Essays on forecasting interest rates, economic activity and inflation, based on factor models of the yield curve

by Efthymios Argyropoulos

A thesis submitted to the Department of Economics
in fulfilment of the requirements for the degree of

Doctor of Philosophy in Economics

June 2013

Athens University of Economics and Business
Athens, Greece
# Contents

Abstract vi

Introduction and summary vii

1 Affine term structure models: literature review 1
   1.1 Introduction ............................................. 1
   1.2 Setup of affine term structure models ................. 3
      1.2.1 Risk-neutral pricing ................................. 3
      1.2.2 Pricing in a risk-free world ....................... 4
      1.2.3 Pricing in the real world .......................... 5
      1.2.4 The market price of risk ........................... 6
      1.2.5 Affine specifications ............................... 8
   1.3 Term premia .............................................. 10
      1.3.1 Empirical evidence from bond returns ............ 14
      1.3.2 Extended models of risk premia .................... 15
   1.4 Real bonds and general equilibrium models .......... 17

2 Term spread regressions of the REHTS allowing for RP effects 23
   2.1 Introduction ............................................. 23
   2.2 Model set up ............................................ 25
   2.3 Empirical analysis ..................................... 28
      2.3.1 Data description .................................... 29
      2.3.2 Retrieving unobserved factors from interest rates data ................................................. 30
      2.3.3 Unit root tests ...................................... 32
      2.3.4 Principal component analysis ...................... 34
      2.3.5 Estimation of the GDTSM and cross-section restriction tests .... 39
      2.3.6 Tests of the REHTS allowing for term premium effects ................................................. 45
   2.4 Conclusions ............................................. 50

3 Real term structure forecasts of consumption growth 52
   3.1 Introduction ............................................. 52
   3.2 Model setup ............................................. 54
      3.2.1 Term structure forecasts of consumption growth ................................................. 57
   3.3 Empirical analysis ..................................... 60
      3.3.1 Unit root tests ...................................... 63
      3.3.2 Principal component analysis ...................... 64
      3.3.3 Estimation of the real term structure model ................................................. 69
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.4</td>
<td>Real term structure forecasts of consumption growth</td>
<td>76</td>
</tr>
<tr>
<td>3.4</td>
<td>Conclusions</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>Forecasting economic activity from yield curve factors</td>
<td>82</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>82</td>
</tr>
<tr>
<td>4.2</td>
<td>Forecasting economic activity based on term spread regressions</td>
<td>84</td>
</tr>
<tr>
<td>4.3</td>
<td>Forecasting economic activity from the slope and curvature factors</td>
<td>89</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Retrieving yield curve factors $s_{it}$ and $c_{it}$</td>
<td>90</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Forecasting economic activity based on yield factors $s_{it}$ and $c_{it}$</td>
<td>97</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Out-of-sample forecasting performance</td>
<td>100</td>
</tr>
<tr>
<td>4.4</td>
<td>Conclusions</td>
<td>102</td>
</tr>
<tr>
<td>5</td>
<td>Forecasting inflation from the term structure and the IRP effects</td>
<td>104</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>104</td>
</tr>
<tr>
<td>5.2</td>
<td>Model setup</td>
<td>106</td>
</tr>
<tr>
<td>5.3</td>
<td>Empirical analysis</td>
<td>112</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Data</td>
<td>113</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Model specification and estimation</td>
<td>116</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Estimation results</td>
<td>118</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Inflation risk premia</td>
<td>125</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Forecasting inflation from the term structure</td>
<td>127</td>
</tr>
<tr>
<td>5.4</td>
<td>Conclusions</td>
<td>130</td>
</tr>
<tr>
<td>5.5</td>
<td>Appendix</td>
<td>131</td>
</tr>
</tbody>
</table>
Preface

"The roots of education are bitter, but the fruit is sweet" - Aristotle

This thesis marks the end of my PhD studies at the Athens University of Economics and Business and I acknowledge that it would not have been completed without the contribution of some significant persons in my life inside and outside academia. First and foremost, I would like to express my greatest and deepest gratitude to the principal supervisor of my thesis, Professor Elias Tzavalis, for his valuable guidance at every stage of the research, the precious knowledge he has transmitted to me and the continuous support and encouragement in every possible way being always available to listen and answer my questions.

I would also like to thank Professor Panagiotis Korliras for his total support during these years of my PhD studies and for opportunity he gave me to work in his office. I also thank the other two members of the supervising committee, Professor Efthymios Tsionas and Assistant Professor Stelios Arvanitis for the knowledge and the inspiration they provided me to successfully continue my research. I am sincerely grateful to all the Professors and fellow Ph.D students of the Department of Economics for their valuable comments, suggestions and the friendly environment I enjoyed. Part of the research presented in this thesis was funded by the Basic Research Funding Program of Athens University of Economics and Business for the period 2011-2012.

Outside academia, I owe my deepest gratitude to my family, and especially, to my parents, Dimitris and Paraskeyi and my sister Stayroula, for their continuous support and encouragement during my studies. Finally and most important of all, I would like to thank my fiancée, Sophia, for her unconditional love, relentless support and encouragement during all my ups and downs.
To my beloved grandfather, Argyris.
Abstract

The aim of the thesis is to provide new insights into the forecasting ability of the term structure of interest rates about future interest rate movements, real consumption growth, economic activity and inflation. After a brief review in term structure modeling, the thesis presents four studies in order to shed some light into these issues.

In particular, it consists of four essays entitled: "Term spread regressions of the rational expectations hypothesis of the term structure allowing for risk premium effects", "Real term structure forecasts of consumption growth", "Forecasting economic activity from yield curve factors", and "Forecasting inflation from the term structure and the inflation risk premia effects".

The thesis provides a number of interesting results for academics and practitioners. First, it shows that term premium effects can explain the puzzles of Expectations Hypothesis to forecast future movements of interest rates. Second, the forecasting ability of the spread and the short-term rate in future real consumption growth can be attributed to two common factors spanning the real term structure. Third, the curvature factor contains important information about the economic growth in short-term, independently of the slope factor. Fourth, the inflation risk premium can affect the forecasting ability of the term spread about future inflation in the short-term.
Introduction and summary

The objective of this thesis is to provide new insights into the ability of yield curve to forecast future interest rate movements, economic activity and inflation. It is positioned in the research area that intersects between finance and macroeconomic theory. The thesis consists of five chapters in total, from which, the first is an introduction to the literature and the other four are studies that make empirical and methodological contributions in the field of dynamic term structure modeling and forecasting.

In particular, the first chapter entitled: "Affine term structure models: literature review", presents a brief review of the vast literature concerning affine term structure models and analyzes their key structural features, starting from the early and simplified setups up to the more recent and elaborated specifications. The latter are implemented in the chapters that follow.

The second chapter entitled: "Term spread regressions of the rational expectations hypothesis of the term structure allowing for risk premium effects", examines the empirical failures of the term spread to forecast future interest rate movements and suggests term spread regression based tests allowing for time-varying term premium effects to answer these failures. To capture the effects of a time-varying term premium on the term spread, a simple and empirically attractive arbitrage-free Gaussian dynamic term structure model is implemented. The model assumes that the term structure of interest rates is spanned by three unobserved factors. To retrieve these factors from the data, we suggest a new empirical methodology which can overcome the effects of measurement (or pricing) errors on the estimates of the unobserved factors, and our tests. The results indicate first, that a three-factor and empirically tractable model can sufficiently explain the cross-section movements of the US term structure of interest rates implied by no-arbitrage conditions in the bond market, second, term spread regressions allowing for a time-varying term premium effects can provide unbiased estimates of future changes of long-term interest rates.
as predicted by the rational expectations hypothesis of the term structure. Finally, the factors
that are priced in the bond market and, thus, cause significant time-varying effects on the term
spread regressions are those which are associated with level and slope shifts in the term structure
of interest rates.

The third chapter entitled: "Real term structure forecasts of consumption growth," proposes
an affine term structure model of real interest rates to predict changes in real consumption
growth. The model is estimated, jointly, by real interest rates and consumption data, and it is
found to be consistent with the consumption smoothing hypothesis. This study contributes into
the above literature on many fronts. First, using real consumption and term structure data,
instead of nominal, it estimates an empirically tractable Gaussian dynamic term structure model
and derives estimates of the underlying unobserved factors spanning the term structure of real
interest rates. Then, it examines if this model fits satisfactorily into the data and tests its
cross-section restrictions implied by no-profitable arbitrage conditions in the bond market. The
real term structure is spanned by two common factors, which can be given the interpretation
of the level and slope factors, respectively. The risks associated with these factors are priced in
the market. Both of them can explain the information content of the short-term real interest
rate and its term spread with longer term interest rate in forecasting future real consumption
growth, over different periods ahead.

The fourth chapter entitled: "Forecasting economic activity from yield curve factors," pro-
vides some new interesting results about the predicting ability of the yield curve and term
spread. The chapter presents clear cut evidence that the slope and curvature factors of the
yield curve contain superior information about future economic activity than the term spread
itself. This is shown for five world leading economies. To extract the slope and curvature factors
of the yield curve, the paper fits the dynamic Nelson-Siegel model. The sign of the predictions
of the slope and curvature factors on future economic activity is different. They imply that an
increase in the slope factor is associated with a slow down in economic activity, while the op-
posite is predicted for an increase in the curvature factor. These results indicate that the slope
factor of the yield curve should reflect future changes in business cycle conditions, which can
last for a few years ahead, while the curvature factor may be associated with short or medium
term changes in the current stance of monetary policy. The effects of these two factors on the
term spread are offset to each other, and thus reduce the ability of the term spread to forecast the correct direction of future changes in economic activity.

Finally, the fifth chapter entitled: "Forecasting inflation from the term structure and the inflation risk premia effects", fits a dynamic Gaussian term structure model into nominal term structure, inflation and real consumption data with the aim of examining how important are inflation risk premium in short run. First, is shown that the model is consistent with the data and can describe sufficiently the dynamics of the nominal term structure and inflation rates, observed in reality. Furthermore, the real interest rates and expected inflation forecasts retrieved by the model are very close to those implied by survey and inflation-indexed bonds over longer horizons. Second, the inflation risk premium is found to be negative, which implies that investors in the bond market require less compensation for holding nominal bonds compared to inflation-indexed bonds. Also the mean and volatility of these premia decline with maturity interval. Third, as inflation risk premium, real interest rates are also volatile in short-term. These together with inflation risk premia can explain the failures of the nominal term spread to forecast future inflation rates over short-term horizons.
Chapter 1

Affine term structure models: literature review

1.1 Introduction

The term structure of interest rates possesses a central role in the modern financial research. Understanding and modeling interest rates has driven academic researchers, financial analysts, practitioners and policy makers to produce an enormous literature consisting of sophisticated term structure models. In this review, we present a brief summary of the most significant achievements in the term structure literature and their evolution until nowadays.

Modeling the term structure is essential for forecasting future interest rates, economic activity and inflation. Such forecasts provide a significant insight into the economy not only for investors, but also for central banks and/or consumers. This was the motivation behind a number of studies that forecast future interest rate movements, (see e.g., Fama and Bliss [59], Campbell and Shiller [24], Piazzesi [94] among others), economic activity (see e.g., Harvey [71], [72], Estrella and Hardouvelis [57], Hamilton and Kim [68], Ang, Piazzesi and Wei [5], Rudebusch and Wu [102], among others) and inflation (see e.g., Fama [58], Mishkin [87], Dewachter and Lyrio [42], Berardi [16], Christensen, Lopez and Rudebusch [30]).

But, what are the key features of the term structure models? Primary models of the term structure have one thing in common. They all assume that the movements of the yield curve are determined by a single state variable, this variable usually is the instantaneous risk-free interest rate. The most famous examples are presented in Vasicek [109] and Cox, Ingersoll and Ross (hereafter CIR) [32]. But, the assumption of a single factor governing yield curve dynamics
is too restrictive. Single factor models lack the ability to capture the observed variability in yields. In fact, the term structure dynamics are much too complex to be summarized in a single source of uncertainty. Nevertheless, the Vasicek and CIR models provide a significant first step into efficient term structure modeling. The term 'efficient' comes from the fact that these early models introduce cross equation restrictions that closely link the cross-section and the dynamics of yields ruling out possible arbitrage opportunities that may exist in the bonds market. After these two initial models, many other followed. Chan, Karolyi, Longstaff, and Sanders [27] provide an empirical comparison of single-factor term structure models, introduced in the literature.

Although single-factor models pave the way for term structure modeling, Litterman and Scheinkman (LS) [81] show substantial empirical evidence that not just one, but three factors explain the observed variability in yields, in total. These factors are taken to reflect changes in the 'level', 'slope', and 'curvature' of the term structure. As a consequence, a number of extensions of the single factor term structure models models triggered, incorporating multiple-state variables. Multifactor models assume that the short rate $r_t$ is a function of state variables collected in vector $X_t$. The multifactor models vary in terms of the number of the state variables, generating interest rates across different maturity intervals $\tau$.

Duffie and Kan [53] provide a complete characterization of multifactor term structure models and introduce the term 'affine'. Affine models imply linear relationships between interest rates $R_t(\tau)$ and/or the short rate $r_t$ and the vector of state variables $X_t$,

$$R_t(\tau) = A(\tau) + D(\tau)X_t$$

This generalized affine type of models nests the Vasicek and CIR famous term structure models, as well as their multifactor extensions, as special cases. It impose the necessary no-arbitrage cross-equation restrictions on coefficients $A(\tau)$ and $D(\tau)$. The vector of state variables $X_t$, capturing the underlying source of uncertainty, is assumed that follows a stochastic process driving the dynamics of the term structure. This can be observable, latent, or it can be linked to macroeconomic variables e.g., the inflation rate. The coefficients $A(\tau)$ and $D(\tau)$ depend on maturity interval $\tau$ and render yield equations consistent not only with each other, but also with the state variable dynamics. The cross-equation restrictions imposed on slope coefficients
D(τ) are essential since they imply the following (see also Piazzesi [96]). First, as previously mentioned, they guarantee the absence of arbitrage conditions. Second, they are needed to model time varying risk premia (see e.g., Duffee [51], Cochrane and Piazzesi [34]). Third, based on them, state variables can be retrieved based on a number of interest rates.

The above affine functional form of bond yields presumes affine state variable dynamics. This means that the risk-adjusted process of \( X_t \) will also have an affine diffusion, i.e., a process with affine instantaneous mean and variance. To make this more clear, consider the following process of \( X_t \):

\[
dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dW_t
\]

where \( \mu_X(X_t) = \mu_1 + \mu_2 X_t \) and \( \sigma_X(X_t) = \sigma_1 + \sigma_2 X_t \). The differential change in the state vector \( dX_t \) is composed of a non random drift term \( \mu_X(X_t)dt \) and a random diffusion term \( \sigma_X(X_t)dW_t \), which includes a differential change in a Brownian motion vector \( dW_t \). The affine specification in yields restricts the drift term \( \mu_X(X_t) \) and the diffusion volatility \( \sigma_X(X_t) \) to be affine in \( X_t \).

The bond yield function \( R_t(\tau) \) is an important link between the state variables \( X_t \) and the term structure of interest rates. The only requirement for the implementation of the models is just the affine formulations.

### 1.2 Setup of affine term structure models

Let \( B_t(\tau) \) be \( t \)-time price of a nominal zero coupon bond that matures \( \tau \) periods from now and pays 1 nominal unit without the risk of default. Then, the nominal interest rate \( R_t(\tau) \) is defined as

\[
R_t(\tau) = -(1/\tau) \log B_t(\tau)
\]

and the instantaneous nominal interest rate, playing the role of short-term rate, is just the limit of \( R_t(\tau) \), as \( \tau \to 0 \), i.e., \( r_t = \lim_{\tau \to 0} R_t(\tau) \)

### 1.2.1 Risk-neutral pricing

Bonds are priced under the risk-neutral probability measure \( \mathbb{Q} \). This means that \( B_t(\tau) \) are the expected values of their future payoffs discounted at the risk-free rate, where the expectation is
taken under $Q$. The price $B_t(\tau)$ of a zero coupon bond is thus given as:

$$B_t(\tau) = E_t^Q \left[ \exp \left( - \int_t^{t+\tau} r_s ds \right) \right],$$  \hspace{1cm} (1.1)

where $E_t^Q$ denotes the expectation taken under $Q$. This means that, in order to price a pure discount bond, we must change from the data-generating, or physical measure $P$, to the risk-neutral measure $Q$ and, then, define the dynamics of the short rate under that measure. In multifactor models, the short rate $r_t$ is a function of a state vector $X$, $f(X_t)$, where $X \in \mathbb{R}^K$, is a time-homogeneous Markov process under $Q$. Thus, the conditional expectation in (1.1) is some function of time to maturity $\tau$ and the state variables $X_t$, defined as:

$$P(X_t, \tau) \equiv B_t(\tau)$$  \hspace{1cm} (1.2)

By Ito’s lemma, we can turn the problem of solving the conditional expectation in (1.1) into the problem of solving a partial differential equation (PDE) for the bond price function $P(X, \tau)$. Computing (1.1) by solving a PDE is called the Feynman–Kac approach.

### 1.2.2 Pricing in a risk-free world

Under risk-neutral measure $Q$, equation (1.1) implies that expected excess returns on bonds are zero, hence, the expected rate of return on a long bond just equals the risk-free rate.

Then, a Markov process $X$ exists in a state space $D \in \mathbb{R}^K$ and solves the following stochastic differential equation (SDE):

$$dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dW_t$$  \hspace{1cm} (1.3)

where $W$ is an $K$–dimensional standard Brownian motion under $P$, $\mu_X : D \to \mathbb{R}^K$ is the drift of $X$ and $\sigma_X : D \to \mathbb{R}^{K \times K}$ is its volatility. Since from equation (1.1) the bond price $B_t(\tau)$ is a function of maturity $\tau$ and the state vector $X$, Ito’s lemma implies,

$$\frac{dP(X_t, \tau)}{P(X_t, \tau)} = \mu_P(X_t, \tau)dt + \sigma_P(X_t, \tau)dW_t,$$  \hspace{1cm} (1.4)

where the instantaneous expected bond return $\mu_P(X, \tau)$ equals

$$\mu_P(X, \tau) = -\frac{P_\tau(X, \tau)}{P(X, \tau)} + \frac{P_X(X, \tau)}{P(X, \tau)}\mu_X(X_t) + \frac{1}{2}tr \left[ \sigma_X(X_t)\sigma_X(X_t)P_{XX}(X, \tau) \right]$$  \hspace{1cm} (1.5)
1.2. SETUP OF AFFINE TERM STRUCTURE MODELS

where \( P_r, P_X \) and \( P_{XX} \) are partial derivatives of \( P \) and \( \text{tr} \) denotes trace.\(^1\)

From the local EH follows that the instantaneous expected bond return \( \mu_P(X, \tau) \) equals the short rate,

\[
\mu_P(X, \tau) = r \text{ for all } X \in \mathcal{D} \text{ and } \tau > 0.
\]

with the initial condition that the price of the bond at maturity is equal to its payoff \( P(X, 0) = 1 \).

1.2.3 Pricing in the real world

In the more realistic case where the physical measure \( \mathcal{P} \) differs from the risk-neutral \( \mathcal{Q} \), the expected rate of return on a long bond \( \mu_P(X, \tau) \) is not equal to the risk-free rate \( r_t \), since, this implies a flat yield curve with zero risk premium. Expected returns on long term-bonds are usually higher than those of shorter ones or eventually of the short rate, because they incorporate a *risk premium* for holding them for short time periods, referred to as liquidity or term premium. In the following sections this premium is presented in detail.

As a consequence, under the physical measure \( \mathcal{P} \), expected returns on bonds will be equal to the riskless rate *only* under the risk-neutral measure \( \mathcal{Q} \). The state vector in this case \( X_t \) solves the following stochastic differential equation:

\[
dX_t = \mu^Q_X(X_t)dt + \sigma^Q_X(X_t)dW^Q_t
\]

where \( \mu^Q_X(X_t) \) and \( \sigma^Q_X(X_t) \) are the risk-neutral measures of \( \mu_X(X_t) \) and \( \sigma_X(X_t) \) and \( W^Q \) is Brownian motion under \( \mathcal{Q} \). Similar to (1.4), the SDE for the bond price now is written as

\[
\frac{dP(X_t, \tau)}{P(X_t, \tau)} = \mu^Q_P(X_t, \tau)dt + \sigma^Q_P(X_t, \tau)dW^Q_t
\]

and the bond’s expected rate of return \( \mu^Q_P(X, \tau) \) will be given as

\[
\mu^Q_P(X, \tau) = -\frac{P_r(X, \tau)}{P(X, \tau)} + \frac{P_X(X, \tau)'}{P(X, \tau)}\mu^Q_X(X_t) + \frac{1}{2} \text{tr} \left[ \sigma^Q_X(X_t)\sigma^Q_X(X_t)^T \frac{P_{XX}(X, \tau)}{P(X, \tau)} \right]
\]

where now the drift \( \mu^Q_X(X_t) \) and the volatility \( \sigma^Q_X(X_t) \) of \( X \) are defined under the risk neutral measure \( \mathcal{Q} \). The risk-neutral expected bond return equals the risk-free rate,

\[
\mu^Q_P(X, \tau) = r \text{ for all } X \in \mathcal{D} \text{ and } \tau > 0
\]

\(^1\)Note also that \( \partial r / \partial \tau = -1 \).
As it is obvious from the above analysis, the risk-neutral measure $Q$ will be different from the data-generating measure $P$.

### 1.2.4 The market price of risk

From the analysis of previous section, it is clear that in order to price a bond we can change from the physical probability measure $P$ to the risk-neutral $Q$. This change of measure is essential since, first, it captures the adjustment for risk and second, the risk-free world is a rather restrictive assumption which is not consistent with reality. This equivalent probability measure, $Q$, solves serious pricing problems like the varying discount rates between investors or the -impossible to identify- individual risk preferences. The way $P$ and $Q$ are linked together is explained below.

Positive martingales play the central role in changing probability measures. Thus, $P$ and $Q$ are linked through them. Recall that process $X$ is the solution of the SDE

$$dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dW_t$$

where $W = (W_t, t \geq 0)$ is the $K$–dimensional standard Brownian motion under the physical probability measure $P$, defined on the probability space $(\Omega, \mathcal{F}_t, P)$ and $\mathcal{F}_t = \sigma(W_s, s \leq t)$ is the Brownian filtration.

Now consider the martingale process $Z_t$ from which the risk-neutral measure $Q$ is constructed. $Z_t$ can also be seen as the Radon-Nikodym derivative $dQ/dP$ of the risk-neutral measure $Q$ with respect to the physical measure $P$ which solves the linear SDE starting at $Z_0 = 1$,

$$\frac{dZ_t}{Z_t} = \Lambda(X_t)'dW_t,$$

The solution has the following exponential form:

$$Z_t = \exp \left\{ \int_0^t \Lambda(X_s)'dW_s - \frac{1}{2} \int_0^t \Lambda(X_s)'\Lambda(X_s)ds \right\} 1$$

(1.10)

in which $\Lambda(X_s)$ is a $K$–dimensional $\mathcal{F}_t$–adapted stochastic process that satisfies the Novikov’s condition

$$E \left( \exp \left\{ \int_0^t \Lambda(X_s)'dW_s - \frac{1}{2} \int_0^t \Lambda(X_s)'\Lambda(X_s)ds \right\} \right) = 1,$$

(1.11)
1.2. SETUP OF AFFINE TERM STRUCTURE MODELS

with

\[ E \left( \exp \left\{ \frac{1}{2} \int_0^t \Lambda(X_s)' \Lambda(X_s) ds \right\} \right) < \infty \]

From the above, \( Z_t \) is a positive martingale with respect to \( \mathcal{F}_t \) for any process \( \Lambda_t \).

It follows from Girsanov’s theorem that the risk-neutral process

\[ W_t^Q = -\int_0^t \Lambda(X_s) ds + W_t \]

is \( K \)–dimensional Brownian motion under the probability measure \( Q \). Note that \( W_t^Q \) represents a stochastic process with predetermined drift \(- \int_0^t \Lambda(X_s) ds \) under \( P \). More specifically, it is exactly the quantity used to change the drift to our Brownian motion, \( W_t \), to create a new Brownian motion \( W_t^Q \) such that all bond returns are equal to the instantaneous risk-free interest rate \( r_t \) under measure \( Q \), i.e.,

\[ dW_t^Q = dW_t - \Lambda(X_t)' dt. \]

By plugging \( dW_t^Q \) into the risk-neutral dynamics (1.7) we can take the stochastic process of \( X \) under \( P \):

\[ dX_t = (\mu_X^Q(X_t) - \sigma_X^Q(X_t) \Lambda(X_t)' dt + \sigma_X^Q(X_t) dW_t. \]

Note that, the volatility of the process \( X \) remains the same under \( P \) and \( Q \) measures, i.e.,

\[ \sigma_X(X_t) = \sigma_X^Q(X_t). \tag{1.12} \]

But, the drift term of stochastic process \( X \) under measure \( P \) is given as

\[ \mu_X(X_t) = \mu_X^Q(X_t) - \sigma_X(X_t) \Lambda(X_t)' \tag{1.13} \]

The vector \( \Lambda(X_t) \) that links the physical measure \( P \) to the risk-neutral measure \( Q \) reflects so-called Market Price of Risk (MPR). In the finance literature, it is the volatility parameter of the Stochastic Discount Factor (SDF) or Pricing Kernel. MPR can have constant or time varying parameters. In fact, as long as \( \Lambda(X_t) \) satisfies the Novikov’s condition, the use of time varying specifications for \( \Lambda(X_t) \) are possible. Most recent affine models of the term structure assume such specifications of \( \Lambda(X_t) \), as we will see in the following sections.
1.2.5 Affine specifications

The two basic assumptions of affine models concern the short rate function \( r_t = f_t(X_t) \) and the state vector’s process under the risk-neutral measure \( Q \). The functional form of the short rate and the state dynamics must be affine. Thus, \( r_t \) is affine function of \( X_t \), and \( X_t \) is an affine diffusion under \( Q \). Also, the risk neutral drift \( \mu_X(X_t) \) and the variance matrix \( \sigma_X(X_t)\sigma_X(X_t)' \) must also be affine in \( X_t \). The functional form of the short rate \( r_t \) is given by

\[
   r_t = \delta_0 + \delta_1 X_t \tag{1.14}
\]

for \( \delta_0 \in \mathbb{R} \) and \( \delta_1 \in \mathbb{R}^K \). In one-factor models, the short rate \( r_t \) depends upon only one factor and hence, \( \delta_0 = 0 \) and \( \delta_1 = 1 \). To conclude, the state vector \( X_t \) solves (1.3)

\[
   dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dW_t, \tag{1.15}
\]

in which, affine specifications for the drift \( \mu_X(X_t) \) and volatility \( \sigma_X(X_t) \) are assumed as follows

\[
   \mu_X(X_t) = k(\theta - X) \\
   \sigma_X(X_t) = \Sigma S(X_t)
\]

where \( S(X_t) \) is a diagonal \( K \times K \) matrix with \( S_i(X_t) = \sqrt{s_{0i} + s_{1i}X} \). Also, \( s_{0i} \in \mathbb{R}, \theta, s_{1i} \in \mathbb{R}^K \), and \( k, \Sigma \in \mathbb{R}^{K \times K} \) are constants.

The intuitive concept of the mean reverting drift term \( \mu_X(X_t) \) is the following: If the current state \( X_t \) is above its long run mean \( \theta \) the change of \( dX_t \) is likely to be negative when the speed of mean reversion parameter is strictly positive \( k > 0 \). Similarly, if the current state \( X_t \) is below its long run mean \( \theta \) the change of \( dX_t \) is likely to be positive. In both cases, the process \( X_t \) is pulled back to its mean with strictly positive speed of mean reversion \( k \). If the speed of mean reversion is zero, \( k = 0 \), the process \( X_t \) is nonstationary.

Brownian motion shocks \( dW_t \) disturb \( X_t \), by moving it from its mean \( \theta \). These are normally distributed with zero mean and variance \( dt \). The effect of these shocks \( dW_t \) on \( X_t \) is determined by the functional form of the volatility. The best known examples of affine diffusions are Gaussian processes and square-root processes. Their difference lies into the volatility term. Gaussian processes imply \( S_i(X_t) = \sqrt{s_{0i}} \), with \( s_{0i} = 1 \), and thus \( \sigma_X(X_t) = \Sigma \), while square-root
1.2. SETUP OF AFFINE TERM STRUCTURE MODELS

processes allow $\sigma_X(X_t)$ to depend on the state, $S_i(X_t) = \sqrt{s_i^0} X$, with $s_{i1} = 1$.

Under the affine assumptions made on the short rate and vector process $X_t$ (see 1.15), Duffie and Kan [53] guess the following solution of $B_t(\tau)$ for PDE (1.5) under risk-neutrality:

$$P(X, \tau) = \exp(A(\tau) + D(\tau)' X)$$  \hspace{1cm} (1.16)

where the coefficients $A(\tau) \in \mathbb{R}$ and $D(\tau) \in \mathbb{R}^K$ solve following the system of ODEs

$$\frac{dA(\tau)}{d\tau} = -\delta_0 + D(\tau)' k \theta + \frac{1}{2} \sum_{i=1}^{K} [D(\tau)' \Sigma_i^0] s_{0i}$$

$$\frac{dD(\tau)}{d\tau} = -\delta_1 - k' D(\tau) + \frac{1}{2} \sum_{i=1}^{K} [D(\tau)' \Sigma_i^0]^2 s_{1i}$$

with initial conditions $A(0) = 0$, $D(0) = 0$. Given the exponential affine form in (1.16) the instantaneous bond return from (1.9) is

$$\mu_P(X, \tau) = -\frac{dA(\tau)}{d\tau} - \frac{dD(\tau)}{d\tau}' X + D(\tau)' \mu_X(X_t) + \frac{1}{2} D(\tau)' \sigma_X(X_t) \sigma_X(X_t)' D(\tau)$$  \hspace{1cm} (1.17)

In closed form solutions the coefficients $A(\tau)$ and $D(\tau)$ can be found in Vasicek [109] and CIR [32] for the one-factor Gaussian and square-root process respectively.\(^2\) Equation (1.16) implies that yields are given by

$$R_t(\tau) = -(1/\tau) \log P(X, \tau) = -(1/\tau) [A(\tau) + D(\tau)' X]$$  \hspace{1cm} (1.18)

On the other hand, when the two probability measures $\mathcal{P}$ and $\mathcal{Q}$ differ, the process $X$ obeys (1.7), with the affine specifications of drift $\mu_X(X_t)$:

$$\mu_X^\mathcal{Q}(X_t) = k^\mathcal{Q} (\theta^\mathcal{Q} - X)$$

and volatility $\sigma_X(X_t)$

$$\sigma_X^\mathcal{Q}(X_t) = \Sigma^\mathcal{Q} S^\mathcal{Q}(X_t),$$

where, as before, $S^\mathcal{Q}(X_t)$ is a diagonal $K \times K$ matrix with $S_i^\mathcal{Q}(X_t) = \sqrt{s_i^\mathcal{Q}} X$, $s_{0i}^\mathcal{Q} \in \mathbb{R}$, $\theta^\mathcal{Q}, s_{1i}^\mathcal{Q} \in \mathbb{R}^K$, and $k^\mathcal{Q}, \Sigma^\mathcal{Q} \in \mathbb{R}^{K \times K}$ are constants. Similarly, in order to obtain affine solutions for bond prices, the drift $\mu_X^\mathcal{Q}(X_t)$ and the variance matrix $\sigma_X^\mathcal{Q}(X_t) \sigma_X^\mathcal{Q}(X_t)'$ must also be affine.

\(^2\)Bolder [19] provides a detailed review of these models including their multifactor extensions.
in \( X_t \). The volatility term under measure \( \mathcal{P} \), needs to be affine, as shown in (1.12). But, the drift of \( X_t \) under \( \mathcal{P} \) may be non-linear in \( X_t \), since it depends on the functional form of \( \Lambda(X_t) \) (see (1.13)).

The ODEs for the bond price coefficients \( A(\tau) \in \mathbb{R} \) and \( D(\tau) \in \mathbb{R}^K \) under measure \( \mathcal{Q} \) now become:

\[
\frac{dA(\tau)}{d\tau} = -\delta_0 + D(\tau)'kQ\theta^Q + \frac{1}{2} \sum_{i=1}^{K} [D(\tau)'\Sigma]^2 \sigma_{i1i} \\
\frac{dD(\tau)}{d\tau} = -\delta_1 - kQ'D(\tau) + \frac{1}{2} \sum_{i=1}^{K} [D(\tau)'\Sigma]^2 \sigma_{i1i}
\]

where the risk-neutral long-run mean and speed of mean reversion parameters replace their data-generating counterparts. Finally, equation (1.17) also holds under the risk-neutral measure \( \mathcal{Q} \):

\[
\mu_P^Q(X, \tau) = -\frac{dA(\tau)}{d\tau} \quad \text{and} \quad \frac{dD(\tau)}{d\tau} = D(\tau)'X + D(\tau)'\mu^Q_X(X_t) + \frac{1}{2} D(\tau)'\sigma^Q_X(X_t)\sigma^Q_X(X_t)'D(\tau) \quad (1.19)
\]

### 1.3 Term premia

Under risk-neutral measure \( \mathcal{Q} \), we have \( E^Q_t \left[ \frac{dB_t(\tau)}{B_t(\tau)} \right] / dt = \mu_P^Q(X_t, \tau) \), where \( \mu_P^Q(X_t) = r_t \). The term (or risk) premium is defined as the difference between the expected return of the bond under the physical measure \( \mathcal{P} \), \( E^P_t \left[ \frac{dB_t(\tau)}{B_t(\tau)} \right] / dt = \mu_P(X_t, \tau) \) minus the risk-free rate \( r_t \). Substracting bond returns under \( \mathcal{P} \) and \( \mathcal{Q} \) measures yields the risk premium:

\[
E_t \left[ \frac{dB_t(\tau)}{B_t(\tau)} \right] / dt - E^Q_t \left[ \frac{dB_t(\tau)}{B_t(\tau)} \right] / dt = \mu_P(X_t, \tau) - r = D(\tau)'(\mu_X(X_t) - \mu_X^Q(X_t)) \quad (1.20)
\]

or by using (1.13)

\[
\mu_P(X_t) - r = -D(\tau)'\sigma_X(X_t)\Lambda(X_t)' \\
\mu_X(X_t) - \mu_X^Q(X_t) = \sigma_X(X_t)\Lambda(X_t)' \quad (1.21)
\]

where \( \mu_X(X_t) - \mu_X^Q(X_t) = \sigma_X(X_t)\Lambda(X_t)' \) or

\[
\mathcal{J}(\tau) \equiv -D(\tau)'\sigma_X(X_t)\Lambda(X_t)' \quad (1.22)
\]

Equation (1.22) gives the closed form formula of the term premia on holding bonds. It states that \( \mathcal{J}(\tau) \) are determined by the covariance of bond returns with the SDF \( \Lambda(X_t) \). In particular, \( \Lambda(X_t) \) contains the market prices of risk for each Brownian motion and expresses the familiar risk-return trade-off relationships. \( D(\tau)'\sigma_X(X_t) \) is the bond’s volatility. The market price of
1.3. TERM PREMIA

risk $\Lambda(X_t)$ describe how the nominal SDF (or the Marginal Rate of Substitution (MRS) for real bonds) responds to the underlying factor shocks and reflects the market values and respond to these shocks. Any asset whose payoffs are driven by shocks $\sigma(\cdot)dW$ must offer a price for the extra risk that investors get by holding them. From a closer view of equation (1.22), we can see that the dynamic properties of excess returns are influenced by three quantities. First, by $\sigma_X(X_t) = \Sigma S(X_t)$, which means that excess returns are influenced by the degree of correlations between the factors, i.e., $\Sigma$, and the magnitude of the factor volatilities $S(X_t)$. Second, by the factor loadings $D(\tau)$ that describe the factor dynamics and, third, by the interaction effects between the factor volatilities $S(X_t)$ and prices of risk $\Lambda(X_t)$. The richer the specification of factor volatilities, the less flexibility there is left in specifying the prices of risk (Dai and Singleton [38]).

Before proceeding into multifactor specifications of $\Lambda(X_t)$, in Table 1.1 we present some of the most popular single-factor term-structure models in the literature. The table also presents the 'traditional' assumptions on the function $\Lambda(X_t)$. The analysis that follows Table 1.1 is focused on how the term premium $\mathcal{J}(\tau)$ is specified in recent multifactor term structure literature. This will make clear how crucial and important is the specification of the market price of risk function $\Lambda(X_t)$ in adding the appropriate flexibility needed to capture stylized facts of bonds' excess returns. As Table 1.1 shows, all models with zero MPR are also defined under the risk-neutral measure $\mathcal{Q}$ implying that there are no term premia affecting yields. The local EH

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Short Rate process</th>
<th>MPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merton</td>
<td>$dX = \mu_r(r)dt + \sigma_r(r)dW_t$</td>
<td>$\Lambda(r)$</td>
</tr>
<tr>
<td>Vasicek</td>
<td>$dr = \theta dt + \Sigma dW$</td>
<td>$\ell_0$ (const.)</td>
</tr>
<tr>
<td>CIR</td>
<td>$dr = k(\theta - r)dt + \Sigma \sqrt{r}dW$</td>
<td>$\ell_0\sqrt{r}$</td>
</tr>
<tr>
<td>Dothan</td>
<td>$dr = kr dt + \Sigma rdW$</td>
<td>$\ell_0$ (const.)</td>
</tr>
<tr>
<td>Duffie &amp; Kan</td>
<td>$dr = k(\theta - r)dt + \Sigma \sqrt{s_0 + s_1 \tau^2}dW$</td>
<td>$\ell_0 \sqrt{s_0 + s_1 \tau}$</td>
</tr>
<tr>
<td>Hull &amp; White (ext. Vasicek)</td>
<td>$dr = (\varphi(t) - kr)dt + \Sigma dW$</td>
<td>0</td>
</tr>
<tr>
<td>Hull &amp; White (ext. CIR)</td>
<td>$dr = (\varphi(t) - r)dt + \Sigma \sqrt{r}dW$</td>
<td>0</td>
</tr>
</tbody>
</table>
thus holds in these models. The volatility of the short rate $\sigma_r(r)$ describes the shocks of a single factor that drives the term structure. Almost all the above models assume that MPR is constant, with parameter $\ell_0$, except for the CIR and DK models. The latter assume that MPR is proportional to the factor volatility.

