 & \gamma_{33}^1 \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}$$

Then the SDE for the state variables X_t under data-generating measure P is given as follows:

$$dX_t = K^P(t)[\theta^P(t) - X_t]dt + \Sigma_t dW_t^P$$

This specification preserves the affine dynamics under the data-generating measure. The connection between the two measures determines the price of the risk. The above equation corresponds to the $A_0(3)$ model in Dai and Singleton (2000). This type of model allows the volatility of state variables to be independent of the state variables. With two ingredients of the affine term structure models at hand, the third ingredient, the risk-free short interest rate is given by

$$i_{t(1)} = \delta_0^X + (\delta_1^X)' X_t$$

CDR follows the Duffie and Kan (1996) framework where the zero-coupon bond prices are exponential-affine functions of the state variables. With appropriate assumptions in SDE of X_t specified in CDR, the zero-coupon bond yields are given by

$$i_{t(\tau)} = X_t^1 + \frac{1 - e^{-\lambda\tau}}{\lambda\tau} X_t^2 + \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) X_t^3 - \frac{C(\tau)}{\tau}$$

where (X_t^1, X_t^2, X_t^3) is the state variables vector corresponding the (L_t, S_t, C_t) in the dynamic Nelson-Siegel model. This is the closest affine no-arbitrage approximation to the dynamic Nelson-Siegel model as discussed in CDR. The difference between these types models is the yield-adjustment term $-\frac{C(\tau)}{\tau}$ (see CDR). The AFDNS model rules out the arbitrage opportunities in the financial market, that is, arbitrage opportunities are immediately traded away in modern well-organized markets. Furthermore, the model inherits the good empirical fit of the dynamic Nelson-Siegel model. Therefore, the AFDNS is theoretically rigorous and empirically appealing.

4.2.2. Data and summary statistics

The U.S. data consist of end-of-month observations of 1, 3, 6, 12, 24, 36, 60, 84, 120 months zero-coupon yields on treasury securities covering the period from January 1985 to March 2008. The data source is econstatsTM. The U.K. zero-coupon yields with maturities of 6, 9, 10, 11, 12, 24, 36, 48, 60, 72, 84, 96, 108, 120 months are retrieved from econstatsTM. It covers the same period as the U.S sample and all data are end-of-month observations. For the Germany zero-coupon government bond yields, the month-end observations with maturities 12, 24, 36, 48, 60, 72, 84, 96, 108, 120 months are retrieved from Deutsche Bundesbank, the central bank of Germany.

The Japanese dataset has two sources. The first sample covering the period from January 1985 to December 1991 is from the Key Economic Statistics Files of the PACAP Database-JapanTM compiled by the Sandra Ann Morsilli Pacific-Basin Capital Markets Research Center at the University of Rhode Island. The end-of-month yields consist of government bond interest rates with maturities of 12, 24, 36, 60, 84, 120 months. The second sample covers the period from January 1992 to March 2008. The dataset is downloaded from Bloomberg. The maturities are 6, 12, 24, 36, 48, 60, 72, 84, 96, 108, 120 months.

The summary statistics including skewness and kurtosis of yields for each maturity and for each country is presented in Table 4.1¹. The 3D plot of term structure of interest rates for each country is graphed in Figure 4.1. One stylized fact of interest rates is they tend to exhibit considerable persistence and are believed to be nonstationary or better approximated by the integrated process. This feature has profound implications for estimation and statistical inference.

The autocorrelation coefficients and augmented Dickey-Fuller tests in Table 3.1 provide evidence of persistence and non-stationarity². However, the yields are usually cointegrated, as implied by the rational expectation hypothesis. The Johansen cointegration analysis presents evidence of common trends in yields³. The cointegration may explain another important stylized fact of the yield curve: spreads are less persistent

¹The statistics for Japan are based on the sample retrieved from Bloomberg. This applies to results for the DNS estimates in Table 2 and the AFDNS estimates in Table 4.

²The ADF test rejects the unit root in the 6-month yield of Japan. Because of the low power of the ADF test, we also conduct the PP test, KPSS test, Ng-Perron tests, the results are mixing.

³The analysis results are available upon request.

than yields. The skewness and kurtosis show yields don't deviate considerably from the normal distribution. The standard deviations in Table 4.1 tells us that short-term yields usually are more volatile than long-term yields with the exception of Japan. In Figure 4.9, we plot average yield curves, for Germany, Japan and U.S., the average yield curves are upward-sloping for the time period under analysis, in constrast, the U.K. average yield curve has S-shape.

Table 4.1 About Here

Figure 4.1 About Here

4.2.3. Country local factors

Non-linear least squares can be employed to estimate the dynamic Nelson-Siegel model in the first equation of section 1.4.2 in chapter 1. In Diebold and Li (2006), they fix the λ_t and set it equal to the value that maximizes the loading on the curvature factor at 30 months. In so doing, one can estimate the dynamic Nelson-Siegel model by ordinary least squares and make the numerical optimization more reliable. We follow this approach to estimate the DNS model and results are presented in Table 4.2 with λ_t fixed at value of 0.0600. The DNS model is capable of replicating a variety of yield curve shapes. Both the three-factors model and two-factors model (without curvature factor) fit yield curves across countries well, but three-factors model has higher explanatory power according to the average R^2 .

Table 4.2 About Here

In order to estimate the AFDNS model, we fix the λ_t at 0.0600. The Kalman filter estimation is initialized by using the unconditional covariance matrix of the state vector and mean vector from the DNS estimates. The starting transition matrix parameters are also from OLS regression of factors extracted by the DNS model. Yield adjustment term⁴ of the AFDNS model bridges the connection between the state variables dynamics under data-generating measure and the volatility matrix Σ under risk-neutral measure. This makes the AFDNS model theoretically consistent and hence differs from

⁴The last item on the right side of equation (49), chapter 2.

the DNS model. The conditional mean-reversion matrix ($\exp(-K^P \Delta t_i)$) determines the mean-reverting rate of state variables. We present the mean-reversion matrix and yield adjustment terms in Table 4.3. In CDR, the yield adjustment terms are trivial for all maturities. However, in our estimation, the yield adjustment terms is significant for long-term interest rates. This may come from the first item on the right hand side of the last equation in section 2.4.2.

Table 4.3 About Here

The extracted level, slope and curvature factors across countries are plotted in figure 4.2-4.4, respectively. For the purpose of comparison, factors from the DNS model and AFDNS model are depicted together in figure 4.5 and figure 4.6 for each country. The dynamics of factors from the AFDNS model mimic the dynamics of factors from the DNS model, the small difference may be induced by yield adjustment terms. The level factor dynamic is homogeneous across countries. In contrast, the curvature evolution over time is heterogeneous across country. The principal component analysis in section 4.2.3 supports the conclusion.

