A Tale of Two Yield Curves: Modeling the Joint Term Structure of Dollar and Euro Interest Rates

Alexei V. Egorov

Department of Economics

West Virginia University, Morgantown, WV 26506

Haitao Li

Stephen M. Ross School of Business

University of Michigan, Ann Arbor, MI 48109

David Ng*

Department of Applied Economics and Management

Cornell University, Ithaca, NY 14853

First version: December, 2004

This version: August, 2006

*Email addresses for the authors are: avegorov@wvu.mail.edu, htli@bus.umich.edu and dtn4@cornell.edu, respectively.

A Tale of Two Yield Curves: Modeling the Joint Term Structure of Dollar and Euro Interest Rates

ABSTRACT

Modeling the joint dynamics of the term structures of interest rates in the U.S. and Europe, the two largest economies in the world, is extremely important in international finance. Such a joint term structure model is essential for pricing and managing interest rate risks for international banks and bond investors. In this paper, we provide both theoretical and empirical analysis of multi-factor joint affine term structure models for dollars and euros. In particular, we provide a systematic classification of multi-factor joint affine term structure models similar to that of Dai and Singleton (2000). A principal component analysis of daily dollar and euro interest rates reveals four factors in the data. We estimate four-factor joint affine term structure models using the approximate maximum likelihood method of Aït-Sahalia (2002a, b) and compare the in-sample and out-of-sample performances of these models using some of the latest nonparametric methods. We find that a new four-factor model with two common and two local factors best captures the joint term structure dynamics in the US and Europe.

Key words: Affine term structure models, International term structure models, Approximate Maximum Likelihood, LIBOR, Euribor, Specification analysis of term structure of interest rates, Out-of-sample model evaluation.

JEL Classification: C4, C5, G1

1 Introduction

In recent years, financial markets have become more globalized. As banks and institutional investors lend, borrow and invest internationally, they take on large bond market positions in different countries. These international bond market positions create exposure to different interest rate risks.[1] Characterizing these risks is important for banks, investors and regulators. Banks and investors are clearly interested in assessing and managing the risks in their portfolios. Regulators are also keen to understand the underlying risk so as to set adequate bank capital requirements and monitor systemic risk. However, the multiple sources of risks involved make it challenging to conduct risk management on these bond portfolios

At the heart of managing these risks is an appropriate model for the joint term structure of interest rates in multiple countries. This paper focuses on the joint term structure of the US dollar and the euro. The euro is now the official currency in twelve European countries and is gaining dominance as one of the two major currencies in the world.[2] Among different bond markets, the dollar and euro bond markets are the two most important. As of the end of 2003, Euro-zone domestic governments and corporations had an outstanding volume of US $5,462 billion worth of euro-denominated bonds issued in their domestic countries.[3] This represents 22.3 percent of outstanding volume of domestic-issued debt among all developed countries and this size is second only to that of the United States. Among international issuers from outside of the Euro-zone, euros and dollars are the favorite currencies. Most international issuers choose to issue their bonds and notes in euros (43.5% of total volume) and in US dollars (40.5% of total volume).[4] The dominance of the Euro and US bond markets means that their joint term structure deserves to be examined seriously.

Recently, a number of academic papers have started to investigate two-country joint term structure models. In this literature, a paper typically develops a term structure model and examines whether it resolves the forward premium puzzle or provides international diversification benefits.[5] While these studies make important contributions, none of them have looked at the euro and dollar term structures jointly. Doing so is the main focus of our paper. In addition, each of these studies uses a different specification of the term structure model.[6] With each different specification, different

[1] Previous literature largely focuses on whether to hedge the exchange rate risk in such portfolios. However, these portfolio values are driven by interest rates in both regions, in addition to the exchange rate risk.

[2] Countries have adopted the euro include Austria, Belgium, Finland, France, Germany, Greece, Italy, Ireland, Luxembourg, The Netherlands, Spain and Portugal.

[3] European Central Bank (2004).

[4] European Central Bank (2004).

[5] Backus, Telmer and Foresi (2001), Inci and Lu (2004), Han and Hammond (2003), and Brennan and Xia (2006) look at the forward premium puzzle. Dewachter and Maes (2001) and Ahn (2004) examine the international diversification benefit. Hodrick and Vassalou (2002) and Tang and Xia (2006) empirically examine term structures in multiple countries. Earlier references include Bakshi and Chen (1997) and Bansal (1997).