Duffie and Kan [53] nest the Vasicek and CIR models given in Table 1.1, as well as some of their affine extensions within a general class of multifactor affine term structure models. This general class of affine models are classified in Dai and Singleton [37] according to the number $n = 1, 2, \ldots, K$ of processes of vector $X_t$ that enter the factor volatility, i.e.

$$S(X_t) = \sqrt{s_{0i} + s_{1i}'X_t},$$

where $X_t$ is of dimension $K \times 1$. Dai and Singleton [37] define as $A_n(K)$ a model with $K$ total factors from which $n = 1, 2, \ldots, K$ enter the volatility matrix. According to this notation, the single factor Vasicek and CIR models are denoted as $A_0(1)$ and $A_1(1)$ respectively, while $A_0(K)$ and $A_K(K)$ are their multifactor extensions.

In particular, the multifactor extension of Vasicek model\(^3\) guess the following solution of the bond price $B_t(\tau)$ satisfying PDE (1.5)

$$P(X_t, \tau) = \exp(A(\tau) - D(\tau)'X_t),$$

where $X_t$ is a vector of Gaussian processes,

$$\sigma_X(X_t) = \Sigma S(X_t) = \Sigma$$

and MPR vector $\Lambda(X_t)$ independent from $X_t$, i.e.

$$\Lambda(X_t) = \ell_0$$

By equation (1.13), the Vasicek’s model implies that the speed of mean reversion is the same under both probability measures, $\mathcal{P}$ and $\mathcal{Q}$, i.e.

$$k = k^\mathcal{Q}.\footnote{Vasicek’s model multifactor extension is given by Langetieg [80].}$$
1.3. TERM PREMIA

Under these two measures, only the long run mean differs, i.e.

$$\theta^Q = \theta - k^{-1}\Sigma\ell_0.$$  

Note that parameters $\theta^Q$ and $\theta$ differ only when market prices of risk $\ell_0$ differ from zero.

The CIR model guess the following solution of PDE (1.5):

$$P(X_t, \tau) = \exp(A(\tau) - D(\tau)'X_t),$$

and assumes a square-root process for the vector of factors $X_t$

$$\sigma_X(X_t) = \Sigma S(X_t) = \Sigma \sqrt{X_t} \quad (1.25)$$

MPR is given as

$$\Lambda(X_t) = \ell_0 \sqrt{X_t}. \quad (1.26)$$

Here, the change of measure affects the long run mean and the mean reversion parameter. A negative $\ell_0$ implies that under the risk neutral measure $Q$, $X_t$ mean reverts more slowly

$$k > k^Q$$

to a higher long-run mean

$$\theta > \theta^Q$$

Dai and Singleton [37] classify all the possible combinations between the physical and risk-neutral measures of $k$ and $\theta$, by guessing the following solution of PDE in (1.5):

$$P(X, \tau) = \exp(A(\tau) - D(\tau)'X)$$

the volatility of the factors as

$$\sigma_X(X_t) = \Sigma S(X_t) \quad (1.27)$$

and

$$\Lambda(X_t) = S(X_t)\ell_0 \quad (1.28)$$

for a constant vector $\ell_0 \in \mathbb{R}^K$. The vector of state variables is affine under both measures $\mathcal{P}$.
and $Q$ and their drift and long-run mean parameters relate to each other as follows:

$$k^Q = k + \Sigma \Phi$$

and

$$\theta^Q = k^{Q-1}(k\theta - \Sigma \psi)$$

in which $\Phi \in \mathbb{R}^{K \times K}$ and $\psi \in \mathbb{R}^K$ with the $i$th row of $\Phi$ given as $\ell_{0i}s'_{1i}$ and the $i$th row of $\psi$ given as $\ell_{0i}q_{0i}$.

### 1.3.1 Empirical evidence from bond returns

The question that arises after the above analysis is whether we need that extra flexibility in specifying the prices of risk $\Lambda(X_t)$. The recent literature presents significant evidence that this is needed. Expected excess returns display rich variation patterns across time and maturities (see e.g., Cochrane and Piazzesi [34], Duffee [51], Dai and Singleton [38]). Also, they switch signs over time. When the term structure is upward sloping, they tend to be positive and when it is downward sloping, they become negative.

Empirically, various studies using term spread regression models like

$$\sum_{i=1}^{n-1} (1 - i/n) \Delta r_{t+i} = a(n) + \beta(n) [R_t(n) - r_t] + u_t(n)$$

(1.29)

and

$$(n - 1) [R_{t+1}(n-1) - R_t(n)] = a(n) + \beta(n) [R_t(n) - r_t] + u_{t+1}(n)$$

(1.30)

find that although the forecasts of future cumulative short rates in (1.29) is in the right direction, the term spread between long and short term interest rates in (1.30) fails to forecast without bias future one-period ahead changes of long term rates, contrary to the predictions of the expectations hypothesis that $\beta(n) = 1$ (see e.g., Fama and Bliss [59], Campbell and Shiller [24], Hardouvelis [70], Cuthbertson [36], Tzavalis and Wickens [108], inter alia).

This puzzling behavior of the term spread can be attributed to the existence of a time-varying risk premium $J(\tau)$ as described above, which investors require as compensation for holding a long term bonds over a short period of time. Tzavalis and Wickens [108] find compelling evidence from yield regressions for such time varying risk premia. To explain the effects of
1.3. TERM PREMIA

these premia on the term spread based tests of the rational expectations hypothesis of the term structure, earlier studies in the literature consider linear factor term premium models (see, e.g., Simon [103], and Tzavalis and Wickens [108]), while more recent studies (see, e.g., Roberds and Whiteman [101], Dai and Singleton [38], Duffee [51]), rely on term premium models implied by affine dynamic term structure models. These more recent studies are mainly interested in investigating if affine dynamic models are consistent with the stylized facts of the term spread regressions mentioned above.

1.3.2 Extended models of risk premia

Evidence in empirical literature indicate that a rich structure can be adopted to capture risk premium effects. The sufficient condition for the market price of risk $\Lambda(X_t)$ to be ‘acceptable’ in the affine framework is that it must satisfy the Novikov’s condition, mentioned before (see (1.11)).

Duffee [51] and Dai and Singleton [38] also have shown that, in order to capture the observed predictability in bond returns, it is necessary to allow for more flexible risk-premia functions $\Lambda(X_t)$. According to Duffee [51], affine term structure models fail to replicate the key empirical relation between expected returns and the spread or slope of the yield curve. They underestimate the expected excess returns of long bonds. This underestimation is largest when the slope of the term structure is steep.

The completely affine models of Dai and Singleton [37], described above, impose the restriction that the compensation for risk is a fix multiple of the variance of the risk. This structure ensures that these models satisfy the no arbitrage requirement. However, this structure restricts the variability of the compensation that investors expect to receive as a reward for bearing a given risk. The compensation is bounded by zero. Therefore, it cannot change sign over time, as evidence suggest.

As is mentioned above, this restriction in the price of risk implies an *unrealistic* behavior of bonds’ excess returns. Duffee [51] breaks the link between the price of risk function and interest rate volatility, by making an extension from the completely affine class to the *essentially* affine. In this extension, the affine time series and cross sectional properties of bond prices are preserved.
CHAPTER 1. AFFINE TERM STRUCTURE MODELS: LITERATURE REVIEW

In particular, Duffee [51] sets the risk premium $J(\tau)$, for all maturity intervals $\tau$ as follows:

$$J(\tau) \equiv -D(\tau)'\sigma_X(X_t)\Lambda(X_t),$$

with

$$\sigma_X(X_t) = \Sigma S(X_t),$$

where, the factor volatility matrix $S(X_t)$ is a diagonal matrix of dimension $K \times K$, i.e.,

$$S_i(X_t) = \sqrt{s_{0i} + s_{1i}'X_t},$$

in which $s_{0i} \in \mathbb{R}, s_{1i} \in \mathbb{R}^K$. The vector of MPR functions $\Lambda(X_t)$ changes to an essentially affine specification, as follows:

$$\Lambda(X_t) = S(X_t)\lambda_0 + S(X_t)^-\lambda_1X_t,$$

with $\lambda_0 \in \mathbb{R}^K, \lambda_1 \in \mathbb{R}^{K \times K}$ and $S(X_t)^-$ defined as,

$$S(X_t)^{-}_{ii} = \begin{cases} (s_{0i} + s_{1i}'X_t)^{-1/2}, & \text{if } \inf(s_{0i} + s_{1i}'X_t) > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Note that for $\lambda_1 = 0$, the vector of MPR functions $\Lambda(X_t)$ corresponds to standard risk-premium specifications, given by (1.28). The tight link between the price of risk vector and the volatility matrix now is broken, since (1.31) allows for independent variation in prices of risk due to their link with the state variables. This is the kind of flexibility needed to fit the empirical behavior of expected excess returns to bonds (see (1.29) and (1.30)).

Dai and Singleton [38] make use of the above essentially affine specification of $\Lambda(X_t)$ to match the stylized facts of the term spread regressions (referred as Linear Projections of Yields (LPY)). Following their notation of term structure models, $A_n(K)$, admissibility requires that the richer the specification of factor volatilities $S(X_t)$, the less flexibility in specifying the prices of risk $\Lambda(X_t)$ there is. The most flexible model regarding the specification of risk premiums $\Lambda(X_t)$ is the Gaussian $A_0(K)$ models and the most constrained one is the square-root CIR $A_K(K)$ model. In the case that $n = 0$, i.e., the vector of Gaussian processes $X_t$ in the essentially affine framework implies $s_{0i} = 1$ and $s_{1i} = 0$, meaning homoscedastic factor volatility

$$\sigma_X(X_t) = \Sigma.$$
For this model, $\Lambda(X_t)$ becomes affine in $X_t$, i.e.,

$$\Lambda(X_t) = \lambda_0 + \lambda_1 X_t$$

with $\lambda_0 \in \mathbb{R}^K$ and $\lambda_1 \in \mathbb{R}^{K \times K}$. Note that in this specification, not only the long run mean, but also the speed of mean reversion parameters differ under the two probability measures $\mathcal{P}$ and $\mathcal{Q}$.

The richer the specification of factor volatilities, the less flexibility there is in specifying $\Lambda(X_t)$. Thus, in the case that all factors enter the volatility function $S(X_t)$, i.e., $n = K$, (like when $X_t$ follows the square-root processes) we have

$$\sigma_X(X_t) = \Sigma S(X_t).$$

Admissibility requires that, in (1.31), we must set $\lambda_1 = 0$ and the MPR is restricted to be proportional to factor volatility $S(X_t)$

$$\Lambda(X_t) = \ell_0 S(X_t)$$

Both of the above cases are specific cases of Duffee [51]. Dai and Singleton [38] show that the Gaussian sub-family of affine models, $A_0(K)$, for $K = 3$, give the requisite flexibility for matching LPY in (1.29) and (1.30) while resolves the Campbell – Shiller [24] rational expectations puzzle related to term spread regressions. On the other hand, the restricted CIR type of models perform relatively poorly (see also Roberds and Whiteman [101]). One could argue that the current workhorse in term structure modeling is the Gaussian essentially affine model, where the factors follow unrestricted Gaussian dynamics and bond risk premia are affine functions of the factors (see also Balduzzi and Chiang [11]).

### 1.4 Real bonds and general equilibrium models

Bond pricing equation based on the martingale approach (risk-neutral measure $\mathcal{Q}$) do not link the yield curve to economic fundamentals. In this section, we will try to make this link. Consider a $\tau$–period real zero coupon bond $B_t^* (\tau)$ which is priced at time $t$ and pays 1 unit of consumption good at maturity $t + \tau$. 
Suppose a representative agent with a separable utility function of the following form

\[ U(t) = E \int_{t=0}^{\infty} e^{-\delta t} u(c_t) dt, \]  

(1.32)

where \( c_t \) denotes consumption at date \( t \), and \( \delta \) is the rate of time preference. Assume that \( u(c_t) \) follows a power utility form

\[ u(c_t) = \frac{c_t^{1-a}}{1-a}, \]

where \(-c_t u''(c_t)/u'(c_t)\) is the risk aversion coefficient which equals the power \( a \).

Assume also that consumption process \( c_t \) is given by

\[ c_t = \exp(\gamma X_t), \]  

(1.33)

where \( X_t \) follows diffusion processes (see also (1.3))

\[ dX_t = \mu_X(X_t) dt + \sigma_X(X_t) dW_t \]

From (1.33) and (1.3), Ito’s Lemma implies

\[ \frac{dc_t}{c_t} = \mu_c(X_t) dt + \sigma_c(X_t) dW_t \]

where the instantaneous expected growth rate in consumption is:

\[ \mu_c(X_t) = \gamma' \mu_X(X_t) + \frac{1}{2} \gamma' \sigma_X(X_t) \sigma_X(X_t)' \gamma, \]

and the volatility of the consumption growth is given as:

\[ \sigma_c(X_t) = \gamma' \sigma_X(X_t). \]

In the equilibrium asset pricing framework, the representative agent smooths her/his intertemporal consumption by choosing a path between current and future consumption flows. The investor finances a purchase of \( \vartheta \) units of the zero coupon bond \( B^*_\tau(\tau) \) by reducing his current consumption from an initial level (endowment) \( \epsilon_t \), to \( c_t \), during a time interval \( dt \), i.e.

\[ c_t = \epsilon_t - \vartheta B^*_\tau(\tau)/dt \]

until (s)he reaches equilibrium in which her/his marginal loss in utility from buying another
zero coupon bond is equal to his marginal gain from it’s future payoff. Thus (s)he maximizes a utility function of the form (1.32). The F.O.C (First Order Conditions) of this maximization problem imply that the price of the $\tau$–period real zero coupon bond $B_t^*(\tau)$ is given as

$$B_t^*(\tau) = E_t \left[ e^{-\delta\tau - \frac{\mu(\tau)}{\mu(\tau)}} \right]$$

Define the continuous time real Stochastic Discount Factor $m_t$ as

$$m_t = e^{-\delta t} u'(c_t) = e^{-\delta t - \gamma'X}.$$ 

Then, the $\tau$–period real zero coupon bond price $B_t^*(\tau)$ can be written as:

$$B_t^*(\tau) = E_t \left[ \frac{m_{t+\tau}}{m_t} \right]$$ \hspace{1cm} (1.34)

Note that the interpretation of the real SDF in discrete time is different than that of the continuous-time. In discrete-time, SDF is the marginal rate of substitution in consumption between two time periods, while in continuous-time SDF refers to the level of marginal utility (See Cochrane [33]). In other words, it measures the growth in marginal utility, not total utility. A high SDF at $t + \tau$ means that you desperately need more consumption at $t + \tau$ and you are willing to sacrifice a lot of consumption in other periods in order to obtain it.

Since the SDF $m_t$ is a function of $X_t$, from Ito’s lemma we have

$$\frac{dm_t}{m_t} = \mu_m(X_t)dt + \sigma_m(X_t)dW_t$$

where drift $\mu_m(X_t)$ equals

$$\mu_m(X_t) = -\delta - \alpha'\mu_X(X_t) + \frac{1}{2} \alpha^2 \gamma' \sigma_X(X_t) \sigma_X(X_t)' \gamma,$$ \hspace{1cm} (1.35)

and volatility $\sigma_m(X_t)$ is given by

$$\sigma_m(X_t) = -\alpha' \sigma_X(X_t).$$

For $\tau \to 0$ equation (1.34) implies that risk-free rate $r_t$ is related to the SDF $m_t$ through

$$r_t dt = -E_t \left[ \frac{dm_t}{m_t} \right].$$
implying that
\[ r_t dt = -\mu_m(X_t) dt, \]
or
\[ \mu_m(X_t) = -r_t. \]
Using (1.35) the last relationship gives:
\[ r_t = \delta + a\gamma'\mu_X(X_t) - \frac{1}{2} a^2 \gamma' \sigma_X(X_t) \sigma_X(X_t)' \gamma. \]
The last relationship shows that, real interest rates are high when impatience \( \delta \) is high and when consumption growth is high; Higher relative risk aversion parameter, \( a \), makes interest rates more sensitive to consumption growth. The above analysis imply that real SDF process takes the following form:
\[
\frac{dm_t}{m_t} = -r_t dt + \sigma_m dW_t, \tag{1.36}
\]
For a risk-free world we have, \( \sigma_m = 0 \) and hence equation (1.36) transforms into
\[
\frac{dm_t}{m_t} = -r_t dt \tag{1.37}
\]
This implies that there is no volatility in the pricing kernel and thus the real discount factor is fully determined by the real short rate. Solving equation (1.37) we have
\[
m_t = m_0 \exp \left( -\int_0^t r_s ds \right)
\]
In similar fashion, solving equation (1.36):
\[
m_t = m_0 \exp \left( -\int_0^t (r_s + \frac{1}{2} \sigma_{m_s}^2) ds + \int_0^t \sigma_{m_s} dW_s \right) = m_0 \exp \left( -\int_0^t r_s ds \right) Z_t
\]
where
\[
Z_t = \exp \left( -\frac{1}{2} \int_0^t \sigma_{m_s}^2 ds + \int_0^t \sigma_{m_s} dW_s \right) \tag{1.38}
\]
Where \( Z_t \) in (1.38) is the martingale from which the risk-neutral measure \( Q \) is constructed and described in previous section (see equation (1.10)).
The last relationship implies that

\[
\frac{Z_{t+\tau}}{Z_t} = \frac{m_{t+\tau}}{m_t} \exp \left( \int_t^{t+\tau} r_s ds \right)
\]

Substituting the real SDF \( m_{t+\tau}/m_t \) into the real bond price (1.34), yields into

\[
B_t^*(\tau) = E_t \left[ \frac{m_{t+\tau}}{m_t} \right] = E_t \left[ \frac{Z_{t+\tau}}{Z_t} \exp \left( -\int_t^{t+\tau} r_s ds \right) \right] = E_t^Q \left[ \exp \left( -\int_t^{t+\tau} r_s ds \right) \right]
\]

in which \( E_t^Q \) is the expectation with respect to the risk-neutral probability process \( Q \). The market prices of risk are given by \( \sigma_m(X_t) = -a'\sigma_X(X_t) \), i.e., minus the volatility of consumption growth \( \sigma_c(X_t) = \gamma'\sigma_X(X_t) \) with relative risk aversion as the constant of proportionality. The larger are the shocks of the factors the higher is the volatility of consumption growth. This volatile consumption is multiplied with the investor’s aversion towards risk meaning that the higher is the coefficient the lower is the tolerance towards risk making investors to turn into safer choices such as bonds.

The above analysis shows how we can price real bonds \( B_t^*(\tau) \). To see how we can price nominal bonds, define as \( B_t(\tau) \) the price of a nominal bond that pays out 1 nominal unit at \( \tau \) periods from \( t \) is:

\[
B_t(\tau) = E_t \left( \frac{m_{t+\tau}}{m_t} \frac{P_t}{P_{t+\tau}} \right) = E_t \left( \frac{M_{t+\tau}}{M_t} \right), \tag{1.39}
\]

where \( M_t \) is the continuous time nominal SDF and \( P_t \) is the price process. Since real bonds \( B_t^*(\tau) \) can be thought as a nominal asset which pays realized inflation upon maturity, the real \( m_t \) and nominal \( M_t \) SDFs are linked through the relationship:

\[
M_t = m_t/P_t.
\]

Note that, the nominal bond price \( B_t(\tau) \) in (1.39), can be decomposed as:

\[
B_t(\tau) = E_t \left( \frac{m_{t+\tau}}{m_t} \frac{P_t}{P_{t+\tau}} \right)
= E_t \left( \frac{m_{t+\tau}}{m_t} \right) \times E_t \left( \frac{P_t}{P_{t+\tau}} \right) + \text{cov} \left( \frac{m_{t+\tau}}{m_t} ; \frac{P_t}{P_{t+\tau}} \right).
\]

The real payoff for an investor who holds a nominal bond maturing at a future period \( T = t + \tau \), depends on how inflation evolves over that period and he will require a premium in order to compensate for that unpredictable risk. The sign of that inflation premium depends on the
covariance between the growth in marginal utility $m_{t+\tau}/m_t$ and the (inverted) inflation rate $P_t/P_{t+\tau}$. Thus, it can be positive or negative implying high or low bond prices.

Finally, if we model the nominal SDF to follow a stochastic process of the form:

$$\frac{dM_t}{M_t} = -r_t dt - \Lambda(X_t)'dW_t,$$

by substituting into (1.39), then the risk-neutral price of the nominal zero-coupon bond, $B_t(\tau)$, is given as in (1.1), i.e.,

$$B_t(\tau) = E_t \left( \exp \left( - \int_t^{t+\tau} r_s + \frac{1}{2} \Lambda(X_s)'\Lambda(X_s)ds - \int_t^{t+\tau} \Lambda(X_s)'dW_s \right) \right)$$

or

$$B_t(\tau) = E_t^Q \left[ \exp \left( - \int_t^{t+\tau} r_s ds \right) \right].$$
Chapter 2

Term spread regressions of the rational expectations hypothesis of the term structure allowing for risk premium effects

2.1 Introduction

There is recently growing interest in explaining the empirical failure of the term spread between long and short-term interest rates to provide unbiased forecasts of future, one-period ahead changes of long-term rates, which is in contrast to the predictions of the expectations hypothesis of the term structure of interest rates (REHTS). See, e.g., Fama and Bliss [59], Campbell and Shiller [24], Hardouvelis [70], Cuthbertson [36], and Driffill et al. [49]. This puzzling behavior of the term spread can be attributed to the existence of a time-varying risk (or term) premium which is required by the bond market investors as a compensation for holding a long term bond over a short period, e.g., one-month. To explain the effects of a time-varying term premium on term spread regression based tests of the REHTS, earlier studies in the literature consider ad hoc linear factor specifications of the term premium (see, e.g., Simon [103], and Tzavalis and Wickens [108]), while more recent studies rely on term premium models implied by affine dynamic term structure models (DTSMs) (see, e.g., Roberds and Whiteman [101], Dai and Singleton [38], Duffee [51], or Hordahl et al [75]). The last category of studies is mainly interested in investigating if affine DTSMs are able to explain the above failures of the term spread. To show this, they calibrate term spread regressions using estimates of DTSMs based on a few interest rates series.
In this paper, we suggest a new empirical methodology to test the predictions of the term spread about future changes of long-term interest rates allowing for term premium effects. The latter are modelled through a simple and empirically tractable affine Gaussian DTSM (GDTSM), which can be estimated based on a very rich data set of interest rates. Our methodology enables us to test not only the dynamic predictions of the term spread implied by the REHTS, but also the cross-section restrictions on the term structure of interest rates implied by non profitable arbitrage conditions in the bond market. This can be done using the same econometric framework. Testing these restrictions is critical in examining if a time-varying term premium can explain the puzzling behavior of the term spread to forecast future interest rates. If they are not satisfied, then DTSMs can not be thought of as the correct economic framework to capture the effects of term premium on tests of the REHTS.

The GDTSM employed by the paper to carry out the above tests, assumes that interest rates are spanned by three common unobserved factors (see, e.g., Ahn [1] and Berardi [16]). This number of factors is chosen based on principal components (PC) analysis, which shows that three factors can explain almost 99.50% of the levels, or first-differences, of interest rates across a very wide spectrum of maturity intervals (see, e.g., Litterman and Scheinkman [81], and Ang and Piazzesi [3]). To retrieve estimates of the unobserved factors spanning the term structure of interest rates, we rely on the approach of Pearson and Sun [92], frequently used in practice (see, e.g. Brown and Schaefer ( [21],[22]), Duffee [52]). According to this approach, a number of zero-coupon bond interest rates are used as instruments to obtain the unobserved interest rates factors. This is done by inverting the pricing equations of zero-coupon bonds implied by DTSMs models. However, this approach relies on the assumption that these zero-coupon bond instruments are priced without measurement errors, which may not be true in practice given that long-term zero-coupon bonds are often calculated based on approximation methods. To overcome this measurement errors problem, instead of observed values, we suggest employing projections (forecasts) of the above interest rate instruments based on the common factors spanning the whole term structure of interest rates. These common factors can be retrieved by PC analysis. Since it is based on a very large set of different maturity interest rates, this analysis can provide term structure factors which constitute well diversified portfolios of interest rates. By construction, the estimates of these factors will be orthogonal to interest rate measurement
The paper provides a number of very useful results which have both academic and practical interest. It shows that adjusting term spread regressions for time-varying term premium effects can indeed provide unbiased forecasts of the future changes of long-term interest rates. These regressions can thus explain the forecasting failures of the term spread, mentioned before. The paper shows that these time-varying term premium effects can be sufficiently captured by the simple GDTSM suggested in the paper. The results of the paper show that this model satisfies the cross-section restrictions on interest rates implied by the no-arbitrage conditions of the bond market. Finally, the paper indicates that the time-varying term premium effects that are priced in the bond market are those associated with level and slope shifts of the term structure. The term premium effects which relate to curvature changes of the term structure are not found to be priced.

The paper is organized as follows. Section 2 presents the GDTSM and a term spread regression allowed for time-varying term premium effects. Section 3 estimates this model and carries out tests of its cross-section (no-arbitrage) restrictions. Then, it tests the REHTS based on the term spread regression adjusted for term premium effects. Finally, Section 4 concludes the paper.

2.2 Model set up

Consider an affine Gaussian Dynamic Term Structure Model (GDTSM) which assumes that the underlying unobserved common factors spanning the term structure of interest rates, denoted as $x_{it}$ for $i = 1, 2, ..., K$, obey the following continuous-time stochastic processes:

$$dx_{it} = k_i(\theta_i - x_{it})dt + \sigma_i dW_{it}, \quad i = 1, 2, ..., K$$

---

1 This is a very difficult computational task for other methods retrieving common interest rates factors from the data like the maximum likelihood or semiparametric latent variables methods (see, e.g., Chen and Scott ([28], [29])). Thus, in practice, these methods rely on very small cross-section sets of interest rates.
CHAPTER 2. TERM SPREAD REGRESSIONS OF THE REHTS ALLOWING FOR RP EFFECTS

where \( \theta_i \) is the long-run mean of \( x_{it} \), \( k_i \) is its mean reversion parameter, \( \sigma_i \) is its volatility parameter and \( W_{it} \) is a Wiener process \(^2\). The pricing kernel, used to price all bonds in the economy, is given as

\[
\frac{dM_t}{M_t} = -r_t dt - \Lambda'_t dW_t,
\]

(2.2)

where \( r_t \) is the instantaneous interest rate, \( W_t \) is a \((KX1)\)-dimension vector of Wiener processes \( W_{it} \) and \( \Lambda_t = (\Lambda_{1t}, \Lambda_{2t}, ..., \Lambda_{Kt})' \) a \((KX1)\)-dimension vector consisting of the market price of risk functions, given as

\[
\Lambda_{it} = \sigma_i^{-1}(\lambda^{(0)}_i + \lambda^{(1)}_i x_{it}), \text{ for all } i.
\]

(2.3)

Ruling out profitable arbitrage conditions in the bond market, the above GDTSM predicts that the price of a zero-coupon bond at time \( t \) with maturity interval \( \tau \), denoted as \( B_t(\tau) \), and its associated interest rate, denoted as \( R_t(\tau) \), have the following \( K \)-factor representation:

\[
B_t(\tau) = e^{-A(\tau) - D(\tau)'X_t}
\]

(2.4)

and

\[
R_t(\tau) = (1/\tau)[A(\tau) + D(\tau)'X_t],
\]

(2.5)

respectively, where \( X_t = (x_{1t}, x_{2t}, ..., x_{Kt})' \) is a \((KX1)\)-dimension vector which collects the unobserved factors \( x_{it} \), \( A(\tau) \) is a scalar function and \( D(\tau) \) is a \((KX1)\)-dimension vector of valued functions, i.e. \( D(\tau) = (D_1(\tau), D_2(\tau), ..., D_K(\tau))' \). For \( \tau \to 0 \), function (2.6) gives the instantaneous interest rate \( r_t \) as

\[
r_t = \delta_0 + \delta_1 X_t
\]

(2.6)

For \( \tau = \tau_1, \tau_2, ..., \tau_N \) maturity intervals, functions \( A(\tau) \) and \( D(\tau) \) are given as the solutions to a set of ordinary differential equations (Duffie and Kan [53]). Recursive relations of these functions can be found in Dai and Singleton [38] or Kim and Orphanides [78]. Given our model assumptions, these relations become analytic, given as follows:

\[
D_i(\tau) = (1 - e^{-k_i^Q \tau})(k_i^Q)^{-1} \delta_1,
\]

\(^2\)As our empirical analysis will show later on, this model fits satisfactorily into the data. It can be thought of as an extension of Vasicek’s [109] term structure model, which constitutes a special case of the general class of affine term structure (ATS) models of interest rates suggested by Duffie and Kan [53]. See also Dai and Singleton [37], Ahn [1], Kim and Orphanides [78], and Berardi [16]).
2.2. MODEL SET UP

with

\[ k_i^Q = k_i + \lambda_i^{(1)}, \]

where \( k_i^Q \) constitutes a risk-neutral measure of the mean reversion parameter \( k_i \) (see, e.g., Dai and Singleton [38]).

The above GDTSM is quite flexible and allows for extra variation in the prices of risk functions \( \Lambda_{it} \). As noted by Duffee [51], or Duarte [50], this variation is necessary in order to explain time-variability in the instantaneous expected excess holding period return of a \( \tau \)-period bond over interest rate \( r_t \), defined as

\[ J_t(\tau) = E_t h_{t+1}(\tau) - r_t = -D(\tau)'\Sigma \Lambda_t, \tag{2.7} \]

where \( \Sigma \) is a \((K \times K)\)-dimension diagonal matrix which consists of the volatility parameters \( \sigma_i \).

This excess return is defined as the term premium for holding a \( \tau \)-period bond over a short period. Substituting (3.7) into (2.7), term premium function \( J_t(\tau) \) can be analytically written as

\[ J_t(\tau) = -\sum_{i=1}^{K} D_i(\tau)\lambda_i^{(0)} - \sum_{i=1}^{K} D_i(\tau)\lambda_i^{(1)}x_{it}. \tag{2.8} \]

Considering the instantaneous rate \( r_t \) as the one-period (e.g. a month) to maturity interest rate and assuming continuously compounded interest rates, implying \( R_t(\tau) = -\frac{1}{\tau} \log B_t(\tau) \), equation (2.7) implies that the term premium, \( J_t(\tau) \), can be written in discrete time as follows:

\[ J_t(\tau) \equiv E_t \left[ \log \left( \frac{B_{t+1}(\tau - 1)}{B_t(\tau)} \right) \right] - r_t = -(\tau - 1)E_t [R_{t+1}(\tau - 1)] + \tau R_t(\tau) - r_t. \]

Rearranging terms, the last equation yields

\[ (\tau - 1)E_t [R_{t+1}(\tau - 1) - R_t(\tau)] = [R_t(\tau) - r_t] - J_t(\tau). \tag{2.9} \]

This equation clearly indicates that term spread \( R_t(\tau) - r_t \) based tests of the REHTS, using the following regression model:

\[ (\tau - 1)[R_{t+1}(\tau - 1) - R_t(\tau)] = a(\tau) + \beta(\tau)[R_t(\tau) - r_t] + u_{t+1}(\tau), \tag{2.10} \]
will lead to downward biased estimates of the slope coefficient $\beta(t)$ from unity, which is its predicted value by the theory. This can be attributed to the contemporaneous correlation between the term spread $R_t(t) - r_t$ and error term $u_{t+1}(t)$. As can be easily seen from relationships (3.7), (2.7) and (2.9), $u_{t+1}(t)$, in addition to expectation errors, defined as $v_{t+1}(t) = R_{t+1}(t) - E_t[R_{t+1}(t)]$, contains term premium $J_t(t)$, i.e. $u_{t+1}(t) = -[J_t(t) + v_{t+1}(t)]$. The latter has a common factor representation which is analogous to that of $R_t(t) - r_t$. The asymptotic bias of the least square (LS) estimator of $(\cdot)$, denoted as $\hat{\beta}(t)$, is given by the following formula:

$$Asybias \hat{\beta}(t) = 1 - \frac{Cov[R_t(t) - r_t; J_t(t)]}{Var[R_t(t) - r_t]}.$$ 

### 2.3 Empirical analysis

In this section, we estimate the affine GDTSM described by equations (2.1)-(2.8) based on a discrete-time econometric framework. Then, we carry out tests of the cross-section and time dimension predictions of this model to see if it is consistent with the data. The first category of these tests will examine if the no-arbitrage conditions of long-term interest rates $R_t(t)$ implied by the above GDTSM are consistent with the data, across different maturity intervals $t$. The second category will investigate if the term spread $R_t(t) - r_t$ adjusted by time-varying term premium effects can predict future changes in long-term interest rates, as predicted by the REHTS.

The section is organized as follows. First, we present the data and the empirical methodology that we follow to retrieve unobserved factors $x_{it}$ from the data. Second, we present unit root tests for interest rates $R_t(t)$, employed in our analysis. These are necessary before specifying the correct econometric framework for estimating and testing GDTSM (2.1)-(2.8). Evidence of unit roots in $R_t(t)$ immediately implies that there will be at least one common factor $x_{it}$ which will contain a unit root in its autoregressive component. This will necessitate appropriate transformations of all series involved in the structural equations of the above model before conducting any inference about its estimates and testing its theoretical predictions. Third, we carry out principal component (PC) analysis to retrieve to determine determine the maximum
number of factors spanning the term structure of interest rates and to study their futures. These should be analogous to those of unobserved factors $x_{it}$. Fourth, we estimate the GDTSM and test its cross-section restrictions. Finally, we carry out tests of the REHTS based on the term spread regressions allowing for time varying-premium effects.

2.3.1 Data description

Our empirical analysis is based on US zero-coupon interest rates series calculated by zero coupon or coupon-bearing bonds. These series cover the period from 1997:7 to 2009:10. They are obtained from the data archive of J. Huston McCulloch (Economics Department, Ohio University). They span a very large cross-section set of different maturity intervals $\tau$, from one month to forty years. Figure 1 presents two and three dimension plots of the above all interest rates series, $R_t(\tau)$, over the whole sample. As can be seen from these plots, interest rates exhibit substantial volatility over both the time and cross-section (maturity) dimensions of our data. This implies a high volatility of term premium $J_t(\tau)$, for all $\tau$, which can substantially obscure bond market’s expectations about future interest rates movements, and it thus makes the REHTS tests a very hard and challenging task.

![Figure 1. The US nominal term structure.](http://www.econ.ohio-state.edu/jhm/ts/ts.html)
2.3.2 Retrieving unobserved factors from interest rates data

To estimate the GDTSM, given by equations (2.1)-(2.8), and test its cross-section and REHTS predictions, we will rely on estimates of the vector of unobserved $K$-factors $X_t$ retrieved from the vector of observed interest rate series, denoted as $R_t = (R_t(\tau_1), R_t(\tau_2), ..., R_t(\tau_N))$, following Pearson’s and Sun [92] approach, denoted as P-S. This approach is modified to cope with the problem of measurement errors in interest rates (or bond) pricing equation (2.5), or (3.8). To explain this modification of the P-S approach, assume that equation (2.5) allows for measurement (or pricing) errors and write it in vector matrix format as follows:\(^4\)

$$R_t = A + DX_t + e_t, \quad (2.11)$$

where $A = ((1/\tau_1)A(\tau_1), (1/\tau_2)A(\tau_2), ..., (1/\tau_N)A(\tau_N))^\prime$ is a $(N \times 1)$-dimension vector of constants $A(\tau)$, $D = [(1/\tau_1)D(\tau_1), (1/\tau_2)D(\tau_2), ..., (1/\tau_N)D(\tau_N)]'$ is a $(N \times N)$-dimension matrix consisting of the loading coefficients of the elements of the vector of unobserved factors $X_t$ on those of vector $R_t$ and $e_t = (e_t(\tau_1), e_t(\tau_2), ..., e_t(\tau_N))^\prime$ is a $(N \times 1)$-dimension vector of $NIID(0, \Sigma_e)$ measurement errors, with variance-covariance matrix $\Sigma_e$. These errors are assumed to be independent of $X_t$.

P-S’ approach assumes that there is $(K \times 1)$-vector of observed interest rates, or transformations of them (collected in vector $Z_t$, with elements $z_{it}$), which are measured with no errors, i.e.,

$$Z_t = A_K + D_KX_t,$$

where $A_K$ and $D_K$ are appropriately defined sub-arrays of vector $A$ and matrix $D$. By inverting the last relationship, we can retrieve values of unobserved factors $X_t$ based on vector $Z_t$, i.e., $X_t = D_K^{-1}(Z_t - A_K)$. Then, equation (2.11) can be written as

$$R_t = A + DD_K^{-1}(Z_t - A_K) + e_t = A - DD_K^{-1}A_K + DD_K^{-1}Z_t + e_t. \quad (2.12)$$

The last relationship enables us to estimate GDTSM (2.1)-(2.8) and test its cross-sectional or REHTS predictions based on GMM or NLLS estimation procedures, which can be easily

\(^4\)See, e.g., Diebold et al [47].
applied. However, if $Z_t$ is measured with errors as someone expects to happen in practice, then this approach will lead to biased estimates of matrix $D$ or vector $A$. Thus, it will render tests of the GDTSM and REHTS biased. To see this more clearly, write vector $Z_t$ as

$$Z_t = Z_t^* + e_{K,t}, \quad \text{where} \quad Z_t^* = A_K + D_K X_t$$

is the component of $Z_t$ which is measured with no errors and $e_{K,t}$ is a $(K \times 1)$-dimension sub-vector of $e_t$. Then, equation (2.12) will be written as

$$R_t = A - DD_K^{-1} A_K + DD_K^{-1} Z_t + e_t - DD_K^{-1} e_{K,t}. \quad (2.13)$$

The last relationship clearly shows that estimating $D$ or $A$ based on a GMM or NLLS procedures will lead to biased estimates of them, since vectors of errors $e_{K,t}$ and $e_t$ are correlated with $Z_t$.