Figure 4.2 About Here

Figure 4.3 About Here

Figure 4.4 About Here

Figure 4.5 About Here

Figure 4.6 About Here

The summary statistics of factors from AFDNS model is presented in Table 4.4. The level factor is more persistent than the slope and curvature factors. The ADF tests show curvature factors across countries may be stationary except Germany, but level factors are nonstationary except U.S. level factor. As to the volatility feature, the level factor can be more or less volatile than the slope and curvature factors. We note that the slope factors are more correlated with curvature factors than level factors. Although the Pearson and likelihood ratio tests reject the independence of level factor and slope factor, in empirical macro-finance model (Diebold, Rudebusch and Aruoba (2006), Tam and Yu (2008)), one factor contains little extra information about other factors or macro-variables. Diebold and Li (2006) provides the empirical evidence of

the interpretation of level, slope and curvature factors as long-term, short-term and medium term factors. Figure 4.7 plots the 10-year yield, 3-month yield minus 10-year yield, and two times 2-year yield minus 10-year and 3-month yields with the level, slope and curvature factors⁵.

Table 4.4 About Here

Figure 4.7 About Here

The fit of the AFDNS model is good. The error terms of the estimation is plotted in figure 4.8. To facilitate the comparison, the scale of the figure is set to be the same as in figure 4.1. Figure 4.9 plots the average fitted yield curves along with the observed yield curves. The AFDNS model replicates the upward sloping yield curves of Germany, Japan and the U.S. and the S-shape of the U.K.. It is important to note that the model fits the middle region of the yield curves better than the end regions. This might be a matter of fact of the model, as pointed out in Diebold and Li (2006): ".....because the maturities are not equally spaced, we implicitly weight the most "active" region of the yield curve most heavily when fitting the model".

Figure 4.8 About here

Figure 4.9 About Here

4.3. Global yield curve factors

4.3.1. Model specification

A number of studies have focused on the international linkages of bond markets. There seems to be a consensus that the bond yields and returns are highly correlated across countries. Hafer et. al. (1997) found that long-term yields seem to be cointegrated across countries, hence there is comovement of international bond markets. Later, Sutton (2000) tried to relate the cointegration in long-term yields with comovement of short-term yields by rational expectation hypothesis. The conclusion is that the comovement of long-term yields come out of the comovement in term premia. Ilmanen (1995) found that a small set of global instruments can forecast a significant fraction of monthly yields variation, and the author concludes that the predictability of global

⁵For other countries, the 3-month yield is not available.

bond returns come from a few global factors. The empirical study of Driessen et. al. (2003) find that world bond markets are correlated by using a linear factor model and principal component analysis, the driving force of the comovement is the level of yields in each country, this is consistent with the matter of fact that the level factor dominates the term structure of interest rates. Engsted and Tanggaard (2007) found that inflation news drive the comovement between the U.S. and Germany bond markets. Barr and Priestley (2004) applied the international CAPM model allowing time-varying market segmentation to investigate global market integration, and they found almost 70% of the variation can be explained by world dynamic beta, and the degree of market integration is stable during the period covered by the sample.

Recently, DLY focused on the entire term structure of interest rates. They used the latent factor dynamic Nelson-Siegel model to fit the yield curve. For a set of country yield curves, they fit them by allowing common global factors and country-specific factors. There are interactions between global factors and country-specific factors, and the loading of country-specific factor on global factors is allowed to vary across countries. The finding is that global factors explain a big fraction of country yield curve. In this paper I use country-specific factors from the AFDNS model to extract the global yield curve factors. The specification is different with DLY. DLY use two-factors model, while the dynamic Nelson-Siegel model estimation⁶ shows that three-factors model significantly improve the goodness-of-fit according to the R-square. Secondly, the level, slope and curvature factors seem to be independent as presented in the macro-finance model of Tam and Yu (2008). CDR shows that the forecasting performance of the independent factors model is no worse than the correlated factors model. Taking into account above points, we use a three-factors model, but assume that the global level (slope, curvature) factor only depends on the domestic level (slope, curvature) factor. This simplifies the estimation and alleviate the local maximum problem associated with the numerical optimization.

For extracting the common factor, principal component analysis is a popular method, although it is difficult to interpret. In this paper, we use Kalman filter to extract global

⁶The OLS estimation of Diebold and Li (2006) model is applied by fixing the λ equal to 0.0600.

factors, and principal component is an interesting benchmark for comparison. First, we decompose the country-specific factors:

$$(4.1) \quad \begin{aligned} l_{it} &= L_t + u_{it}^l \\ s_{it} &= S_t + u_{it}^s \\ c_{it} &= C_t + u_{it}^c \end{aligned}$$

where l_{it} , s_{it} , c_{it} are country-specific factors from the independent AFDNS estimation, L_t , S_t , C_t are global level, slope and curvature factors. The u_{it}^l , u_{it}^s , u_{it}^c are country idiosyncratic level, slope and curvature factors. The i denotes one of four countries: the U.S., the U.K., Germany and Japan. As aforementioned, the assumption of independent level, slope and curvature dynamics are reasonable, therefore, we extract three global factors independently. we assume country idiosyncratic factors follow an AR(1) process:

$$(4.2) \quad \begin{pmatrix} l_{1t} - L_t \\ l_{2t} - L_t \\ \vdots \\ l_{kt} - L_t \end{pmatrix} = \begin{pmatrix} \beta_1 & 0 & 0 & 0 \\ 0 & \beta_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \beta_k \end{pmatrix} \begin{pmatrix} l_{1t-1} - L_{t-1} \\ l_{2t-1} - L_{t-1} \\ \vdots \\ l_{kt-1} - L_{t-1} \end{pmatrix} + \begin{pmatrix} w_{1t} \\ w_{2t} \\ \vdots \\ w_{kt} \end{pmatrix}$$

k is number of countries. The specification assumes the country idiosyncratic factors are independent, with the diagonal variance-covariance matrix. This make sense economically if there are no regional factors in the hierachical model. The global factor dynamics are given by an AR(1) process:

$$(4.3) \quad L_t = \alpha + \rho L_{t-1} + \epsilon_t$$

The Eqs.(4.2) and (4.3) can be used to extract the global factors.

4.3.2. Global factors and idiosyncratic factors

The Kalman filter estimation of system equation (4.2) and (4.3) is applied to extract the global factors of yield curves. The unrestricted vector autoregression (VAR) estimation

shows that one factor has little extra information about the dynamic of other factor, plus empirical evidence from Diebold and Li (2006), Tam and Yu (2008), therefore we extract the global level, slope and curvature factors by independently iterating the Kalman filter. This simplifies the extraction of global factors significantly. We initialize the Kalman filter with the unconditional covariance matrix and a mean vector from the average of country-specific factors. The estimated parameters are reported in the upper panel of Table 4.5.

Table 4.5 About Here

The global factors essentially are one common component of country-specific factors. Two interesting questions before we scrutinize the global factors are: what is the explanatory power of one common component at most? What is the relationship of the components extracted by the Kalman filter and components from principal component analysis? To answer questions, the principal component analysis results are presented in the lower panel of Table 4.5. As we mentioned, the cross-correlation of level factors is higher than slope and curvature factors. The first principal component can explain 91% of variation of the country-specific level factors. For slope and curvature factors, only 57% and 48% of variation can be interpreted by the first principal component. There is strong interactions between global factors from the Kalman filter and the first principal component. The adjusted first principal component and global factors are plotted in the figure 4.12. The correlations of level, slope and curvature factors and the corresponding first principal components are respectively 0.99, 0.87 and 0.75. Table 4.6 gives the descriptive statistics of global factors. As in the country model, the level factor is more persistent than the slope and curvature factors. The skewness and kurtosis of curvature factor is not in favor of the normality distribution.