[6] A concurrent working paper, Mosburger and Schneider (2005), also compares several different models on joint term structure. However, they lack the proper tools to compare the performances of non-nested models and they only go up

empirical results are drawn. For example, Backus, Telmer and Foresi (2001) propose a two-country model that could potentially resolve the forward premium puzzle, but the long term yield could go up as high as 80% due to implausibly high market price of risk. Using a different model specification with an extra factor, Han and Hammond (2003) find that such high market price of risk is not necessary. Inci and Lu (2004) present yet another model and find similar results. Clearly, there is a need for a comprehensive study examining multiple term structure specifications in the context of Euro and US dollar bond markets.

In this paper, we provide a thorough analysis of multi-factor joint affine term structure models for dollar and euro. Our study makes both theoretical and empirical contributions. Theoretically, our paper systematically examines, decomposes and classifies joint term structure models with up to four local and common factors. In a domestic setting, Dai and Singleton (2000) provide an important contribution by classifying affine term structure models into subfamilies.[7] Within each subfamily of two- or three-factor models, Dai and Singleton derive the maximal model that nests existing models. However, international joint term structure models add another layer of complexity due to the presence of local and common factors. We classify all three-factor or four-factor international term structure models within the maximally admissible classification schemes. Ours is the first paper that provides a comprehensive classification for international term structure models.

Empirically, we provide new evidence on the joint term structure using daily data in LIBOR and Euribor from July 1999 to June 2003. Just like Litterman and Sheinkman (1991) who conclude that three factors are needed to capture the US term structure, we find that a four-factor model best captures the joint US-Euro term structure dynamics. One class of four-factor term structure model, with two common and two other individual country factors, is particularly promising in terms of in-sample goodness of fit and out-of-sample forecasting ability. The model that works the best is a four-factor model in which only one of the common factors drives volatilities.

We conduct our empirical tests in three stages. First, using principal component analysis, we examine the total number of factors and the numbers of common vs local factors that should be included in the term structure models. Domestic term structure models usually use up to three-factors (e.g. Dai and Singleton (2000)). But in an international setting, it remains an open question how many factors should be included and how many of them should be common or local. Most international papers use two or three factors with different combinations of local and common factors (e.g. Inci and Lu (2004), Hodrick and Vassalou (2002), Tang and Xia (2006), Backus, Telmer and Foresi (2001)). Motivated by empirical evidence from principal component analysis, we go beyond the usual three-factor models to examine four-factor models.

to three factors.

[7]Traditionally, interest rates were assumed to follow a univariate model, e.g. Vasicek (1977) and Cox, Ingersoll and Ross (1985). Over time, researchers like Litterman and Sheinkman (1991) realized the importance of using multifactor models. Among these, Duffie and Kan (1996)'s class of multifactor affine models is particularly popular due to its flexibility and tractibility.

Second, we estimate the model utilizing the approximate maximum likelihood estimation, a powerful method developed recently by Aït-Sahalia (1996, 2002a, and 2002b) and Aït-Sahalia and Kimmel (2002). The absence of a closed-form solution for transition density of affine models makes maximum likelihood estimation infeasible. Different authors estimating international term structure models resort to other estimation methods that are either inaccurate in small sample (Quasi-maximum likelihood or Efficient Methods of Moment) or computationally burdensome (Simulated Methods of Moment). The approximate maximum likelihood estimation provides extremely fast and accurate estimations for affine models. We show that Aït-Sahalia's approximate maximum likelihood estimation techniques are applicable in estimating international term structure models.

Third, we utilize the latest advance in non-parametric modeling to compare non-nested models. Dai and Singleton (2000) point out the difficulties in testing across non-nested affine models. Likewise, Tang and Xia (2006) note that "a suitable test in the current [non-nested] setting is... still not available in the literature." Most papers use the Akaike Information Criteria and Schwartz criteria as suggestive evidence of model comparison. New developments in non-parametrics allow us to overcome this difficulty. Hong and Li (2005) and Hong, Li and Zhao (2006) recently developed powerful nonparametric tests which are applicable even to non-nested models. Making use of these newly developed tests, we are able to compare the in-sample and out-of-sample forecasting performances across different international affine models.

Our paper opens the door to a wide range of research possibilities. First of all, our framework makes it possible to classify and compare new classes of term structure models. Recent literature in domestic term structure proposes various new specifications like the essentially affine term structure models of Duffee (2002) and the market price of risk specification of Cheridito, Filipović and Kimmel (2006). Our framework can be easily extended to incorporate these models.