One way of overcoming the above estimation problem is to rely on linear projections of vector $Z_t$ on a set of instruments in (2.13), instead of using observed values of $Z_t$. These projections will be taken to be orthogonal to vector of errors $e_{K,t}$, or $e_t$. They can be obtained by regressing $Z_t$ on well diversified portfolios of interest rates $R_t(\tau)$, which are net of measurement errors. A natural choice of such portfolios can be taken to be the $K$ orthogonal principal components spanning the whole term structure of interest rates. These can be retrieved through principal component (PC) analysis, in first step.

Let us denote the $(K \times 1)$-dimension vector of the principal component factors spanning the term structure of interest rates $R_t(\tau)$, at time $t$ as $PC_t$, with elements $pc_{jt}$, for $j=1,2,...,K$. Then, it can be safely assumed that vector $PC_t$ is net of measurement (or pricing) errors, i.e.

$$PC_t = WR_t = WA + WDX_t + We_t = WA + WDX_t, \quad \text{with} \quad We_t = 0,$$

where $W$ is a $(K \times N)$-dimension matrix of the weights (loading coefficients) of interest rates $R_t(\tau)$ on principal component factors $pc_{jt}$. The assumption that $We_t=0$ means that vector $PC_t$ does not suffer from measurement errors $e_t$ (or $e_{K,t}$). This can be attributed to the fact that these errors are diversified away by forming principal component portfolios for a sufficiently large number of interest rates, $N$ (see, e.g. Joslin et al [77])\(^5\). The linear projections of $Z_t$ on

\(^5\)For large $N$, this means that $We_t$ is zero.
PC_t can be written as the following conditional mean:

\[ E(Z_t|PC_t) = G_0 + GPC_t, \]  

(2.14)

where \( E(e_t|PC_t) = 0 \) and \( E(e_{K,t}|PC_t) = 0 \). Rotation of the weights of \( PC_t \)'s or the estimates of the term-structure loading coefficients, collected in matrix \( D \), will not affect the estimates of conditional mean \( E(Z_t|PC_t) \), which can provide estimates (or forecasts) of vector \( Z^*_t \) net of measurement error effects. These can be employed to retrieve estimates of the vector of unobserved factors \( X_t \) from interest rates data, by inverting the following relationship:

\[ X_t = D_K^{-1}(E(Z_t|PC_t) - A_K). \]

### 2.3.3 Unit root tests

To test for unit root tests in interest rates \( R_t(\tau) \), we will carry out a second generation of ADF tests, known as efficient ADF (E-ADF) test (see, e.g., Elliott et al. [55], Elliott [54], and Ng and Perron [91]).6 These tests are designed to have maximum power against stationary alternatives which are local to unity. Thus, they can improve the power performance of the standard ADF statistic, often used in practice to test for a unit root in \( R_t(\tau) \).

Next, we carry out the E-ADF test suggested by Elliott et al. [55]. This requires to define first the quasi-differences of \( R_t(\tau) \) as follows:

\[ d(R_t(\tau)|\phi) = \begin{cases} R_t(\tau) & \text{if } t = 1 \\ R_t(\tau) - \phi R_{t-1}(\tau) & \text{if } t > 1, \end{cases} \]

where \( \phi \) denotes the local parameter of the alternative hypothesis against which the unit root hypothesis is tested. Then, we will estimate the following regression of the quasi-differenced interest rates series \( d(R_t(\tau)|\phi) \) on the quasi-differences of the vector of deterministic components

---

6Evidence provided in the literature on unit root tests of interest rates series is mixed. Earlier studies of this literature based on single time series unit root tests, such as the standard ADF test, can not reject the null hypothesis of a unit root (see, e.g., Hall et al. [67]). On the other hand, more recent studies based on panel data tests or Bayesian panel data methods tend to reject this hypothesis (see, e.g. Constantini and Lupi [35] and Meligotsidou et al [86]).
2.3. **EMPIRICAL ANALYSIS**

\( D_t = [1], \) or \( D_t = [1, t], \) denoted as \( d(D_t|\phi) \):

\[
d(R_t(\tau)|\phi) = d(D_t|\phi)'\delta(\phi) + \nu_t,
\]

based on the least squares (LS) estimation procedure. To estimate the last regression model, we need to fix a value for local parameter \( \phi \), i.e. \( \phi = \overline{\phi} \). Depending on the specification of \( D_t \), \( \phi \) must be fixed to the following values which maximize the local power of the E-ADF test:

\[
\overline{\phi} = \begin{cases} 
1 - 7/T & \text{if } D_t = [1] \\
1 - 13.5/T & \text{if } D_t = [1, t],
\end{cases}
\]

where \( T \) denotes the total number of time series observations of our sample. The estimates of slope coefficients of the last regression \( \delta(\phi) \), denoted as \( \hat{\delta}(\phi) \), will be used to detrend interest rates series \( R_t(\tau) \) as follows:

\[
R_t(\tau)^d \equiv R_t(\tau) - D_t\hat{\delta}_{\text{GLS}}(\overline{\phi})
\]

Then, the E-ADF unit root test can be carried out based on the following ADF auxiliary regression:

\[
\Delta R_t(\tau)^d = (\phi - 1)R_{t-1}(\tau)^d + \sum_{j=1}^{p} \vartheta_j \Delta R_{t-j}(\tau)^d + \varepsilon_t,
\]

(2.15)

based on detrended interest rates series \( R_t(\tau)^d \). This model allows for serial correlation of \( R_t(\tau)^d \) of maximum lag order \( p \). In particular, the unit root hypothesis can be tested by examining if the following hypothesis is true: \( 1 - \phi = 0 \) (or, \( \phi = 1 \)). This can be done based on the t-ratio test statistic. When \( D_t = [1] \), the asymptotic distribution of this statistic is the same to that of the standard ADF test. However, this is not true for the case that \( D_t = [1, t] \). For this case, critical values of the distribution of this statistic are provided by Elliott et al. \[55\].

In addition to the above efficient ADF unit root test, Elliott et al. \[55\] have also proposed another efficient unit root test statistic known as point optimal test. This test statistic is defined as follows:

\[
P_T = \frac{(SSR(\overline{\phi}) - \overline{\phi}SSR(1))}{\hat{\omega}^2},
\]

where \( SSR(\phi) = \sum \hat{\varepsilon}_t^2(\phi) \) is the sum of squared residuals \( \hat{\varepsilon}_t(\phi) = d(R_t(\tau)|\phi) - d(D_t|\phi)'\delta(\phi) \)

\[\text{See Elliott et al. (1996).}\]
and $\hat{\omega}^2$ is an estimator of the residual spectrum at frequency zero. Note that test statistic $P_T$ has the same asymptotic distribution as the E-ADF statistic, described before.

Table 1 reports the values of E-ADF and $P_T$ unit root test statistics for a set of different maturity interest rates $R_t(\tau)$ used in our empirical analysis, i.e., $\tau = \{5, 10, 15, 20, 25, 30\}$ years. To capture a possible linear deterministic trend $t$ in the levels of $R_t(\tau)$, occurring during our sample (see Figure 1), we assume that $D_t$ contains a deterministic trend, i.e. $D_t = [1, t]$. The results of the table clearly indicate that, despite the fact that the values of the autoregressive coefficients $\phi$ are very close to unity, the unit root hypothesis is rejected against its stationary alternative, for all maturity intervals $\tau$ considered. This is true at 5%, or 1% significance levels. The estimates of the autoregressive coefficient $\phi$ reported in the table imply that interest rates $R_t(\tau)$ exhibit a very slow mean reversion towards their long run mean, especially these of the shorter maturity intervals of 5 or 10 years. As noted before, evidence of stationarity of interest rates $R_t(\tau)$ implies that the common factors spanning interest rates $R_t(\tau)$ must be stationary, too. This will be also confirmed in the next section, based on our PC analysis.

Table 1: Efficient Unit Root Tests for Interest Rates

<table>
<thead>
<tr>
<th>$R_t(\tau)$, $\tau =$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi - 1$</td>
<td>-0.07</td>
<td>-0.14</td>
<td>-0.21</td>
<td>-0.27</td>
<td>-0.22</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.93</td>
<td>0.86</td>
<td>0.79</td>
<td>0.73</td>
<td>0.78</td>
<td>0.84</td>
</tr>
<tr>
<td>E-ADF</td>
<td>-2.35</td>
<td>-3.07*</td>
<td>-3.85**</td>
<td>-4.28**</td>
<td>-3.75**</td>
<td>-3.10*</td>
</tr>
<tr>
<td>$P_T$</td>
<td>5.11*</td>
<td>4.02*</td>
<td>2.37**</td>
<td>1.77**</td>
<td>2.57**</td>
<td>4.081*</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. $D_t$ is defined as $D_t=[1,t]$. The lag order $p$ of the auxiliary regressions used to carry out the unit root tests are chosen based on the AIC criterion. This lag found to be $p = 3$ for most maturity intervals $\tau$ considered (*) and (**) mean significance at 5% and 1% levels, respectively.

2.3.4 Principal component analysis

According to principal component (PC) analysis, the common principal component factors spanning the term structure of interest rates (or their first differences $\Delta R_t(\tau)$), denoted as $pc_{jt}$, for $j = 1, 2, \ldots, K$, can be obtained by the spectral decomposition of the variance-covariance matrix of vector of interest rates $R_t$. This can be done efficiently, based on a sufficiently large set of different maturity intervals $N$, where $N > K$. 
2.3. EMPIRICAL ANALYSIS

This matrix is defined as:

$$\Sigma_R = \Omega'\Theta\Omega,$$

where $\Theta$ is a diagonal matrix of dimension $(N \times N)$ whose elements are the eigen values of matrix $\Sigma_R$ and $\Omega$ is a $(N \times N)$-dimension orthogonal matrix whose columns are the eigen vectors corresponding to the eigen values of $\Sigma_R$. Given estimates of $\Omega$ and $\Theta$, the $(K \times 1)$-dimension vector of principal component series $PC_t = (pc_{1,t}, pc_{2,t}, ..., pc_{K,t})'$ can be retrieved from the $(N \times 1)$-dimension vector of interest rates series $R_t(\tau)$, denoted as $R_t$, as follows:

$$PC_t = \Omega'(R_t - \bar{R}),$$

where $\bar{R}$ is the sample mean of vector of interest rates series $R_t$.

Our PC analysis is based on a very large set of different maturity interest rates $R_t(\tau)$, i.e., $N = 468$. This covers maturity intervals $\tau$ from one to forty years. This set also contains the one-month interest rate $r_t$, which is considered in our analysis as the short-term interest rate. This large cross-sectional set of different maturity interest rates guarantees that the retrieved
PC\_jt factors from the data will diversify away any measurement (or pricing) errors in interest rates \(R_t(\tau)\), which is independent of unobserved factors \(x_{it}\)\(^8\). Figure 2 presents plots of the first three PC\_jt, for \(j = 1, 2, 3\), retrieved from our data. These correspond to the first three largest in magnitude eigen values of matrix \(\Sigma_R\), which are found to explain 99.50% (or 97.81%) of the total variation of the levels (or first differences) of all interest rates \(R_t(\tau)\), employed in our analysis. See the following table:\(^9\)

<table>
<thead>
<tr>
<th>Total Number of PCs</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>% variation explained in (\Delta R_t(\tau))</td>
<td>82.94</td>
<td>94.56</td>
<td>97.81</td>
</tr>
<tr>
<td>% variation explained in (R_t(\tau))</td>
<td>80.00</td>
<td>98.27</td>
<td>99.50</td>
</tr>
</tbody>
</table>

The above results of our PC analysis are consistent with those reported by other studies in the literature (see, e.g., Litterman and Scheinkman [81], or Bliss [18]). They show that three PC\_jt factors explain almost all of the variation of the term structure of interest rates. The first factor, PC\_1t, explains the largest part of this variation (e.g., 80% for the levels of \(R_t(\tau)\)). Together with the second factor, denoted as PC\_2t, they explain the 98.27% of this variation. The remaining percentage, which is very small, is explained by the third factor, PC\_3t. To interpret PC\_jt factors and study their stochastic properties, in Figure 3 we graphically present estimates of their loading coefficients on the first difference of interest rates series \(\Delta R_t(\tau)\), while in Tables 2A and 2B we report some useful descriptive statistics of them, as well as the E-ADF and P\_T unit root test statistics. Table 2A also reports estimates of the correlation coefficients of PC\_jt with observed interest rate variables (or transformations of them), often used in the literature as proxies of unobserved interest rates factors \(x_{it}\) (see, e.g., Ang and Piazzesi [3]). This set of variables include: the level interest rate with maturity \(\tau = 10\) years, defined as \(z_{1t} \equiv R_t(10)\), the term spread between the short and the long-term interest rates, defined as \(z_{2t} \equiv R_t - R_t(40)\), and,

\(^8\)As shown in Bai and Ng [9], and Bai [8] consistent estimates of common factors spanning economic series can be obtained by PC analysis when the following condition holds: \(\frac{\sqrt{T}}{N} \rightarrow 0\), where \(T\) is the total number of time series observations of each series and \(N\) is the cross-section dimension (here, maturity intervals \(\tau\)).

\(^9\)Note that these factors correspond to the first three largest in magnitude eigenvalues of variance-covariance matrix \(\Sigma\). The relative variation of the three factors is calculated as

\[
\frac{\sum_{i=1}^{3} \psi_i}{\text{tr}(\Sigma)},
\]

for \(i = 1, 2, 3\), where \(\psi_i\) is the eigen value of matrix \(\Sigma\) and \(\text{tr}(\cdot)\) stands for the trace of a matrix.
finally, variable $z_{3t} \equiv r_t - 2R_t(6) + R_t(40)$. Variables $z_{2t}$ and $z_{3t}$ constitute linear transformations of short-term interest rate $r_t$ and longer term interest rates. As can be seen from the results of Table 2A, there is almost one-to-one correspondence between variables $z_{jt}$ and series $pc_{jt}$. Since variables $z_{jt}$ are observed, in our analysis they will be taken to play the role of interest rate instruments. These will be projected on principal component factors $pc_{jt}$ to obtain estimates of $z_{jt}$ net of measurement errors.

Turning into a discussion on the interpretation of the three principle component factors $pc_{jt}$, the results of Tables 2A-2B and Figure 3 clearly indicate that these factors share similar features to those found in other empirical studies (see, e.g., Litterman and Scheinkman [81]). In particular, $pc_{1t}$ plays the role of a "level" factor, which causes almost parallel shifts to the whole maturity spectrum of interest rates. $pc_{2t}$, referred to as "slope" factor, determines the slope of the term structure, while $pc_{3t}$ constitutes a "curvature" factor since its loading coefficients on interest rates have a U-shape. Note that, in contrast to $pc_{1t}$ and $pc_{3t}$, the effect of $pc_{2t}$ on interest rates $R_t(\tau)$ is always declining and increases in terms of magnitude with maturity interval $\tau$. Since $pc_{2t}$ is positively and highly correlated with term spread $r_t - R_t(40)$, the last result means that a positive value of term spread $R_t(40) - r_t$ (i.e. a negative of $r_t - R_t(40)$) will have a positive effect on the slope of the term structure. For interest rates of maturity

![Figure 3. The loadings of the first three principal components.](image-url)
interval $\tau > 15$ years, this effect will be offset by the "curvature" effects, captured by $pc_{3t}$.

Finally note that, due to their high and positive correlation with principal component factors $pc_{jt}$, analogous interpretations to the above can be also given to observed variables $z_{jt}$, which are used as instruments to retrieve unobserved factors $x_{it}$ in our analysis.

Table 2A: Summary Statistics of Interest Rates PCs

<table>
<thead>
<tr>
<th>Factors</th>
<th>$pc_{1t}$</th>
<th>$pc_{2t}$</th>
<th>$pc_{3t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.027</td>
<td>-0.029</td>
<td>0.000</td>
</tr>
<tr>
<td>Max. Value</td>
<td>33.71</td>
<td>14.39</td>
<td>6.01</td>
</tr>
<tr>
<td>Min. Value</td>
<td>-43.30</td>
<td>-17.91</td>
<td>-5.74</td>
</tr>
<tr>
<td>Variance</td>
<td>268.79</td>
<td>61.27</td>
<td>4.09</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>0.93</td>
<td>0.95</td>
<td>0.74</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>0.87</td>
<td>0.92</td>
<td>0.48</td>
</tr>
<tr>
<td>$\rho_3$</td>
<td>0.83</td>
<td>0.91</td>
<td>0.40</td>
</tr>
<tr>
<td>$\rho_4$</td>
<td>0.78</td>
<td>0.87</td>
<td>0.28</td>
</tr>
<tr>
<td>$\rho_5$</td>
<td>0.74</td>
<td>0.84</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>$z_{1t}$</th>
<th>$z_{2t}$</th>
<th>$z_{3t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_{1t}$</td>
<td>0.98</td>
<td>0.84</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Notes: Max stands for maximum, while Min. for minimum. $\rho_s$ are the autocorrelations of PCs $pc_{jt}$, $j = 1, 2, 3$, of lag order $s = 1, 2, \ldots, 5$.

Table 2B: Unit Root Tests for Interest Rates PCs

<table>
<thead>
<tr>
<th>PCs:</th>
<th>$pc_{1t}$</th>
<th>$pc_{2t}$</th>
<th>$pc_{3t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1 - 1$</td>
<td>-0.15</td>
<td>-0.05</td>
<td>-0.20</td>
</tr>
<tr>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.05)</td>
<td></td>
</tr>
<tr>
<td>$\phi_i$</td>
<td>0.85</td>
<td>0.95</td>
<td>0.80</td>
</tr>
<tr>
<td>$p$</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>E-ADF</td>
<td>-3.43*</td>
<td>-2.04*</td>
<td>-3.46*</td>
</tr>
<tr>
<td>$P_T$</td>
<td>4.41*</td>
<td>1.96**</td>
<td>3.44**</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses, $D_t$ is defined as $D_t=[1,t]$. The tests reported in the table are based on E-ADF autoregression: $\Delta x_{it} = (\phi_1 - 1)x_{it-1} + \sum_{t=1}^{p} \vartheta y\Delta x_{it}(\tau) + \nu_{it}$. The lag order $p$ of the dynamic (first difference) terms of the E-ADF regressions chosen are based on the SIC criterion (except for $x_{2t}$). (*) and (**) mean significance at 5% and 1% levels.

Finally, the results of unit root tests reported in Table 2B indicate that all the three principal component factors $pc_{jt}$ constitute stationary series. These results are consistent with those on
interest rates series $R_t(\tau)$, reported in Table 1. From the econometric methodology point of view, they imply that the key equations of our term structure model (2.1), (2.5) and (2.8) can be estimated as a system of structural equations based on standard asymptotic estimation and inference procedures, as this system consists of stationary series. The estimates of the autoregressive coefficients $\phi$ reported in the table indicate that, among three factors $pc_{jt}$, the second (i.e. $pc_{2t}$), which is highly correlated with spread $r_t - R_t(40)$, is the most persistent one. This result can be also confirmed by the values of autocorrelation coefficients $\rho_s$, reported in Table 2A.

2.3.5 Estimation of the GDTSM and cross-section restriction tests

Having established good grounds to support that the three orthogonal and stationary factors explain almost all variation of the term structure of interest rates $R_t(\tau)$, in this section we estimate the structural equations of the GDTSM (2.1), (2.5) and (2.8). To retrieve estimates of the unobserved factors $x_{jt}$ from the data, we rely on our empirical methodology, presented in subsection 3.2. As vector of observed instruments $Z_t$, we will use the (3X1)-dimension vector $Z_t = (z_{1t}, z_{2t}, z_{3t})'$, where $z_{1t} \equiv R_t(10)$, $z_{2t} \equiv r_t - R_t(40)$ and $z_{3t} \equiv r_t - 2R_t(6) + R_t(40)$. As shown in the previous section, this vector consists of interest rate series or transformations of them which are highly correlated with the three principal component factors $pc_{jt}$, spanning the term structure of interest rates. Thus, the expectation of this vector conditional on that of three principal component series $PC_t = (pc_{1t}, pc_{2t}, pc_{3t})'$, i.e. $E(Z_t|PC_t) = G_0 + GPC_t$, which is employed to retrieve $X_t$ from the data through relationship $X_t = D_K^{-1}(E(Z_{t+1}|PC_{t+1}) - A_K)$, will be accurately estimated and net of measurement errors.

More specifically, the system of equations that is employed to estimate the model is given as follows:

$$R_{t+1} = A + DX_{t+1} + e_{t+1}, \quad \text{with} \quad X_{t+1} = D_3^{-1}(E(Z_{t+1}|PC_{t+1}) - A_3), \quad (2.16)$$

$$\Delta X_{t+1} = \Phi_0 + (\Phi - I)X_t + \omega_{t+1}, \quad (2.17)$$
CHAPTER 2. TERM SPREAD REGRESSIONS OF THE REHTS ALLOWING FOR RP EFFECTS

\[ Z_{t+1} = G_0 + GPC_{t+1} + \varepsilon_{t+1}, \quad \text{and} \]

\[ EHR_{t+1} = \Gamma_0 + \Gamma X_t + \eta_{t+1}, \]

where \( EHR_{t+1} \) is a \((N \times 1)\)-dimension vector of excess holding period returns \( h_{t+1}(\tau) - r_t \), for \( \tau = \tau_1, \tau_2, \ldots, \tau_N \), observed at time \( t + 1 \). Equation (2.17) gives the vector of discretized continuous-time processes (2.1), for \( i = 1, 2, 3 \), where \( \Phi \) is a \((3 \times 3)\) diagonal matrix whose elements are given as \( \Phi_{ii} = e^{-k_i \Delta} \). \( \varepsilon_{t+1}, \omega_{t+1}, \varepsilon_{t+1} \) and \( \eta_{t+1} \) constitute vectors of zero-mean error terms.

The above system of equations can be employed to retrieve from the data estimates of the mean reversion and price of risk premium parameters \( k_i \) and \( \lambda_i^{(1)} \), respectively. These are the key parameters of the GDTSM, which are important in forecasting future interest rate changes \( \Delta R_{t+1}(\tau) \) and in capturing time-varying risk premium effects \( J_t(\tau) \), respectively. To retrieve their estimates, we will impose the following cross-section restrictions on the coefficients of the above system implied by the no-arbitrage conditions of the GDTSM (see Section 2):

\[ D_i(\tau) = (1 - e^{-k_i^{Q\tau}})(k_i^{Q\tau})^{-1} \delta_1, \]

with \( k_i^Q = k_i + \lambda_i^{(1)} \), and \( \Gamma_i(\tau) = -\tau D_i(\tau) \lambda_i^{(1)} \), for \( \tau = \tau_1, \tau_2, \ldots, \tau_N \), where \( D_i(\tau) \) and \( \Gamma_i(\tau) \), \( i = 1, 2, 3 \), constitute elements of the row vectors of matrices \( D \) and \( \Gamma \), for rows \( \tau = \tau_1, \tau_2, \ldots, \tau_N \), and \( \delta_1 \) are the load coefficients of factors \( x_{it} \) on short-term interest rate \( r_t \), for \( i = 1, 2, 3 \). Note that the latter are the elements of the first row of matrix \( D \).

Compared to other econometric models used in the literature to estimate DTSMs (see, e.g., Dai and Singleton ([37], [38])), the above structural system of equations, in addition to equations (2.16) and (2.17), also includes excess holding period returns equations (2.19). These equations provides useful information about the factors determining time variation of price of

\^10 Following empirical literature (see, e.g., Dai and Singleton [38]), we do not impose any cross-section restriction on the intercepts of structural equations (2.16) and (2.19). These intercepts can also capture a fixed component of bond pricing errors or imperfections of the bond market.
risk functions \( \Lambda_{it} \). As can be seen from equation (2.8), this set of equations can help to identify the coefficients of price of risk functions \( \Lambda_{it} \lambda_{i}^{(1)} \), under the physical probability measure. They can be also used to estimate term premium \( J_{t}(\tau) \), at time \( t \), and, thus, to control its effects on term spread regressions (3.14).

The system of equations (2.16)-(2.19) is estimated subject to cross-section restrictions (2.20). This is done for the set of maturity intervals \( \tau = \{0,5,10,15,20,30,25,30\} \) years, where \( \tau = 0 \) stands for the short-term interest rate \( r_{t} \) maturity interval.\(^\text{11}\) The estimation of this system is carried out based on the Generalized Method of Moments (\( \text{GMM} \)), using as set of instruments current and past time values of the vector of observed variables \( Z_{t} \) (see bottom of Table 3). \( \text{GMM} \) can provide asymptotically efficient estimates of the vector of structural parameters of the system which are robust to possible heteroscedasticity and/or missing serial correlation of error terms vectors \( \epsilon_{t+1}, \omega_{t+1}, \varepsilon_{t+1} \) and \( \eta_{t+1} \). Furthermore, since the number of orthogonality conditions used in the \( \text{GMM} \) estimation procedure (i.e., the number of instruments multiplied by the number of the individual equations) is bigger than the number of the parameters estimated, the set of overidentified conditions implied in the estimation procedure can be employed to test if our GDTSM constitutes a correct specification of the data. To this end, we will employ Sargan’s overidentified restrictions test, denoted as \( J \). This is distributed as \( \chi^{2} \), with degrees of freedom, which equal the number of overidentified restrictions (i.e. 7X18-33=93, in our case). Table 3 presents the \( \text{GMM} \) estimation results of the key parameters of system (2.16)-(2.19) \( k_{i} \) and \( \lambda_{i}^{(1)} \), as well as the elements of projection matrix \( G \), denoted \( G_{ij} \). A number of interesting conclusions emerge from the results of the table. First, the GDTSM is found to be consistent with the data, and satisfies the no-arbitrage restrictions (2.20). This can be justified by the probability values (\( p \)-values) of \( J \) test statistic, reported in the table. The values of the correlation coefficients of future interest rates \( \Delta R_{t+1}(\tau) \) and excess returns \( h_{t+1}(\tau) - r_{t} \) with their predicted values, denoted as \( \text{corr}(\Delta R,\Delta \hat{R}) \) and \( \text{corr}(HR,H \hat{R}) \), indicate that the GDTSM has significant forecasting power on variables \( \Delta R_{t+1}(\tau) \) and \( h_{t+1}(\tau) - r_{t} \), for all \( \tau \). These variables are very difficult to forecast (see, e.g. Duffee, 51), Almeida and Vicente (2), or

\(^{11}\)Note that, for this set of maturity intervals \( \tau \), the biggest failures of the expectations hypothesis of the term structure are reported in the literature (see, e.g., Dai and Singleton, 38).
Carriero et al [25], more recently).

Table 3: GMM Estimates of system (2.16)-(2.19)

<table>
<thead>
<tr>
<th></th>
<th>$x_{1t}$</th>
<th>$x_{2t}$</th>
<th>$x_{3t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{i}$</td>
<td>0.32</td>
<td>0.22</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.07)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$\lambda_{i}^{(1)}$</td>
<td>-0.31</td>
<td>-0.12</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.06)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

Matrix $G = [G_{ij}]$

$G_{1j}$

<table>
<thead>
<tr>
<th></th>
<th>$G_{1j}$</th>
<th>$G_{2j}$</th>
<th>$G_{3j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.03</td>
<td>0.09</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.002)</td>
<td>(0.006)</td>
</tr>
<tr>
<td></td>
<td>-0.02</td>
<td>0.13</td>
<td>-0.68</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.005)</td>
<td>(0.01)</td>
</tr>
<tr>
<td></td>
<td>0.11</td>
<td>0.09</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.004)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>

$J(93) = 115.2$ (p-value=0.06),

Instruments: 1, $z_{1t}$, $z_{2t-1}$, $z_{2t-2}$, $z_{2t-3}$, $z_{2t-4}$ and $z_{3t}$

B: Correlation coefficients

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$corr(\Delta R, \Delta R)$</td>
<td>0.05</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
<td>0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>$corr(\hat{HR}, \hat{HR})$</td>
<td>0.13</td>
<td>0.13</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Notes: Data are monthly from 1997:07 to 2009:10. The Newey - West heteroscedasticity and autocorrelation consistent standard errors shown in parentheses. $J$ is Sargan’s overidentified restrictions test statistic. This is chi-squared distributed with degrees of freedom which are equal to the number of orthogonality conditions employed in the estimation procedure (7X18) minus the number of the parameters estimated (33). The instruments used are lags and current values of the following variables: $z_{1t} = R_{t}(10)$, $z_{2t} = r_{t} - R_{t}(40)$ and $z_{3t} = R_{t}(40) - 2R_{t}(6)$. $corr(.)$ are the correlation coefficients of the fitted values of the first difference of interest rates $\Delta R_{t+1}(\tau)$ and excess holding period returns $h_{t+1}(\tau) - r_{t}$, respectively denoted as $\Delta \hat{R}$ and $\hat{HR}$, with their observed values, denoted as $\Delta R$ and $h$, respectively.

Second, the estimates of mean reversion parameters $k_{i}$ are found to be significant, for all unobserved factors $x_{it}$. They imply very high persistency of all factor innovations, collected in vector $\omega_{t}$, on the level of interest rates $R_{t}(\tau)$, for all $\tau$. The implied by them estimates of autoregressive coefficients $\Phi_{ii}$ (i.e. the diagonal elements of matrix $\Phi$) are as follows: $\Phi_{11} =$
0.97, $\Phi_{22} = 0.98$ and $\Phi_{33} = 0.96$. These estimates are more close to unity than those of the autoregressive coefficients of principal component factors $pc_{jt}$, reported in Table 2B. This result means that there is no one-to-one correspondence between factors $x_{it}$ and $pc_{jt}$. This can be also confirmed by the inspection of plots of estimates of $x_{it}$, implied by the estimates of system (2.16)-(2.19). These are graphically presented in Figures 4A-C against those of factors $pc_{it}$. Their graphs indicate that the estimates of all three unobserved factors $x_{it}$ are more volatile than $pc_{jt}$. In fact, the estimates of $x_{it}$ seem to capture substantial and persistent shifts in the term structure of interest rates. These seem to occur in years 1998, 2003, 2005 and 2008, and they are associated with world financial crises or US monetary policy regime changes. The estimates of $pc_{jt}$ smooth out these shifts. These results imply that approximating unobserved factors $x_{it}$ with principal component factors $pc_{jt}$ may not provide accurate representations of the former, and thus may lead to inaccurate estimates of the GDTSM' parameters. An analogous conclusion can be also drawn for observed series $z_{it}$, used also in practice in place of $x_{it}$. The estimates of the elements of matrix $G$, $G_{ij}$, reported in the table, clearly show that there is no one-to-one correspondence between $z_{it}$ and $pc_{jt}$, or $x_{it}$.

Regarding the price of risk function parameters $\lambda_{i}^{(1)}$, the results of Table 3 indicate that the estimates of them associated with the first two factors $x_{1t}$ and $x_{2t}$ are significant at 5%, or 1%, level. The estimate of $\lambda_{3}^{(1)}$, associated with factor $x_{3t}$, is not significant. These results imply that only time-varying effects of factors $x_{1t}$ and $x_{2t}$ are priced in the bond market. According to relationship (2.8), these two factors are responsible for the time-variation of term premium $J_{t}(\tau)$. Thus, they can explain the bias of a time-varying term premium effects on the REHTS based on term spread $R_{t}(\tau) - r_{t}$. This result can be also confirmed by Figure 5, which graphically presents estimates of $J_{t}(\tau)$ against the observed values of excess returns $h_{t+1}(\tau) - r_{t}$, for $\tau = 5$ years. Inspection of this figure indicates that persistent shifts in term premium $J_{t}(\tau)$, like those of financial crisis of years 1998 or 2008 seem to be mainly associated with shifts in term structure factors $x_{1t}$ and $x_{2t}$ in these years. It must be noted at this point that the insufficiency of price of risk coefficient $\lambda_{3}^{(1)}$ does not mean that changes in factor $x_{3t}$ do not determine the curvature of the term structure interest rates $R_{t}(\tau)$. As can be seen by equation (2.20), the

\[ \text{See, e.g., Bansal and Zhou [12], Dai et al. [39], Cochrane and Piazzesi [34].} \]
\[ \text{Analogous graphs are taken for other maturity intervals } \tau. \]
curvature of $R_t(\tau)$ depends also the values of mean reversion parameters $k_i$, which is found to be different than zero for factor $x_{3t}$. Finally, note that both the sign and magnitude of the price of risk premium function coefficients $\lambda_1^{(1)}$ and $\lambda_2^{(1)}$ are consistent with that found by other studies (see, e.g., Duffie [52]).

Figure 4.A. Estimates of the first factor and the first principal component.

Figure 4.B. Estimates of the second factor and the second principal component.
2.3. EMPIRICAL ANALYSIS

2.3.6 Tests of the REHTS allowing for term premium effects

Having established good grounds to support that the GDTSM, given by equations (2.1)-(2.8), constitutes a correct specification of the data, in this section we formally examine if the term premium \( J_t(\tau) \) implied by this model can explain the failure of term spread \( R_t(\tau) - r_t \) to forecast future changes in long-term interest rates \( R_{t+1}(\tau - 1) - R_t(\tau) \), which is against the predictions of the REHTS. To this end, we will estimate the following term spread regression model:
Chapter 2. Term Spread Regressions of the REHTS Allowing for RP Effects

\[(\tau - 1)[R_{t+1}(\tau - 1) - R_t(\tau)] = a(\tau) + \beta(\tau)[R_t(\tau) - r_t] - J_t(\tau) + \nu_{t+1}(\tau) \quad \text{or} \]

\[(\tau - 1)[R_{t+1}(\tau - 1) - R_t(\tau)] = a(\tau) + \beta(\tau)[R_t(\tau) - r_t] + \sum_{i=1}^{K} \tau D_i(\tau) \lambda_i^{(1)} x_{it} + \nu_{t+1}(\tau) \quad (2.21)\]

where \(a^*(\tau) \equiv a(\tau) - \sum_{i=1}^{K} \tau D_i(\tau) \lambda_i^{(0)}\). This model is adjusted by the term premium effects \(J_t(\tau)\), given by equation (2.8). By including the time-varying component of term premium \(J_t(\tau)\), in the right hand side (RHS) of regression model (2.21), we can capture bias effects of \(J_t(\tau)\) on the slope of the term spread \(R_t(\tau) - r_t\), \(\beta(\tau)\), about future movements in \(R_{t+1}(\tau - 1) - R_t(\tau)\).\(^{14}\)

Regression model (2.21) is estimated jointly with the systems of equations (2.16)-(2.18), for all \(\tau\), based on the GMM procedure used before to obtain estimates of \(k_i\) and \(\lambda_i^{(1)}\). This model has replaced that of excess return equations (2.19) in our previous system of equations, (2.16)-(2.19). Regression model (2.21) enables us to estimate freely from the data the term spread slope parameter \(\beta(\tau)\) and, then, to examine if this coefficient is equal to unity, which is its theoretical value predicted by the REHTS. The above estimation is carried out under the set of no-arbitrage restrictions given by (2.20). This will help us to identify and estimate precisely from the data the structural parameters of the GDTSM \(k_i\) and \(\lambda_i^{(1)}\), which determine time-variation of term premium \(J_t(\tau)\). Estimating these parameters based only on regression model (2.21) will run into multicollinearity problems due to the very high correlation between term spread \(R_t(\tau) - r_t\) and the estimates of unobserved factors \(x_{it}\), used as independent regressors in (2.21).

GMM estimates of model (2.21) together with the system of equations (2.16)-(2.18) are reported in Table 5, \(\tau = \{5, 10, 15, 20, 25, 30\}\).\(^{15}\) In Table 4, we report estimates of slope

\(^{14}\)Analogous regressions have been suggested in the literature by Simon [103], Tzavalis and Wickens [108] and Tzavalis [106], or Psaradakis et al [98] for forward exchange rate. These regression models however use ad hoc specifications of the term premium effects. See Baillie [10], for a survey.

\(^{15}\)Note that in the estimation of regression model (3.14) the regressand, namely \(R_{t+1}(\tau - 1) - R_t(\tau)\), is not approximated by \(R_{t+1}(\tau) - R_t(\tau)\), as often made in many studies due to the lack of monthly maturity interval (see, e.g., Campbell and Shiller [24]). As was noted by Bekaert et al. [15] and was confirmed by our empirical analysis, this approximation can cause further biases to the LS estimates of \(\beta\) than those implied by the time-varying term
coefficient $\beta(\tau)$ which ignore term premium effects. These estimates are based on the least squares (LS) method.\textsuperscript{16} They confirm the severity of the biases of the estimates of $\beta(\tau)$. They should be compared to those of Table 5, which allow for time-varying term premium effects.

Table 4: Term spread regression model without term premium effects

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a(\tau)$</td>
<td>0.020</td>
<td>0.257</td>
<td>0.473</td>
<td>0.711</td>
<td>0.951</td>
<td>1.203</td>
</tr>
<tr>
<td>(0.162)</td>
<td>(0.326)</td>
<td>(0.446)</td>
<td>(0.580)</td>
<td>(0.716)</td>
<td>(0.859)</td>
<td></td>
</tr>
<tr>
<td>$\beta(\tau)$</td>
<td>-2.726</td>
<td>-4.187</td>
<td>-4.884</td>
<td>-6.094</td>
<td>-7.766</td>
<td>-9.541</td>
</tr>
<tr>
<td>(1.813)</td>
<td>(2.440)</td>
<td>(2.580)</td>
<td>(2.994)</td>
<td>(3.647)</td>
<td>(4.379)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Newey - West heteroscedasticity and autocorrelation consistent standard errors are shown in parenthesis.

Comparison of the results of Tables 4 to those of Table 5 clearly indicates that allowing for a time-varying term premium in term spread regression based tests of the REHTS can indeed save this hypothesis. The estimates of intercepts $a(\tau)$ are very close to zero and slope coefficient $\beta(\tau)$ becomes very close to unity, for all $\tau$, as predicted by the REHTS. With the exception of $\tau = 5$, the estimates of coefficients $a(\tau)$ and $\beta(\tau)$ reported in Table 5 can not reject the joint hypothesis that $a(\tau) = 0$ and $\beta(\tau) = 1$. Another interesting conclusion which can be drawn from the results of Table 5 is that the estimates of the mean reversion and price of risk parameters $k_i$ and $\lambda_i^{(1)}$, respectively, estimated by this version of the system of equations the GDTSM, i.e., (2.16)-(2.18) and (2.21), hardly changes compared with those of its previous

\textsuperscript{16}Note that GMM estimation of slope coefficients $\beta(\tau)$ based on the unadjusted for term premium effects term spread regression using as instruments values of variables $z_{1t}$, $z_{2t}$ and $z_{3t}$ can not save the REHTS since these variables are strongly correlated with the term premium $J_t(\tau)$.
version (2.16)-(2.19), reported in Table 3.