Figure 4.12 About Here

Table 4.6 About Here

To investigate the correlation and explanatory power of global factors on country specific factors, we run the following regression:

$$f_{it} = \alpha + \beta f_{wt} + \varepsilon_{it}, i = GM, JP, UK, US;$$

$$f = level, slope, curvature; f_{wt} : global\ factor$$

Table 4.7 presents the results. All country-specific factors have positive loadings on global factors. The global level factor has the most significant explanatory power judged by R^2 of the regression. The global slope and curvature factors have lower but still significant power of explanation. The regression implies that the level factor has the highest degree of integration.

Table 4.7 About Here

The previous analysis provides the static correlation of global factors and country-specific factors. The purpose of the paper is to investigate the time-varying bond markets integration, hence the dynamic correlation is our interest. Engle (2002) proposes the dynamic conditional correlation (DCC-GARCH) model, and it is appropriate for the purpose here. The model is extensively applied because it preserves the simplicity of univariate model in a multivariate setting. The DCC-GARCH model for factors is as follows:

$$\mathbf{f}_t | \Omega_t \sim N(0, H_t)$$

$$H_t = D_t R_t D_t$$

$\mathbf{f}_t$ is the vector of the level, slope or curvature factors (Germany, Japan, The U.K, the U.S., world). The maximum likelihood method can be used to estimate the DCC-GARCH model. Figures 4.13, 4.14 and 4.15 present the dynamic conditional correlation of level, slope and curvature factors across countries. In general, the global level factor is positively correlated with the country-specific factors. The dynamic conditional correlations of slope and curvature factors shift more frequently. Anyway, all factors are highly correlated, although the correlation may be positive or negative.

Figure 4.13 About Here

Figure 4.14 About Here

Figure 4.15 About Here

4.4. Global market interactions and integration

4.4.1. The integration model

There are interactions and linkages of the government bond markets across countries. However, market integration is a stricter restriction in the sense that it implies the comovement or interaction in the bond markets across countries, but not vice versa. Interactions are empirical phenomena, the market integration should be theoretically consistent in addition to being empirically correlated. In the international CAPM model (Barr and Priestley (2004)), government bond markets are integrated if the world beta price the excess return. In contrast, the AFDNS model is for describing the yield curve level dynamics, therefore we define bond market integration as a situation in which the movement in global yield factors determines the movement in yields with different maturities in each country's market. Otherwise, the markets are segmented if the movement of yields is determined by the movement of idiosyncratic factors.

The notion of integration is challenging and controversial. The definition here requires bond markets to fluctuate together. The accuracy of the measure relies on the performance of underlying yield curve models. Once the AFDNS is called into question, so is the dynamic measure of integration. Empirically, the AFDNS model provides the necessary accuracy.

The definition focuses on the change of interest rates instead of the levels. Because the level factor dominates the yield curve dynamics, the change of factors eliminates the level-factor dominance effect. This allows a stable interest rate differential between two markets that is consistent with the efficient market hypothesis if there is also a stable inflation wedge. With this definition, more attention is directed to the interactions of three latent factors from the AFDNS model. These interactions have important information about the market integration. For example, Sutton (2000) finds that the comovement of long-term yields can't be explained by the interactions of short-term yields. This is the evidence of heterogeneous factor dynamics. The cross-section

maturity structure of the term structure contains useful information for forecasting future interest rates and the macroeconomic dynamics (Ang, Piazzesi, and Wei (2006), Campbell and Shiller (1991), Diebold, Rudebusch, and Aruoba(2006), Estrella and Hardouvelis (1991), Hamilton and Kim (2002)). Tan and Yu (2008) offers the further evidence of heterogeneous factors dynamics using the dynamic conditional correlation analysis. Most of previous studies focus on the time series properties of comovement, with above definition, the properties of yield curves in cross-section are investigated. According to our decomposition of country-specific factors in (4.1), even the world factors are constant, they still play a role in explaining the yields in each market, while the change of factors circumvent the problem. This provides another reason.

In completely segmented markets, the change of yields in each market is determined by the country idiosyncratic factors

$$(4.4) \quad \begin{pmatrix} \Delta y_{it(\tau_1)} \\ \Delta y_{it(\tau_2)} \\ \vdots \\ \Delta y_{it(\tau_N)} \end{pmatrix} = A \begin{pmatrix} \Delta u_{it}^l \\ \Delta u_{it}^s \\ \Delta u_{it}^c \end{pmatrix} + \begin{pmatrix} w_{it(\tau_1)} \\ w_{it(\tau_2)} \\ \vdots \\ w_{it(\tau_N)} \end{pmatrix}$$

where i = Germany, Japan, UK, US, and N is the the observations in cross-section at each point of time, and

$$(4.5) \quad A = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{pmatrix}$$

In a completely integrated global market, the change of yields in each market is determined by global factors

$$(4.6) \quad \begin{pmatrix} \Delta y_{1t(\tau_1)} \\ \Delta y_{1t(\tau_2)} \\ \vdots \\ \Delta y_{kt(\tau_N)} \end{pmatrix} = A \begin{pmatrix} \Delta L_t \\ \Delta S_t \\ \Delta C_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t(\tau_1)} \\ \varepsilon_{1t(\tau_2)} \\ \vdots \\ \varepsilon_{kt(\tau_N)} \end{pmatrix}$$

This equation is consistent with uncovered interest rate parity and the AFDNS model if the expected change of exchange rate is a martingale process. Given the uncovered interest rate parity is

$$(4.7) \quad E_t(ex_{t+1} - ex_t)/ex_t = y_{1t(\tau_j)} - y_{2t(\tau_j)}$$

where ex is the exchange rate, $y_{it(\tau_j)}$ is the yield for country 1 with maturity τ_j . Because the expected change of exchange rate is zero (martingale process), the yields differential should be zero for all maturities according to the law of one price. Since the yield curve is given by the AFDNS model, the zero differential holds in time series when

$$(4.8) \quad \begin{aligned} l_{1t} &= l_{2t} \\ s_{1t} &= s_{2t} \\ c_{1t} &= c_{2t} \end{aligned}$$

therefore the uncovered interest rate parity is a polar case of our model.