Also, our framework can be adopted to study the forward premium puzzle and exchange rate dynamics. Previous literature on joint term structure focuses on the forward premium puzzle, while previous literature on the international bond portfolio focuses on the benefit for exchange rate risk hedging.[8] This paper focuses on capturing the joint term structure risk and not the exchange rate dynamics. Our approach is supported by our principal component evidence that the exchange rate is driven primarily by a factor that is separate from the term structures and by similar results shown in Inci and Lu (2004) and Han and Hammond (2003). As the exchange rate risk is to a large extent orthogonal, the joint term structure risk can be examined separately. In practice, this focus is important for investors in both markets. For instance, investors in a Euribor portfolio would normally just assess the factors driving the Euribor term structure. With our framework, we can relate the impact of LIBOR term structure on that of Euribor while abstracting from the second-order effect of exchange rate. Nevertheless, our framework can be extended easily to incorporate exchange rate dynamics and allows comparison of different models.

[8]See for example Levich (2001), Filatov and Rappaport (1992), and Glen and Jorion (1993).

Finally, our framework can be extended to calculate the Value-at-Risk (VaR) measure of risk inherent in an international bond portfolio. Using the approximate maximum likelihood estimation method, we can characterize the potential gains and losses and calculate the VaR of such a portfolio.

Given the importance and applicability of the topic, the two-country joint affine term structure is likely to attract a substantial amount of research in the future. Our goal is to understand the quantitative properties of these models in capturing the dynamics of joint term structures in the US and Europe.

The rest of this paper is organized as follows. Section 2 introduces two-country affine models of interest rate term structure. Section 3 presents the specification analysis of three-factor joint term structure models. Section 4 describes the estimation and ranking methodologies. Section 5 describes the data and reports the empirical results for joint dynamics of LIBOR and Euribor. Section 6 concludes the paper.

2 Joint Affine Term Structure Models

In single-country affine models, it is assumed that the spot rate $r(t)$ is an affine function of N latent state variables $X(t) = [X_1(t), X_2(t), ..., X_n(t)]'$:

$$r(t) = \delta_0 + \delta' X(t), \tag{1}$$

where δ_0 is a scalar and δ is an $N \times 1$ vector. In the absence of arbitrage opportunities, the time t price of a zero-coupon bond maturing at $t + \tau_m$ ($\tau_m > 0$) equals

$$P(t, \tau_m) = E_t^Q \left[\exp \left(- \int_t^{t+\tau_m} r(s) ds \right) \right],$$

where the expectation E_t^Q is taken under the risk-neutral measure Q. Thus, the whole yield curve is determined by $X(t)$, which is assumed to follow an affine diffusion under the physical measure:

$$dX(t) = \kappa \left[\vartheta - X(t) \right] dt + \Sigma S_t dW(t), \tag{2}$$

and under the risk-neutral measure:

$$dX(t) = \widetilde{\kappa} \left[\widetilde{\vartheta} - X(t) \right] dt + \Sigma S_t d\widetilde{W}(t). \tag{3}$$

$\widetilde{W}(t)$ is an $N \times 1$ independent standard Brownian motion under measure Q, and $\widetilde{\kappa}, \Sigma$ are $N \times N$ parameter matrices, and $\widetilde{\theta}$ is an $N \times 1$ parameter vector. The matrix S_t is diagonal with (i, i)-th elements

$$S_{t(ii)} \equiv \sqrt{\alpha_i + \beta_i' X(t)}, \quad i = 1, ..., N, \tag{4}$$

where α_i is a scalar parameter and β_i is an $N \times 1$ parameter vector.

Under the above assumptions, bond prices have the exponential affine form:

$$P(X(t), \tau_m) = \exp\left(-A(\tau_m) - B(\tau_m)' X(t)\right).$$

The yields of zero coupon bonds (denoted by $Y(X(t), \tau_m) = -\ln(P(X(t), \tau_m))/\tau_m$) are an affine function of the state variables,[9]

$$Y(X(t), \tau_m) = A(\tau_m)/\tau_m + \left(B(\tau_m)'/\tau_m\right) X(t) \tag{5}$$

where the scalar function $A(\cdot)$ and the $N \times 1$ vector-valued function $B(\cdot)$ either have a closed-form or can be easily solved via numerical methods.