Table 5: Estimates of the term spread model with term premium effects

<table>
<thead>
<tr>
<th></th>
<th>$x_{1t}$</th>
<th>$x_{2t}$</th>
<th>$x_{3t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_i$</td>
<td>0.30</td>
<td>0.24</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.07)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>$\lambda_i^{(1)}$</td>
<td>-0.29</td>
<td>-0.15</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.07)</td>
<td>(0.15)</td>
</tr>
</tbody>
</table>

Matrix $G \equiv [G_{ij}]$

| $G_{1j}$ | 0.03 | -0.11 | 0.20 |
|          | (0.001) | (0.003) | (0.006) |
| $G_{2j}$ | -0.02 | 0.15 | -0.65 |
|          | (0.001) | (0.003) | (0.01) |
| $G_{3j}$ | 0.11 | 0.09 | 0.74 |
|          | (0.001) | (0.004) | (0.01) |

Estimates of $a^*(\tau)$ and $\beta^*(\tau)$ coefficients

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a^*(\tau)$</td>
<td>-0.25</td>
<td>-0.34</td>
<td>-0.36</td>
<td>-0.38</td>
<td>-0.44</td>
<td>-0.52</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.19)</td>
<td>(0.26)</td>
<td>(0.31)</td>
<td>(0.36)</td>
<td>(0.44)</td>
</tr>
<tr>
<td>$\beta^*(\tau)$</td>
<td>0.87</td>
<td>1.05</td>
<td>1.06</td>
<td>0.97</td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>

$J(87) = 110.2$ (p-value = 0.05), Instruments: 1, $z_{1t}$, $z_{2t-1}$, $z_{2t-2}$, $z_{2t-3}$, $z_{2t-4}$ and $z_{3t}$

Notes: Data are monthly from 1997:07 to 2009:10. The Newey - West heteroscedasticity and autocorrelation consistent standard errors shown in parentheses. $J$ is Sargan’s overidentified restrictions test statistic. This is chi-squared distributed with degrees of freedom which are equal to the number of orthogonality conditions employed in the estimation procedure (7×18) minus the number of the parameters estimated (39), i.e. 87. The instruments used are lags and current values of the following variables: $z_{1t} = R_t(10)$, $z_{2t} = r_t - R_t(40)$ and $z_{3t} = r_t + R_t(40) - 2R_t(6)$.

The above results together with the value of $J$ test statistic reported in Table 5, which can not reject the overidentified restrictions implied by the GMM estimation procedure of structural system (2.16)-(2.18) augmented by term spread regression model (2.21), provide strong support of the view that our GDTSM can consistently explain both the cross-section (no arbitrage) and dynamic predictions of the rational expectations theory of the term structure. These results have important forecasting policy implications. They mean that adjusting term spread $R_t(\tau) - r_t$
with the term premium effects predicted by our GDTSM can provide forecasts of future changes of long term interest rates $R_{t+1}(\tau - 1) - R_t(\tau)$ which are in the right direction.

To examine the stability and forecasting ability of regression model (2.21), in Table 6 we present values of some forecasting performance metrics and statistics for it, over different maturity intervals $\tau$. This is done for an out-of-sample exercise which relies on a recursive estimation of the structural parameters of model (2.21) after period 2001:12, based on the system of equations (2.16)-(2.18) and (2.21), by adding to our chosen initial window of our sample 1997:07-2001:12 one observation at a time. The forecasting performance metrics include the mean square and mean absolute values of forecasting errors, denoted as MSE and MAE, respectively, while the test statistics include those of Diebold and Mariano [45] (denoted as DM) and Giacomini and Rossi [63] (denoted as GR). The DM test statistic compares the forecasting performance of model (2.21) to the random walk (RW) model of interest rates, which is often considered in the literature to describe movements in long-term interest rates $R_t(\tau)$ (see, e.g., Mankiw and Miron [85], and Duffee [51]).

The GR statistic examines the out-of-sample forecasting performance of model (2.21), by testing if its forecasts break down due to unforeseen structural breaks occurred during our sample. In this case, the out-of-sample forecasts will not be consistent with the in-sample ones.

The results of Table 6 indicate that the forecasting performance of model (2.21) is more satisfactory than that of the RW. This can be supported by the values of both the MSE and MAE metrics reported in the table, for all $\tau$. The values of DM statistic indicate that the superiority of model (2.21) relatively to the random walk can be also inferred from our data for most of the maturity intervals $\tau$ considered, i.e. $\tau = \{5, 10, 15, 20\}$ years. For $\tau = \{25, 30\}$, the two models are found to performed equally, according to the DM statistic. This can be

\[ DM = \frac{3}{\tilde{\sigma}_T^2} \left( \frac{\tilde{\sigma}_T^2}{\tilde{\sigma}_T^2} \right)^{1/2}, \]

where $\tilde{\sigma}_T^2$ is a consistent estimate of the asymptotic (long-run) variance of $\sqrt{T \tilde{d}}$. The GR statistic is based on the testing principle that, if the forecast performance of a model does not break down, then there should be no difference between its expected out -of-sample and in-sample performance. It is defined as

\[ GR_{m,n,t} \equiv m \frac{\tilde{L}_{m,n}}{\tilde{\sigma}_{m,n}/\sqrt{n}} \]

where $\tilde{L}_{m,n}$ is the average surprise loss given as

\[ \tilde{L}_{m,n} = n^{-1} \sum_{t=m}^{T} \left( L(Y_{t+\tau}) - \hat{f}(\tilde{\beta}_t) \right) - m^{-1} \sum_{j=t-m+1}^T L(Y_j, \hat{y}_j(\tilde{\beta}_t)) \]

for $t = m, \ldots, T - \tau$, where $Y_{t+\tau}$ is the forecasted variable and $n = T - \tau - m + 1$. $\tilde{\sigma}_{m,n}$ is given in Corollary 4 of Giacomini and Rossi [63]. $GR_{m,n,t}$ converges in distribution to a Standard Normal $N(0,1)$ as $m, n \to \infty$. 

\[ GR_{m,n,t} \rightarrow N(0,1), \quad m, n \to \infty. \]
attributed to the fact that the forecasting ability of spread \( R_t(\tau) - r_t \) cease at the very end of the term structure, due to mean reverting properties of the underlying term structure common factors \( x_{it} \), especially those with slower mean reversion. Finally, the values of GR test statistic indicate that the forecasting performance of model (2.21) is stable over sample.

Table 6: Out-of-sample forecasting performance

<table>
<thead>
<tr>
<th>( \tau )</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.10</td>
<td>0.10</td>
<td>0.08</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.09</td>
<td>0.09</td>
<td>0.07</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>MAE</td>
<td>0.24</td>
<td>0.23</td>
<td>0.21</td>
<td>0.19</td>
<td>0.18</td>
<td>0.19</td>
<td>0.23</td>
<td>0.21</td>
<td>0.19</td>
<td>0.18</td>
<td>0.17</td>
<td>0.18</td>
</tr>
<tr>
<td>DM</td>
<td>-1.87</td>
<td>-1.73</td>
<td>-1.76</td>
<td>-1.67</td>
<td>-0.94</td>
<td>-0.75</td>
<td>0.28</td>
<td>0.43</td>
<td>0.06</td>
<td>0.11</td>
<td>0.16</td>
<td>0.25</td>
</tr>
<tr>
<td>GR</td>
<td>0.09</td>
<td>0.07</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.23</td>
<td>0.19</td>
<td>0.18</td>
<td>0.17</td>
<td>0.18</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Notes: The table presents values of metrics and statistic assessing the one period ahead (t+1) out-of-sample forecast performance of regression model (2.21) and the random walk (RW) model with drift. MSE and MAE denote the mean square and absolute error metrics, while DM and GR denote Diebold and Mariano [45] and Giacomini and Rossi [63] test statistics, respectively. These statistics follow the standard normal distribution. Our out-of-sample forecasts are based on a recursive estimation of model (2.21), jointly with the system of equations (2.16)-(2.18), and the random walk model. This is done after period after period 2002:01, by adding one observation at a time until the end of sample. The number of the out-of-sample observations used in our forecasting exercise is \( n \equiv T - \tau - m + 1 = 93 \), where \( m \) in our initial sample window of observations.

2.4 Conclusions

This paper suggests new term spread based regression tests allowing for time-varying premium effects with the aim of examining if a time-varying term premium can explain the puzzling behavior of the spread between the long and short-term interest rates to fail to forecast future movements in the former. This is against the predictions of the rational expectations hypothesis of the term structure of interest rates. To capture the time-varying term premium effects on the tests of this hypothesis, the paper employs a simple and empirically tractable Gaussian Dynamic Term Structure Model (GDTSM). To estimate this model, the paper suggests a new empirical methodology, which retrieves estimates of the unobserved factors spanning the term structure of interest rates net of measurement errors. This is done by projecting interest rates series (or transformations of them), used as instruments in inverting the interest rates relationships implied by affine term structure models, on well diversified portfolios of zero-coupon bond
interest rates. The latter are net of measurement error effects and are estimated from a very large set of interest rates data based on principal component analysis.

The paper provides a number of interesting results, which have important policy implications. First, it shows that a three-factor and empirically tractable GDTSM can sufficiently explain the cross-section movements of the US term structure of interest rates implied by no-arbitrage conditions in the bond market. This model provides estimates of the unobserved factors of the term structure of interest rates which are persistent and can capture substantial movements in interest rates, observed during our sample. Second, the paper shows that adjusting term spread regressions by the term premium effects implied by the above GDTSM can explain the empirical failures of the term spread to forecast future movements in long-term interest rates. This result means that our model can be successfully employed to forecast the correct direction of future long-term interest rate changes, net of term premium effects, as predicted by the expectations hypothesis. Finally, the paper shows that the factors that are priced in the bond market and, thus, cause significant time-varying effects on the term spread regressions are those which are associated with "level" and "slope" shifts in the term structure of interest rates.
Chapter 3

Real term structure forecasts of consumption growth

3.1 Introduction

There are few studies in the literature estimating term structure models of real interest rates, in contrast to the vast amount of studies on the nominal term structure models (see, e.g., Dai and Singleton [38], Ang et al [5], for a survey). This may be attributed to the lack of availability of real interest rates, for different maturity intervals. Estimating real term structure models is useful for two main reasons. First, it can indicate the number of factors spanning the real term structure and it will estimate their mean reversion and associated prices of risk. The results of this analysis can be compared to those on the nominal term structure of interest rates. Second, it can explain if the real term structure contains information about future real consumption growth.

The information content of the real term structure of interest rates about real consumption growth has been studied in a number of studies in the literature (see, e.g., Harvey [71], Plosser and Rouwenhorst [97], Chapman [26], Rendu de Lint and Stolin [99], Berardi and Torous [17], Tsang [105]). These studies show that the term spread between long and short-term real (or nominal) interest rates appear to contain information about future real consumption growth and economic activity, at short or long horizons. As noted by Harvey [71], this information

---

1Note that there is also a close related literature which studies and confirms the leading indicator property of the term structure for real economic activity and consumption growth and, in particular, of the term spread between the long and short term interest rates (see, e.g., Stock and Watson [104], Estrella [56] and Jardet [76]). These papers however rely on the term spread between nominal interest rates, following Donaldson et al [48].
of the term spread can be attributed to the desire of investors to smooth their consumption over time. This is consistent with the predictions of the consumption capital asset pricing. In addition to the term spread, evidence suggests that the level of real (or nominal) short-term real interest rate also contains information about the future real consumption growth beyond that implied by the term spread.

This paper contributes into the above literature on many fronts. First, using real consumption and term structure data, instead of nominal, it estimates an empirically tractable Gaussian dynamic term structure model and derives estimates of the underlying unobserved factors spanning the term structure of real interest rates. Then, it examines if this model fits satisfactorily into the data and tests its cross-section restrictions implied by no-profitable arbitrage conditions in the bond market. This is done based on an econometric framework, which apart from real interest rates and excess holding period returns, it also includes consumption data. Second, it employs the model to investigate if the short-term real interest rate and its spread with real long-term interest rates can predict future real consumption growth, over different horizons ahead. To this end, the paper derives closed form formulas of the slope coefficients of these two variables in consumption growth regressions, where they are regressors.

The results of the paper lead to a number of interesting conclusions. First, they show that our term structure model is consistent with the consumption smoothing hypothesis. Second, it finds that there are two common factors which can explain almost all of variation of the term structure of real interest rates. These factors are closely correlated with their estimates retrieved from the data based on principal component analysis. The parameter estimates of the real term structure model indicate that the first of the two factors spanning the real term structure exhibits very slow mean reversion, while its associated price of risk is very small. This factor can explain level shifts in real interest rates. The opposite happens for the second factor, which determines the slope of the term structure. This factor has a much higher mean-reversion and price of risk than the first factor, and it can explain the ability of the real term spread to forecast future real consumption growth over different horizons ahead.

The paper is organized as follows. Section 2 presents the real term structure of interest rates and derives closed form solutions of the slope coefficients of the consumption growth regression model, using the short-term interest rate and its spread with a long-term interest rate
as regressors. Section 3 carries out the empirical analysis. This section also includes unit root tests and principal component analysis for real interest rates. The unit root tests can confirm if real interest rates constitute stationary series, as it is assumed by affine term structure models. The principal component analysis can indicate the number of unknown factors spanning the real term structure of interest rates. Our empirical analysis is based on data from the US economy. Section 4 concludes the paper and summarizes some of its more important results.

3.2 Model setup

Consider an economy with production and stochastic investment opportunity sets (see, e.g., Cox, Ingersoll and Ross (CIR) ([31], [32]), or Longstaff and Schwartz [82]). The investment opportunity set consists of contingent claims (e.g., zero-coupon bonds), a riskless asset and a stochastic production process. We assume that this economy is characterized by $K$ state variables at time $t$, $x_{it}$, stacked into a $K$-dimensional vector $X_t$. These variables obey the following uncorrelated Gaussian processes:

$$dX_t = k(\theta - X_t)dt + \sigma dW_t$$

(3.1)

where $W_t$ denotes a $K$-dimensional Wiener process, $k$ and $\sigma$ are $K \times K$ matrices of the speed of mean-reversion and the volatility of the factors respectively and $\theta$ is a $K$-dimensional vector of the factor long-run mean parameters.

The instantaneous real interest rate $r^*_t$ it is assumed to be affine in the state vector

$$r^*_t = \rho_0 + \rho'_1 X_t$$

(3.2)

where $\rho_0$ is a constant and $\rho_1$ is a $K \times 1$ vector. These state variables constitute common factors which determine real consumption $C_t$ in the economy. If inflation is a constant rate (see, e.g., Harvey [71], it can be proved that real consumption growth $\frac{dC_t}{C_t}$ obeys the following process:

$$\frac{dC_t}{C_t} = \left( \sum_{i=1}^{K} x_{it} - c \right) dt + \sigma_d dW_{ct},$$

(3.3)

where $c$ is a constant which depends on inflation rate and the proportion between consumption

\footnote{See also Vasicek [109], Dai and Singleton [38], Ahn [1], Berardi and Torous [17].}
and wealth, and $W_{ct}$ is a Wiener process. By solving forward equation (3.3), it can be shown that the expected growth rate of real consumption from current period $t$ to a future period $t + \tau$ (i.e., $\tau$-horizons ahead) is given as

$$E_t [\Delta \tau C_{t+\tau}] = \psi_0(\tau) + \psi_1(\tau)'X_t$$

(3.4)

where $\Delta \tau C_{t+\tau} = \ln(C_{t+\tau}/C_t)$ and $\psi_1(\tau) = (I - e^{-k'\tau})(k')^{-1}$. 

In the above economy, the real price of a zero-coupon bond with a $\tau$-period maturity, denoted as $B_t^*(\tau)$ and, hence, its associated real interest rate, denoted as $R_t^*(\tau)$, can be derived by solving the following pricing kernel relationship:

$$B_t^*(\tau) = E_t \left( \frac{mt+z}{mt} \right),$$

(3.5)

where $\frac{mt+z}{mt}$ is the real pricing kernel. This is assumed that is given as

$$\frac{dm_t}{mt} = -r_t^*dt - \Lambda_t^*dW_t,$$

(3.6)

where $r_t^*$ is the instantaneous real interest rate or short-term rate and the $1 \times K$ vector $\Lambda_t^*$ contains the risk pricing functions, $\Lambda_{it}^*$, for each factor $i = 1, 2, \ldots, K$. The risk pricing functions $\Lambda_t^*$ evaluate the $K$-independent sources of risk associated with factors $x_{it}$. Following Duffee [51], we will assume that functions $\Lambda_t^*$ are linear in factors $x_{it}$, i.e.

$$\Lambda_t^* = \sigma^{-1}(\lambda_0 + \lambda_1X_t)$$

(3.7)

Substituting equations (5.1), (3.6), (3.2) and (3.7) into pricing kernel equation (3.5) yields the following zero-coupon real bond pricing formula:

$$B_t^*(\tau) = e^{-A(\tau)D(\tau)'X_t},$$

(3.8)

where $X_t$ is a $(KX1)$-dimension vector collecting all state variables (factors) $x_{it}$, i.e. $X_t = (x_{1t}, x_{2t}, \ldots, x_{Kt})'$. $A(\tau)$ is a scalar function and $D(\tau)$ is a $(KX1)$-dimension vector of valued functions, defined as $D(\tau) = (D_1(\tau), D_2(\tau), \ldots, D_K(\tau))'$, which collects the loading coefficients of factors $x_{it}$ on bond pricing formula (3.8). From this, we can obtain a pricing formula for real interest rates of zero-coupon bonds $R_t^*(\tau)$, with maturity interval $\tau$, as
CHAPTER 3. REAL TERM STRUCTURE FORECASTS OF CONSUMPTION GROWTH

\[ R_t^* (\tau) = \frac{1}{\tau} \left[ A(\tau) + D(\tau)'X_t \right], \text{ for } \tau = 1, 2, ..., N \]  

(3.9)

Closed form solutions of value functions \( A_i(\tau) \) and \( D_i(\tau) \) can be obtained by solving a set of ordinary differential equations under no arbitrage profitable conditions (see Duffie and Kan [53]). For our Gaussian dynamic term structure model (GDTSM), described above, these solutions for \( D_i(\tau) \) are analytically given as follows: \(^3\)

\[ D_i(\tau) = \left(1 - e^{-k_i^Q \tau}\right) (k_i^Q)^{-1} \rho_{it}, \text{ with } k_i^Q = k_i + \lambda_i^{(1)}, \]  

(3.10)

where \( k_i^Q \) constitutes a risk-neutral measure of mean-reversion parameter \( k_i \). These solutions imply a set of cross-section restrictions on the term structure loading coefficients \( D_i(\tau) \), for all \( i \), which can be tested, in practice.

The above GDTSM of real interest rates implies that the expected excess holding period real return of a \( \tau \)-period to maturity zero-coupon bond over short-term interest rate \( r_t^* \), referred to as term premium (see, e.g., Tzavalis and Wickens [108], Bolder [19] and Duffee [51]), is given as follows:

\[ E_t[h_{t+1}(\tau) - r_t^*] = -\sum_{i=1}^{K} D_i(\tau) \sigma_i \Lambda_{it} \]  

(3.11)

\[ = -\sum_{i=1}^{K} D_i(\tau) \lambda_i^{(0)} - \sum_{i=1}^{K} D_i(\tau) \lambda_i^{(1)} x_{it}. \]

Joint estimation of the last relationship and interest rates formula (3.9) (with, or without, cross-section restrictions (5.17)) will enable us to identify the price of risk slope coefficients \( \lambda_i^{(1)} \), which determine the time-varying part of the term premium. To calculate excess return \( h_{t+1}(\tau) - r_t^* \) in discrete-time, we consider the one-period (e.g., one-month) interest rate as short-term interest rate, \( r_t^* \), and we assume continuously compounded interest rates, implying \( R_t^* (\tau) = -\frac{1}{\tau} \log B_t(\tau) \). Then, \( h_{t+1}(\tau) - r_t^* \) can be written as follows:

\[ h_{t+1}(\tau) - r_t^* = \log \left( \frac{B_{t+1}(\tau - 1)}{B_t(\tau)} \right) - r_t^* = -(\tau - 1) \left[ R_{t+1}^* (\tau - 1) \right] + \tau R_t^*(\tau) - r_t^*. \]

\(^3\)See, e.g., Dai and Singleton [38], Kim and Orphanides [79].
3.2. MODEL SETUP

3.2.1 Term structure forecasts of consumption growth

The forecasting implications of the term structure of real interest rates $R_t^*(\tau)$ about future consumption growth $\tau-$periods ahead, defined as $\Delta_{\tau} c_{t+\tau}$ where $c_t = \log C_t$, can be investigated by equations (3.4), (3.9) and (3.2). These equations show that both $\Delta_{\tau} c_{t+\tau}$ and $R_t^*(\tau)$, for all $\tau$, are driven by $K$ common unobserved factors $x_{it}$, for $i = 1, 2, ..., K$. Substituting out these factors from relationships (3.4) and (3.9) implies that $c_{t+\tau}$ can be written as a linear function of short-term rate $r_t^*$ and its spreads with long-term interest rates, defined as $S_{p_t}(\tau_L) \equiv R_t^*(\tau_L) - r_t^*$, where $R_t^*(\tau_L)$ denotes a long-term interest rate with maturity interval $\tau_L$. To see this more rigorously, assume that the number of common factors $x_{it}$ are $K = 2$, as will be confirmed by our empirical analysis in the next section. Then, equations (3.9) and (3.2) imply the following system of equations for short-term interest rate $r_t^*$ and spread $S_{p_t}(\tau_L)$:

$$Z_t = A^* + D^*X_t,$$  

(3.12)

where $A^* \equiv \left[ \begin{array}{c} \rho_0 \\ (1/\tau_L)A(\tau_L) - \rho_0 \end{array} \right]$, $D^* \equiv \left[ \begin{array}{cc} \rho_{11} & \rho_{12} \\ (1/\tau_L)D_1(\tau_L) - \rho_{11} & (1/\tau_L)D_2(\tau_L) - \rho_{12} \end{array} \right]$ and $X_t = (x_{1t}, x_{2t})^\prime$.

Based on equation (3.12), we can derive the following relationship:

$$E_t[\Delta_{\tau} c_{t+\tau}] = \psi_0(\tau) - \Psi_1(\tau)^\prime D^{*-1} A^* + \Psi_1(\tau)^\prime D^{*-1}Z_t,$$  

(3.13)

where $\Psi_1(\tau)^\prime \equiv (\psi_{11}(\tau), \psi_{12}(\tau))$ (see (3.4)). This can be done by writing equation (3.4) in a matrix form and substituting out the vector of common factors $X_t$ from it. This relationship indicates that term spread $S_{p_t}(\tau_L)$ and short-term rate $r_t^*$ contains information about future consumption growth $\Delta_{\tau} c_{t+\tau}$, over different horizons $\tau$. Thus, it can theoretically justify the use of the following linear regression model to forecast consumption growth $\Delta_{\tau} c_{t+\tau}$:

$$r_{t+\tau} = a_{t+\tau} + b_{t+\tau}S_{p_t}(\tau_L) + e_{t+\tau},$$  

(3.14)
\[ \Delta r_{t+\tau} = \text{const} + \gamma_1(\tau)r_t^* + \gamma_2(\tau)Sp_t(\tau_L) + u_{t+\tau}, \quad (3.14) \]

where \( u_{t+\tau} \) denotes a disturbance (error) term. According to (3.13), the slope coefficients of this regression model \( \gamma_1(\tau) \) and \( \gamma_2(\tau) \) are given in closed form as

\[
\begin{align*}
\gamma_1(\tau) &= \frac{\psi_{11}(\tau)[D_2(\tau_L) - \rho_{12}] - \psi_{12}(\tau)[D_1(\tau_L) - \rho_{11}]}{\rho_{11}[D_2(\tau_L) - \rho_{12}] - \rho_{12}[D_1(\tau_L) - \rho_{11}]} \quad (3.15) \\
\gamma_2(\tau) &= \frac{\psi_{12}(\tau)\rho_{11} - \psi_{11}(\tau)\rho_{12}}{\rho_{11}[D_2(\tau_L) - \rho_{12}] - \rho_{12}[D_1(\tau_L) - \rho_{11}]} \quad (3.16)
\end{align*}
\]

where \( \psi_{1i}(\tau) = (1-e^{-k_i\tau})(k_i)^{-1}\rho_{1i}, D_i(\tau_L) = (1/\tau_L)D_i(\tau_L) \) and \( D_i(\tau_L) = (1-e^{-k_i^Q\tau_L})(k_i^Q)^{-1}\rho_{1i}. \)

The analytical solutions of slope coefficients \( \gamma_1(\tau) \) and \( \gamma_2(\tau) \), given by formulas (3.15) and (3.16), indicate that the degree of information contained in \( r_t^* \) and \( Sp_t(\tau_L) \) about \( \Delta r_{t+\tau} \) does not depend on the values of loading coefficients \( \rho_{1i} \), since these coefficients cancel each other out. On the contrary this information depends on the mean-reversion and price of risk parameters \( k_i \) and \( \lambda_i^{(1)} \), respectively, or the risk-neutral mean reversion parameter \( k_i^Q \), as well as maturity interval \( \tau \). This can be seen from the following after the substitution of \( \psi_{1i}(\tau) \) and \( D_i(\tau_L) \) into (3.15) and (3.16):

\[
\begin{align*}
\gamma_1(\tau) &= \frac{((1-e^{-k_1\tau})/k_1)[((1-e^{-k_2\tau_L})/k_2\tau_L) - 1] - ((1-e^{-k_2\tau_L})/k_2)[((1-e^{-k_1\tau_L})/k_1\tau_L) - 1]}{((1-e^{-k_2\tau_L})/k_2\tau_L) - ((1-e^{-k_1\tau_L})/k_1\tau_L)} \\
\gamma_2(\tau) &= \frac{((1-e^{-k_2\tau_L})/k_2) - ((1-e^{-k_1\tau_L})/k_1)}{((1-e^{-k_2\tau_L})/k_2\tau_L) - ((1-e^{-k_1\tau_L})/k_1\tau_L)}
\end{align*}
\]

Next, we in Figure 1, we show the effects of \( k_i \) on \( \gamma_1(\tau) \) and \( \gamma_2(\tau) \), assuming for analytic convenience that \( \lambda_i^{(1)} = 0 \). The results of this analysis can be used to explain the pattern of the estimates of \( \gamma_1(\tau) \) and \( \gamma_2(\tau) \) with \( \tau \), observed in practice. In particular, in Figure 1 we plot the slope coefficients \( \gamma_1(\tau) \) and \( \gamma_2(\tau) \) as shown in (3.15) and (3.16), calculated for different mean reversion coefficient values of the underlying factors. In the upper-left graph we set a slow speed of mean reversion for both factors i.e., \( k_i \to 0 \) for \( i = 1, 2 \). In particular, we set \( k_1 = 0.05 \) and \( k_2 = 0.051 \). As the graph shows, in such a case, coefficient \( \gamma_1(\tau) \) approaches...
maturity, $\gamma_1(\tau) \rightarrow \tau$, and coefficient $\gamma_2(\tau)$ is almost zero for all maturities until $\tau = 9$. At $\tau = 12$, $\gamma_2(\tau) = 0.23$. These results mean that the forecasting ability of $r_t^i$ about consumption growth $\Delta_r c_{t+\tau}$ increases linearly with $\tau$, when $k_i \rightarrow 0$ for $i = 1, 2$, while the forecasting ability of the spread $Sp_t(\tau_L)$ is much weaker up to 9 months ahead. This can be attributed to the fact that shocks to factors $x_{it}$ tend to have permanent effects on the level of interest rates and real consumption, due to their high persistency. In the upper-right and the lower-left graph we set only one of the two factors to be fast mean reverting, i.e., $k_1 > 0$ or $k_2 > 0$. In particular we set $k_1(\text{or } k_2) = 0.05$ and $k_2(\text{or } k_1) = 2$. Independently of which factor we choose, the two graphs provide the same results. As both graphs show, in such a case, coefficient $\gamma_1(\tau)$ again approaches maturity, $\gamma_1(\tau) \rightarrow \tau$, independently of the choice for $k_1 = 0.05$ or 2, and coefficient $\gamma_2(\tau)$ is significantly positive for maturities of longer than 6 months. At $\tau = 12$, $\gamma_2(\tau) = 0.70$. Finally, in the lower-right graph we allow for both factors to be fast mean reverting, i.e., $k_i > 0$ for $i = 1, 2$. In particular we set $k_1 = 2$ and $k_2 = 2.01$. In that case, coefficients $\gamma_1(\tau)$ and $\gamma_2(\tau)$ take their highest values and indicate that $r_t^i$ and $Sp_t(\tau_L)$ have significant forecasting power on $\Delta_r c_{t+\tau}$. In that case $\gamma_1(\tau)$ and $\gamma_2(\tau)$ increase with maturity. These results imply that mean reversion increases the forecasting power of both of $r_t^i$ and $Sp_t(\tau_L)$ on $\Delta_r c_{t+\tau}$, as is expected by the theory.
3.3 Empirical analysis

Based on data on real interest rates and real consumption, in this section we estimate and test the GDTSM presented in the previous section and we examine if this term structure model can explain the pattern of the slope coefficients of short-term interest rate $r_t^*$ and spread $Sp_t(\tau L)$, $\gamma_1(\tau)$ and $\gamma_2(\tau)$, respectively, observed in reality.

Our empirical analysis is based on monthly US real interest rates of zero-coupon bonds covering the period from 1997:07 to 2009:10. These series are taken from the archive of J. Huston McCulloch (Economics Department, Ohio University). Our consumption data set consists of monthly observations of seasonally adjusted personal real consumption expenditures for the

---

4 http://www.econ.ohio-state.edu/jhm/ts/ts.html
3.3. EMPIRICAL ANALYSIS

above period, taken from the Federal Reserve Economic Data (FRED) (see code PCEC96). Figure 2 plots all real interest rates series used in our empirical analysis, covering a wide spectrum of maturity intervals from one month to five years (i.e. $\tau = 60$ months). This very broad set of real interest rates will be used in our analysis to examine the number of factors spanning the real term structure and to consider alternative maturity interval long-term spreads $Sp_t(\tau_L)$ as regressors in (3.14). As it can be seen from Figure 2, the real term structure of interest rates does not exhibit substantial volatility neither over their cross-section (maturity) dimension nor over the time interval of our data, with the exception of period 2001-2003 and year 2008. Between January 3, 2001 and June 25, 2003, which covers the first period, the Fed lowered its lending, short-term interest rate from 6.5% to 1.0%, which constitutes its lowest level since year 1996. This is done by the Fed to avoid a further slowing down of the US economy and to boost the stock market. Note that, during this period, the US stock market was in the bear regime due to the terrorist attack of September 11, 2001 and the collapses of the Enron and WordCom companies in year 2002 (see, e.g., Ghosh and Constantinides [61], and Dendramis et al [41]). In year 2008, which is the second period of a bond market turmoil during our sample, the recent financial crisis, associated with the collapse of Lehman Brothers in September 16, 2008, began. In this year, the US nominal interest rates increased substantially to reflect the higher credit risk levels of the US economy, compared to those in previous years.
Our empirical analysis has the following order. First, we carry out unit root tests for all real interest rates $R_t^r(\tau)$ employed in our estimation and testing procedures. These tests are critical in choosing the correct econometric framework for estimating and testing our GDTSM, avoiding any spurious regression effects. The latter can appear in estimating (3.14), if interest rate $r_t^* \text{ or spread } Sp_t(\tau_L)$ are integrated series of order one. Second, we conduct principal component (PC) analysis with the aim of determining the maximum number of common factors (state variables) $K$ which explain the total variation of $R_t^r(\tau)$ in our sample. Since principal component factors constitute well diversified portfolios of interest rates which are net of measurement or pricing errors effects in $R_t^r(\tau)$, they can be employed as instruments in the estimation of the GDTSM to minimize the bias effects of the above errors on the parameter estimates of the model (see, e.g., Argyropoulos and Tzavalis [7]). Third, we estimate and test the GDTSM, with and without consumption data. The estimates of the model are then used to examine if the pattern of slope coefficients $\gamma_1(\tau) \text{ and } \gamma_2(\tau)$ observed in reality is consistent with that predicted by the theory. In this part of the paper, we also examine if the random walk model of real consumption constitutes a better forecasting model than (3.14).
3.3. EMPIRICAL ANALYSIS

3.3.1 Unit root tests

To test for a unit root in the level of real interest rates $R_t^s(\tau)$, we carry out a second generation ADF unit root test, known as efficient ADF (E-ADF) test (see, e.g., Elliott et al. [55], Elliott [54], and Ng and Perron [91]). This test is designed to have maximum power against stationary alternatives to unit root hypothesis which are local to unity. Thus, it can improve the power performance of the standard ADF statistic often used in practice to test for a unit root in $R_t^s(\tau)$.

Values of E-ADF unit root test statistic are reported in Table 1. This is done for real interest rates $R_t^s(\tau)$, with maturity intervals $\tau = \{1, 3, 6, 12, 24, 36, 48, 60\}$ months. Note that, in addition to E-ADF, the table also presents values of $P_T$ unit root test statistic, suggested by Elliott et al. [55] as alternative to E-ADF. To capture a possible linear deterministic trend in the levels of $R_t^s(\tau)$ during our sample, both statistics E-ADF and $P_T$ assume that the vector of deterministic components $D_t$ employed to detrend $R_t^s(\tau)$ contains also a deterministic trend, i.e. $D_t = [1, t]$. The results of the table clearly indicate that, despite the fact that the values of the autoregressive coefficients $\phi$ are close to unity, the unit root hypothesis for $R_t^s(\tau)$ is rejected against its stationary alternative, for all $\tau$ considered. This is true at 5%, or 1% significance levels. The estimates of the autoregressive coefficient $\phi$ reported in the table indicate that $R_t^s(\tau)$ exhibit a slow mean reversion towards their long-run mean, especially those of longer maturity intervals of 36 and 60 months.

![Table 1: Efficient unit root tests of real interest rates $R_t^s(\tau)$](image-url)

<table>
<thead>
<tr>
<th>$\tau$:</th>
<th>1</th>
<th>3</th>
<th>6</th>
<th>12</th>
<th>24</th>
<th>36</th>
<th>48</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi - 1$</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.13</td>
<td>-0.10</td>
<td>-0.09</td>
<td>-0.08</td>
<td>-0.09</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.87</td>
<td>0.90</td>
<td>0.91</td>
<td>0.92</td>
<td>0.91</td>
</tr>
<tr>
<td>E-ADF</td>
<td>-4.22**</td>
<td>-4.18**</td>
<td>-4.10**</td>
<td>-3.92**</td>
<td>-3.47*</td>
<td>-3.05*</td>
<td>-2.89*</td>
<td>-2.95*</td>
</tr>
<tr>
<td>$P_T$</td>
<td>2.56**</td>
<td>2.61**</td>
<td>2.70**</td>
<td>2.94**</td>
<td>3.71**</td>
<td>4.78*</td>
<td>5.32*</td>
<td>5.14*</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. E-ADF and $P_T$ are the efficient unit root test statistics suggested by Elliott et al. [55]. Critical values of test statistics E-ADF and $P_T$ are provided by Elliott.

---

5Evidence of high persistency in real interest rates can be found in Neely and Rapach [89].

6Evidence provided in the literature on unit root tests for interest rates series is mixed. Earlier studies based on single time series unit root tests, such as the standard ADF test, can not reject the null hypothesis of a unit root (see, e.g., Hall et al. [67]). On the other hand, more recent studies based on panel data tests or Bayesian panel data methods tend to reject this hypothesis (see, e.g. Costantini and Lupi [35] and Meligotsidou et al [86]).
The lag-order of the auxiliary regressions $p$ used to carry out the tests are chosen based on the SIC criterion. This is found to be $p = 1$, for all maturity intervals $\tau$ examined. (*) and (**) mean significance at 5% and 1% levels, respectively.

### 3.3.2 Principal component analysis

Principal components (PC) analysis can retrieve a $K$ number of common factors spanning the term structure of real interest rates $R^*_t(\tau)$ (or their first differences $\Delta R_t(\tau)$), denoted as $pc_{iit}$, for $i = 1, 2, ..., K$. This can be done by the spectral decomposition of the variance-covariance matrix of $R^*_t(\tau)$, for $\tau = 1, 2, ..., N$, denoted as $\Sigma_R$, i.e.

$$\Sigma_R = \Omega \Theta \Omega',$$

where $N > K$, $\Theta$ is a diagonal matrix of dimension $(N \times N)$.$^7$ The elements of $\Theta$ are the eigenvalues of matrix $\Sigma_R$. $\Omega$ is a $(N \times N)$-dimension orthogonal matrix whose columns are the eigen vectors corresponding to the eigenvalues of matrix $\Sigma_R$. Given estimates of $\Omega$ and $\Theta$, the $(K \times 1)$-dimension vector of principal component factors $pc_{iit}$, defined as $PC_t = (pc_{1it}, pc_{2it}, ..., pc_{Kit})'$, can be retrieved from the $(N \times 1)$-dimension vector of interest rates series $R^*_t(\tau)$, denoted as $R_t$, as follows:

$$PC_t = \Omega'(R_t - \bar{R}),$$

where $\bar{R}$ is the sample mean of vector $R_t$. Note that, due to the rotation problem of PC analysis, $pc_{iit}$ may not correspond one-to-one to unobserved factors $x_{iit}$, for all $i$. However, they will be very highly correlated with $x_{iit}$, as they constitute portfolios of $R^*_t(\tau)$. Furthermore, as is noted in Joslin and et al. [77], their estimates will diversify away any measurement or pricing error in $R^*_t(\tau)$.