In the real world, government bond markets across countries are expected to be neither perfectly integrated nor completely segmented. The degree of segmentation might be time-varying, it is even expected there is a trend of increasing degree of integration due to deregulations. The switch of regimes may be caused by common shocks in both financial markets and real economy. It could be a surprise or partially expected. The Markov-switching Nelson-Siegel model allows time-varying dynamic evolution of market integration,

$$(4.9) \quad \begin{pmatrix} \Delta y_{1t(\tau_1)} \\ \Delta y_{1t(\tau_2)} \\ \vdots \\ \Delta y_{kt(\tau_N)} \end{pmatrix} = \phi_t A \begin{pmatrix} \Delta L_t \\ \Delta S_t \\ \Delta C_t \end{pmatrix} + (1 - \phi_t) A \begin{pmatrix} \Delta u_{it}^l \\ \Delta u_{it}^s \\ \Delta u_{it}^c \end{pmatrix} + \begin{pmatrix} \epsilon_{1t(\tau_1)} \\ \epsilon_{1t(\tau_2)} \\ \vdots \\ \epsilon_{kt(\tau_N)} \end{pmatrix}$$

where ϕ_t is the probability of the market integration. $\phi_t = 1$ implies the perfect market integration, $\phi_t = 0$ means markets are completely segmented. The regime probability ϕ_t follows a Markov chain process and the EM algorithm is an efficient estimator (Hamilton 1994), the optimal inference and forecast of regime is given by iterating the following equations

$$\hat{\xi}_{t|t} = \frac{\hat{(\xi_{t|t-1} \odot \eta_t)}}{1'(\hat{\xi_{t|t-1} \odot \eta_t})}$$

and

$$\hat{\xi}_{t+1|t} = P \hat{\xi}_{t|t}$$

with P being the transition probability matrix. Moreover, the disturbance vector is given by

$$\begin{pmatrix} \epsilon_{1t(\tau_1)} \\ \epsilon_{1t(\tau_2)} \\ \vdots \\ \epsilon_{kt(\tau_N)} \end{pmatrix} = \phi_t \begin{pmatrix} \varepsilon_{1t(\tau_1)} \\ \varepsilon_{1t(\tau_2)} \\ \vdots \\ \varepsilon_{kt(\tau_N)} \end{pmatrix} + (1 - \phi_t) \begin{pmatrix} w_{it(\tau_1)} \\ w_{it(\tau_2)} \\ \vdots \\ w_{it(\tau_N)} \end{pmatrix}$$

This allows regime-dependent heteroskedasticity.

4.4.2. Integration measure and interpretations

The starting point of time of the sample used for integration analysis is January 1985. It is then that financial deregulation has become a global phenomenon. Based on prior knowledge, one might expect that the global bond markets are close to full integration. However, in the last two decades world financial market has gone through a turbulent age, so bond market integration may be subject to the turbulences. It is reasonable to postulate ex ante that market integration is time-varying. We are also suspicious about

the full integration due to the reasons enumerated in the introduction, such as, home bias, tax treatment difference, exchange rate risk, liquidity risk, among many others.

The Markov-switching Nelson-Siegel model allows us to measure the dynamic evolution of the government bond market integration. The estimated transition probabilities and the log likelihood are presented in Table 4.8. Here state 1 has natural interpretation of market integration and state 2 represents market segmentation. In this model, the transition probabilities are fixed to be constants. As the Markov-switching model is highly nonlinear, it may be subject to local maximum and corner solution. In estimation, we use the parameter vector from the regression of country-specific yields on global yields as the initial parameters for state 1. For state 2, the parameters vector from regression of country-specific yields on the idiosyncratic yields is used as the starting parameters. The global yields are defined as the first item on the right hand side of Equation (4.9), they are plotted in figure 4.16, and the idiosyncratic yields are the second item on the right hand side of the same equation.

Table 4.8 About Here

Figure 4.16 About Here

4.4.2.1. Germany. Two major changes may affect the government bond market in Germany for the period covered by the sample. One is the monetary union marked by the Deutsche mark becoming legal tender in East Germany, this rise the government funding needs. The other is the introduction of the Euro. There are also some structural changes in the Germany bond market, for example, issuing technique changes from the underwriting procedure to combined with auctions in July 1997, introduction of Bund futures and options on Bund futures in late 1980s. The estimated result for Germany is plotted in upper panel of Figure 4.17. The market integration is quite volatile by the Markov-switching Nelson-Siegel model. This is not surprising because we use the first moment to measure the integration, as long as world yields are rising (falling) while Germany yields are falling (rising), market is segmented. In contrast, the second moment (volatility) is applied in the capital asset pricing model. Instead of filtered probabilities, we may look at the predicted probabilities ($\xi_{t+1|t}$), that is the expected

degree of market integration ex ante, the expected integration is stable and between the interval of 0.3 and 0.45.

Figure 4.17 About Here

4.4.2.2. Japan. The deregulation beginning in 1970 has reconstructed the Japanese financial markets, up to 1985, restrictions, for instance, interest rate ceiling, capital moving to and from overseas, have been removed. The financial markets are freed of strict regulations. After that, in 1997 the Bank of Japan Law was revised and the Bank of Japan acquired a more independent legal statue. The ongoing financial reform, "Japanese Big Bang", has far-reaching consequences in financial markets. These events may change the state of integration of the Japanese government bond market with the world bond market. The middle panel of figure 4.17 presents the results for Japan. The market integration is also volatile. If we look at the expected degree of integration, it is relatively stable and in the interval 0.25 — 0.50.

4.4.2.3. U.K.. Our prior expectation is the UK would have high degree of integration because financial market in the UK are free of regulations. The results for the UK is in the lower panel of figure 4.17. The volatile integration is all the same as Germany and Japan. However, the dynamic degree of integration is in the interval 0.225 — 0.3, it is the lowest among four countries investigated.

4.4.2.4. U.S.. The US is the single biggest and most important market in the world. The dynamic expected degree of integration is in the interval 0.2 — 0.5, it is more volatile than other markets. Measured by the mean, it has the highest degree of integration. The filtered probabilities are still volatile due to the aforementioned reason. The upper panel of figure 4.18 plots the filtered and predicted probabilities of being in the integration state.

Figure 4.18 About Here

4.4.2.5. World. The integration of world bond market as a whole is main interest. The Markov-switching model allow us to measure the world market integration dynamically. We choose one long-term yield and one short-term yield for each country (Germany: 2-year, 9 year, Japan: 1-year, 10-year, the UK: 6-month, 8-year, the US: 3-month, 7-year), it consists of the dependent variables. The world factors and corresponding idiosyncratic factors are independent factors. We can't include all yields,

otherwise, the coefficient matrix for the global factors is singular. The results are in the lower panel of figure 4.18. The expected dynamic integration is in the interval 0.1 — 0.4. This is not suprising given the degree of integration for each country. The world integration is a stricter restriction because it requires the global factors to explain yields across countries at the same time.

Figure 4.18 About Here

The stable dynamic predicted probabilities of integration implies that the market expectation of integration is stable. This is consistent with the finding in BP where the authors reject the time-varying bond market integration. However, our results imply a lower degree of integration. It is not suprising because we take into account the maturity structure of the yield curve in cross-section. Combined with the results of the principal component analysis, the idiosyncratic regression and DCC-GARCH analysis, the market segmentation stems mainly from low degree of integration on slope and curvature factors. Another reason is that the heterogeneous dynamics of the factors, the integration of one factor may accompany the segmentation of some other factor. Because the level factor represents the long-term factor, the long end of the term structure is more likely to be integrated than short end of the term structure that are represented by the curvature factors. This finding is consistent with Sutton (2000). This may suggest that the short end of the bond market is the impediment to market integration.

Empirically, the level factor is highly correlated with inflation. The slope factor is associated with real activity (Diebold et. al (2006)), and curvature is associated with uncertainty (Zhu 2008). In previous macro-finance model (Ang and Piazzesi 2003), the macro factors are found to affect the short end of the yield curve, but leave the long end not accounted for. Therefore, the segmentation is more likely coming from the dynamics of real economy than from the nominal dynamics.

4.5. Conclusions

It is difficult to measure the integration of the government bond market into the world bond market. Some previous studies investigate the long end integration of yield curves, while others investigate the short end integration of yield curves. In this

paper, we propose a measure that take into account the maturity structure of bond yields. This measure also allows for time-varying conditional market integration. We use a theoretically consistent latent factors model, the AFDNS model, to describe the dynamics of the yield curves in both time series and cross-section. Then the global factors are extracted from the country factors. Finally, the market integration is measured by the Markov-switching Nelson-Siegel model.

Some results are consistent with previous studies, but we shed new light on bond market integration. The new finding is that market segmentation are from two aspects, one is the integration asymmetry of bond markets. The level factor is more integrated than the global slope and curvature factors. The other reason is the heterogeneous dynamics of latent factors. The integration of factors are not simultaneous. This tells us that market integration is from the short- and medium end of yield curves.

A number of extensions deserve further exploration. Why is the short end of the market more segmented than the long end of the market? It is widely believed that the central banks can take control of the short-term interest rates. However, can monetary policy shocks explain partially the short end segmentation of the market? This question is associated with the monetary policy transmission mechanism. The rational expectations hypothesis plays a pivotal role in the term structure of interest rates. Sutton (2000) hence tried to relate the market integration with the rational expectation hypothesis. In this framework of integration, what role do the rational expectations play?

From a financial economics perspective, does this imply gain of portfolio diversification? If not, what is the risk of portofolio diversification. There are a number of risks, such as exchange rate risk, liquidity risk, but which one is the dominant one? Associated with the risk, one also has to explain the price of risk. These are some possible future extensions.

Table 4.1: Summary Statistics for Bond Yields

Germany								
maturity (months)	Mean	Std. Dev.	skew- ness	kurto- sis	$\hat{\rho}_{(1)}$	$\hat{\rho}_{(12)}$	$\hat{\rho}_{(30)}$	ADF
12	4.6981	1.9934	0.8865	2.8322	0.9925	0.8239	0.4654	-1.482
24	4.8881	1.9001	0.8248	2.7951	0.9911	0.8091	0.5000	-1.623
60	5.4378	1.6805	0.4953	2.4366	0.9908	0.8245	0.6302	-1.476
120	5.9188	1.4821	0.1114	2.0354	0.9903	0.8484	0.7597	-1.126
Japan								
maturity (months)	mean	Std. Dev.	skew- ness	kurto- sis	$\hat{\rho}_{(1)}$	$\hat{\rho}_{(12)}$	$\hat{\rho}_{(30)}$	ADF
6	0.6474	0.9599	1.7709	4.9275	0.9636	0.5883	0.2093	-3.183
12	0.7299	0.9855	1.6646	4.5191	0.9639	0.6016	0.2462	-1.366
60	1.4694	1.1333	1.2654	3.4682	0.9648	0.6535	0.3677	-1.757
120	2.2334	1.1761	1.0684	2.9223	0.9692	0.7132	0.4459	-1.405
U.K.								
maturity (months)	mean	Std. Dev.	skew- ness	kurto- sis	$\hat{\rho}_{(1)}$	$\hat{\rho}_{(12)}$	$\hat{\rho}_{(30)}$	ADF
6	7.1674	2.9581	0.8249	2.4826	0.9635	0.5883	0.2093	-1.972
12	7.1004	2.7671	0.7402	2.3789	0.9878	0.8186	0.6054	-1.438
60	7.1984	2.3906	0.3329	1.7688	0.9878	0.8602	0.7724	-1.435
120	7.2003	2.3584	0.2000	1.5532	0.9893	0.8968	0.8223	-1.450
U.S.								
maturity (months)	mean	Std. Dev.	skew- ness	kurto- sis	$\hat{\rho}_{(1)}$	$\hat{\rho}_{(12)}$	$\hat{\rho}_{(30)}$	ADF
6	4.9810	2.0397	-0.1685	2.4790	0.9929	0.6960	0.2327	-1.711
12	5.1507	2.0483	-0.1063	2.4441	0.9911	0.7047	0.2952	-1.712
60	6.0391	1.9105	0.3038	2.5432	0.9811	0.7164	0.6114	-2.416
120	6.4194	1.8052	0.5292	2.6213	0.9907	0.8443	0.7225	-2.692

(1) The summary statistics for Japan is based on Bloomberg sample.

Notes: (2) $\hat{\rho}_{(\tau)}$ is the τ -th autocorrelation coefficient.

(3) The lag length of ADF test is selected by SIC.

Table 4.2: OLS Estimates of Dynamic Nelson-Siegel Model

Germany							
model	R^2	level	t-value	Std.Dev.	slope	t-value	Std.Dev
three factors	0.9823	6.4474	2396	1.4431	-1.7399	-433.7	1.6689
		curvature	t-value	Std.Dev.			
		-2.2329	-143.8	2.3655			
two factors	0.8426	6.1004	867.8	1.5138	-2.1903	-75.52	1.9845
Japan							
model	R^2	level	t-value	Std.Dev.	slope	t-value	Std.Dev
three factors	0.9804	2.9365	170.3	1.8958	-2.0608	-117.7	1.1887
		curvature	t-value	Std.Dev.			
		-3.5256	-45.8	1.8454			
two factors	0.8287	2.1938	80.63	1.2506	-2.1318	-34.63	0.7634
U.K							
model	R^2	level	t-value	Std.Dev.	slope	t-value	Std.Dev
three factors	0.9425	7.3024	1231	2.4339	-0.1741	-41.63	2.0108
		curvature	t-value	Std.Dev.			
		-0.3378	-4.226	1.9574			
two factors	0.7958	6.1004	1094	2.3130	-2.1903	-6.27	2.0090
U.S.							
model	R^2	level	t-value	Std.Dev.	slope	t-value	Std.Dev
three factors	0.9183	6.7331	488.5	1.7350	-2.1327	-193.8	1.6050
		curvature	t-value	Std.Dev.			
		-0.3017	-22.24	2.0716			
two factors	0.8387	6.6523	507.1	1.9768	-2.0859	-83.46	1.4816

- Notes:
- (1) Statistics for Japan are based on the sample retrieved from Bloomberg.
 - (2) The two-factors model doesn't have curvature factor.
 - (3) The reported statistics are mean of cross-section OLS estimation.
 - (4) For regression details, refer to Diebold and Li (2006); Here $\lambda = 0.0600$.