Our PC analysis relies on a set of $N = 60$ real interest rates $R^*_t(\tau)$, spanning a very wide maturity spectrum from one month to five years (sixty months). This is a large cross-section set of $R^*_t(\tau)$ which guarantees that the retrieved by the PC analysis common factors $pc_{iit}$ will

---

$^7$As shown in Bai and Ng [9], and Bai [8] consistent estimates of principal component factors can be obtained by PC analysis of interest rates $R_t(\tau)$, if the following condition holds: $\sum_{i=1}^{T} \frac{\tau^2}{T} \rightarrow 0$, where $T$ is the total number of interest rates $R_t(\tau)$ observations and $N$ is their cross-section dimension across different maturity intervals $\tau$. 
efficiently span the real term structure of interest rates and its unobserved factors $x_{it}$. Figure 3 graphically presents the estimates of $pc_{it}$, for $i = 1, 2$, retrieved by our PC analysis. These correspond to the first two largest in magnitude eigen values of matrix $\Sigma_R$, which are found to explain 99.96% (or 99.87%) of the total variation of the levels of $R^*_t(\tau)$ (or their first differences $\Delta R_t(\tau)$), over all $\tau$, as shown in the following table:\footnote{Note that the relative variation of the first two $pc_{it}$ is calculated as \[ \frac{\sum_{i=1}^{2} v_i}{tr(V)}, \] where $v_i$ is the eigen value of matrix $\Sigma_R$ and $tr(V)$ stands for the trace of the matrix of the eigen values of $\Sigma_R$, denoted as $V$.}

<table>
<thead>
<tr>
<th>total number of PCs</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>% variation explained in $R^*_t(\tau)$</td>
<td>94.83</td>
<td>99.96</td>
</tr>
<tr>
<td>% variation explained in $\Delta R_t(\tau)$</td>
<td>92.83</td>
<td>99.87</td>
</tr>
</tbody>
</table>

The results of this table clearly indicate that the first two principal component factors $pc_{1t}$ and $pc_{2t}$, obtained through our PC analysis, explain almost all the variation of the term structure of real interest rates $R^*_t(\tau)$.\footnote{Note that these two factors explain a total variation in real interest rates $R_t(\tau)$ which is analogous in magnitude to that of nominal interest rates captured by three factors (see, e.g., Litterman and Scheinkman [81], and, more recently, Argyropoulos and Tzavalis [7]).} The first factor $pc_{1t}$ explains the largest part of this variation, which is about 95% of the levels of $R^*_t(\tau)$, or 93% of their differences $\Delta R_t(\tau)$. Its remaining part, which is about 5% for $R^*_t(\tau)$ (or 7%, for $\Delta R_t(\tau)$), is explained by the second factor $pc_{2t}$. Although the proportion of the total variation of $R^*_t(\tau)$ explained by $pc_{2t}$ is very small, this can explain the slope of the real term structure. As can be seen by the graphs of $pc_{1t}$ and $pc_{2t}$, given by Figure 3, most of the variation of $pc_{1t}$ can be attributed to the turmoils of the US bond market in period 2001-2003 and year 2008. These turmoils have caused shifts in the levels of $R^*_t(\tau)$ (see also Figure 2). The second factor $pc_{2t}$ has been also affected by these
events, but at a less extent. This factor oscillates less than \( pc_{1t} \) over the whole sample.

![Figure 3. The real term structure and the first two principal components.](image)

To gain some economic insight of principal component factors \( pc_{1t} \) and \( pc_{2t} \), in Figure 4 we graphically present estimates of their loading coefficients on the first differences of interest rates \( \Delta R_t(\tau) \). In Tables 2A and 2B we report some useful descriptive statistics for them, including E-ADF and \( P_T \) unit root test statistics.\(^{10}\) The results of these tables allow us to investigate stochastic features of \( pc_{it} \), which have economic meaning. Table 2A also reports values of the correlation coefficients of \( pc_{1t} \) and \( pc_{2t} \) with the level of the two year long-term interest rate \( R_t(24) \) and the spread between the five-year and one-month interest rates, denoted \( Sp_t(60) \equiv R_t(60) - r_t^* \). These two variables are found to have the maximum degree correlation with \( pc_{1t} \) and \( pc_{2t} \), respectively.

\(^{10}\)Similar graphs of the loading coefficients of the first two \( pc_{it} \) factors are obtained for the levels of interest rates \( R_t(\tau) \).
3.3. **EMPirical analysis**

Figure 4. Loading coefficients of principal components on the real term structure.

Table 2A: Summary statistics of interest rates PCs

<table>
<thead>
<tr>
<th>Factors</th>
<th>$pc_{1t}$</th>
<th>$pc_{2t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Max. Value</td>
<td>26.05</td>
<td>15.42</td>
</tr>
<tr>
<td>Min. Value</td>
<td>-25.67</td>
<td>-7.39</td>
</tr>
<tr>
<td>Variance</td>
<td>128.69</td>
<td>6.96</td>
</tr>
<tr>
<td>Correlation</td>
<td>$\rho_1$</td>
<td>0.93</td>
</tr>
<tr>
<td>$R_t(24)$</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>$Sp_t(60)$</td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>0.82</td>
<td>0.67</td>
</tr>
<tr>
<td>$\rho_3$</td>
<td>0.73</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Notes: Max stands for maximum, while Min. for minimum. $\rho_j$ are the autocorrelations of the principal components $pc_{it}, i = 1, 2$, of lag order $j = 1, 2, 3$. 
CHAPTER 3. REAL TERM STRUCTURE FORECASTS OF CONSUMPTION GROWTH

Table 2B: Unit root tests for interest rates PCs

<table>
<thead>
<tr>
<th></th>
<th>pc_{1t}</th>
<th>pc_{2t}</th>
</tr>
</thead>
<tbody>
<tr>
<td>φ_i - 1</td>
<td>-0.11</td>
<td>-0.22</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>φ_i</td>
<td>0.89</td>
<td>0.78</td>
</tr>
<tr>
<td>p</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>E-ADF</td>
<td>-3.58**</td>
<td>-5.05**</td>
</tr>
<tr>
<td>P_T</td>
<td>3.49**</td>
<td>1.90**</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. E-ADF and P_T are the efficient unit root test statistics suggested by Elliott et al. [55]. See Table 1. The lag order p of the dynamic (first difference) terms of the E-ADF regressions chosen are based on the SIC criterion. (*) and (**) mean significance at 5% and 1% levels.

The results of Table 2B clearly indicate that both principal component factors pc_{1t} and pc_{2t} constitute stationary series. These results are consistent with those on unit root tests for real interest rates R_t(\tau), reported in Table 1. Figure 4 indicates that the loading coefficients of the first factor pc_{1t} on \Delta R_t(\tau) decays with maturity interval \tau, but with a very slow rate. On the other hand, the loading coefficients of the second factor pc_{2t} on \Delta R_t(\tau) increases with maturity interval \tau, but with a much faster rate than that of first factor pc_{1t}. These patterns of the loading coefficients of pc_{1t} and pc_{2t} on \Delta R_t(\tau) are similar to those found in the empirical literature of the nominal term structure (see, e.g., Litterman and Scheinkman [81]). Thus, first principal component factor pc_{1t} can be interpreted as a "level" factor. This can explain almost parallel shifts in the whole term structure of real interest rates. This factor is found to be highly correlated with the levels of real interest rates, e.g., the two-year interest rate R_t(24). The second principal component factor p_{2t} can be given the interpretation of a "slope" factor, since it determines the slope of the real term structure. This factor is found to have maximum correlation with long-term spread S_p_t(60) (see Table 2A). The very high values of correlation coefficients of pc_{1t} and pc_{2t} with R_t(24) and S_p_t(60), respectively, reported in Table 2A, means that, in the estimation of the GDTSM, variables R_t(24) and S_p_t(60) can be employed as appropriate instruments (vehicles) to retrieve estimates of unobserved factors x_{it} from our data, by inverting pricing relationship (3.9).

The close relationship between x_{it} and pc_{it}, expected by the theory of Section 2, means that pc_{it} can be also employed to provide forecasts of consumption growth rate \Delta_{\tau} c_{t+\tau}. Table
3 presents least squares (LS) estimates of the slope coefficients of the regression of $\Delta_{\tau}c_{t+\tau}$ on principal component factors $pc_{it}$, for $\tau = 1, 2, 3, 6, 9$ and 12 months. This can be thought of as an alternative consumption forecasting model to (3.13). Note that, in addition to the first two principal component factors $pc_{1t}$ and $pc_{2t}$, in this regression model we also include the third principal component factor $pc_{3t}$, as a regressor. This is done in order to examine if this factor, whose effect on real term structure variation is almost zero, has any significant information about $\Delta_{\tau}c_{t+\tau}$.

The results of Table 3 clearly indicate that $pc_{1t}$ and $pc_{2t}$ contain significant information about $\Delta_{\tau}c_{t+\tau}$, as expected by the theory. This information tends to increase with $\tau$. Consistently with the results of our PC analysis, the third factor $pc_{3t}$ is found to have no information about $\Delta_{\tau}c_{t+\tau}$, for all $\tau$. Summing up, the results of this section imply that two factors can sufficiently explain almost all the variation of the real term structure. The first two principal component factors of the real term structure obtained by our PC analysis are found to have substantial forecasting power on future consumption growth up to one year ahead.

<table>
<thead>
<tr>
<th>Model: $\Delta_{\tau}c_{t+\tau} = c_0(\tau) + c_1(\tau)pc_{1t} + c_2(\tau)pc_{2t} + c_3(\tau)pc_{3t} + \varepsilon_{t+\tau}$</th>
<th>$\tau$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1(\tau)$</td>
<td></td>
<td>-0.006</td>
<td>0.01</td>
<td>0.02</td>
<td>0.04</td>
<td>0.05</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.007)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>$c_2(\tau)$</td>
<td></td>
<td>0.02</td>
<td>0.05</td>
<td>0.08</td>
<td>0.13</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.07)</td>
<td>(0.11)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>$c_3(\tau)$</td>
<td></td>
<td>-0.18</td>
<td>-0.32</td>
<td>-0.33</td>
<td>-0.41</td>
<td>-0.01</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.15)</td>
<td>(0.26)</td>
<td>(0.34)</td>
<td>(0.76)</td>
<td>(1.20)</td>
<td>(1.93)</td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td>0.06</td>
<td>0.13</td>
<td>0.18</td>
<td>0.21</td>
<td>0.23</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. These are corrected for heteroscedasticity and forward-looking moving average serially correlated errors based on Newey–West method.

### 3.3.3 Estimation of the real term structure model

Having established good grounds to support that two common factors can explain almost all the variation of the real term structure of interest rates, in this section we estimate the GDTSM presented in Section 2, assuming $K = 2$. This model consists of the following structural equations:
\[ \Delta R_{t+1}(\tau) = const + D_1(\tau)E_t[\Delta x_{1t+1}] + D_2(\tau)E_t[\Delta x_{2t+1}] + \eta_{t+1}(\tau), \quad \tau = 0, 1, 2, \ldots \] (3.17)

\[ h_{t+1}(\tau) = const - \lambda_1^{(1)} D_1(\tau)x_{1t} - \lambda_2^{(2)} D_2(\tau)x_{2t} + \epsilon_{t+1}(\tau), \quad \tau = 1, 2, \ldots N \] (3.18)

\[ \Delta x_{it+1} = const + (e^{-k_i \Delta t} - 1)x_{it} + \omega_{it+1}, \quad i = 1, 2 \] (3.19)

\[ \Delta c_{t+1} = const + \psi_1(\tau)x_{1t} + \psi_2(\tau)x_{2t} + \nu_{t+1} \] (3.20)

These correspond to the theoretical formulas (3.9), (3.11), (5.1) and (3.4) of the GDTSM, presented in Section 2. Note that, for the real short-term rate \( r_t \), equation (3.17) assumes that \( \Delta r_{t+1} = const + \rho_{11} E_t[\Delta x_{1t}] + \rho_{12} E_t[\Delta x_{2t}] \), which corresponds to equation (3.2). The expectation terms \( E_t[\Delta x_{1t}] \) and \( E_t[\Delta x_{2t}] \) are estimated through equation (3.19) of the system. Apart from any possible mispecification errors, the error term of equation (3.17) \( \eta_{t+1}(\tau) \) can reflect measurement or pricing errors (see, e.g., Diebold et al. [47]). These can be attributed to the fact that long-term zero coupon bond prices constitute approximations of coupon-bearing bond prices.

The above system of equations, in addition to equations (3.17) and (3.19) often used to estimate affine term structure models of nominal interest rates (see, e.g., Dai and Singleton [38], and Ang, Piazzesi and Wei [5]), also includes the set of excess return equations (3.18). As mentioned in Section 2, this set of equations helps to identify key parameters of the term structure model from the data, like the mean reversion and price of risk parameters \( k_i \) and \( \lambda_i^{(1)} \), respectively. The latter determines the time-varying component of the term premium, as shown by equation (3.11).

To estimate the system of equations (3.17)-(3.20), we employ the Generalized Method of Moments (GMM) (see Hansen [69]). This method can provide asymptotically efficient estimates of the vector of parameters of the systems which are robust to possible heteroscedasticity and/or serial correlation of error terms \( \eta_{t+1}(\tau), \epsilon_{t+1}(\tau), \omega_{it+1} \) and \( \nu_{t+r}(\tau) \). In the estimation procedure, we impose the no-arbitrage restrictions given by equation (5.17) on loading coefficients \( D_i(\tau) \).

The values of unobserved factors \( x_{it} \) involved in the system will be obtained by inverting the following interest rates pricing relationship (3.12), i.e.,
\[ X_t = D^{*-1} (Z_t - A^*) , \]

following Pearson’s and Sun [92] approach, where \( D^* \) is defined by (3.12) and vector of series \( Z_t \) consists of \( z_{1t} = R_t(24) \) and \( z_{2t} = Sp_t(60) \). As shown in Table 2A, \( R_t(24) \) and \( Sp_t(60) \) are found to have the maximum degree of correlation with the principal component factors \( pc_{1t} \) and \( pc_{2t} \), respectively, and thus may be less affected by measurement errors. All constants of the system are left unrestricted in the estimation procedure, as they can reflect possible imperfections of the bond market. As \( R_t^e(\tau) \), real consumption growth \( \Delta_1 c_{t+1} \) is given in percentage terms, i.e. \( \Delta_1 c_{t+1} \equiv 100 \ln[c(t + 1)/c(t)] \), and is also annualized.

Tables 4 and 5 present GMM estimates of the mean-reversion and price of risk parameters of system (3.17)-(3.20) \( k_i \) and \( \lambda_i^{(1)} \), with and without including consumption growth equation (3.20) in it, respectively. Comparison of these two different sets of estimates for \( k_i \) and \( \lambda_i^{(1)} \) can show if they remain robust to the inclusion of consumption data, when estimating the GDTSM. To estimate both specifications of the above system based on the GMM, we use lagged values of \( R_t(24) \) and \( Sp_t(60) \) as instrumental variables.

Each of Tables 4 and 5 present two different sets of estimates of \( k_i \) and \( \lambda_i^{(1)} \). The first relies on values of unobserved factors \( x_{it} \) retrieved from the data through relationship \( X_t = D^{*-1} (Z_t - A^*) \), often used in practice. See Panels A of the tables. The second set is based on a procedure which slightly modifies the above procedure, suggested by Argyropoulos and Tzavalis [7]. This replaces the observed values of vector \( Z_t \) in relationship \( X_t = D^{*-1} (Z_t - A^*) \) with their projected values of the elements \( Z_t \) on principal component factors \( pc_{1t} \) and \( pc_{2t} \). These are obtained based on the following regressions:

\[ z_{it} = const_i + d_{i1} pc_{1t} + d_{i2} pc_{2t} + \varsigma_t, \text{ for } i = 1, 2, \]

which are estimated simultaneously with our system of equations (3.17)-(3.20). This procedure minimizes the effects of pricing, or measurement, errors of interest rates \( R_t^e(\tau) \) on retrieving estimates of unobserved factors \( x_{it} \) through \( X_t = D^{*-1} (Z_t - A^*) \). This can be attributed to the fact that principal component factors \( pc_{it} \) constitute well diversified portfolios of interest rates \( R_t^e(\tau) \), if a large set of \( R_t^e(\tau) \) is used to retrieve them. Thus, they can eliminate the effects...
of measurement or pricing errors in $R_t^*(\tau)$, or $Sp_t(\tau)$, on the estimates of $x_{it}$.

Table 4: GMM estimates of system (3.17)-(3.20)

<table>
<thead>
<tr>
<th></th>
<th>Panel A</th>
<th>Panel B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$x_{1t}$</td>
<td>$x_{2t}$</td>
</tr>
<tr>
<td>$\rho_{1i}$</td>
<td>0.52</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>$(4 \times 10^{-4})$</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$k_i$</td>
<td>0.12</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$\lambda_{i(1)}$</td>
<td>-0.07</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$d_{1i}$</td>
<td>0.20</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>$(5 \times 10^{-5})$</td>
<td>$(6 \times 10^{-5})$</td>
</tr>
<tr>
<td>$d_{2i}$</td>
<td>0.05</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>$(1 \times 10^{-4})$</td>
<td>$(1 \times 10^{-4})$</td>
</tr>
</tbody>
</table>

$J(102) = 126.50$ (p-value=0.05)  
$J(112) = 135.32$ (p-value=0.07)

Notes: The table presents GMM estimates of parameters $k_i$ and $\lambda_i$ of the system of equations (3.17)-(3.20), including consumption growth equation. Panel A presents estimates of $k_i$ and $\lambda_i$, for $i=1,2$, based on the observed values of the variables of vector $Z_t$ in inverting relationship (3.12), while Panel B presents GMM estimates of these parameters based on projected values of vector $Z_t$ on principal component factors $pc_{1t}$ and $pc_{2t}$ (see equation (3.21)). The estimates of the slope coefficients $d_{1i}$ and $d_{2i}$ of this regression are also given in the table. Heteroscedasticity and autocorrelation consistent (Newey-West) standard errors are shown in parentheses. $J(.)$ is Sargan’s overidentifying restriction test, distributed as chi-squared with degrees of freedom given in parentheses. These are equal to the number of orthogonality conditions employed in the GMM estimation procedure minus that of the parameters estimated.
Table 5: GMM estimates of system (3.17)-(3.19)

<table>
<thead>
<tr>
<th>ΔR_{t+1}(τ) = const + \sum_{i=1}^{2} D_i(τ)E_t[Δx_{it+1}] + \eta_{t+1}(τ), \ τ = {0, 1, 2, \ldots, N}</th>
</tr>
</thead>
<tbody>
<tr>
<td>h_{t+1}(τ) - r_i^* = const - \sum_{i=1}^{2} \lambda_i^{(1)} D_i(τ)x_{it} + e_{t+1}(τ)</td>
</tr>
<tr>
<td>Δx_{it+1} = const + (e^{-k_iΔt} - 1)x_{it} + ε_{it+1}</td>
</tr>
</tbody>
</table>

where $D_i(τ) = (1 - e^{-k_i Q}) (k_i Q)^{-1}$, with $k_i Q = k_i + \lambda_i^{(1)}$

Panel A

<table>
<thead>
<tr>
<th>$\rho_{1i}$</th>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.88</td>
<td>-0.63</td>
<td></td>
</tr>
<tr>
<td>(5 x 10^{-4})</td>
<td>(3 x 10^{-4})</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$k_i$</th>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.06</td>
<td>2.18</td>
<td></td>
</tr>
<tr>
<td>(3 x 10^{-4})</td>
<td>(0.002)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda_i^{(1)}$</th>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0005</td>
<td>-0.38</td>
<td></td>
</tr>
<tr>
<td>(1 x 10^{-4})</td>
<td>(8 x 10^{-4})</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$d_{1i}$</th>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>(5 x 10^{-5})</td>
<td>(1 x 10^{-4})</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$d_{2i}$</th>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>(7 x 10^{-5})</td>
<td>(1 x 10^{-4})</td>
<td></td>
</tr>
</tbody>
</table>

$J(96) = 121.37$ (p-value: 0.05)

Panel B

<table>
<thead>
<tr>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>-0.73</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.0008)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.26</td>
<td>2.10</td>
</tr>
<tr>
<td>(2 x 10^{-4})</td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.06</td>
<td>-0.18</td>
</tr>
<tr>
<td>(1 x 10^{-4})</td>
<td>(5 x 10^{-4})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20</td>
<td>-0.27</td>
</tr>
<tr>
<td>(5 x 10^{-5})</td>
<td>(1 x 10^{-4})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$x_{1i}$</th>
<th>$x_{2i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.45</td>
</tr>
<tr>
<td>(7 x 10^{-5})</td>
<td>(1 x 10^{-4})</td>
</tr>
</tbody>
</table>

$J(106) = 132.40$ (p-value: 0.05)

Notes: Panel A presents GMM estimates of parameters $k_i$ and $\lambda_i$ of the system of equations (3.17)-(3.19), without including in it consumption growth equation (3.20). These are based on observed values of vector $Z_t$ when inverting relationship (3.12). Panel B presents GMM estimates of $k_i$ and $\lambda_i$ of the above system based on projected values of vector $Z_t$ on principal component factors $pc_{it}$ (see equation (3.21)). The estimates of the slope coefficients $d_{1i}$ and $d_{2i}$ of this regression are also given in the table. Heteroscedasticity and autocorrelation consistent (Newey-West) standard errors are reported in parentheses. $J(.)$ is Sargan’s overidentifying restriction test statistic, distributed as chi-squared with degrees of freedom given in parentheses. These are equal to the number of orthogonality conditions employed in the GMM estimation procedure minus that of the parameters estimated.

The results of Tables 4 and 5 lead to a number of interesting conclusions. First, they show that the specification of our two factor GDTSM, presented in Section 2, is consistent with the data, which supports the consumption smoothing hypothesis. This is true independently of weather consumption growth equation (3.20) is included in the estimation of system of equations (3.17)-(3.20), or not. This result can be justified by the value of Sargan’s overidentifying restrictions test statistic, denoted as $J(.)$, reported in the table. At 5% significance level, $J$ statistic can not reject the orthogonality conditions implied by the system of structural equations (3.17)-(3.20) and the instruments used by the GMM estimation procedure of it. This implies that the cross-section restrictions imposed on loading coefficients $D_i(τ)$ of the GDTSM can not
reject the no-arbitrage conditions (5.17) implied by the theory.

The estimates of parameters $k_i$ and $\lambda^{(1)}_i$ reported in Tables 4 and 5 are close to those found in many studies estimating GDTSM based on nominal interest rates (see, e.g., Ang et al. [3], and Duffee [52]). In particular, the estimates of $k_i$ imply a very slow mean reversion for the first unobserved factor $x_{1t}$, which is very close to zero, and a much faster for the second factor $x_{2t}$. The reported values of mean-reversion parameter $k_2$ imply values of the autoregressive coefficient $\phi$ of the discretized process (5.1) for $x_{2t}$, which are much smaller than those implied by the estimates of unit root auxiliary autoregressive models for $R^*_t(\tau)$ and $pc_{it}$. This can be obviously attributed to the fact that $R^*_t(\tau)$ and $pc_{it}$ constitute linear transformations of unobserved factors $x_{1t}$ and $x_{2t}$, which exhibit different degree of mean reversion.

Regarding the estimates of price of risk parameters $\lambda^{(1)}_i$, the results of the tables indicate that these are significant for both factors $x_{1t}$ and $x_{2t}$. This result means that time-varying risk premia effects associated with both factors $x_{1t}$ and $x_{2t}$ are priced in the bond market. According to (3.11), the negative values of $\lambda^{(1)}_1$ and $\lambda^{(1)}_2$ imply that term premium embodied the real term structure is positive. Note that the estimate of $\lambda^{(1)}_2$, related to the second factor $x_{2t}$ is bigger in absolute value than that of factor $x_{1t}$. As will be seen latter on, this factor captures the slope of the term structure. Its higher price in absolute terms reduces the mean-reversion parameter $k_2$ of factor $x_{2t}$ under the risk neutral measure, due to risk aversion effects.

The different sets of values of parameters $k_i$ and $\lambda^{(1)}_i$ reported in Tables 4 and 5 are quite close between the alternative systems of equations estimated, with and without consumption growth equation (3.20) (i.e., Tables 4 and 5), and across the two methods employed to retrieve estimates of unobserved factors $x_{it}$ (see Panels A and B), i.e. based on observed values of $Z_t$ or projected values of them on principal component factors $pc_{it}$. The last set of estimates provides more robust estimates of $k_i$ and $\lambda^{(1)}_i$ for both specifications of system (3.17)-(3.20), with and without consumption growth equation. This may be attributed to the fact that the estimates of $x_{it}$ based on the projected values of $Z_t$ on $pc_{it}$ are smoother than those based on the actual values of $Z_t$.

To see if $x_{1t}$ and $x_{2t}$ are closely related to principal component factors $pc_{1t}$ and $pc_{2t}$, in Figures 5 and 6 we graphically present estimates of them vis-a-vis those of $pc_{1t}$ and $pc_{2t}$ presented in Figure 3. The estimates of $x_{1t}$ and $x_{2t}$ presented in the figures are based on the parameter
3.3. EMPIRICAL ANALYSIS

estimates of system (3.17)-(3.20) relying on the projected values of $Z_t$, reported by Panel B of Table 4. Inspection of the graphs of the above figures clearly indicate that, as was expected, there is a very close relationship between the estimates of $x_{1t}$ and $pc_{1t}$, and between $x_{2t}$ and $pc_{2t}$. However, there is no one-to-one correspondence between $x_{it}$ and $pc_{it}$, for $i=\{1, 2\}$. The estimates of $x_{it}$ are smoother than those of $pc_{it}$, especially for factor $x_{1t}$. These results imply that, in estimating GDTSMs, replacing unobserved factors $x_{it}$ with estimates of principal component factors may lead to inaccurate estimates of the parameters of these models.

Figure 5. Estimates of factor $x_{1t}$ versus principal component factor $pc_{1t}$.

Figure 6. Estimates of factor $x_{2t}$ versus principal component factor $pc_{2t}$.
3.3.4 Real term structure forecasts of consumption growth

In this section, we examine if the forecasting ability short-term real rate $r_t$ and term spread $Sp_t(\tau_L)$ about future consumption growth $\Delta_c c_{t+\tau}$, found in the literature (see related studies in the introduction), is in accordance with the theory. Our analysis is mainly interested in examining if the estimates of the key parameters of the GDTSM $k_i$ and $\lambda_{it}$ can match the pattern of the LS estimates of the slope coefficients of the consumption forecasting regression model (3.14), $\gamma_1(\tau)$ and $\gamma_2(\tau)$, observed in practice. In addition to this, we also examine the out-of-sample forecasting ability of model (3.14) relative to that implied by the random walk model of real consumption with drift, suggested by Hall [66]. As is noted in the literature (see, e.g. Duffee [52]), the latter is a hard model to beat in forecasting real consumption level, or its growth rate.

Table 6A presents LS estimates of the slope coefficients of regression model (3.14),

$$\Delta_c c_{t+\tau} = const + \gamma_1(\tau)r_t + \gamma_2 Sp_t(\tau_L) + u_{t+\tau}.$$ 

This is done for two different spreads of interest rates: $Sp_t(60) = R_t(60) - r_t$ and $Sp_t(36) = R_t(36) - r_t$, and for $\tau = 1, 3, 6, 9, 12$ months ahead. Table 6B presents values of some metrics and test statistics evaluating the in-sample and out-of-sample forecast performance of the above model for $\Delta_c c_{t+\tau}$ and that implied by the RW model with drift for $c_{t+\tau} = \log C_{t+\tau}$. These metrics include the mean square and absolute errors, denoted as MSE and MAE, respectively. The test statistics employed are those of Diebold and Mariano [45], denoted DM,\(^{11}\) and Giacomini and Rossi [62], denoted as GR (See also Giacomini and Rossi [63]). The latter is an out-of-sample forecast performance statistic which can test if the forecasts of a model can break down, due to unforeseen breaks-events.\(^{12}\) To calculate the out-of-sample values of the above metrics

\(^{11}\)DM test statistic is based on the loss difference $d_t = L(u_{t+\tau}^{Model (3.14)}) - L(u_{t+\tau}^{RW})$. It is defined as $DM = \frac{2}{\bar{a}_T^{\gamma_2(\tau)}}$, where $\bar{a}_T^{\gamma_2(\tau)}$ is a consistent estimate of the asymptotic (long-run) variance of $\sqrt{T}d$.

\(^{12}\)The GR statistic is based on the testing principle that, if the forecast performance of a model does not break down, then there should be no difference between its expected out-of-sample and in-sample performance. It is defined as $GR_{m,n,t} = \frac{\overline{SL}_{m,n}}{\bar{a}_T^{\gamma_2(\tau)}}$, where $\overline{SL}_{m,n}$ is the average surprise loss given as $\overline{SL}_{m,n} = n^{-1} \sum_{t=m}^{T} \{ \sum_{j=1}^{m} L(\Delta_c c_{t+\tau}, \hat{\gamma}_1(\tau), \hat{\gamma}_2(\tau)) - m^{-1} \sum_{j=1}^{m+1} L(\Delta_c c_{t+\tau}, \hat{\gamma}_1(\tau), \hat{\gamma}_2(\tau)) \}$, for $t = m, ..., T - \tau$, where $n \equiv T - \tau - m + 1$ is the number of out-of-sample observations and $m$ is the sample window of our initial estimates. $\bar{a}_T^{\gamma_2(\tau)}$ is given in Corollary 4 of Giacomini and Rossi [62]. $GR_{m,n,t}$ converges in distribution to a Standard Normal
and statistics, we rely on recursive estimates of model (3.14) and the RW model for consumption after period 2004:01, by adding one observation at a time and, then, re-estimating the models by the LS method until the end of sample. The number of the out-of-sample observations used to calculate the above metrics and test statistics are given as $n \equiv T - \tau - m + 1$, where $m$ is our sample window. Note that, for model (3.14), the tables presents two sets of results. The first employs spread $Sp_t(60) \equiv R_t(60) - r^*_t$ as regressor, while the second uses $Sp_t(36) \equiv R_t(36) - r^*_t$.

Table 6A: Real consumption growth forecasts

| Model: $\Delta c_{t+\tau} = \text{const} + \gamma_1(\tau)r^*_t + \gamma_2(\tau)Sp_t(\tau) + u_{t+\tau}$ |
|---|---|---|---|---|---|---|
| $\tau$ | 1 | 2 | 3 | 6 | 9 | 12 |
| $Sp_t(60) \equiv R_t(60) - r^*_t$ | | | | | | |
| $\gamma_1(\tau)$ | 0.08 | 0.17 | 0.26 | 0.50 | 0.76 | 0.97 |
| (0.03) | (0.05) | (0.07) | (0.19) | (0.34) | (0.43) |
| $\gamma_2(\tau)$ | 0.10 | 0.21 | 0.33 | 0.60 | 0.90 | 1.06 |
| (0.05) | (0.07) | (0.10) | (0.26) | (0.44) | (0.53) |
| $R^2$ | 0.04 | 0.11 | 0.16 | 0.20 | 0.23 | 0.24 |
| $Sp_t(36) \equiv R_t(36) - r^*_t$ | | | | | | |
| $\gamma_1(\tau)$ | 0.08 | 0.16 | 0.24 | 0.46 | 0.69 | 0.87 |
| (0.02) | (0.04) | (0.06) | (0.17) | (0.31) | (0.40) |
| $\gamma_2(\tau)$ | 0.13 | 0.27 | 0.40 | 0.73 | 1.06 | 1.23 |
| (0.05) | (0.08) | (0.12) | (0.31) | (0.52) | (0.63) |
| $R^2$ | 0.05 | 0.12 | 0.17 | 0.21 | 0.23 | 0.23 |

Notes: Newey-West standard errors corrected for heteroscedasticity and moving average errors up to $\tau - 1$ periods ahead are reported in prostheses. $R^2$ is the coefficient of determination.

The results of Tables 6A and 6B indicate that short-term real interest rate $r^*_t$ and spread $Sp_t(\tau)$ contains significant information about future consumption growth $\Delta c_{t+\tau}$, for all $\tau$. This is true for both cases of term spread $Sp_t(\tau)$ considered. The values of $R^2$, reported in Table 6A, imply that the forecasting ability of model (3.14) increases with $\tau$. As was expected, the values of $R^2$ are similar to those of principal component factors model forecasting future consumption growth $\Delta c_{t+\tau}$, reported in Table (3). The values of the MSE and MAE reported in Table 6B clearly indicate that the forecast performance of model (3.14) is better than that of the RW model, for all $\tau$. This is true for both the in-sample and out-of-sample exercises. The better performance of model (3.14) than the RW model is also supported by the values of DM statistic.

$N(0,1)$ as $m, n \rightarrow \infty$. 


CHAPTER 3. REAL TERM STRUCTURE FORECASTS OF CONSUMPTION GROWTH

The negative values of this statistic indicate that this model provides smaller in magnitude forecast errors than the RW model, especially as \( \tau \) increases. These values clearly reject the null hypothesis that the two models have the same forecasting ability, at 5% significance level.

Further support for model (3.14) in forecasting \( \Delta_t c_{t+\tau} \) can be obtained by the GR test statistic. The values of this statistic reported in the table indicate that this model can produce out-of-sample consumption growth forecasts which are stable and consistent with its in-sample forecasts up to three-months ahead.

Table 6B: Forecasting performance

<table>
<thead>
<tr>
<th>( \tau )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
<th>In-sample MSE</th>
<th>0.17</th>
<th>0.27</th>
<th>0.39</th>
<th>1.12</th>
<th>2.20</th>
<th>3.42</th>
<th>0.14</th>
<th>0.32</th>
<th>0.53</th>
<th>1.75</th>
<th>3.47</th>
<th>5.52</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Out-of-sample MSE</td>
<td>0.17</td>
<td>0.27</td>
<td>0.39</td>
<td>1.12</td>
<td>2.20</td>
<td>3.42</td>
<td>0.14</td>
<td>0.32</td>
<td>0.53</td>
<td>1.75</td>
<td>3.47</td>
<td>5.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.30</td>
<td>0.40</td>
<td>0.49</td>
<td>0.83</td>
<td>1.22</td>
<td>1.54</td>
<td>0.29</td>
<td>0.42</td>
<td>0.56</td>
<td>0.97</td>
<td>1.37</td>
<td>1.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DM</td>
<td>-1.46</td>
<td>-2.18</td>
<td>-2.71</td>
<td>-2.72</td>
<td>-2.75</td>
<td>-3.02</td>
<td>-1.47</td>
<td>-2.18</td>
<td>-2.71</td>
<td>-2.71</td>
<td>-2.74</td>
<td>-3.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>-0.41</td>
<td>1.34</td>
<td>2.16</td>
<td>5.94</td>
<td>8.89</td>
<td>10.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
<th>In-sample MSE</th>
<th>0.17</th>
<th>0.27</th>
<th>0.49</th>
<th>1.12</th>
<th>2.20</th>
<th>3.49</th>
<th>0.14</th>
<th>0.31</th>
<th>0.53</th>
<th>1.72</th>
<th>3.46</th>
<th>5.60</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Out-of-sample MSE</td>
<td>0.17</td>
<td>0.27</td>
<td>0.49</td>
<td>1.12</td>
<td>2.20</td>
<td>3.49</td>
<td>0.14</td>
<td>0.31</td>
<td>0.53</td>
<td>1.72</td>
<td>3.46</td>
<td>5.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.30</td>
<td>0.40</td>
<td>0.49</td>
<td>0.83</td>
<td>1.22</td>
<td>1.54</td>
<td>0.28</td>
<td>0.42</td>
<td>0.55</td>
<td>0.96</td>
<td>1.37</td>
<td>1.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>-0.44</td>
<td>1.29</td>
<td>2.06</td>
<td>5.88</td>
<td>8.84</td>
<td>11.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
<th>In-sample MSE</th>
<th>0.18</th>
<th>0.31</th>
<th>0.47</th>
<th>1.42</th>
<th>2.86</th>
<th>4.62</th>
<th>0.16</th>
<th>0.38</th>
<th>0.54</th>
<th>2.20</th>
<th>4.51</th>
<th>7.28</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Out-of-sample MSE</td>
<td>0.18</td>
<td>0.31</td>
<td>0.47</td>
<td>1.42</td>
<td>2.86</td>
<td>4.62</td>
<td>0.16</td>
<td>0.38</td>
<td>0.54</td>
<td>2.20</td>
<td>4.51</td>
<td>7.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.31</td>
<td>0.43</td>
<td>0.53</td>
<td>0.88</td>
<td>1.23</td>
<td>1.55</td>
<td>0.31</td>
<td>0.47</td>
<td>0.62</td>
<td>1.06</td>
<td>1.48</td>
<td>1.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>2.81</td>
<td>2.76</td>
<td>2.80</td>
<td>6.56</td>
<td>9.35</td>
<td>11.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents values of the MSE and MAE metrics and DM and GR statistics assessing the forecasting performance of model \( \Delta t c_{t+\tau} = const + \gamma_1(\tau) r_t^* + \gamma_2(\tau) S_p(t, L) + u_{t+\tau} \) and that implied by the random walk (RW) model of the level of real consumption with drift. DM and GR denote the Diebold-Mariano and Giacomini-Rossi test statistics, respectively. These statistics follow the standard normal distribution. Note that the GR test statistic is an out-of-sample test statistic, which can test the stability of the out-of-sample forecasts compared to the in-sample ones. To calculate the out-of-sample values of the above metrics and statistics, we rely on recursive estimates of model (3.14) and the RW model for consumption after period 2004:01, adding one observation at a time and, then, re-estimating the models until the end of sample. The total number of observations used in our out-of-sample forecasting exercise is \( n = T - \tau - m + 1 = 69 \), where \( m \) is our sample sample window.

Another interesting conclusion that can be drawn from the results Table 6A is that the LS estimates of the slope coefficients \( \gamma_1(\tau) \) and \( \gamma_2(\tau) \) of model (3.14) increase with \( \tau \). To examine
if these estimates of $\gamma_1(\tau)$ and $\gamma_2(\tau)$ can match those implied by the parameter estimates of the GDTSM, over different $\tau$, in Table 7 we present estimates of the latter against the LS estimates. The estimates of $\gamma_1(\tau)$ and $\gamma_2(\tau)$ implied by the GDTSM are derived based on relationship (3.15) and the estimates of parameters $k_i$ and $\lambda_i^{(1)}$ reported in Panels A and B of Table 4.