Table 4.3: Estimates of the Diagonal AFDNS Model Parameters

Germany						
$\exp(-K^P \frac{1}{12}) = \begin{pmatrix} 0.9905 & 0 & 0 \\ 0 & 0.9681 & 0 \\ 0 & 0 & 0.9579 \end{pmatrix}$						
Yield-term	-0.0250	-0.0795	-0.1476	-0.2183	-0.2857	-0.3480
	-0.4054	-0.4590	-0.5101	-0.5600		
Japan						
$\exp(-K^P \frac{1}{12}) = \begin{pmatrix} 0.9796 & 0 & 0 \\ 0 & 0.9645 & 0 \\ 0 & 0 & 0.9164 \end{pmatrix}$						
Yield-term	-0.0020	-0.0076	-0.0293	-0.0619	-0.09976	-0.1390
	-0.1781	-0.2167	-0.2553	-0.3341		
U.K.						
$\exp(-K^P \frac{1}{12}) = \begin{pmatrix} 0.9887 & 0 & 0 \\ 0 & 0.9705 & 0 \\ 0 & 0 & 0.9213 \end{pmatrix}$						
Yield-term	-0.0086	-0.0180	-0.0218	-0.0259	-0.0303	-0.1032
	-0.2009	-0.3070	-0.4113	-0.5100	-0.6028	-0.6912
	-0.7769	-0.8617				
U.S.						
$\exp(-K^P \frac{1}{12}) = \begin{pmatrix} 0.9820 & 0 & 0 \\ 0 & 0.9828 & 0 \\ 0 & 0 & 0.9481 \end{pmatrix}$						
Yield-term	-0.0001	-0.0008	-0.0032	-0.0119	-0.0428	-0.0868
	-0.1885	-0.2908	-0.4467			

- Notes:
- (1) Reported parameters for Japan based on Bloomberg sample.
 - (2) Yield-term: yield-adjustment terms in equation (??) associated with corresponding maturities (CDR for details),
 - (3) $\exp(-K^P \frac{1}{12})$: one-month conditional mean-reversion matrix

Table 4.5: Extraction of Global Yield Curve Factors

Kalman Filter*						
	$\hat{\alpha}$	$\hat{\rho}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
level	0.0190 (0.036)	0.9953 (0.005)	0.9408 (0.024)	0.9803 (0.008)	0.9606 (0.021)	0.9998 (0.003)
slope	-0.045 (0.063)	0.9735 (0.032)	0.9874 (0.010)	0.9694 (0.014)	0.9677 (0.016)	0.9643 (0.019)
curv	-0.303 (0.183)	0.9276 (0.046)	0.9822 (0.014)	0.9557 (0.024)	0.9510 (0.019)	0.9314 (0.023)
Principal Component Analysis						
level	eigenvalue		3.6536	0.2150	0.1025	0.0288
	variance Prop.		0.9134	0.0537	0.0256	0.0072
	cumulative Prop.		0.9134	0.9172	0.9928	1.0000
slope	eigenvalue		5.9135	2.6078	1.5672	0.2611
	variance Prop.		0.5714	0.2520	0.1514	0.0252
	cumulative Prop.		0.5714	0.8233	0.9748	1.0000
curv	eigenvalue		8.1583	4.4534	2.6453	1.7494
	variance Prop.		0.4797	0.2619	0.1555	0.1029
	cumulative Prop.		0.4797	0.7416	0.8971	1.0000

- Notes:
- (1) *: Equation system (4.2) and (4.3) in the text body.
 - (2) The statistic in the parentheses is Std. Error.
 - (3) The Japan factors used for extracting global factors consist of estimates of two samples from the PACAP and Bloomberg.

Table 4.6: Summary Statistics for Global Factors

	mean	Std.Dev.	skewness	kurtosis	$\hat{\rho}_{(1)}$	$\hat{\rho}_{(12)}$	$\hat{\rho}_{(30)}$
level	7.3025	1.4169	0.2342	1.9836	0.9901	0.8029	0.8607
slope	-1.9008	0.3899	0.1629	2.3144	0.9701	0.4762	-0.0159
curv	-3.9876	0.6066	1.3589	6.3131	0.9278	0.2208	-0.1473

Note: $\hat{\rho}_{(\tau)}$ is the autocorrelation coefficient with lag length τ periods

Table 4.7: Idiosyncratic Regression

$f_{it} = \alpha + \beta f_{wt} + \varepsilon_{it}, i = GM, JP, UK, US;$ $f = level, slope, curvature; f_{wt} : global \text{ factor}$			
Germany			
	$\hat{\beta}$	Std. Error	R^2
Level	0.5837	0.1936	0.8171
Slope	0.3348	0.1827	0.4990
Curvature	2.7673	0.2030	0.4016
Japan			
	$\hat{\beta}$	Std. Error	R^2
Level	1.4114	0.0286	0.8980
Slope	1.3795	0.1130	0.3499
Curvature	1.9063	0.1180	0.4853
U.K.			
	$\hat{\beta}$	Std. Error	R^2
Level	1.6501	0.0313	0.9093
Slope	3.7154	0.2151	0.5186
Curvature	0.7813	0.1882	0.0586
U.S.			
	$\hat{\beta}$	Std. Error	R^2
Level	1.1945	0.0182	0.9395
Slope	2.4012	0.2065	0.3281
Curvature	1.2116	0.1675	0.1588

Table 4.8: Bond Market Integration:
Estimates of Markov-switching Nelson-Siegel Model

Germany		
transition matrix $\begin{pmatrix} 0.4202 & 0.3321 \\ 0.5798 & 0.6679 \end{pmatrix}$	t-values of transition matrix $\begin{pmatrix} 1.3662 & \\ & -3.5387 \end{pmatrix}$	Log Likelihood 2024.08
Japan		
transition matrix $\begin{pmatrix} 0.4689 & 0.2630 \\ 0.5311 & 0.7364 \end{pmatrix}$	t-values of transition matrix $\begin{pmatrix} 0.3332 & \\ & -3.4716 \end{pmatrix}$	636.70
U.K.		
transition matrix $\begin{pmatrix} 0.2977 & 0.2149 \\ 0.7023 & 0.7851 \end{pmatrix}$	t-values of transition matrix $\begin{pmatrix} 2.6776 & \\ & -6.5729 \end{pmatrix}$	1697.70
U.S.		
transition matrix $\begin{pmatrix} 0.4956 & 0.2141 \\ 0.5044 & 0.7859 \end{pmatrix}$	t-values of transition matrix $\begin{pmatrix} 0.9600 & \\ & -6.2117 \end{pmatrix}$	1524.11
World		
transition matrix $\begin{pmatrix} 0.4178 & 0.0896 \\ 0.5822 & 0.9104 \end{pmatrix}$	t-values of transition matrix $\begin{pmatrix} 0.8397 & \\ & -8.3693 \end{pmatrix}$	1059.62

Notes: (1) The model is given in equation (9) in the text body
(2) The State 1 represents the market integration.

Figure 4.1. Yield Curves Across Countries

Figure 4.2. AFDNS Model Estimates of Level Factor Across Countries

Figure 4.3. AFDNS Model Estimates of Slope Factor Across Countries

Figure 4.4. AFDNS Model Estimates of Curvature Factor Across Countries

Figure 4.5. Factors from AFDNS Estimates and DNS Estimates: Germany and Japan

Figure 4.6. Factors from AFDNS Estimates and DNS Estimates: U.K. and U.S.

Figure 4.7. US level, slope and curvature factors and empirical factors

Figure 4.8. Error Terms of AFDNS Estimates across Countries

Figure 4.9. Average and AFDNS-fitted Yield Curves across Countries and Time:

Figure 4.10. The ACF and PACF of Errors from the AFDNS Estimation:
Part 1

Figure 4.11. The ACF and PACF of Errors from the AFDNS Estimation:
Part 2

Figure 4.12. Global Factors from Kalman Filter Estimates and Principle Component Analysis(with appropriate adjustment)

Figure 4.13. Dynamic Conditional Correlations: Level Factor

Figure 4.14. Dynamic Conditional Correlations: Slope Factor

Figure 4.15. Dynamic Conditional Correlations: Curvature Factor

Figure 4.16. Global Yield Curves

Figure 4.17. Filtered and Predicted Probabilities of Integration for Germany, Japan and the UK

Figure 4.18. Filtered and Predicted Probabilities of Integration: the US and World

CHAPTER 5

Conclusions and Recommendations for Future Study

I have examined the dynamic evolution of interest rates with an emphasis on bidirectional and simultaneous linkages between the term structure of interest rates and macroeconomic factors. In essay I, I allow a role for both macroeconomic factors and regime-switching in the testing of the expectations hypothesis. I found some evidences in favor of the expectations hypothesis. My interpretation for the new findings is that the yield factors at least partially capture the time-varying risk premiums. Essay II provided some new evidences on the close relationship between the term structure of interest rates and macroeconomic factors. Furthermore, the macroeconomic factors are important in identifying regimes that are related to business cycles. Essay III explored the integration and interaction of global government bond markets employing an arbitrage-free dynamic model. My contribution includes developing an econometric model to empirically support the expectations hypothesis of the term structure and examine the joint dynamics and feedback mechanism between the yield curve and macroeconomic factors.

There are several different directions future research might take. Empirical macroeconomic research has not been able to establish if and how government deficits affect the term structure dynamics. However, this is a question of crucial importance for policy making and for academic research. On the other hand, recent theoretical and empirical research in finance has led to a better understanding of the dynamic properties of the term structure of interest rates. I feel that the modern term structure

models provide us a good framework to incorporate fiscal policy variables and identify the effects of fiscal policy shocks. Furthermore, a more fundamental model, such as a general equilibrium term structure model, will highlight the transmission mechanism between the yield curve and fiscal policy.

Perhaps the most natural extension of Essay II is to develop an arbitrage-free regime-switching macro-finance model of the term structure. Because nominal bonds are traded in well-organized markets, the theoretical restrictions that rule out riskless arbitrage opportunities across maturities hold a powerful appeal, and they provide the foundation for a large finance literature on arbitrage-free models. The extension that rule out arbitrage opportunities will be interesting.

From the perspective of bond portfolio diversification, understanding the dynamic correlation structure of international bond markets is a high priority question. It also relates to a fundamental question of international market efficiency. An integrated world capital market implies increased cross-country correlations in the sense that investors respond to any new information throughout the world. If international markets are not perfectly correlated, we need both world and local factors in international asset pricing models. At the practical level, the benefit of international diversification will be nil in a perfectly correlated international market.

Another approach is to consider more macroeconomic fundamentals. Bond returns are found to be predictable in the long-run by yield spreads and previous excess bond returns. However, we are not clear about the macroeconomic fundamentals where the predictability stems from. This is an important question for bond portfolio management and monetary-policy making.

References

- [1] Akdogan, H., 1995, The integration of international capital markets. Edward Elgar.
- [2] Ang, A., Bekaert, G., 2002. Regime switches in interest rates. *Journal of Business and Economic Statistics* 20, 163-182.
- [3] Ang, A., Piazzesi, M., 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent factors. *Journal of Monetary Economics* 50, 745-787.
- [4] Ang, A., Piazzesi, M., and Sen Dong, 2007, No-arbitrage Taylor rules. NBER Working Paper 13448.
- [5] Ang, Andrew, Piazzesi, M., and Wei, M., 2006, What does the yield curve tell us about GDP growth? *Journal of Econometrics* 131, 359-403.
- [6] Bansal, R., Tauchen, G., and Zhou, H., 2004. Regime shifts, risk premium in the term structure, and business cycles. *Journal of Business and Economic Statistics* 20, 396-409.
- [7] Bansal, R., Zhou, H., 2002, Term structure of interest rates with regime shifts. *Journal of Finance* 57, 1997-2043.
- [8] Barr, D.G., Priestley, R., 2004, Expected returns, risk and the integration of international bond markets. *Journal of International Money and Finance* 23, 71-97.
- [9] Bekaert, G., Harvey, C.R., 1995, Time-varying world market integration. *Journal of Finance* 50, 403-444.
- [10] Bekaert, G., Hodrick, R.J., 2001. Expectation hypotheses tests. *Journal of Finance* 56, 1357-1394.
- [11] Bekaert, G., Robert J.H., and David A.M., 2001, Peso problem explanations for term structure anomalies, *Journal of Monetary Economics* 48, 241-270.
- [12] Campbell, J.Y., Andrew W. Lo, and Mackinlay, A.C., 1997, *The econometrics of financial markets*. Princeton University Press.

- [13] Campbell, J.Y., 1986, A defense of the traditional hypotheses about the term structure of interest rates, *Journal of Finance* 41, 183-193.
- [14] Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. *Journal of Political Economy* 95, 1062-1088.
- [15] Campbell, J.Y., and Robert, J.S., 1991, Yields spreads and interest rate movements: a bird's eye view, *Review of Economics Studies* 58, 494-514.
- [16] Carriero, A., Favero, C.A., and Kaminska, I., 2006, Financial factors, macroeconomic information and the expectations theory of the term structure of interest rates, *Journal of Econometrics* 131, 339-358.
- [17] Cater, C.K., Kohn, R., 1994. On Gibbs sampling for state space models. *Biometrika* 81, 541-553.
- [18] Chang, R., 1997. Financial integration with and without international policy coordination. *International Economic Review* 38, 547-564.
- [19] Chen, Z., Knez, P.J., 1995, Measurement of Market Integration and Arbitrage. *Review of Financial Studies* 8, 287-325.
- [20] Chib, S., Greenberg, E., 1994. Bayes inference in regression models with ARMA(p,q) errors. *Journal of Econometrics* 64, 183-206.
- [21] Christensen, J.H.E., Diebold, F.X., Rudebusch, G.D., 2007, The affine arbitrage-free class of Nelson-Siegel term structure models. Manuscript. Department of Economics, University of Pennsylvania.
- [22] Cochrane, J.H., Piazzesi, M., 2005, Bond risk premia. *American Economic Review* 95, 138-160.
- [23] Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 385-407.
- [24] Dai, Q., Singleton, K.J., 2002, Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* 63, 415-441.
- [25] Dai, Q., Singleton, K.J., and Yang, W., 2006, Are regime shifts priced in U.S. treasury markets? *Review of Financial Studies* 20, 1669-1702.
- [26] Dewachter, H., Lyrio, M., 2003, Macro factors and the term structure of interest rates, *Journal of Money Credit and Banking* 38, 119-140.
- [27] Diebold, F.X., Li, C., 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* 130, 337-364.
- [28] Diebold, F.X., Li, C., Yue, V.Z., 2006, Global yield curve dynamics and interactions: a generalized Nelson-Siegel approach. *Journal of Econometrics*, forthcoming.