The results of Table 7 clearly indicate that the pattern of the LS estimates of coefficients $\gamma_1(\tau)$ and $\gamma_2(\tau)$ with maturity horizon $\tau$, reported in Table 6A, is consistent with that implied by the estimates of our GDTSM. The implied by the GDTSM estimates of $\gamma_1(\tau)$ and $\gamma_2(\tau)$ are close to their LS estimates, even for the forecasting period of $\tau = 12$ months. These lie within the two standard deviations confidence interval of the LS estimates of them. As is predicted by the analysis of subsection 2.1, the estimates of slope coefficient $\gamma_1(\tau)$ are close to time to maturity and the estimates of $\gamma_2(\tau)$ are increasing with $\tau$. This is true for both sets of implied values of $\gamma_1(\tau)$ and $\gamma_2(\tau)$, reported in the table. These results are also consistent across the two different spreads $Sp_t(\tau,)$, i.e. $Sp_t(60) = R_t(60) - r_t^*$ and $Sp_t(36) = R_t(36) - r_t^*$, considered in our analysis.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\gamma_1(\tau)$</th>
<th>$\gamma_2(\tau)$</th>
<th>$\gamma_1(\tau)$</th>
<th>$\gamma_2(\tau)$</th>
<th>$\gamma_1(\tau)$</th>
<th>$\gamma_2(\tau)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.08</td>
<td>0.01</td>
<td>0.08</td>
<td>0.02</td>
<td>0.08</td>
<td>0.10</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
<td>0.03</td>
<td>0.18</td>
<td>0.05</td>
<td>0.17</td>
<td>0.21</td>
</tr>
<tr>
<td>3</td>
<td>0.26</td>
<td>0.05</td>
<td>0.29</td>
<td>0.10</td>
<td>0.26</td>
<td>0.33</td>
</tr>
<tr>
<td>6</td>
<td>0.54</td>
<td>0.21</td>
<td>0.62</td>
<td>0.34</td>
<td>0.50</td>
<td>0.60</td>
</tr>
<tr>
<td>9</td>
<td>0.81</td>
<td>0.41</td>
<td>0.97</td>
<td>0.65</td>
<td>0.76</td>
<td>0.90</td>
</tr>
<tr>
<td>12</td>
<td>1.10</td>
<td>0.66</td>
<td>1.30</td>
<td>0.99</td>
<td>0.97</td>
<td>1.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\gamma_1(\tau)$</th>
<th>$\gamma_2(\tau)$</th>
<th>$\gamma_1(\tau)$</th>
<th>$\gamma_2(\tau)$</th>
<th>$\gamma_1(\tau)$</th>
<th>$\gamma_2(\tau)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.08</td>
<td>0.01</td>
<td>0.08</td>
<td>0.01</td>
<td>0.08</td>
<td>0.13</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
<td>0.03</td>
<td>0.17</td>
<td>0.04</td>
<td>0.16</td>
<td>0.27</td>
</tr>
<tr>
<td>3</td>
<td>0.26</td>
<td>0.06</td>
<td>0.27</td>
<td>0.08</td>
<td>0.24</td>
<td>0.40</td>
</tr>
<tr>
<td>6</td>
<td>0.52</td>
<td>0.21</td>
<td>0.56</td>
<td>0.30</td>
<td>0.46</td>
<td>0.73</td>
</tr>
<tr>
<td>9</td>
<td>0.79</td>
<td>0.42</td>
<td>0.85</td>
<td>0.57</td>
<td>0.69</td>
<td>1.06</td>
</tr>
<tr>
<td>12</td>
<td>1.05</td>
<td>0.67</td>
<td>1.14</td>
<td>0.86</td>
<td>0.87</td>
<td>1.23</td>
</tr>
</tbody>
</table>
CHAPTER 3. REAL TERM STRUCTURE FORECASTS OF CONSUMPTION GROWTH

Notes: The table reports theoretical values of slope coefficients $\gamma_1(\tau)$ and $\gamma_2(\tau)$ based on relationship (3.15) and the estimates of the mean-reversion and risk price function parameters of our GDTSM $k_i$ and $\lambda_i^{(1)}$, reported in Panels A and B of Table 4, respectively. This is done against the LS estimates of these coefficients, reported in Table 6A.

3.4 Conclusions

This paper suggests a Gaussian dynamic real term structure model to explain the ability of the short-term real interest rate and its term spread with longer term real interest rates to forecast future changes in real consumption growth. The paper fits the model into real term structure and consumption data from the US economy, and it provides a number of interesting results which are consistent with the consumption smoothing hypothesis.

First, it shows that two stationary common factors can explain most of the variation of the real term structure of interest rates. The first of these two factors, which exhibits very slow mean reversion, can explain persistent shifts in the levels of real interest rates. This factor is found to be affected more strongly by the recent financial crisis and the stock market crises of period 2001-2003, which also affected the US bond market. The second factor, which has higher degree of mean reversion, can explain the slope of the real term structure.

Second, the estimates of the price of risk parameters reported by the paper indicate that both of the above factors are priced in the market and, thus, they can explain time variation of excess holding period returns of the market. The estimates of the price of risk and mean-reversion parameters of the two term structure factors retrieved by our data are also found to be consistent with the cross-section restrictions of the real term structure model suggested by the paper. These restrictions arise by ruling out profitable arbitrage conditions of the market. They are tested based on a structural system of equations consisting of real interest rates, excess holding period real returns, reflecting term premia effects, and real consumption growth.

Finally, the paper rigorously shows that the forecasting ability of the short-term real interest rate and its spread with long-term real interest rates about future real consumption growth, over different periods ahead, can be consistently explained by the common factor representation of the real term structure and consumption growth. This forecasting model of consumption growth is found to perform better than that implied by the random walk model of real consumption
level. The ability of the term spread to forecast future consumption growth can be attributed to the high degree of mean reversion of the second common factor driving the real term structure of interest rates.
Chapter 4

Forecasting economic activity from yield curve factors

4.1 Introduction

There is recently growing interest in examining empirically the information context of the term structure of interest rates about future economic activity (see, e.g., Harvey [71], [72], Estrella and Hardouvelis [57], Plosser and Rouwenhorst [97], Ang and Piazzesi [3], Rendu de Lint and Stolin [99], Rudebusch and Wu [102], Piazzesi [95], Diebold et al [47], Ang et al [5]). Most of these studies rely on a regression model which employs the term spread between long and short-term interest rates, referred to as the slope of yield curve, as a regressor and output (or industrial production index) growth rate as a regressand. According to the theory, the short-term interest rate directly depends on the central bank’s decisions, while the long-term is determined by the expectations of the bond market participants. A zero or negative term spread (which means a flat, or inverted, yield curve) is often associated with a decline in future economic activity, or as a predictor of economic recessions. This can be explained as follows. Consider, for instance, a tight monetary policy, which increases the short-term interest rate. This policy will also decrease long-term rates and, thus, will flatten (or invert) the yield curve, since bond market’s expectations about future recessionary conditions will increase the current demand for savings in the economy.

As is well known in the empirical term structure literature, the yield curve is spanned by three factors, referred to as level, slope and curvature factors (see, e.g., Litterman and Scheinkman [81] and Bliss [18]). Given that the level factor explains parallel shifts of interest rates independently
of maturity intervals, often related to changes in long-term expectations about inflation in the economy, the term spread should be mainly determined by the two other factors, the slope and curvature. This paper empirically examines if these two factors constitute independent sources of information about future economic activity and contain more information about it than the term spread itself. The results of our analysis can also shed light on recent macroeconomic studies asserting that the slope factor of the yield curve reflects future business cycle (BC) conditions, while the curvature factor captures policy actions related to short or medium-term adjustments of the current stance of monetary policy (see, e.g., Bekaert [14], Dewachter et al. [43], Dewachter and Lyrio [42], Hordahl et al., [75] and Moench [88]). That is, the fact that, if for instance, economic growth is considered to be undesirably rapid, a restrictive monetary policy will be undertaken by the central bank, and conversely. To retrieve the unobserved factors driving the yield curve and, hence, the term spread, the paper fits into term structure data coming from five leading economies of world, the dynamic Nelson Siegel [90] model (DNSM) (see also Diebold et al. [47] and Diebold and Li [44], inter alia). This model is popular among market and central bank practitioners, as it has been found that fits adequately into yield curves (see, e.g., Diebold et al. [46]).

The results of the paper lead to a number of interesting conclusions. First, they show that the slope and curvature factors of the yield curve constitute independent sources of information about future economic activity. Together, these two factors have superior information about future economic activity than the term spread itself. Using the term spread to forecast future economic activity will thus undermine this information, as the slope and curvature factors are loaded into the term spread with opposite signs and can thus be offset to each other. For most of the countries examined, the paper finds that the slope factor contains significant information about future economic activity up to two-years ahead, while the curvature factor for shorter or medium, horizons. An increase in the slope factor is found to predict a slow down in future economic activity, as it affects negatively the term spread. That is, it implies a flat, or inverted, term spread. This is consistent with evidence provided in the literature by term spread regressions forecasting future economic activity (see above). On the other hand, an increase in the curvature factor is found to be positively associated with future economic activity, as this factor is positively associated with the term spread. These results are in accordance with the
theory predicting that the slope factor of the yield curve reflects future business cycle conditions, while the curvature factor captures independent changes in the current monetary policy which last over shorter horizons.

The paper is organized as follows. In section 2, we describe the data and re-estimate term spread regressions forecasting future economic activity. In section 3, we fit the DNSM into our data set and retrieve the slope and curvature factors of the yield curve. Then, we examine if these two factors contain significant information about future economic activity, by conducting regression analysis. Section 4 concludes the paper.

4.2 Forecasting economic activity based on term spread regressions

Our data set consists of 265 monthly observations of zero-coupon yields from 1987:05 to 2009:05, with maturity intervals (denoted as $\tau$) varying from 3 to 120 months. This set covers the following five developed countries: the United States (US), Canada (CA), the United Kingdom (UK), Germany (DE) and Japan (JP). To approximate the economic activity of these countries, we rely on their Industrial Production Index (IPI) from 1989:01 to 2009:05, obtained from the OECD database. For every country $i$, we measure the cumulative annualized economic growth from the current $t$-period to $k$-periods ahead, denoted as $g_{it,t+k}$, in percentage terms, i.e.,

$$g_{it,t+k} = 100(12/k)(\ln g_{it+k} - \ln g_t).$$

Figure 1 presents graphs of the term spread between the 5-years and the 3-months zero-coupon interest rate of our data set, which plays the role of the short-term interest rate. This spread is denoted as $spr_{it}(5y) \equiv r_{it}(5y) - r_{it}(3m)$, for all countries $i$. The shaded areas of the graphs indicate recession periods, announced by the official authorities of the countries. Inspection of the graphs of Figure 1 indicates that a flat (or inverted) yield curve, where short-term rate $r_{it}(3m)$ is almost the same with long-term rate $r_{it}(5y)$ (or it takes higher values than it, implying that $spr_{it}(5y)$ takes negative values) precedes economic slow downs, or recessions, as discussed in the introduction.

The graphs of the figure indicate that term spread $spr_{it}(5y)$ precedes economic slow downs
for most of the countries examined, with the exceptions Germany for period 1990-1993, where the yield curve inverts during this recessionary period, and Japan for periods 1997-1999 and 2000-2002. For the case of Germany, this can be attributed to the tight monetary policy of the Bundesbank followed Germany’s unification in October of year 1990 in order to avoid inflationary pressures. For Japan, it can be attributed to the very tight regulation of Japanese financial markets by the government during the above recessionary periods, which limited the role of market expectations in determining long-term interest rates. Finally, another interesting conclusion which can be drawn from the graphs of Figure 1 is that almost all the countries examined (especially the US, UK and Canada) tend to simultaneously enter into the recessions occurring during our sample. This can be obviously attributed to common economic policies
followed by the countries examined over our sample.

Figure 1. Term spreads and recessionary periods (shaded areas).

Table 1.a presents least squares (LS) estimates of the following regression model used in the
4.2. **FORECASTING ECONOMIC ACTIVITY BASED ON TERM SPREAD REGRESSIONS**

In this section, we use term spread regressions from the literature to forecast economic activity:

\[ g_{it,t+k} = \text{const} + \beta^{(k)}_i \text{spr}_{it}(5y) + \varepsilon_{it+k}, \text{ for all countries } i. \]  

(4.1)

This is done for forecasting horizons \( k = \{3, 6, 12, 24\} \) months ahead. The results of Table 1.a are consistent with previous evidence reported in the literature (see references mentioned in the introduction). Term spread \( \text{spr}_{it}(5y) \) has significant power in predicting future economic activity. This tends to increase with forecasting horizon \( k \). The only exception is Japan, where \( \text{spr}_{it}(5y) \) has forecasting power on \( g_{it,t+k} \) only for short-term horizons, i.e., \( k = 3 \). As was expected by the theory, the estimates of the term spread slope coefficients \( \beta^{(k)}_i \) are positive, implying that a positive (negative) value of \( \text{spr}_{it}(5y) \) predicts an increase (decrease) in future economic activity. To see if the term spread holds its predictive ability on marginal changes of growth rate \( g_{it+k-j,t+k} \), between two different future periods \( t+k-j \) and \( t+k \) where \( j < k \), in Table 1.b we present LS estimates of term spread regression models, for \( j = \{12, 24\} \) and \( k = \{24, 36\} \) months ahead (see also Estrella and Hardouvelis [57]), using the following spreads: \( [r_{it}(2y) - r_{it}(1y)] \) and \( [r_{it}(3y) - r_{it}(1y)] \) as regressors, respectively. The results of this table clearly indicate that the term spread contains also important information about marginal changes in future economic activity, for all \( k \) and \( j \) examined. The estimates of slope coefficients \( \beta^{(j,k)}_i \), reported in the tables, have the correct sign and are significant, for all \( j \) and \( k \) considered. Note that, for some countries (i.e., Germany and Japan), the forecasting power of these term spread regression models is higher than that of model (4.1), predicting cumulative growth changes \( g_{it,t+k} \).
CHAPTER 4. FORECASTING ECONOMIC ACTIVITY FROM YIELD CURVE FACTORS

Table 1.a: Forecasting economic activity from term spread $spr_{it}(5y)$

<table>
<thead>
<tr>
<th>Horizon $k$ (months)</th>
<th>US</th>
<th>CA</th>
<th>DE</th>
<th>UK</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_{i}^{(k)}$</td>
<td>$\beta_{i}^{(k)}$</td>
<td>$\beta_{i}^{(k)}$</td>
<td>$\beta_{i}^{(k)}$</td>
<td>$\beta_{i}^{(k)}$</td>
</tr>
<tr>
<td>$g_{it,t+k} = \text{const} + \beta_{i}^{(k)} spr_{it}(5y) + \varepsilon_{it,k}$, with $spr_{it}(5y) = r_{it}(5y) - r_{it}(3m)$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>3</td>
<td>0.53</td>
<td>0.007</td>
<td>2.01</td>
<td>0.12</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.48)</td>
<td>(0.86)</td>
<td>(0.35)</td>
<td>(1.18)</td>
</tr>
<tr>
<td>6</td>
<td>0.67</td>
<td>0.02</td>
<td>1.92</td>
<td>0.16</td>
<td>1.72</td>
</tr>
<tr>
<td></td>
<td>(0.62)</td>
<td>(0.45)</td>
<td>(0.89)</td>
<td>(0.37)</td>
<td>(1.47)</td>
</tr>
<tr>
<td>12</td>
<td>1.18</td>
<td>0.07</td>
<td>1.71</td>
<td>0.24</td>
<td>2.07</td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(0.32)</td>
<td>(1.09)</td>
<td>(0.27)</td>
<td>(1.40)</td>
</tr>
<tr>
<td>24</td>
<td>1.37</td>
<td>0.22</td>
<td>1.21</td>
<td>0.23</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(0.24)</td>
<td>(0.56)</td>
<td>(0.11)</td>
<td>(0.89)</td>
</tr>
</tbody>
</table>

Notes: The table presents estimates of the slope coefficients of term spread forecasting regressions (4.1), for the United states (US), Canada (CA), Germany (DE), the United Kingdom (UK) and Japan (JP). Term spread $spr_{it}(5y)$ is defined as $spr_{it}(5y) = r_{it}(5y) - r_{it}(3m)$ and $g_{it,t+k}$ as $g_{it,t+k} = 100(12/k)(\ln g_{it+k} - \ln g_{it})$, where $k$ denotes a forecasting horizon ahead. Newey-West standard errors corrected for heteroscedasticity and moving average errors up to $k$- periods ahead are reported in parentheses. $R^2$ is the coefficient of determination.

Table 1.b: Forecasting future marginal changes in economic activity from spreads

<table>
<thead>
<tr>
<th>Horizon $k$ (months)</th>
<th>US</th>
<th>CA</th>
<th>DE</th>
<th>UK</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_{i}^{(j)}$</td>
<td>$\beta_{i}^{(j)}$</td>
<td>$\beta_{i}^{(j)}$</td>
<td>$\beta_{i}^{(j)}$</td>
<td>$\beta_{i}^{(j)}$</td>
</tr>
<tr>
<td>$g_{it+j,t+k} = \text{const} + \beta_{i}^{(j,k)} [r_{it}(2y) - r_{it}(1y)] + \varepsilon_{t+j}$, for $k = 24, j = 12$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>24</td>
<td>0.48</td>
<td>0.17</td>
<td>0.24</td>
<td>0.07</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.10)</td>
<td>(0.26)</td>
<td>(0.05)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$g_{it+j,t+k} = \text{const} + \beta_{i}^{(j,k)} [r_{it}(3y) - r_{it}(1y)] + \varepsilon_{t+j}$, for $k = 36, j = 24$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>36</td>
<td>0.03</td>
<td>0.10</td>
<td>0.05</td>
<td>0.14</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>

Notes: The table presents estimates of the slope coefficients of term spread regressions forecasting marginal changes of economic activity $g_{it+j,t+k}$ between two future periods $t+k-j$ and $t+k$, for $j = \{12, 24\}$ and $k = \{24, 36\}$ months, based on the following term spreads: $r_{it}(2y) - r_{it}(1y)$ and $r_{it}(3y) - r_{it}(1y)$, respectively. Newey-West standard errors corrected for heteroscedasticity and moving average errors up to $j$- period ahead are reported in parentheses. $R^2$ is the coefficient of determination.
4.3 Forecasting economic activity from the slope and curvature factors

The term spread $spr_{it}(5y)$, used as regressor in model (4.1), contains composite information about future economic activity. As it is argued in many recent macroeconomic studies (see introduction), movements in future economic activity may be independently related to current changes in the slope and curvature factors of the yield curve.

To address the above questions, we first need to retrieve estimates of the slope and curvature factors from the yield curve. To this end, in this section we fit the dynamic term structure model of Nelson and Siegel [90], denoted as (DNSM), into the yield curve, for all countries $i$. This model enables us to decompose term spread $spr_{it}(5y)$ into the slope and curvature factors of the yield curve, by writing interest rates $r_{it}(\tau)$ in state-space form as follows:\footnote{See also Diebold et al [47] and Diebold and Li [44].}

$$r_{it}(\tau) = l_{it} + s_{it} \left( \frac{1 - e^{-\lambda_i \tau}}{\lambda_i \tau} \right) + c_{it} \left( \frac{1 - e^{-\lambda_i \tau}}{\lambda_i \tau} - e^{-\lambda_i \tau} \right), \text{ for all } i,$$

where $\tau = \{\tau_1, \tau_2, ..., \tau_n\}$ denote maturity intervals, and $l_{it}, s_{it}$ and $c_{it}$ are latent variables which denote the three factors spanning the term structure of interest rates $r_{it}(\tau)$, for all $\tau$. In particular, $l_{it}$ denotes the level factor of the yield curve (referred to as term structure of interest rates $r_{it}(\tau)$) causing parallel shifts to $r_{it}(\tau)$, for all $\tau$, which are often attributed to changes in long-run expectations about inflation. $s_{it}$ denotes the slope factor of the yield curve. This factor converges to unity, as $\tau \to 0$, and to zero, as $\tau \to \infty$, for all $t$. Thus, it can reflect the effects of changes in future business cycle conditions on $r_{it}(\tau)$. These die out in the long run. $c_{it}$ denotes the curvature factor of the term structure. This component of $r_{it}(\tau)$ converges to zero as $\tau \to 0$ and $\tau \to \infty$, which means that it is concave in $\tau$. Its effects on $r_{it}(\tau)$ are more profound for short and medium term interest rates (see also Christensen et al. [30], inter alia). Finally, parameter $\lambda_i$ determines the exponentially decaying effects of factors $s_{it}$ and $c_{it}$ on $r_{it}(\tau)$.

Taking the spread between two different maturity interest rates, i.e., $r_{it}(\tau_l)$ and $r_{it}(\tau_s)$, where $\tau_l$ and $\tau_s$ stand for the long and short-end maturity intervals, respectively, equation (4.2) implies that term spread $spr_{it}(\tau_l)$ is determined by the slope and curvature factors $s_{it}$ and $c_{it}$,
respectively, i.e.,

$$spr_{it}(\tau_i) \equiv r_{it}(\tau_i) - r_{it}(\tau_s) = \gamma_s i_t + \gamma_c c_{it},$$  \hspace{1cm} (4.3)$$

for all \(i\), where

$$\gamma_s i_t = \left[ \frac{1 - e^{-\lambda_i \tau_1}}{\lambda_i \tau_l} - \left( \frac{1 - e^{-\lambda_i \tau_s}}{\lambda_i \tau_s} \right) \right]$$

and

$$\gamma_c c_{it} = \left[ \frac{1 - e^{-\lambda_i \tau_1}}{\lambda_i \tau_l} - e^{-\lambda_i \tau_1} - \left( \frac{1 - e^{-\lambda_i \tau_s}}{\lambda_i \tau_s} - e^{-\lambda_i \tau_s} \right) \right].$$

The level factor \(l_t\) is cancelled out from term spread \(spr_{it}(\tau_l)\). The slope coefficients \(\gamma_s i_t\) and \(\gamma_c c_{it}\) of the last relationship depend on maturity intervals \(\tau_s\) and \(\tau_l\), and parameter \(\lambda_i\). The patterns of \(\gamma_s i_t\) and \(\gamma_c c_{it}\) with respect to \(\tau_s\), \(\tau_l\) and \(\lambda_i\) will be studied latter on, after estimating \(\lambda_i\) from the data. These can indicate how fast the effects of a change in factors \(s_{it}\) and \(c_{it}\) on term spread \(spr_{it}(\tau_l)\) slow down.

### 4.3.1 Retrieving yield curve factors \(s_{it}\) and \(c_{it}\)

To retrieve estimates of factors \(s_{it}\) and \(c_{it}\), next we fit the DNSM into our term structure of interest rates data. This is done through the application of the Kalman filter, by writing measurement equation (4.2) as follows:

$$r_{it} = \Gamma_i(\lambda_i)x_{it} + \varepsilon_{it},$$  \hspace{1cm} (4.4)$$

where \(r_{it} = (r_{it}(\tau_1), ..., r_{it}(\tau_N))'\), \(N = 17\) denote the different maturity intervals used in our estimation, for all \(i\)^4 \(\Gamma_i(\lambda_i)\) is an \((N \times 3)\)-dimension matrix of loading coefficients, defined as

$$\Gamma_i(\lambda_i) = 
\begin{bmatrix}
1 & \frac{1 - e^{-\lambda_i \tau_1}}{\lambda_i \tau_l} & \frac{1 - e^{-\lambda_i \tau_1}}{\lambda_i \tau_l} - e^{-\lambda_i \tau_1} \\
1 & \frac{1 - e^{-\lambda_i \tau_2}}{\lambda_i \tau_l} & \frac{1 - e^{-\lambda_i \tau_2}}{\lambda_i \tau_l} - e^{-\lambda_i \tau_2} \\
... & ... & ... \\
1 & \frac{1 - e^{-\lambda_i \tau_N}}{\lambda_i \tau_l} & \frac{1 - e^{-\lambda_i \tau_N}}{\lambda_i \tau_l} - e^{-\lambda_i \tau_N}
\end{bmatrix}.$$  

^4In particular, we estimate the DNSM model based on the following maturities \(\tau = \{3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, 120\}\) months, for all countries \(i\).
4.3. FORECASTING ECONOMIC ACTIVITY FROM THE SLOPE AND CURVATURE FACTORS

where \( \varepsilon_{it} \sim NIID(0, \Sigma_\varepsilon) \) and \( x_{it} = (l_{it}, s_{it}, c_{it})' \) is the vector of state variables. Vector \( x_{it} \) is assumed that follows a vector autoregressive process of lag order one, i.e.,

\[
\begin{bmatrix}
l_{it} \\
s_{it} \\
c_{it}
\end{bmatrix} =
\begin{bmatrix}
\mu_l \\
\mu_s \\
\mu_c
\end{bmatrix} +
\begin{bmatrix}
\phi_{11} & \phi_{12} & \phi_{13} \\
\phi_{21} & \phi_{22} & \phi_{23} \\
\phi_{31} & \phi_{32} & \phi_{33}
\end{bmatrix}
\begin{bmatrix}
l_{i t-1} \\
s_{i t-1} \\
c_{i t-1}
\end{bmatrix} +
\begin{bmatrix}
\eta_{l t} \\
\eta_{s t} \\
\eta_{c t}
\end{bmatrix},
\]  
(4.5)

or

\[
x_{it} = \mu + \Phi x_{it-1} + \eta_{it},
\]  
(4.6)

where \( \eta_{it} = (\eta_{l t}^l, \eta_{s t}^s, \eta_{c t}^c)' \), with \( \eta_{it} \sim NIID(0, \Sigma_\eta) \). Equations (4.4) and (4.6) constitute a state space system, which can be written in a more compact form as follows:

\[
r_{it} = \Gamma_{i}(\lambda_{i})x_{it} + \varepsilon_{it}, \quad \text{with} \quad x_{it} = \mu + \Phi x_{it-1} + \eta_{it},
\]

where

\[
\begin{bmatrix}
\varepsilon_{it} \\
\eta_{it}
\end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_\varepsilon & 0 \\ 0 & \Sigma_\eta \end{bmatrix} \right),
\]

\( \Sigma_\varepsilon \) and \( \Sigma_\eta \) are the variance-covariance matrices of error terms \( \varepsilon_{it} \) and \( \eta_{it} \), respectively. Note that error terms \( \varepsilon_{it} \) and \( \eta_{it} \) are assumed to be uncorrelated. This is a standard assumption made in the empirical literature (see, e.g., Diebold et al. [47]). In Table 2, we report estimates of the parameters of the DNSM, for all countries \( i \). Estimates of factors \( l_{it}, s_{it} \) and \( c_{it} \) are graphically presented in Figures 2.a-2.c. In Tables 3.a and 3.b, we present descriptive statistics and correlation coefficients among the estimates of \( l_{it}, s_{it} \) and \( c_{it} \), respectively. The results of these tables can be used to investigate stochastic properties of the three yield curve factors which have economic interest.
<table>
<thead>
<tr>
<th></th>
<th>United States (US)</th>
<th></th>
<th>Canada (CA)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi$</td>
<td>0.96</td>
<td>-0.003</td>
<td>0.03</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.007)</td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>0.97</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.008)</td>
<td>(0.006)</td>
<td>(0.01)</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.02</td>
<td>0.88</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$\Sigma_\eta$</td>
<td>0.12</td>
<td>-0.10</td>
<td>-0.13</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td></td>
<td>0.16</td>
<td>0.13</td>
<td>0.94</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.01)</td>
<td>(0.06)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>6.87</td>
<td>-2.99</td>
<td>-1.12</td>
<td>6.56</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.70)</td>
<td>(0.46)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.04</td>
<td></td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents estimates of (4.4) and (4.6) for the United States (US), Canada (CA), Germany (DE), United Kingdom (UK) and Japan (JP). Our sample consists of 265 monthly observations from 1987:05 to 2009:05. Standard errors are reported in parentheses.
### Table 2 (continued): Kalman filter estimates of (4.4) and (4.6)

<table>
<thead>
<tr>
<th>Germany (DE)</th>
<th>United Kingdom (UK)</th>
<th>Japan (JP)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Φ</strong></td>
<td><strong>Φ</strong></td>
<td><strong>Φ</strong></td>
</tr>
<tr>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>-0.004</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>0.01</td>
<td>0.01</td>
<td>0.005</td>
</tr>
<tr>
<td>0.94</td>
<td>0.98</td>
<td>0.93</td>
</tr>
<tr>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>0.02</td>
<td>-0.03</td>
<td>-0.04</td>
</tr>
<tr>
<td>0.05</td>
<td>-0.05</td>
<td>0.11</td>
</tr>
<tr>
<td>0.87</td>
<td>0.88</td>
<td>0.80</td>
</tr>
<tr>
<td>0.03</td>
<td>(0.03)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>0.01</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>0.02</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>0.02</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td><strong>Σ_η</strong></td>
<td><strong>Σ_η</strong></td>
<td><strong>Σ_η</strong></td>
</tr>
<tr>
<td>0.08</td>
<td>0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>-0.07</td>
<td>-0.08</td>
<td>-0.07</td>
</tr>
<tr>
<td>-0.12</td>
<td>-0.08</td>
<td>-0.04</td>
</tr>
<tr>
<td>(0.006)</td>
<td>(0.009)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>0.13</td>
<td>0.26</td>
<td>0.10</td>
</tr>
<tr>
<td>0.09</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>0.02</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>0.76</td>
<td>0.70</td>
<td>0.43</td>
</tr>
<tr>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.03)</td>
</tr>
<tr>
<td><strong>μ</strong></td>
<td><strong>μ</strong></td>
<td><strong>μ</strong></td>
</tr>
<tr>
<td>6.20</td>
<td>6.39</td>
<td>3.55</td>
</tr>
<tr>
<td>-2.57</td>
<td>-0.96</td>
<td>-1.78</td>
</tr>
<tr>
<td>-2.24</td>
<td>0.28</td>
<td>-1.99</td>
</tr>
<tr>
<td>(0.81)</td>
<td>(1.77)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>(0.97)</td>
<td>(1.20)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>(0.79)</td>
<td>(0.76)</td>
<td>(0.17)</td>
</tr>
<tr>
<td><strong>λ</strong></td>
<td><strong>λ</strong></td>
<td><strong>λ</strong></td>
</tr>
<tr>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.006)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>
CHAPTER 4. FORECASTING ECONOMIC ACTIVITY FROM YIELD CURVE FACTORS

Figure 2a: Estimates of level factors $l_{it}$.

Figure 2b: Estimates of slope factors $s_{it}$.

Figure 2c: Estimates of curvature factors $c_{it}$. 
4.3. FORECASTING ECONOMIC ACTIVITY FROM THE SLOPE AND CURVATURE FACTORS

Table 3.a: Descriptive statistics of the estimates of yield curve factors

<table>
<thead>
<tr>
<th>Country</th>
<th>US</th>
<th>CA</th>
<th>UK</th>
<th>DE</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( l_{it} )</td>
<td>mean</td>
<td>6.80</td>
<td>6.90</td>
<td>6.90</td>
<td>6.34</td>
</tr>
<tr>
<td></td>
<td>st. dev</td>
<td>1.48</td>
<td>2.12</td>
<td>1.57</td>
<td>1.57</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>3.91</td>
<td>3.12</td>
<td>3.49</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>9.78</td>
<td>11.37</td>
<td>12.37</td>
<td>7.35</td>
</tr>
<tr>
<td></td>
<td>( \rho(1) )</td>
<td>0.97</td>
<td>0.98</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>( \rho(12) )</td>
<td>0.84</td>
<td>0.91</td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>( s_{it} )</td>
<td>mean</td>
<td>-2.32</td>
<td>-1.50</td>
<td>-1.88</td>
<td>3.81</td>
</tr>
<tr>
<td></td>
<td>st. dev</td>
<td>2.08</td>
<td>1.99</td>
<td>1.87</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>-6.54</td>
<td>-5.56</td>
<td>-5.73</td>
<td>-5.47</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>1.00</td>
<td>3.64</td>
<td>2.76</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>( \rho(1) )</td>
<td>0.97</td>
<td>0.96</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>( \rho(12) )</td>
<td>0.50</td>
<td>0.46</td>
<td>0.50</td>
<td>0.46</td>
</tr>
<tr>
<td>( c_{it} )</td>
<td>mean</td>
<td>-1.20</td>
<td>-1.23</td>
<td>-1.81</td>
<td>-2.04</td>
</tr>
<tr>
<td></td>
<td>st. dev</td>
<td>2.13</td>
<td>1.84</td>
<td>2.11</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>-11.10</td>
<td>-7.22</td>
<td>-6.32</td>
<td>-5.47</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>2.69</td>
<td>2.96</td>
<td>4.10</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>( \rho(1) )</td>
<td>0.88</td>
<td>0.79</td>
<td>0.91</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>( \rho(12) )</td>
<td>0.34</td>
<td>0.22</td>
<td>0.19</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Notes: The table presents descriptive statistics of the estimates of yield curve factors \( l_{it} \), \( s_{it} \) and \( c_{it} \), namely their mean, standard deviation, minimum and maximum values, and autocorrelation coefficients of one month, one and two years.

Table 3.b: Correlation among yield curve factors and term spread \( spr_{it} \)

<table>
<thead>
<tr>
<th>Country</th>
<th>( s_{it} )</th>
<th>( c_{it} )</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>( spr_{it}; s_{it} )</td>
<td>-0.86</td>
<td>-0.90</td>
<td>-0.96</td>
</tr>
<tr>
<td>( spr_{it}; c_{it} )</td>
<td>-0.02</td>
<td>0.48</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Notes: The table presents values of correlation coefficients between slope \( s_{it} \) and curvature \( c_{it} \) factors, as well as between these factors and term spread \( spr_{it} = r_{it}(5y) - r_{it}(3m) \), for all countries examined.

The results of Tables 3.a-3.b and Figures 2.a-2.c indicate that, as was expected, the level factor \( l_{it} \) takes positive values which are very highly correlated among all countries \( i \), thus implying common shifts in the levels of interest rates \( r_{it} \), for all \( i \). The slope factor \( s_{it} \) is also substantially correlated, for all countries \( i \), with the exception of Germany and Japan with the
US. These results indicate that possible future business cycles conditions reflected in the slope factor exhibit significant similarities across all countries examined. The exceptions for Germany and Japan can be attributed to the recessions of these two countries occurred in the nineties, due to Germany’s unification and Japan’s financial markets regulations mentioned before. The curvature factor \( c_{it} \) is less correlated among the countries, compared to \( s_{it} \). Thus, it may be affected more from domestic factors influencing, separately, the yield curve of each country. The estimates of coefficients \( \lambda_i \), reported in Table 2, are very small in magnitude, for all countries \( i \), varying between 0.04 and 0.06. These values of \( \lambda \) imply that the loading coefficients \( \gamma_{si} \) and \( \gamma_{ci} \) of factors \( s_{it} \) and \( c_{it} \) into term spread \( spr_{it}(\tau_i) \equiv r_{it}(\tau_i) - r_{it}(\tau_s) = \gamma_{si}s_{it} + \gamma_{ci}c_{it} \) will be quite persistent with respect to maturity interval \( \tau_i - \tau_s \). To see this more clearly, in Figure 3 we graphically present estimates of \( \gamma_{si} \) and \( \gamma_{ci} \) with respect to different maturity intervals \( \tau_s \) of spread \( r_{it}(\tau_i) - r_{it}(\tau_s) \), for \( \tau_s = \{3m, 6m, 1y, 3y, 4y\} \), keeping fix the long-term maturity interval \( \tau_l \) to \( \tau_l = \{5y\} \).

![Figure 3. Loading coefficients \( \gamma_{si} \) and \( \gamma_{ci} \) with respect maturity interval \( \tau_l - \tau_s \), for \( \tau_l = 5y \) and \( \tau_s = \{3m, 6m, 1y, 2y, 3y, 4y\} \)](image)

The results of this figure clearly show that changes in the slope factor \( s_{it} \) have more persistent effects on term spread, compared to those of the curvature factor. Changes in \( s_{it} \) determine the
4.3. FORECASTING ECONOMIC ACTIVITY FROM THE SLOPE AND CURVATURE FACTORS

Slope of the yield curve even at its long-end, i.e., \( r_{it}(5y) - r_{it}(3y) \). In contrast, the effects of changes in \( c_{it} \) on the yield slope cease more shortly, i.e., after one (or two) years. These results mean that the forecasting ability of term spread \( r_{it}(5y) - r_{it}(3y) \) on future marginal changes of economic activity at long term horizons, \( g_{it+k-j,t+k} \), implied by the results of Table 1.b can be mainly attributed to slope factor \( s_{it} \). This will be investigated more formally in the next section.

4.3.2 Forecasting economic activity based on yield factors \( s_{it} \) and \( c_{it} \)

Having obtained estimates of factors \( s_{it} \) and \( c_{it} \) from our term structure data, in this section we estimate the following regression model forecasting future economic growth rate \( g_{it+k} \):

\[
g_{it+k} = \text{const} + \beta_s s_{it} + \beta_c c_{it} + \varepsilon_{it+k}, \text{ for all } i,
\]

(4.7)

using yield factors \( s_{it} \) and \( c_{it} \) as independent regressors. This is done for the same forecasting horizons \( k \), considered in the estimation of the term spread regression (4.1), i.e., \( k = \{3, 6, 12, 24\} \) months (see Tables 1.a-1.b). Table 4.a presents LS estimates of the above regression model. Newey-West standard errors, correcting for MA errors (due to the overlapping nature of the data) and heteroscedasticity are reported in parentheses.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>US</th>
<th>CA</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k )</td>
<td>( \beta_s^{(k)} )</td>
<td>( \beta_c^{(k)} )</td>
<td>( R^2 )</td>
</tr>
<tr>
<td>3</td>
<td>-0.38</td>
<td>1.68</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.36)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>6</td>
<td>-0.41</td>
<td>1.43</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.42)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>12</td>
<td>-0.55</td>
<td>1.12</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.53)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>24</td>
<td>-0.64</td>
<td>0.53</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.26)</td>
<td>(0.22)</td>
</tr>
</tbody>
</table>
### Table 4.a: Forecasts of economic activity from slope and curvature factors (cont’d)

<table>
<thead>
<tr>
<th>Horizon $k$</th>
<th>UK</th>
<th></th>
<th></th>
<th>JP</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_s^{(k)}$</td>
<td>$\beta_c^{(k)}$</td>
<td>$R^2$</td>
<td>$\hat{\beta}_s^{(k)}$</td>
<td>$\hat{\beta}_c^{(k)}$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>3</td>
<td>-0.19</td>
<td>0.49</td>
<td>0.03</td>
<td>-1.62</td>
<td>1.33</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.28)</td>
<td>(0.60)</td>
<td>(0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-0.38</td>
<td>0.26</td>
<td>0.07</td>
<td>-1.32</td>
<td>0.97</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.29)</td>
<td>(0.71)</td>
<td>(0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-0.52</td>
<td>0.12</td>
<td>0.17</td>
<td>-1.09</td>
<td>0.88</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.30)</td>
<td>(0.76)</td>
<td>(0.65)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>-0.51</td>
<td>-0.18</td>
<td>0.30</td>
<td>-0.30</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.11)</td>
<td>(0.60)</td>
<td>(0.42)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents LS estimates of the slope coefficients $\beta_s^{(k)}$ and $\beta_c^{(k)}$ of regression model (4.7), forecasting $g_{it,t+k}$ from yield curve factors $s_{it}$ and $c_{it}$, for US, Canada, Germany, UK and Japan. The sample is from 1989:1 to 2009:5. Newey-West standard errors corrected for heteroscedasticity and moving average errors up to $k$—periods ahead are reported in parentheses. $R^2$ is the coefficient of determination.

The results of Table 4.a indicate that both factors $s_{it}$ and $c_{it}$ contain independent information about future economic activity. The values of the coefficient of determination $R^2$, reported in the table, indicate that $s_{it}$ and $c_{it}$ have higher forecasting power on future economic growth rate $g_{it,t+k}$, compared to term spread $spr_{it}(5y)$ (see Table 1.a) This is true for all countries $i$ examined. The slope factor $s_{it}$ contains significant information about future economic growth rate $g_{it,t+k}$ for all forecasting horizons examined. Its slope coefficient, $\beta_s$, has negative sign, for all $i$ and $k$, which is consistent with the macroeconomic interpretation given to factor $s_{it}$ that reflects changes in future business cycle conditions. The negative sign of $\beta_s$ implies that a flattened, or negative, term spread (or yield curve) will be followed by a slow down in economic activity, after a few periods ahead.

In contrast to $s_{it}$, the curvature factor $c_{it}$ is found to contain important information about future economic activity $g_{it,t+k}$ only for short horizons ahead, i.e., for $k = \{3, 6\}$ months. For forecasting horizons higher than 12 months months ahead, this factor does not seem to contain significant information about future levels of $g_{it,t+k}$, with the only exception of the US. The sign of the slope coefficient of this factor, $\beta_c$, is positive for all forecasting horizons up to $k = 12$
4.3. FORECASTING ECONOMIC ACTIVITY FROM THE SLOPE AND CURVATURE FACTORS

months ahead. This means that a positive (or negative) shock to this factor is associated with future economic growth (or slow down), which is opposite to what happens with a positive (or negative) shock in $s_{it}$. In term spread regressions like (4.1), note that these effects of $c_{it}$ on $g_{it+k}$ are offset by those of $s_{it}$. This happens because they have opposite sign, as the analysis of the previous section has shown.

The more temporary in nature and different in sign forecasting ability of $c_{it}$ about future economic activity than $s_{it}$ is consistent with the macroeconomic interpretation given to this factor by Dewachter and Lyrio [42], inter alia. It is considered that captures policy actions beyond the endogenous responses of monetary authorities to inflation and output gap deviations from their target rates, which the business cycle factor $s_{it}$ summarizes. For example, changes in $c_{it}$ can be associated with changes in the current stance of monetary policy with the aim of tightening monetary policy in the short and medium terms, if economic growth or inflation are undesirably high. These changes in the stance of monetary policy can anchor expectations about future inflation and output pressures, and will thus reduce the term premia effects embodied in the yield curve. This will result in an increase of interest rates of intermediate maturities relative to the short-term rate, as also noted by Moench [88]. Thus, a positive shock in curvature factor $c_{it}$ will be associated with an increase in future economic activity in short and medium horizons.

The above interpretation of curvature factor $c_{it}$ means that the ability of term spread $spr_{it}(5y) = r_{it}(5y) - r_{it}(3m)$ to forecast future marginal changes in output growth rate $g_{it+k-j,t+k}$, between future periods $t + k - j$ and $t + k$ (see Table 1.b), can be solely attributed to slope factor $s_{it}$. To see if this is true, in Table 4.b. we present LS estimates of regression model (4.7), using $g_{it+k-j,t+k}$ as dependent variable. As in Table 1.b, this is done for horizons $k = \{24, 36\}$ and $j = \{12, 24\}$ ahead. The results of this table are consistent with the above macroeconomic interpretation of factor $c_{it}$. For all countries examined, $c_{it}$ does not have any forecasting power on $g_{it+k-j,t+k}$. In contrast, slope factor $s_{it}$ successfully forecasts future changes in economic activity between future periods $t + k - j$ and $t + k$. The results of Table 4.b are in accordance to those of Table 1.b and Figure 3, which imply that any predictive power of term spread $spr_{it}(5y)$ on future economic activity at longer horizons (i.e., higher than one year ahead) lies in its slope factor $s_{it}$. 
### Table 4.b: Marginal forecasts of economic activity from slope and curvature factors

Model: \( g_{it+k-j,t+k} = \text{const} + \beta_s^{(k)} s_{it} + \beta_c^{(k)} c_{it} + \varepsilon_{it+j}, \) for \( k = 24, j = 12 \)

<table>
<thead>
<tr>
<th>( \beta_s^{(k)} )</th>
<th>( \beta_c^{(k)} )</th>
<th>( R^2 )</th>
<th>( \beta_s^{(k)} )</th>
<th>( \beta_c^{(k)} )</th>
<th>( R^2 )</th>
<th>( \beta_s^{(k)} )</th>
<th>( \beta_c^{(k)} )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>CA</td>
<td>DE</td>
<td>US</td>
<td>CA</td>
<td>DE</td>
<td>US</td>
<td>CA</td>
<td>DE</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.04</td>
<td>0.17</td>
<td>-0.05</td>
<td>-0.01</td>
<td>0.08</td>
<td>-0.06</td>
<td>-0.02</td>
<td>0.11</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>JP</td>
<td></td>
<td>UK</td>
<td>JP</td>
<td></td>
<td>UK</td>
<td>JP</td>
<td></td>
</tr>
<tr>
<td>-0.04</td>
<td>-0.02</td>
<td>0.14</td>
<td>0.07</td>
<td>-0.06</td>
<td>0.02</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model: \( g_{it+k-j,t+k} = \text{const} + \beta_s^{(k)} s_{it} + \beta_c^{(k)} c_{it} + \varepsilon_{it+j}, \) for \( k = 36, j = 24 \)

<table>
<thead>
<tr>
<th>( \beta_s^{(k)} )</th>
<th>( \beta_c^{(k)} )</th>
<th>( R^2 )</th>
<th>( \beta_s^{(k)} )</th>
<th>( \beta_c^{(k)} )</th>
<th>( R^2 )</th>
<th>( \beta_s^{(k)} )</th>
<th>( \beta_c^{(k)} )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>CA</td>
<td>DE</td>
<td>US</td>
<td>CA</td>
<td>DE</td>
<td>US</td>
<td>CA</td>
<td>DE</td>
</tr>
<tr>
<td>-0.06</td>
<td>0.03</td>
<td>0.22</td>
<td>-0.03</td>
<td>0.002</td>
<td>0.08</td>
<td>-0.04</td>
<td>-0.02</td>
<td>0.21</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>JP</td>
<td></td>
<td>UK</td>
<td>JP</td>
<td></td>
<td>UK</td>
<td>JP</td>
<td></td>
</tr>
<tr>
<td>-0.02</td>
<td>-0.02</td>
<td>0.12</td>
<td>0.01</td>
<td>-0.06</td>
<td>0.05</td>
<td>(0.007)</td>
<td>(0.01)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>(0.007)</td>
<td>(0.01)</td>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents LS estimates of the slope coefficients \( \beta_s^{(k)} \) and \( \beta_c^{(k)} \) of regression model (4.7), forecasting marginal changes of economic growth rate \( g_{it+k-j,t+k} \), between two future periods \( t+k-j \) and \( t+k \), for \( j = \{12, 24\} \) and \( k = \{24, 36\} \) months. The sample is from 1989:1 to 2009:5. Newey-West standard errors corrected for heteroscedasticity and moving average errors up to \( k \)-periods ahead are reported in parentheses. \( R^2 \) is the coefficient of determination.

### 4.3.3 Out-of-sample forecasting performance

In this section we investigate if the superior information contained in factors \( s_{it} \) and \( c_{it} \) about future economic activity than term spread \( \text{spr}_{it}(5y) \), found by our in-sample estimates in the previous section, also holds for out-of-sample. To this end, we compare the out-of-sample
4.3. **Forecasting Economic Activity from the Slope and Curvature Factors**

Forecasting performance of yield factor model (4.7) to that of term spread model (4.1).

Table 5 presents values of the mean square error (MSE) and mean absolute error (MAE) metrics for the above two models. It also reports values of Diebold-Mariano [45], denoted as (DM), and Giacomini and Rossi [62], denoted as GR, test statistics. A negative and significantly different than zero value of DM statistic means that model (4.7) provides smaller in magnitude errors than (4.1), and thus it rejects the null hypothesis that the two models have the same forecasting ability. The GR statistic can test if the two models can produce consistent forecasts with their in-sample ones, which means that they do not suffer from structural breaks problems.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>MSE</th>
<th>MAE</th>
<th>DM</th>
<th>GR</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>0.31</td>
<td>0.63</td>
<td>-2.56**</td>
<td>0.16</td>
</tr>
<tr>
<td>MAE</td>
<td>3.98</td>
<td>5.67</td>
<td>-2.31*</td>
<td>0.27</td>
</tr>
<tr>
<td>DM</td>
<td>-2.96**</td>
<td>-0.61</td>
<td>-2.53**</td>
<td></td>
</tr>
<tr>
<td>GR</td>
<td>0.16</td>
<td>0.27</td>
<td>0.16</td>
<td>0.09</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>0.27</td>
<td>0.45</td>
<td>-2.55**</td>
<td>0.09</td>
</tr>
<tr>
<td>MAE</td>
<td>3.56</td>
<td>4.67</td>
<td>-1.96*</td>
<td>0.10</td>
</tr>
<tr>
<td>DM</td>
<td>-2.35**</td>
<td>1.26</td>
<td>-3.27**</td>
<td></td>
</tr>
<tr>
<td>GR</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.12</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>0.17</td>
<td>0.15</td>
<td>-2.25*</td>
<td>-0.03</td>
</tr>
<tr>
<td>MAE</td>
<td>2.80</td>
<td>3.16</td>
<td>-1.00</td>
<td>-0.14</td>
</tr>
<tr>
<td>DM</td>
<td>3.00**</td>
<td>-1.42</td>
<td>-1.95*</td>
<td>-0.34</td>
</tr>
<tr>
<td>GR</td>
<td>-0.12</td>
<td>-0.34</td>
<td>-0.10</td>
<td>-0.53</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>0.07</td>
<td>0.08</td>
<td>-2.05*</td>
<td>-0.26</td>
</tr>
<tr>
<td>MAE</td>
<td>1.88</td>
<td>2.41</td>
<td>-1.40</td>
<td>-0.35</td>
</tr>
<tr>
<td>DM</td>
<td>2.50</td>
<td>3.85</td>
<td>-1.35</td>
<td>-1.36</td>
</tr>
<tr>
<td>GR</td>
<td>0.25</td>
<td>0.35</td>
<td>-0.40</td>
<td>-1.80</td>
</tr>
</tbody>
</table>

**Table 5: Out-of-sample forecasting performance for (4.7) and (4.1)**

Notes: The table presents values of the MSE and MAE metrics, and of the DM and GR test statistics assessing the forecasting performance of regression models (4.7) and (4.1). DM and GR denote the Diebold-Mariano and Giacomini-Rossi test statistics, respectively. These statistics follow the standard normal distribution. Note that the GR test statistic is an out-of-sample test statistic, which can test the stability of the out-of-sample forecasts of the above models compared to their in-sample one. To
calculate the out-of-sample values of the above metrics and statistics, we rely on recursive estimates of models (4.7) and (4.1) of economic activity after period 1999:04, by adding one observation at a time and, then, re-estimating the models until the end of sample. The total number of observations used in our out-of-sample forecasting exercise is \( n \equiv T - k - m + 1 \), where \( T = 245 \), the forecasting horizon is \( k = \{3, 6, 12, 24\} \) months and our in-sample window is \( m = 120 \) observations. All values concerning MSE and MAE are expressed in basis points. (*) and (**) mean significance at 5% and 1% level, respectively.

To carry out our out-of-sample forecasting exercise and calculate the values of the above metrics and statistics, we have recursively estimated regression models (4.7) and (4.1) after period 1999:04, by adding one observation at a time and, then, re-estimating the two models until the end of sample. The total number of observations used in our out-of-sample forecasting exercise is \( n \equiv T - k - m + 1 \), where \( T = 245 \) denotes the total sum of our sample observations, \( k = \{3, 6, 12, 24\} \) denotes the forecasting periods (months) ahead and \( m \) denotes our sample window. The latter is set to \( m = 120 \) observations. All reported values of the MSE and MAE are in percentage terms.

The results of Table 5 clearly indicate that regression model (4.7) provide better forecasts about future economic activity than term spread model (4.1), especially for short and medium horizons \( k \) ahead. For all cases of \( k \) and \( i \) (countries) examined, the values of MSE and MAE metrics, reported in the table, are smaller for model (4.7) than model (4.1), with the exception of Germany (DE) and United Kingdom (UK) for \( k = 12 \) and \( k = 6 \), respectively. The reported values of DM test statistic are consistent with the above results. These confirm the better forecasting performance of model (4.7) than model (4.1) confirmed at 1% and 5% significance levels. Finally, the values of the GR test indicate that the out-of-sample forecasts of model (4.7) are consistent with those in-sample. Thus, they are robust to possible structural breaks occurred during our sample.

4.4 Conclusions

Many recent studies use the term spread between the long and short-term interest rates to forecast future economic activity, or economic recessions. In this paper, we provide some new interesting results about the predicting ability of the yield curve and term spread. We indicate that the slope and curvature factors spanning the yield curve contain superior information
about future economic activity than the term spread itself. This is shown for five world leading economies. To extract the slope and curvature factors of the yield curve, the paper fits the dynamic model of Nelson-Siegel into term structure data of the above countries.

The paper presents clear cut evidence that the slope factor of the yield curve contain significant information about future economic activity over much longer horizons than the curvature factor, for all countries examined. The latter seems to affect the short (or medium) end of the yield curve. The sign of the predictions of the slope and curvature factors on future economic activity is different. They imply that an increase in the slope factor is associated with a slow down in economic activity, while the opposite is predicted for an increase in the curvature factor. These results are consistent with the theoretical predictions of recent macroeconomic studies asserting that the slope factor of the yield curve should reflect future changes in business cycle conditions, which can last for a few years ahead, while the curvature factor may be associated with short or medium term changes in the current stance of monetary policy. The fact that the slope and curvature yield factors have opposite in sign effects on the term spread can explain why the latter becomes less successful in predicting future economic activity over shorter, or medium, horizons, compared to a regression model using these two factors as independent variables. The effects of these two factors on the term spread are offset to each other, and thus reduce the ability of the term spread to forecast the correct direction of future changes in economic activity. The above results are also confirmed by out-of-sample tests.
Chapter 5

Forecasting inflation from the term structure and the inflation risk premia effects

5.1 Introduction

There is recently growing interest in the literature in retrieving market expectations about inflation and inflation risk premia based on the term structure of nominal and real interest rates (see, e.g., Christensen et al. [30], D’Amico et al. [40], Grishchenko and Huang [64]) or inflation swap rates (Haubrich et al. [73]). The real interest rates are obtained from inflation-indexed bonds such as the treasury inflation protected securities (TIPS) or inflation swap rates. Since inflation-indexed bonds are available for long-term maturities (i.e., for five years, or longer maturity intervals) and data on inflation swap rates start from 2003, the above studies are focused on retrieving inflation expectations and inflation risk premia from term structure data over long-term horizons. Thus, little is known about the market participants’ expected inflation and the importance of inflation risk premia over short-term horizons (i.e., up to one-year ahead), which is of great interest for monetary authorities and forecasting inflation. Furthermore, estimates of expected future inflation rates obtained from inflation-indexed bond markets are not net of the inflation risk premia effects.

To provide estimates of expected future inflation rates and analyze how important are inflation risk premia effects, especially over short-term horizons, in this paper we estimate an arbitrage-free, affine Gaussian dynamic term structure model (GDTSM) based on nominal interest rates, real consumption growth and inflation rate. Our model estimates the real term
structure of interest rates by fitting the GDTSM into the nominal term structure of interest rates and real consumption growth rate, simultaneously. Exploring information from real consumption data can help in better capturing the dynamics of the real term structure of interest rates (see, e.g., Berardi and Torous [17]). As in the empirical literature (see, e.g., Litterman and Scheinkman [81]) and, more recently, Argyropoulos and Tzavalis [7], our affine GDTSM assumes that nominal interest rates are spanned by three common factors. Two of them are unobserved and also span the real term structure of interest rates and real consumption growth. The third factor which spans the nominal term structure is assumed to be inflation rate. These assumptions are often made in the term structure literature (see Ang and Piazzesi [3], Dewachter and Lyrio [42], Ang et al. [4]). Although there is little macroeconomic structure in our model\(^1\), we specify our factor dynamics in a general way which allows for feedback and/or contemporaneous correlation effects between inflation and real interest rates. These specifications are in line with those assumed by Diebold et al. [47] and Christensen et al. [30].

To retrieve estimates of the two unobserved factors spanning the nominal and real term structure of interest rates, we rely on the approach of Pearson and Sun [92]. According to this, a number of zero-coupon interest rates are used as instruments to obtain the unobserved factors. This is done by inverting the pricing equations of zero-coupon bonds implied by the GDTSM. However, this approach relies on the assumption that these zero-coupon bond instruments are priced without measurement errors, which may not be true in practice. To overcome this problem, instead of observed values, we suggest employing projections (or forecasts) of the above interest rate instruments based on principal component factors spanning the term structure of interest rates. The latter are retrieved by principal components (PC) analysis. Since it is based on a very large set of different maturity interest rates, PC analysis can provide term structure factors which constitute well diversified portfolios of interest rates (see also Joslin et al. [77]), net of interest rate measurement errors.

The results of the paper lead to a number of interesting conclusions. First, they show that our model can provide estimates of real interest rates and expected inflation rates which are

\(^{1}\)Models with more structural macroeconomic specification in the literature are found in Hordahl et al [75], Rudebusch and Wu [102], among others. Also, standard new keynesian macro-finance models which encompass financial and macro variables can be found in Hordahl and Tristani [74].
very close to those provided in the literature based on survey data or inflation indexed bonds, for long-term horizons. Second, inflation risk premia are found to be negative and more volatile over short-term horizons, compared to long-term ones. This result can challenge empirical studies using the difference between nominal and real yields (based on inflation-indexed bonds) to retrieve market expectations about future inflation rates. These expectations are not net of risk premia effects. The negative sign of the inflation risk premium implies that investors would prefer to hold nominal bonds compared to inflation-indexed bonds. Third, as the maturity horizon increases, both the volatility of inflation risk premium are found to decline, considerably. A similar conclusion can be drawn for real interest rates. These results can explain the failure of the term spread between nominal interest rates to forecast future changes in inflation rates in short-run (see, e.g., Mishkin [87], Tzavalis and Wickens [107], Malliaropulos [84]). Fourth, we have found that our GDTSM model enables us to adjust term spread regressions for inflation risk premium and/or real interest rate effects in successfully forecasting future inflation rates, according to the predictions of the expectation hypothesis.

The paper is organized as follows. Section 2 presents the GDTSM and provides the closed form solution of inflation risk premia, implied by the theory. In Section 3, we fit the model into the data and present estimates of the real yields, expected inflation and inflation risk premia. This section also examines if the nominal term structure can successful forecast future inflation rates, net of inflation risk premium effects. Finally, Section 4 concludes the paper.

5.2 Model setup

State Variable Dynamics

In this section we setup a Gaussian dynamic term structure model (GDTSM) in order to describe the bond prices. Consider that all nominal bonds in the economy are characterized by $K$—state variables at time $t$, $x_{it}$, stacked into a $K$—dimensional vector $X_t$. These state variables obey the following uncorrelated Gaussian processes:

\[ dX_t = k(\theta - X_t)dt + \sigma dW_t \quad (5.1) \]

\[^2\text{See also Vasicek [109], Dai and Singleton [38], Ang and Piazzesi [3], Ahn [1], Berardi and Torous [17].}\]
where $W_t$ denotes a $K$-dimensional Wiener process, $k$ and $\sigma$ are $K \times K$ matrices of the speed of mean-reversion and the volatility of the factors respectively and $\theta$ is a $K$-dimensional vector of the factor long-run mean parameters.

**Inflation and Consumption Processes**

In our model we further specify at time $t$, real consumption and an exogenous price level for the consumption good which we denote as $C_t$ and $P_t$ respectively. The real consumption $C_t$ and the price level $P_t$ evolve according to Gaussian process (see e.g., Boudoukh [20], Veronesi [110] Bansal Yaron [13], Berardi and Torous [17], Berardi [16])

$$\frac{dP_t}{P_t} = \pi_t dt + \sigma_P dW_t$$

and

$$\frac{dC_t}{C_t} = \vartheta_t dt + \sigma_c dW_t$$

where, $\pi_t$, is the instantaneous expected inflation rate and $\vartheta_t$ is the drift of the growth rate of real consumption. In equilibrium, the drift of the growth rate of real consumption $\vartheta_t$ equals the drift of the dividend’s growth rate of the asset that the investor holds (see Lucas [83], Veronesi [110]). In our case, the asset is a real bond paying 1 unit of consumption, thus the growth rate of real consumption is the instantaneous real interest rate, $r_t^*$. The instantaneous expected inflation $\pi_t$ and instantaneous real interest rate, $r_t^*$ are both affine in the state variables:

$$\pi_t = \delta_\pi + \delta_\pi^t X_t$$

$$r_t^* = \delta_0^r + \delta_1^r X_t$$

where $\delta_0^*$ is a constant and $\delta_1^*$ is a $K \times 1$ vector. The expected inflation rate and the real consumption growth from current period $t$ to a any future period $t + \tau$, are also affine in the state variables and can be obtained by (Risa [100], Berardi and Torous [17]):

$$E_t[\Delta_\tau P_{t+\tau}] = g_0(\tau) + g_1(\tau)^t X_{it}$$

$$E_t[\Delta_\tau C_{t+\tau}] = \psi_0(\tau) + \psi_1(\tau)^t X_{it}$$

where $E_t[\Delta_\tau P_{t+\tau}] = E_t[\ln(P_{t+\tau}/P_t)]$, $E_t[\Delta_\tau C_{t+\tau}] = E_t[\ln(C_{t+\tau}/C_t)]$ and $g_0(\tau)$, $\psi_0(\tau)$ are
constants and \( g_1(\tau), \psi(\tau) \) are \((K \times 1)\)-dimension vectors defined as 
\[
g_1(\tau) = (I - e^{-k'\tau})(k')^{-1}\delta_{\pi}
\]
and 
\[
\psi_1(\tau) = (I - e^{-k'\tau})(k')^{-1}\delta_{\tau}.
\]

The Real and Nominal Term Structures

The time \( t \) price of a zero-coupon bond, \( B^*_t(\tau) \), paying one unit of consumption at time \( t + \tau \) is the conditional expectation value of the marginal rate of substitution between \( t \) and \( t + \tau \):
\[
B^*_t(\tau) = E_t \left( \frac{m_{t+\tau}}{m_t} \right),
\]
where \( m_t \) is the continuous time real stochastic discount factor (or pricing kernel). The real stochastic discount factor measures the growth in marginal utility and follows:
\[
dm_t = -r^*_t dt - \Lambda^*_t dW_t,
\]
where \( r^*_t \) is the instantaneous real interest rate or short-term rate and the \( 1 \times \) \( K \) vector \( \Lambda^*_t \) contains the risk pricing functions, \( \Lambda^*_t \), for each factor \( i = 1, 2, ..., K \). The dollar price \( B_t(\tau) \), of a nominal bond that pays out one dollar at periods from \( t \) is:
\[
B_t(\tau) = E_t \left( \frac{m_{t+\tau}}{m_t} P_t P_{t+\tau} \right) = E_t \left( \frac{M_{t+\tau}}{M_t} \right),
\]
where \( M_t \) is the continuous time nominal stochastic discount factor. As we show in (5.9), since real bonds can be thought as a nominal asset which pays realized inflation upon maturity, the real and nominal stochastic discount factors are linked through the no-arbitrage relationship:
\[
M_t = m_t / P_t
\]
The pricing kernel \( M_t \) follows the Gaussian process:
\[
\frac{dM_t}{M_t} = -r_t dt - \Lambda'_t dW_t \tag{5.10}
\]
where \( r_t \) is the instantaneous nominal interest rate or short-rate and \( \Lambda_t \) is the vector of the risk pricing functions, for each state variable as mentioned above. The short rate is affine to the \( K \) state variables:
\[
r_t = \delta_0 + \delta_1 X_t \tag{5.11}
\]
5.2. MODEL SETUP

where $\delta_0$ is a constant and $\delta_1$ is a $K \times 1$ vector. Since the risk pricing functions evaluate the $K$ independent sources of risk associated with the factors we model $\Lambda_t$. Following Duffee [51] for maximal flexibility in our model we set $\Lambda_t$ as:

$$\Lambda_t = \sigma^{-1}(\lambda_0 + \lambda_1 X_t)$$  \hspace{1cm} (5.12)

Where $\lambda_0$ is a $K \times 1$ vector and $\lambda_1$ is a diagonal $K \times K$ matrix. The risk neutral dynamics of the state vector (5.1) under the equivalent probability measure $Q$, are given by

$$dX_t = k^Q(\theta^Q - X_t)dt + \sigma dW^Q_t$$  \hspace{1cm} (5.13)

where $k^Q = k + \lambda_1$ and $\theta^Q = k^{Q-1}(k\theta - \lambda_0)$. Substituting (5.10), (5.11) and (5.12) into (5.9), and setting prices on nominal bonds to be exponentially affine to the $K$ state variables (see e.g., Dai and Singleton [38], Fisher [60]) we get that the price of a nominal bond at any maturity $\tau$ is:

$$B_t(\tau) = e^{-A(\tau)-D(\tau)'X_t}, \text{ for } \tau = 1, 2, ..., N$$  \hspace{1cm} (5.14)

where $X_t$ is a ($K \times 1$)-dimension vector collecting all the state variables $x_{it}$, $A(\tau)$ is a scalar function and $D(\tau)$ is a ($K \times 1$)-dimension vector, defined as $D(\tau) = (D_1(\tau), D_2(\tau), ..., D_K(\tau))'$, which collects the loading coefficients of factors $x_{it}$ on bond pricing formula (5.14). From this, we can obtain the corresponding pricing formula for zero-coupon yields $R_t(\tau)$, with maturity interval $\tau$, as

$$R_t(\tau) = (1/\tau) \left[ A(\tau) + D(\tau)'X_t \right], \text{ for } \tau = 1, 2, ..., N$$  \hspace{1cm} (5.15)

Note, that the above affine relationship also holds for real yields too, with the difference that the number of the state variables governing real yields is likely to be less than $K$ (see e.g., Argyropoulos and Tzavalis [7]):

$$R_t^*(\tau) = (1/\tau) \left[ a(\tau) + d(\tau)'X_t \right], \text{ for } \tau = 1, 2, ..., N^*$$  \hspace{1cm} (5.16)

Closed form solutions of value functions $A(\tau)$, $D(\tau)$ and $a(\tau)$, $d(\tau)$ can be obtained by solving a set of ordinary differential equations under no arbitrage profitable conditions (see Duffie and Kan [53]). For our Gaussian dynamic term structure model (GDTSM), described above, these
solutions for $D(\tau)$ are analytically given as follows:\(^3\)

$$D(\tau) = (I - e^{-k^Q\tau})(k^Q\tau)^{-1}\delta_{t1}$$ (5.17)

These solutions imply a set of arbitrage free cross-section restrictions on the term structure loading coefficients $D(\tau)$.

The above GDTS model implies that the expected excess holding period return of a $\tau$-period to maturity zero-coupon bond over the short-term interest rate $r_t$, referred to as term premium (see, e.g., Tzavalis and Wickens [108], Bolder [19] and Duffee [51]), is given as follows:

$$E_t[h_{t+1}(\tau) - r_t] = -D(\tau)'\sigma\Lambda_t$$ (5.18)

Joint estimation of the last relationship (5.18) and interest rates formula (5.15) (with, or without, cross-section restrictions (5.17)) will enable us to identify the price of risk slope coefficients matrix $\lambda_1$, which determine the time-varying part of the term premium. To calculate excess return $h_{t+1}(\tau) - r_t$ in discrete-time, we consider the one-period (e.g., one-month) interest rate as short-term interest rate, $r_t$, and we assume continuously compounded interest rates, implying

$$R_t(\tau) = -(1/\tau)\log B_t(\tau).$$

Then, $h_{t+1}(\tau) - r_t$ can be written as follows:

$$hpr_{t+1}(\tau) \equiv h_{t+1}(\tau) - r_t = \log \left( \frac{B_{t+1}(\tau - 1)}{B_t(\tau)} \right) - r_t = -(\tau - 1)\left[ R_{t+1}(\tau - 1) + \tau R_t(\tau) - r_t \right]$$ (5.19)

**Nominal Bonds and Inflation Risk Premia**

From equations (5.8) and (5.9) we can show the implied relationship between real and

---

\(^3\)We show the closed form solution only for the nominal yield factor loading $D(\tau)$ since it is similar to the real $d(\tau)$. For the detailed closed form solutions for all factor loadings see Risa [100], Dai and Singleton (2002) [38], Kim and Orphanides (2012) [79].
nominal zero coupon yields:

\[ B_t(\tau) = E_t \left( \frac{m_{t+\tau}}{m_t} \frac{P_t}{P_{t+\tau}} \right) \]

\[ = E_t \left( \frac{m_{t+\tau}}{m_t} \right) \times E_t \left( \frac{P_t}{P_{t+\tau}} \right) + cov \left( \frac{m_{t+\tau}}{m_t}; \frac{P_t}{P_{t+\tau}} \right) \]

\[ = B'_t(\tau) \times E_t \left( \frac{P_t}{P_{t+\tau}} \right) \times \left( 1 + \frac{cov \left( \frac{m_{t+\tau}}{m_t}; \frac{P_t}{P_{t+\tau}} \right)}{E_t \left( \frac{m_{t+\tau}}{m_t} \right) \times E_t \left( \frac{P_t}{P_{t+\tau}} \right)} \right) \]

Alternatively, we can write for nominal bond prices \( B_t(\tau) \):

\[ B_t(\tau) = B'_t(\tau) E_t(P_t/P_{t+\tau}) IP_t(\tau) \] (5.20)

where the inflation bond risk premia \( IP_t(\tau) \) for any maturity \( \tau \) are:

\[ IP_t(\tau) = 1 + \frac{cov \left( \frac{m_{t+\tau}}{m_t}; \frac{P_t}{P_{t+\tau}} \right)}{E_t \left( \frac{m_{t+\tau}}{m_t} \right) \times E_t \left( \frac{P_t}{P_{t+\tau}} \right)} \] (5.21)

converting (5.20) into yields by taking logs and multiplying by \(-1/\tau\) we have a first flavor of the Fisher equation, that is, for any maturity \( \tau \), the nominal yield should be equal to the real yield plus an inflation compensation. This inflation compensation is defined as the sum of the expected inflation rate plus an inflation risk premium:

\[ R_t(\tau) = R'_t(\tau) + \frac{1}{\tau} E_t \left[ \ln \frac{P_{t+\tau}}{P_t} \right] - \frac{1}{\tau} \ln \left( 1 + \frac{cov \left( \frac{m_{t+\tau}}{m_t}; \frac{P_t}{P_{t+\tau}} \right)}{E_t \left( \frac{m_{t+\tau}}{m_t} \right) \times E_t \left( \frac{P_t}{P_{t+\tau}} \right)} \right) \]

Alternatively, we can write for nominal bond yields \( R_t(\tau) \):

\[ R_t(\tau) = R'_t(\tau) + \pi'_t(\tau) + \varphi_t(\tau) \] (5.22)

in which, obviously the expected inflation \( \pi'_t(\tau) \) at any maturity \( \tau \) is:

\[ \pi'_t(\tau) = (1/\tau) E_t \left[ \ln(\Delta_t P_{t+\tau}) \right] = (1/\tau) E_t \left( \frac{P_{t+\tau}}{P_t} \right) \]

(5.23)
and the inflation risk premia for the nominal bond yields $\varphi_t(\tau)$ are defined as:

$$\varphi_t(\tau) = -(1/\tau) \ln[1 + \text{cov} \left(\frac{m_{t+\tau}/m_t}{E_t(P_{t+\tau}/P_t)}\right)]$$

Subtracting real from nominal yields in (5.20) we get the Break-Even Inflation Rate (BEI) at any maturity $\tau$ and BEI equals to the inflation compensation i.e., $BEI(\tau) \equiv R_t(\tau) - R_t^*(\tau) = \pi_t(\tau) + \varphi_t(\tau)$. From (5.22) and (5.23) we can compute the inflation risk premium $\varphi_t(\tau)$ at any maturity $\tau$ From (5.15), (5.16), and (5.6), we have that nominal bond yields, real bond yields, and expected inflation rate are all affine in the state variables $X_t$:

$$\varphi_t(\tau) = R_t(\tau) - R_t^*(\tau) - \pi_t^*(\tau)$$

this leads to

$$\varphi_t(\tau) = (1/\tau) \left[ A(\tau) + D(\tau)'X_t \right] - (1/\tau) \left[ a(\tau) + d(\tau)'X_t \right] - (1/\tau)[g_0(\tau) + g_1(\tau)'X_t]$$

Summarizing, the system of equations that totally describes our model is (5.1), (5.6), (5.7), (5.15) and (5.18).

### 5.3 Empirical analysis

In this section, we carry out our empirical analysis. This is organized as follows. First, we present the data and carry out principal component (PC) analysis to retrieve from the data the unknown common factors (state variables) spanning the term structure of interest rates. This analysis can determine the maximum number of these factors. Next, we present efficient unit root tests for the interest rates and principal component factors to examine if these series contain a unit root in their autoregressive component. These tests are crucial in setting up the appropriate econometric framework of estimating the GDTSM presented in the previous section from available data. Third, we estimate and test this model based on a rich set of data, which consist of nominal interest rates, real consumption growth rate, inflation rates and holding returns. To retrieve the unobserved factors determining the dynamics of our model, we extend Pearson’s and Sun [92] approach, denoted as P-S, retrieving these factors from observed nominal interest rates by inverting the zero-bond yield equations implied by the GDTSM. Our
extension minimizes the effects of measurement errors of interest rates on the estimates of the unobserved factors. Finally, we compare the estimates of the real interest rates obtained by our model to those implied by inflation-indexed bonds, for long-term maturities, and we assess how important are these and inflation risk premia in forecasting future inflation rates over short and long-term horizons.

5.3.1 Data

Our analysis is based on zero-coupon interest rates of the US economy, calculated by zero coupon or coupon-bearing bonds. These series cover the period from 1997:7 to 2009:10. They have monthly frequency, as all other economic series employed in our analysis, and they span a very large cross-section set of different maturity intervals $\tau$, from one month to five years. Inflation rate is calculated as the seasonally adjusted 12-month percentage change of Consumer Price Index for All Urban Consumers (CPI-U), as is assumed when pricing TIPS rates. Our real consumption growth series is calculated based on the seasonally adjusted annual rate of real personal consumption expenditures (PCEC96) taken from the federal reserve economic data archive (FRED).

**Principal component (PC) analysis**

Our PC analysis is based on a large set of different maturity nominal interest rates $R_t(\tau)$, i.e., $N = 60$. This covers maturity intervals $\tau$ from one month to five years (60 months), with monthly maturity frequency. The results of our analysis are presented in Table 1. This shows that three principal components, denoted as $pc_{it}$, for $i = 1, 2, 3$, explain 99.99% of the total variation of the levels (or first differences) of $R_t(\tau)$.

<table>
<thead>
<tr>
<th>Table 1.: Total Number of PCs</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>% variation explained in $\Delta R_t(\tau)$</td>
<td>93.20</td>
<td>99.17</td>
<td>99.99</td>
</tr>
<tr>
<td>% variation explained in $R_t(\tau)$</td>
<td>98.48</td>
<td>99.95</td>
<td>99.99</td>
</tr>
</tbody>
</table>

Notes: The total variation in nominal yields explained by principal components.

The results of our PC analysis are consistent with those reported by Litterman and Scheinkman [81]. They indicate that three common factors explain almost all of the variation of the term

---

6They are obtained from the data archive of J. Huston McCulloch, http://www.econ.ohio-state.edu/jhm/ts/ts.html
structure of interest rates. The first factor, $p_{c1t}$, explains the largest part of the total variation of the levels of $R_t(\tau)$, i.e., 98.48%. Together with the second factor, denoted as $p_{c2t}$, they explain the 99.95% of this variation. The remaining percentage, which is actually, very small is explained by the third factor, $p_{c3t}$.

To interpret $p_{cit}$ factors, in Figure 1 we graphically present estimates of the loading coefficients of the first two $p_{cit}$ factors ($i = 1, 2$) on interest rates series $R_t(\tau)$, while in Table 2 we report some useful descriptive statistics for them. In particular, the table presents estimates of correlation coefficients between the first two $p_{cit}$ factors and observed interest rate variables (or transformations of them) often used in the literature as proxies of unobserved interest rates factors $x_{it}$ (see, e.g., Ang and Piazzesi [3]). These variables are: the long term interest rate with maturity interval $\tau = 5$ years, defined as $z_{it} \equiv R_t(60)$, and the term spread between this interest rate and the short-term one, which is taken to be the one-month interest rate $r_t$. 