- [29] 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, 209-338.
- [30] Driessen, J., Melenberg, B., Nigman, T., 2003, Common factors in international bond returns. *Journal of International Money and Finance* 22, 629-656.
- [31] Duffee, G., 2002, Term premia and interest rate forecasts in affine models. *Journal of Finance* 57, 405-443.
- [32] Duffie, D., 2001, *Dynamic asset pricing theory*. Princeton University Press.
- [33] Engle, R., 2002, Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic statistics* 20, 339-350.
- [34] Engsted, T., Tanggaard, C., 2007, The comovement of US and Germany bond market. *International Review of Financial Analysis* 16, 172-182.
- [35] Estrella A., Hardouvelis, G.A., 1991, The term structure as a predictor of real economy activity. *Journal of Finance* 46, 555-576.
- [36] Estrella, A., Mishkin, F.S., 1998. Predicting U.S. recessions: financial variables as leading indicators. *Review of Economics and Statistics* 80, 45-61.
- [37] Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 65, 269-282.
- [38] Garcia, R., Perron, P., 1996. An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics* 78, 111-125.
- [39] Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches: the calculation of posterior moments. *Bayesian statistics 4: Proceedings of the fourth Valencia international meeting*, 169-193.
- [40] Gray, S.F., 1996, Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics* 42, 27-62.
- [41] Hafer, R.W., Kutan, A.M., Zhou, S., 1997, Linkage in EMS term structure: evidence from common trends and transitory components. *Journal of International Money and Finance* 16, 595-607.
- [42] Hamilton, J.D., 1988, Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates, *Journal of Economic Dynamics and Control* 12, 385-423.
- [43] Hall, M.J.B., 2003, *The international handbook on financial reform*. Edward Elgar.

- [44] Hamilton, J.D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycles, *Econometrica* 57, 357-384.
- [45] Hamilton, J.D., 1994, Time series analysis. Princeton University Press.
- [46] Hamilton, J.D., 1996, Specification testing in Markov-switching time-series models. *Journal of Econometrics* 70, 127-157.
- [47] Hamilton, J.D., and Dong Heon Kim, 2002, A reexamination of the predictability of economic activity using yield spread, *Journal of Money, Credit and Banking* 34, 340-360.
- [48] Harvey, C.R., The real term structure and consumption Growth, *Journal of Financial Economics* 22, 305-333.
- [49] Harvey, A., Rossi, G.D., 2006, Signal extraction. *Palgrave Handbook of Econometrics* vol 1, Palgrave Macmillan.
- [50] Johannes, M., Polson, N., 2009, MCMC methods for continuous-time financial econometrics. *Handbook of Financial Econometrics*, Elsevier.
- [51] Jorion, P., Mishkin, F., 1991, A multicountry comparison of term-structure forecasts at long horizons, *Journal of Financial Economics* 29, 59-80.
- [52] Kamara, A., 1997, The relation between default-free interest rates and expected economic growth is stronger than you think, *Journal of Finance* 52, 1681-1694.
- [53] Kim., C.-J., 1994. Dynamic linear models with Markov-switching. *Journal of Econometrics* 60, 1-22.
- [54] Kim, C.-J., Nelson, C.R., 1998. State-space models with regime-switching: classical and Gibbs-sampling approaches with applications. Cambridge: MIT press.
- [55] Kozicki, S., Tinsley, P.A., 2001, Shifting endpoints in the term structure of interest rates, *Journal of Monetary Economics* 47, 613-652.
- [56] Krolzig, Hans. -M., Markov-Switching Vector Autoregressions. *Modelling, Statistical Inference and Application to Business Cycle Analysis, Lecture Notes in Economics and Mathematical Systems*, Volume 454, Berlin: Springer.
- [57] Nelson, C.R., Siegel, A.F., 1987, Parsimonious modeling of yield curves. *The Journal of Business* 60, 473-489.
- [58] Ilmanen, A., 1995, Time-varying expected returns in international bond markets. *Journal of Finance* 50, 481-506.
- [59] Longstaff, F.A., 2000a, Arbitrage and the expectations hypothesis. *Journal of Finance* 55, 989-994.

- [60] Longstaff, F.A., 2000b, The Term structure of very short term rates: new evidence for the expectations Hypothesis, *Journal of Financial Economics* 58, 397-415.
- [61] Piazzesi, M., 2009, Affine term structure models. *Handbook of Financial Econometrics*, Elsevier.
- [62] Raftery, A.E., Lewis, S., 1992. How many iterations in the Gibbs Sampler? in J. M. Bernardo, J. Berger, A. P. David, and A. F. M. Smith, ed., *Bayesian Statistics 4*, Oxford University Press, Oxford, pp. 763-773.
- [63] Sarno, L, Thornton, D.L., Valente, G., 2007. The empirical failure of the expectations hypothesis of the term structure of bond yields. *Journal of Financial and Quantitative Analysis* 42, 81-100.
- [64] Shiller, R.J., 1979, The volatility of long-term interest rates and expectations models of the term structure. *Journal of Political Economy* 87, 1190-1219.
- [65] Sill, K., 2000, Understanding asset values: stock prices, exchange rates and the Peso problem. *Business Review*, September/October, 3-13.
- [66] Sola, M., Driffill, J., 1994, Testing the term structure of interest rates using a stationary vector autoregression with regime switching, *Journal of Economic Dynamics and Control* 18, 601-628.
- [67] Sutton, G.D., 2000, Is there excess comovement of bond yields between countries? *Journal of International Money and Finance* 19, 363-376.
- [68] Tam, C.S., Yu, I.W., 2008, Modelling sovereign bond yield curves of the US, Japan and Germany. *International Journal of Finance and Economics* 13, 82-91.
- [69] Taylor, J.B., 1993. Discretion versus policy rule in practice. *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.
- [70] Vasicek, Oldrich A., 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177-188.
- [71] Walsh, C.E., 2003, *Monetary theory and policy*. 2nd edition, the MIT Press.
- [72] Xiaoneng, Zhu, 2007, Testing the present value model under regime shift with bootstrap-based Wald test: the Japanese term structure. Working Paper.
- [73] Yu, B., Mykland, P., 1998. Looking at Markov samplers through cusum path plots: a simple diagnostic idea. *Statistics and computing* 8, 275-286.
- [74] Zhu, X., 2008, Global yield curves and international business cycles. Working Paper.