![Figure 1. The loadings of the first two principal components across maturity.](image-url)
5.3. EMPIRICAL ANALYSIS

\[ z_{2t} \equiv R_t(60) - r_t. \]

Table 2: Summary Statistics of Interest Rates PCs

<table>
<thead>
<tr>
<th>Factors</th>
<th>( pc_{1t} )</th>
<th>( pc_{2t} )</th>
<th>( pc_{3t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Variance</td>
<td>152.14</td>
<td>2.28</td>
<td>0.25</td>
</tr>
<tr>
<td>Min.</td>
<td>-22.87</td>
<td>-3.07</td>
<td>-0.52</td>
</tr>
<tr>
<td>Max.</td>
<td>21.92</td>
<td>5.38</td>
<td>1.70</td>
</tr>
<tr>
<td>( \rho_1 )</td>
<td>0.98</td>
<td>0.91</td>
<td>0.62</td>
</tr>
<tr>
<td>( \rho_2 )</td>
<td>0.96</td>
<td>0.79</td>
<td>0.17</td>
</tr>
<tr>
<td>( \rho_3 )</td>
<td>0.94</td>
<td>0.70</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>( z_{1t} )</th>
<th>( z_{2t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( z_{1t} )</td>
<td>0.97</td>
<td>0.23</td>
</tr>
<tr>
<td>( z_{2t} )</td>
<td>-0.80</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.04</td>
</tr>
</tbody>
</table>

Notes: Max stands for maximum, while Min. for minimum. \( \rho_s \) are the autocorrelations of PCs \( pc_{it} \), \( i = 1, 2, 3 \), of lag order \( s = 1, 2, 3 \). The variables \( z_{1t}, z_{2t} \) are defined as follows: \( z_{1t} \equiv R_t(60) \) and \( z_{2t} \equiv R_t(60) - r_t \), where \( r_t \equiv R_t(1) \).

Inspection of Figure 1 clearly indicates that \( pc_{1t} \) plays the role of a "level" factor, which causes almost parallel shifts to the whole maturity spectrum of interest rates. This factor dominates the volatility of interest rates \( R_t(\tau) \). On the other hand, \( pc_{2t} \), referred to as "slope" factor, determines the slope of the term structure. Thus, its effects on the term structure decays over time. As the estimates of correlation coefficients reported in Table 2 reveal, there is very close relationship between variables \( z_{it} \) and series \( pc_{it} \), for \( i = 1, 2 \). That is, the five-year interest rate \( R_t(60) \) can be thought of as a very good proxy of \( pc_{1t} \) factor and term spread \( R_t(60) - r_t \) of \( pc_{2t} \). Thus, in our analysis variables \( z_{it} \) will be taken to play the role of interest rate instruments from which unobserved term structure factors (state variables) \( x_{it} \) will be retrieved. As is explained in more details below, these variables will be projected on principal component factors \( pc_{it} \) to obtain estimates of \( z_{it} \) net of measurement errors (see also Argyropoulos and Tzavalis [6]).

Tests for unit roots

To test for a unit root in the level of nominal interest rates \( R_t(\tau) \) and principal component factors \( pc_{it} \), we carry out a second generation ADF unit root test, known as efficient ADF (E-ADF) test (see, e.g., Elliott et al. [55] and Ng and Perron [91]). This test is designed to have maximum power against stationary alternatives to unit root hypothesis which are local to unity. Thus, it can improve the power performance of the standard ADF statistic, often used in practice to test for a unit root in \( R_t(\tau) \).

Values of E-ADF unit root test statistic are reported in Table 3. This is done for nominal
interest rates $R_t(\tau)$, with maturity intervals $\tau = \{1, 3, 6, 12, 24, 36, 48, 60\}$ months. Note that, in addition to E-ADF, the table also presents values of $P_T$ unit root test statistic, suggested by Elliott et al. [55] as alternative to E-ADF. To capture a possible linear deterministic trend in the levels of $R_t(\tau)$ during our sample, both E-ADF and $P_T$ statistics assume that the vector of deterministic components $D_t$ employed to detrend series $R_t(\tau)$ contains also a deterministic trend.

Table 3.: Efficient unit root tests of nominal interest rates $R_t(\tau)$ and $pc_{it}$.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>1</th>
<th>3</th>
<th>6</th>
<th>12</th>
<th>24</th>
<th>36</th>
<th>48</th>
<th>60</th>
<th>$pc_{1t}$</th>
<th>$pc_{2t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>0.98</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>E-ADF</td>
<td>-2.13</td>
<td>-2.12</td>
<td>-2.10</td>
<td>-2.26</td>
<td>-2.21</td>
<td>-2.11</td>
<td>-2.09</td>
<td>-2.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P_T$</td>
<td>4.95*</td>
<td>5.04*</td>
<td>5.19*</td>
<td>3.53**</td>
<td>3.84**</td>
<td>4.69*</td>
<td>5.13*</td>
<td>5.11*</td>
<td>2.28**</td>
<td>1.28**</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. E-ADF and $P_T$ are the efficient unit root test statistics suggested by Elliott et al. [55]. Critical values of test statistics E-ADF and $P_T$ are provided by Elliott et al. [55]. (*) and (**) mean significance at 5% and 1% levels, respectively. The results of Table 3 clearly indicate that, despite the fact that the values of the autoregressive coefficients $\phi$ are found to be very close to unity, the unit root hypothesis for $R_t(\tau)$ is rejected against its stationary alternative, for all $\tau$ considered. This is true at 5%, or 1% significance levels. The estimates of the autoregressive coefficient $\phi$ reported in the table indicate that $R_t(\tau)$ exhibit a very fast mean reversion towards their long-run mean, especially those of shorter maturity intervals of 1, 3 and 6 months. The two principal component factors $pc_{it}$ also constitute stationary series. These results are consistent with those on interest rates series $R_t(\tau)$.

5.3.2 Model specification and estimation

Given the results of PC analysis that $K = 3$ stationary factors span the term structure of interest rates, next we specify the econometric framework which will be used to estimate our GDTSM, presented in Section 2. We assume that the number of unobserved factors jointly spanning nominal interest rates $R_t(\tau)$ and real consumption growth rate (or real term structure $R_t^*(\tau)$), collected in vector $X_t^* = (x_{1t}^*, x_{2t}^*)'$, As in Ang and Piazzesi [3], Diebold et al. [47], the third state variable $x_{3t}$ is taken to be inflation rate $\pi_t$, which is an observed variable. Thus, the vector of our state variables $X_t$ is defined as $X_t \equiv (X_t^*, \pi_t)$ This specification of $X_t$ allows us to capture any feedback and/or contemporaneous effects between vector of factors $(x_{1t}^*, x_{2t}^*)'$, determining real consumption growth, and inflation rate $\pi_t$, and to provide forecasts of future inflation rate $\pi_t$ exploiting information in real and/or nominal variables, jointly.
5.3. **EMPIRICAL ANALYSIS**

The system of equations employed to estimate the GDTSM is given from the equations (5.1), (5.6), (5.7), (5.15) and (5.18), presented in Section 2. We write these equations in econometric form as follows:

\[ \Delta R_{t+1} = \text{const} + DE_t(\Delta X^*_{t+1}) + D^\pi E_t(\Delta \pi_{t+1}) + e_{t+1}(\tau) \]  
(5.27)

\[ \Delta c_{t+1} = \text{const} + \psi_1 X^*_t + \xi_{t+1} \]  
(5.28)

\[ hpr_{t+1}(\tau) = \text{const} + \Gamma X^*_t + \Gamma^\pi \pi_t + \varsigma_{t+1}(\tau) \]  
(5.29)

\[ \Delta X_{t+1} = \text{const} + (\Phi - I) X_t + \omega_{t+1} \]  
(5.30)

where

\[ \Delta X_{t+1} \equiv \begin{bmatrix} \Delta x^*_{1t+1} \\ \Delta x^*_{2t+1} \\ \Delta \pi_{t+1} \end{bmatrix}, \quad \Phi \equiv \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}, \]

with the diagonal elements of matrix \( \Phi \) defined in terms of their continuous-time mean-reversion parameters as \( \phi_{ii} = e^{-h_i \Delta t} \), for all \( i \), and \( I \) is the identity matrix of dimension \((3 \times 3)\).

The above system of equations consist of four key relationships. The vector of state variables \( \Delta X_{t+1} \) of lag order one (see (5.30)), where \( X_t \equiv (X^*_t, \pi_t) \). Note that the matrix of autoregressive coefficients \( \Phi \) of this vector is not assumed to be diagonal, which means that it allows for possible feedback effects between all variables of vector \( X_t \). Second, the affine relationships between nominal interest rates and vector of state variables \( X_t \), for all \( \tau \) (see (5.15)), augmented with a specification, or measurement, error term \( e_{t+1}(\tau) \) (see (5.27)). These relationships are given in first differences, contemporaneously with the vector of state variables \( \Delta X_{t+1} \). Thus, expected values of \( \Delta X^*_{t+1} \) and \( \Delta \pi_{t+1} \), at time \( t \), are entered into their RHS. These values are the estimates of \( E_t(\Delta X_{t+1}) \) simultaneously obtained by the estimates of (see 5.30). Note that, for \( \tau = 1 \) month, (5.27) gives the relationship determining the short-term nominal interest rate, \( r_t \), given by (5.11). Third, the real consumption growth rate equation augmented by error term \( \xi_{t+1} \), for one-period ahead (see (5.28)). Fourth, the set of excess holding period return equations (5.29), based on their theoretical relationships (5.18). As argued in Section 2, adding this set of equations into the above system of structural equations will help in identifying key parameters from the data such as the price of risk parameters \( \lambda_{it} \) of risk pricing functions \( \Lambda_t \).

To estimate the above system of equations (5.27)-(5.30), we will employ the Generalized
Method of Moments (GMM) (see Hansen [69]). This method can provide asymptotically efficient estimates of the vector of parameters of the system which are robust to possible heteroscedasticity and/or serial correlation of error terms $e_{t+1}, \omega_{t+1}, \xi_{t+1}$ and $\varsigma_{t+1}$. In this estimation procedure, we will impose the no-arbitrage restrictions given by equation (5.17) on the loading coefficients of relationships (5.27), (5.28) and (5.29), i.e., the elements of matrix $D$ and vector $\psi^1$. These constitute a set of cross-section restrictions on the parameters of the system which can be tested by our data based on Sargan’s overidentifying restrictions test statistic. Finally, to retrieve estimates of the vector of unobserved state variables $X_t^*$ from our data, we will rely on the GMM estimates of the vector of observed instruments $Z_t$, where $Z_t = (z_{1t}, z_{2t})'$ and $z_{1t} \equiv R_t(60)$ and $z_{2t} \equiv R_t(60) - r_t$, on the principal components $pc_{it}$, for $i = 1, 2, 3$ (See Appendix). These regressions are estimated simultaneously with system of equations (5.27)-(5.30). Variables $z_{1t}$ and $z_{2t}$ are found to be highly correlated with $pc_{it}$, and thus may better capture the dynamics of state vector $X_t^*$ spanning the term structure of interest rates. By construction, their estimated values, obtained by their above projections of $z_{it}$ on $pc_{it}$, will be also orthogonal to any measurement errors in interest rates $R_t(\tau)$. From the estimated values of $Z_t$, we can retrieve estimates of $X_t^*$ by inverting interest rates relationships (5.15), as in Pearson and Sun (PS) [92]. More details above the benefits of the above method compared with the standard PS method can be found in the appendix (see also Argyropoulos and Tzavalis [6]).

### 5.3.3 Estimation results

GMM estimates of the key parameters of the system of equations (5.27)-(5.30), namely loading coefficients of state variables $x_{it}$ on short-term interest rate $r_t$, $\delta_{1i}$, mean reversion and price of risk parameters $k_i$ and $\lambda_{1i}$, as well as the elements of the matrix of autoregression coefficients $\Phi$ and the correlation matrix of their residual terms, denoted as $\hat{\omega}_{it}$, are given in Table 4. These estimates are obtained using interest rates $R_t(\tau)$ and excess holding period returns $hpr_t(\tau)$ with maturity intervals of $\tau = \{3, 6, 9, 24, 36\}$ months. As instruments, we have used lagged values of the ten year nominal interest rate, the spread between the two year and one month nominal interest rates, and inflation rate (see Table 4). The table also presents estimates of Sargan’s overidentifying restrictions test statistics, denoted as $J$, and the variance-covariance/correlation matrix of the estimates of the vector of error terms of the vector autoregression of the state.
variables $\omega_{t+1}$.

Table 4: GMM estimates of the system (5.27)-(5.30)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$x_{1t}^*$</th>
<th>$x_{2t}^*$</th>
<th>$\pi_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{1i}$</td>
<td>1.46</td>
<td>-1.05</td>
<td>0.002</td>
</tr>
<tr>
<td>$k_i$</td>
<td>0.13</td>
<td>0.35</td>
<td>2.50</td>
</tr>
<tr>
<td>$\lambda_{1i}$</td>
<td>-0.009</td>
<td>-0.05</td>
<td>-0.86</td>
</tr>
<tr>
<td>$\phi_{1i}$</td>
<td>-0.01</td>
<td>0.001</td>
<td>-0.0003</td>
</tr>
<tr>
<td>$\phi_{2i}$</td>
<td>0.001</td>
<td>-0.03</td>
<td>-0.0005</td>
</tr>
<tr>
<td>$\phi_{3i}$</td>
<td>0.12</td>
<td>-0.33</td>
<td>-0.18</td>
</tr>
</tbody>
</table>

Variance covariance matrix of the residuals

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\omega}_{1t+1}$</th>
<th>$\hat{\omega}_{2t+1}$</th>
<th>$\hat{\omega}_{3t+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\omega}_{1t+1}$</td>
<td>0.38</td>
<td>0.93</td>
<td>-0.10</td>
</tr>
<tr>
<td>$\hat{\omega}_{2t+1}$</td>
<td>0.72</td>
<td>0.93</td>
<td>-0.08</td>
</tr>
<tr>
<td>$\hat{\omega}_{3t+1}$</td>
<td>0.61</td>
<td>0.93</td>
<td>0.61</td>
</tr>
</tbody>
</table>

$J = 118.84$ (0.11)

Instruments: 1; $R_{t-1}(10y)$ for $i = 0, 1; Sp_t(2y)$; $\pi_{t-i}$ for $i = 1, 2, 3, 4$

Notes: The table presents GMM estimates of parameters of the system of equations (5.27)-(5.30). We also include into the system projected values of vector $Z_t$ on principal component factors $pc_{1t}$ and $pc_{2t}$ (see equation (5.36)). The estimates of Matrix $D_W^{-1}$ and vector $D_W^n$ are also given in the table. Heteroscedasticity and autocorrelation consistent (Newey-West) standard errors are shown in parentheses. $J$ is Sargan’s overidentifying restriction test.

The results of the table indicate that the GDTSF estimated based on system of equations (5.27)-(5.30) is consistent with the data. The values of $J$ statistic reported in the table indicate that no-arbitrage restrictions (5.17) imposed on the parameters of the system can not be rejected.
at 1% or 5% probability values. The estimates of mean-reversion and price of risk parameters $k_i$ and $\lambda_i$, respectively are significant at 5% level, for all factors $i$. These estimates indicate that the first two state variables $x_{1t}^*$ and $x_{2t}^*$ underlying the term structure of $R_t(\tau)$ and real consumption growth $\Delta c_{t+1}$ have high persistency and they are priced in the bond market. On the other hand, the effects of inflation rate changes (or shocks) on $R_t(\tau)$ seem to decay more rapidly than those of $x_{1t}^*$ and $x_{2t}^*$. Another interesting conclusion that can be drawn from the results of Table 4 is that there are significant feedback effects from $x_{1t}^*$ and $x_{2t}^*$ to future inflation rate $\pi_{t+1}$, but not inversely. The standard errors of the estimates of $\phi_{31}$ and $\phi_{32}$ reported in the table imply that they are different than zero at 5% level. Taken these results together with those of the estimates of the correlation matrix between the residual terms $\hat{\omega}_{it}$ of (5.30), someone can conclude that inflation surprises cause future changes in state variables $x_{1t}^*$ and $x_{2t}^*$, driving the real term structure of interest rates. These result obviously implies that there is no risk-neutrality of money is short-term. The values of state variables $x_{1t}^*$ and $x_{2t}^*$ retrieved through the estimation of our system (5.27)-(5.30) are graphically presented in Figure 2 together with those of the first two $pc_{it}$ factors, for $i = 1, 2$. In addition, in Table 5 we present the correlation coefficients for components and factors.
5.3. EMPIRICAL ANALYSIS

Figure 2. Estimates of the factors and the first two principal components from the data.

Inspection of these figures indicate that the estimates of $x_{it}$ do not correspond one-to-one to those of $pc_{it}$. This can be also confirmed by our regression estimates of $x^*_{it}$ on $pc_{it}$, not reported for reasons of space, and the correlation matrix between $pc_{it}$, factors and the state variables $x^*_1$, $x^*_2$ and $\pi_t$ given below

Table 5.: Correlation coefficients for PCs and factors

<table>
<thead>
<tr>
<th></th>
<th>$pc_{1t}$</th>
<th>$pc_{2t}$</th>
<th>$pc_{3t}$</th>
<th>$x^*_{1t}$</th>
<th>$x^*_{2t}$</th>
<th>$\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$pc_{1t}$</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.38</td>
<td>-0.46</td>
<td>0.10</td>
</tr>
<tr>
<td>$pc_{2t}$</td>
<td>1.00</td>
<td>0.00</td>
<td>0.88</td>
<td>0.92</td>
<td>-0.21</td>
<td></td>
</tr>
<tr>
<td>$pc_{3t}$</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>$x^*_{1t}$</td>
<td>1.00</td>
<td>0.63</td>
<td></td>
<td></td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>$x^*_{2t}$</td>
<td></td>
<td>1.00</td>
<td>-0.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents correlation coefficients for PCs and factors estimates of the system of equations (5.27)-(5.30).

These results were expected, since $pc_{it}$ factors constitute linear transformations of $x^*_1$, $x^*_2$
and \( \pi_t \). They imply that employing principal component factors to proxy state variables may not correctly capture state variables \( x_{1t} \), driving the nominal and real term structure of interest rates.

Figures 3a and 3b graphically present the estimates of real interest rates \( R_t^*(\tau) \) implied by the parameter estimates of our GDTSM and estimates of state variables \( x_{1t}^* \) and \( x_{2t}^* \), driving real consumption growth \( \Delta c_{t+1} \). In particular, these are estimated based on real term structure relationship (5.16) and the estimates of state variables \( x_{1t}^* \). To see how closely are these estimates of \( R_t^*(\tau) \) to those based on survey data and/or inflation indexed bonds, Figure 2a also presents values of \( R_t^*(\tau) \) taken from the Cleveland fed survey (see also Haubrich et al. [73]), which are available for \( \tau = 12 \) months.\(^7\) Figure 2b also presents values of \( R_t^*(\tau) \) implied by the 5-year zero coupon TIPS rate. These are taken from Gürkaynak et al. [65].\(^8\).

\[ \text{Figure 3.a. Survey based estimates of the 1-year real interest rate, against model estimates.} \]

\(^7\)http://www.clevelandfed.org/research/data/inflation_expectations
\(^8\)http://www.federalreserve.gov/pubs/feds/2008/200805/200805abs.html
5.3. **EMPIRICAL ANALYSIS**

Furthermore, in Table 6 we present the correlation coefficients between model implied real yields and survey based / TIPS measures, i.e., $\text{Corr}(\hat{R}_t^R(\tau); R_t^R(\tau))$.

<table>
<thead>
<tr>
<th>Maturity in months $\tau$</th>
<th>1</th>
<th>12</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Corr}(\hat{R}_t^R(\tau); R_t^R(\tau))$</td>
<td>0.77</td>
<td>0.76</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Notes: The real yield estimates compared to our model with maturities of 1-month and 1-year are survey based measures taken from Cleveland Fed (see also Haubrich et al. [73]). The 5 year real rate is realized TIPS yield taken from Gürkaynak et al. [65].

Inspection of Figures 3a-3b and Table 6, indicate that our estimates of $R_t^R(\tau)$ are very closed to those implied by the survey and TIPS’ market term structure data. The correlation coefficients between our model estimates of $R_t^R(\tau)$ and those implied by the above data are found to be 0.76 and 0.70, respectively. The only period that seems to be a significant variation between our estimates of $R_t^R(\tau)$ and those of the other sources is that during 2008-2009. This can be obviously attributed to the effects of the recent financial crisis began with the collapse of Lehman brothers. Fears of credit and liquidity risks, triggered by this financial crisis, may have driven the yields of TIPS at higher rate, given that these are less liquid assets than nominal bonds.

As was expected, similar conclusions to the above can be drawn for the expected future
inflation rates obtained by our model, based on relationship (5.6), and those given by the Cleveland fed survey and the TIPS real interest rates. Note that the estimates of expected inflation implied by the TIPS real term structure data are not net of inflation risk premia effects.\textsuperscript{9} Values of correlation coefficients between the estimates of expected future inflation rates by our model and those taken from Cleveland fed survey, as well as from TIPS data are presented in Table 7 for maturities $\tau = \{12, 36, 60\}$ months, while in Figure 4 we graphically present inflation expectations for $\tau = 36$ months.

The correlation coefficients between our model’s expected future inflation values and those of the two other data sources are found to be 0.92 and 0.96, respectively. Again the lower values of expected inflation rates implied by these two sources of data compare to those of our model can be attributed to the effects of the recent financial crisis, as was explained before.

\textsuperscript{9}The values of expected future inflation taken from the TIPS data are calculated by the difference between nominal minus real interest rates, implied by these securities.
5.3.4 **Inflation risk premia**

In this section, we assess how important are the inflation risk premia. Recall that inflation risk premium is defined as

\[
\varphi_t(\tau) \equiv R_t(\tau) - R^*_t(\tau) - \pi^*_t(\tau) = -\frac{1}{\tau} \ln \left( 1 + \frac{\text{cov}_t \left( m_{t+\tau}/m_t; P_t/P_{t+\tau} \right)}{E_t \left( m_{t+\tau}/m_t \right) E_t(P_t/P_{t+\tau})} \right). \tag{5.31}
\]

It can be easily calculated for any maturity interval \( \tau \) based on its state variables relationship (5.26) implied by our GDTSM and its parameter estimates, presented in Table 4, and the estimates of the vector of state variables \( X_t \) retrieved by our data.

Table 7 presents correlation coefficients, \( \text{Corr}(\hat{\varphi}_t(\tau); \varphi_t(\tau)) \), between the estimates of inflation risk premium obtained by our model estimates, \( \hat{\varphi}_t(\tau) \), and those implied the TIPS’ market real interest rates, for \( \tau = \{12, 36, 60\} \) months. The latter employ the Cleveland’s fed survey data on expected future inflation rates to retrieve estimates of \( \pi^*_t(\tau) \) net of inflation expectations, as is required by the definition of \( \varphi_t(\tau) \) (see relationship (5.31)).

<table>
<thead>
<tr>
<th>Maturity in months ( \tau ):</th>
<th>12</th>
<th>36</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation expectations (IE) Corr( (\pi_t^*(\tau); \pi_t(\tau)) )</td>
<td>0.92</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>Inflation risk premia (IRP) Corr( (\hat{\varphi}_t(\tau); \varphi_t(\tau)) )</td>
<td>0.65</td>
<td>0.68</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Notes: Inflation expectations (IE), \( \pi_t^*(\tau) \), compared to our model estimates with maturities of 1,3 and 5 years are survey based measures taken from Cleveland Fed (see also Haubrich et al. [73]). Inflation risk premia (IRP), \( \varphi_t(\tau) \), compared to our model estimates are calculated as survey based inflation expectations subtracted from actual break-even rates (BEI). As actual BEI rate we define the difference between nominal and real yields, where the latter are taken from either surveys or inflation-indexed bonds (TIPS).

In Figure 5 we graphically present \( \hat{\varphi}_t(\tau) \) and \( \varphi_t(\tau) \) estimates of inflation risk premium for

---

10Note that similar graphs of \( \varphi_t(\tau) \) can be also taken by using our model’s expected future inflation rates, instead of the Cleveland fed survey data.
CHAPTER 5. FORECASTING INFLATION FROM THE TERM STRUCTURE AND THE IRP EFFECTS

\( \tau = 36 \) months.

![3-Year Inflation Risk Premium](image)

**Figure 5**: Comparisons of 3-year inflation risk premia between survey-based and model estimates.

Summary statistics of the estimates of \( \phi_t(\tau) \) obtained by our model and those approximated by the survey and TIPS’ data are given in Table 8, for different \( \tau \).

<table>
<thead>
<tr>
<th>Maturity in months ( \tau ) :</th>
<th>3</th>
<th>6</th>
<th>12</th>
<th>36</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-1.53</td>
<td>-1.51</td>
<td>-1.47</td>
<td>-1.31</td>
<td>-1.17</td>
</tr>
<tr>
<td>Volatility</td>
<td>1.32</td>
<td>1.14</td>
<td>0.94</td>
<td>0.71</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Notes: Summary statistics of the model estimates of inflation risk premia for various maturities.

The results of Tables 7 and 8 and inspection of Figure 5 indicate that our model estimates of \( \phi_t(\tau) \) are closely related to those based on the survey and the TIPS’ market data. The correlation coefficients between these two alternative measures of \( \phi_t(\tau) \) vary between 0.65 and 0.67 values. Both of them vary between negative and positive values, but they take negative values for most periods of the sample and, especially, during the recent financial crisis. From relationship (5.31), it can be seen that negative values of \( \phi_t(\tau) \) can be attributed the positive values of the covariance between marginal utility growth \( m_{t+\tau}/m_t \) of investors in the bond market and the inverted price level change \( P_t/P_{t+\tau} \), over investment horizon \( \tau \). A positive covariance matrix between these two variables implies that nominal interest rates are below
the sum of real rates $R_t^r(\tau)$ and expected future inflation rates $\pi_t^e(\tau)$, predicted by the Fisher equation. A negative inflation risk premium means that investors would prefer to hold nominal bonds rather than inflation-indexed bonds. This may be attributed to hedging or liquidity preferences effects.

Another interesting conclusion that can be drawn from the results of Table 8, is that both the mean and volatility (standard deviation) of inflation risk premium $\varphi_t(\tau)$ declines with maturity intervals $\tau$. This is also consistent with evidence provided by Grishchenko and Huang [64] based on market and survey data. As can be seen from the state variables representation of $\varphi_t(\tau)$, given by equation (5.26), the declining estimates of $\varphi_t(\tau)$, or their volatility, with $\tau$ can be attributed to the fact that the persistent factors (state variables) affecting this risk premium are offsetting and they are also scaled by the maturity interval $\tau$.

### 5.3.5 Forecasting inflation from the term structure

Having obtained estimates of risk premium $\varphi_t(\tau)$ and real interest rates $R_t^r(\tau)$ through our GDTSME, in this section we examine if time variation of these two variables can explain the failures of the nominal spread $R_t(\tau) - r_t$, or its marginal definition $R_t(\tau) - R_t(s)$, for $\tau > s$, to provide forecasts of future inflation in the right direction, i.e., according to the predictions of the rational expectations hypothesis of the term structure (REHTS). This regression model of forecasting future inflation rate was first suggested in the literature by Mishkin [87]). It is written as

$$\pi_t(\tau) - \pi_t(s) = \alpha_{\tau,s} + \beta_{\tau,s}(R_t(\tau) - R_t(s)) + \varepsilon_t(\tau, s), \quad (5.32)$$

where $\pi_t(\tau) - \pi_t(s)$ is the change of inflation rates between future periods $t + \tau$ and $t + s$. If real interest rates $R_t(\tau)$ and risk premium $\varphi_t(\tau)$ are constant, then the REHTS predicts that $\beta_{\tau,s}=1$, for all $\tau \neq s$.

Table 9 presents GMM estimates of regression model (5.32), for different $\tau$ and $s$. In the estimation procedure, as instrument we employ lagged values of nominal and real spreads, as well as inflation risk premia based on the Cleveland fed survey (see the table). The results of the table are consistent with the findings of Mishkin [87]. They show that the nominal term spread $R_t(\tau) - R_t(s)$ contain information about future inflation rate changes only at the long-
end of the term structure of nominal interest rates, i.e., for the following pairs of maturity intervals: \((\tau, s)=(36,12), (60,12)\). For the pairs of maturity intervals \((\tau, s)=(12,3), (36,3)\), which include short-term short-term investment horizons, model (5.32) fails to predict the future changes of inflation rates \(\pi_t(\tau) - \pi_t(s)\). In this case, the slope coefficient of this regression is far away from unity, and for \((\tau, s)=(12,3)\) takes negative values.

Table 9.: Estimates of inflation change forecasting equations (5.32)

<table>
<thead>
<tr>
<th>Maturity in months ((\tau, s))</th>
<th>(c_{\tau, s})</th>
<th>(b_{\tau, s})</th>
<th>signif. (J)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(12,3)</td>
<td>0.24</td>
<td>-1.86</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(1.80)</td>
<td></td>
</tr>
<tr>
<td>(36,3)</td>
<td>0.01</td>
<td>0.17</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.31)</td>
<td></td>
</tr>
<tr>
<td>(36,12)</td>
<td>-0.20</td>
<td>0.88</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.55)</td>
<td></td>
</tr>
<tr>
<td>(60,12)</td>
<td>-0.36</td>
<td>0.67</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.30)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents GMM estimates of inflation change forecasting equations (5.32). Heteroscedasticity and autocorrelation consistent (Newey-West) standard errors are shown in parentheses. \(J\) is Sargan’s overidentifying restriction test.

To examine if the above puzzling behavior of term spread \(R_t(\tau) - R_t(s)\) can be attributed to the time variation of inflation risk premium and/or that real term spread \(R_t^*(\tau) - R_t^*(s)\), next we have estimated the following version of model (5.32) adjusting for the above time varying effects:

\[
\pi_t(\tau) - \pi_t(s) = c_{\tau, s} + b_{\tau, s}(R_t(\tau) - R_t(s)) + \theta_t(\tau, s) + u_t(\tau, s),
\]

where

\[
\theta_t(\tau, s) = (R_t^*(\tau) - R_t^*(s)) - (\varphi_t(\tau) - \varphi_t(s))
\]
captures the inflation risk premia and real term structure effects. This regression model stems from theoretical relationship ((5.22)). Under the REHTS, it implies that \(b_{\tau, s}=1\). GMM esti-
5.3. **EMPIRICAL ANALYSIS**

mates of the above regression model is given in Table 10. Panel A of the table presents estimates of (5.33) where nominal spread $R_t^\tau - R_t^s$ is adjusted only for the real term structure effects (i.e., $\theta_t^\tau - \theta_t^s \equiv R_t^\tau(m) - R_t^s(n)$), while Panel B adjusts $R_t^\tau - R_t^s$ for both the real term structure and the inflation risk premia effects (i.e., $\theta_t^\tau - \theta_t^s \equiv (R_t^\tau(m) - R_t^s(n)) - (\varphi_t^\tau(m) - \varphi_t^s(n))$).

Table 10A.: Estimates of inflation change forecasting equations allowing for IRP effects (5.33)

<table>
<thead>
<tr>
<th>Maturity in months $(\tau, s)$</th>
<th>$c_{\tau,s}$</th>
<th>$b_{\tau,s}$</th>
<th>signif. $J$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(12,3)</td>
<td>-0.50</td>
<td>1.40</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.52)</td>
<td></td>
</tr>
<tr>
<td>(36,3)</td>
<td>0.15</td>
<td>0.45</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.22)</td>
<td></td>
</tr>
<tr>
<td>(36,12)</td>
<td>0.23</td>
<td>0.22</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(0.80)</td>
<td></td>
</tr>
<tr>
<td>(60,12)</td>
<td>0.43</td>
<td>0.85</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.09)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents GMM estimates of inflation change forecasting equations (5.33). Heteroscedasticity and autocorrelation consistent (Newey-West) standard errors are shown in parentheses. $J$ is Sargan’s overidentifying restriction test.

Table 10B.: Estimates of inflation change forecasting equations allowing for IRP effects (5.33)

<table>
<thead>
<tr>
<th>Maturity in months $(\tau, s)$</th>
<th>$c_{\tau,s}$</th>
<th>$b_{\tau,s}$</th>
<th>signif. $J$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(12,3)</td>
<td>-0.22</td>
<td>0.72</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.35)</td>
<td></td>
</tr>
<tr>
<td>(36,3)</td>
<td>0.35</td>
<td>0.81</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.18)</td>
<td></td>
</tr>
<tr>
<td>(36,12)</td>
<td>-0.42</td>
<td>0.83</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.14)</td>
<td></td>
</tr>
<tr>
<td>(60,12)</td>
<td>0.66</td>
<td>0.80</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.16)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents GMM estimates of inflation change forecasting equations (5.33). Heteroscedas-
icity and autocorrelation consistent (Newey-West) standard errors are shown in parentheses. $J$ is Sargan’s overidentifying restriction test.

The results of Table 10 indicate that adjusting the nominal term spread $R_t(\tau) - R_t(s)$ both for time-varying real term structure effects and inflation risk premia can explain the puzzling behavior of it to fail to forecast future inflation rates in short run. As the analysis of our previous sections has shown, this failure can be attributed to the volatility of these two components of nominal spread in short run. The estimates of slope coefficients $b_{r,s}$ become close to unity and, thus, they are in the right direction with the predictions of the REHTS. Note that comparison between the results of Tables 9 and 10 indicate that responsible for the failure of model (5.32) to forecast future inflation rate changes is not only the volatility of inflation risk premia, but also that of the real term structure. As forecasting horizons $\tau$ and $s$ increase, the volatility of these two factors decline and, thus, their effects on the inflation predictability of nominal spread become negligible.

5.4 Conclusions

This paper has employed a Gaussian dynamic term structure model with the aim of examining how important are inflation risk premium effects and, if these effects, can explain the failure of the spread between nominal interest rates of different maturity to forecast future change in inflation rates. The paper fits the model into the nominal term structure of interest rates, inflation and real consumption data. The latter can captured movements of the real term structure of interest rates, especially in short-term where inflation-indexed real interest rates are not available and/or are less liquid.

The paper provides a number of interesting findings, which can be proved very useful to monetary authorities in forecasting future inflation rates. First, it shows that the model is consistent with the data and can describe sufficiently the dynamics of the nominal term structure and inflation rates, observed in reality. Furthermore, the real interest rates and expected inflation forecasts retrieved by the model are very close to those implied by survey and inflation-indexed bonds over longer horizons. Second, the inflation risk premium is found to be negative for most intervals of our sample and very volatile, especially for short-term horizons. This
means that investors require less compensation for holding nominal bonds compared to real
(inflation-indexed) ones. This attitude of investors may be related to liquidity effects of the
inflation-indexed market. Third, as inflation risk premium, real interest rates are also volatile
in short-term. These together with inflation risk premia can explain the failures of the nominal
term spread to forecast future inflation rates over short-term horizons.

5.5 Appendix

In this appendix, we describe how we retrieve the unobserved factors from observable variables.
Using the affine relationship for bond yields (5.15) we can extract an estimate for the vector of
unobserved factors $X_t^*$ since we can write:

$$ R_t = A + DX_t + e_t $$

or

$$ R_t = A + DX_t^* + D^x \pi_t + e_t $$

where $A = ((1/\tau_1)A(\tau_1), (1/\tau_2)A(\tau_2), ..., (1/\tau_N)A(\tau_N))^\prime$ is a $(NX1)$-dimension vector of con-
stants $A(\tau)$, and $D = [(1/\tau_1)D(\tau_1), (1/\tau_2)D(\tau_2), ..., (1/\tau_N)D(\tau_N)]^\prime$ is a $(NX(K-1))$-dimension
matrix consisting of the loading coefficients of the elements of the vector of factors $X_t^*$ on those of
vector $R_t$, and the $(NX1)$-dimension vector $D^x = [(1/\tau_1)D^x(\tau_1), (1/\tau_2)D^x(\tau_2), ..., (1/\tau_N)D^x(\tau_N)]^\prime$.

Then, there exists a $(MX1)$-vector of observed interest rates, or transformations of them, namely $Z_t$, with elements $z_{it}$, that holds:

$$ Z_t = A_Z + D_Z X_t^* + D_Z^\pi \pi_t + e_t $$  \hspace{1cm} (5.34)

where $A_Z$, $D_Z$ and $D_Z^\pi$ are appropriately defined sub-arrays of vector $A$ and matrices $D$ and $D^x$. Note that vector $Z_t$ is also measured with errors. Relying on linear projections of vector $Z_t$ on
a set of instruments, instead of using observed values of $Z_t$ itself, is one way of overcoming the
measurement errors problem. We diversify away any possible measurement errors by forming
principal component portfolios for a sufficiently large number of interest rates, $N$ (see also e.g.
Joslin et al [77], Argyropoulos and Tzavalis [6]). In particular,

\[ PC_t = WR_t = WA + WDX_t + We_t = WA + WDX_t^* + WD^x\pi_t + We_t \]

in which \( W \) is a \((K \times N)\)-dimension matrix of the weights (loading coefficients) of interest rates \( R_t(\tau) \) on principal component factors \( pc_{jt} \). The assumption that \( We_t = 0 \) means that vector \( PC_t \) does not suffer from measurement errors \( e_t \). Hence,

\[ PC_t = AW + D_W X_t^* + D_W^x\pi_t \]

Then

\[ X_t^* = D_W^{-1}(PC_t - D_W^x\pi_t - AW) \]

Substituting (5.35) into (5.34), we get linear projections of \( Z_t \) on \( PC_t \) and \( \pi_t \).

\[ Z_t = A^* + D_Z D_W^{-1}PC_t - (D_Z D_W^{-1}D_W^x - D_Z^x)\pi_t \]

where \( A^* \equiv A_Z - D_Z D_W^{-1}AW \). For example, if we set \( Z_t = (R_t(\tau_a), sp_t(\tau_L)) \), where \( R_t(\tau_a) \) denotes an interest rate with maturity \( \tau_a \) and \( sp_t(\tau_L) \) denotes a spread between a long term interest rate \( R_t(\tau_L) \) with maturity interval \( \tau_L \), and the short term rate, then \( D_Z \) and \( D_Z^x \) become known functions of the structural parameters:

\[ D_Z \equiv \begin{bmatrix} D_1(\tau_a) & D_2(\tau_a) \\ (1/\tau_L)D_1(\tau_L) - \delta^1_l & (1/\tau_L)D_2(\tau_L) - \delta^2_l \end{bmatrix}, \quad D_Z^x \equiv \begin{bmatrix} D^x(\tau_a) \\ (1/\tau_L)D^x(\tau_L) - \delta^x_l \end{bmatrix} \]
Bibliography


BIBLIOGRAPHY