The completely affine models of Dai and Singleton (2000) assume that the market price of risk is

$$\Lambda(t) = S_t \lambda \tag{6}$$

where λ is an $N \times 1$ parameter vector. This implies that the compensation for risk is a fixed multiple of the variance of risk and that the market prices of risk cannot change sign over time. These restrictions make it difficult to replicate some stylized facts of historical excess bond returns. Duffee (2002) extends completely affine models to essentially affine models.

Dai and Singleton (2000) greatly simplify the econometric analysis of affine models by providing a systematic scheme that classifies all admissible N-factor affine models into $N+1$ subfamilies[10], denoted as $\mathbf{A}_m(N)$, where $m \in \{0, 1, ..., N\}$ is the number of state variables that affect the instantaneous variance of $X(t)$. They also introduce a canonical representation for $\mathbf{A}_m(N)$, which has the most flexible specification within each subfamily, as it either nests or is equivalent (via an invariant transform) to all the models in $\mathbf{A}_m(N)$.[11] In the canonical representation, Σ is normalized to the identity matrix and the state vector $X(t)$ is ordered so that the first m elements of $X(t)$ affect the instantaneous variance of $X(t)$. Setting $\alpha_i = 0$ for $i = 1, 2, ..., m$, and $\alpha_i = 1$ for $i = m+1, ...N$, we have $S_{t(ii)} = X_i(t)^{1/2}$ for $i = 1, ..., m$, and $S_{t(ii)} = [1 + \beta_i' X_i(t)]^{1/2}$ for $i = m+1, ..., N$, where $\beta_i = (\beta_{i1}, ..., \beta_{im}, 0, ..., 0)'$.

For a two-country model we need to define expressions for both domestic spot rate $r(t)$ and foreign spot rate $r^*(t)$. As in the one-country model (1), the spot rates are assumed to follow affine structures:

$$r(t) = \delta_0 + \delta' X(t), \quad r^*(t) = \delta_0^* + \delta^{*\prime} X(t), \tag{7}$$

[9] See, e. g., Dai and Singleton (2000) and references therein.

[10] Admissibility means that $\alpha_i + \beta_i' X(t) \geq 0$ for all i and all possible values of $X(t)$.

[11] See Dai and Singleton (2000) for their complete set of restrictions on model parameters. Aït-Sahalia and Kimmel (2002) discuss the limitations of these restrictions and provide a comprehensive set of existence, stationarity and boundary restrictions for affine models with up to three factors.

where δ_0 and δ_0^* are scalars, δ and δ^* are $N \times 1$ vectors and N is the total number of factors in the joint term structure. $X(t)$ denotes all factors including both domestic and common factors. On the one hand, if all elements of vectors δ and δ^* are non-zero, then all factors affect both spot rates $r(t)$ and $r^*(t)$. On the other hand, if only those elements of δ are non-zero for which the corresponding elements of δ^* are zero, then only domestic factors affect spot rates $r(t)$ and $r^*(t)$. We call factors that enter expressions for both $r(t)$ and $r^*(t)$ common factors. We call all other factors in vector $X(t)$ country-specific factors. This setup is general so that the joint affine term structure model can be decomposed into two country-specific affine models, each of which can still depend on common factors.

Under the physical structure, the affine dynamics of the factors follow:

$$dX(t) = \kappa \left[\vartheta - X(t)\right] dt + S_t dW(t), \tag{8}$$

For the joint affine term structure, we further assume the canonical representation $\mathbf{A}_m(N)$, where $m \in \{0, 1, ..., N\}$ is the number of state variables that affect the instantaneous variance of $X(t)$. In particular, for $i = m + 1, ...N$, we have $S_{t(ii)} = X_i(t)^{1/2}$ for $i = 1, ..., m$, and $S_{t(ii)} = [1 + \beta_i' X_i(t)]^{1/2}$ for $i = m + 1, ..., N$, where $\beta_i = (\beta_{i1}, ..., \beta_{im,} 0, ..., 0)'$.[12]

The country risk premiums are assumed to follow completely affine specifications. Hence, the domestic country risk premium is defined by $\Lambda(t) = S_t \lambda$, where λ is an $N \times 1$ parameter vector with zero components corresponding to the foreign country-specific factors. Similarly, we assume that the foreign country risk premium is defined by $\Lambda^*(t) = S_t \lambda^*$, where λ^* is an $N \times 1$ parameter vector with zero components corresponding to the domestic country-specific factors. Our decomposition from joint to single-country term structure models can be applied in a similar way towards decomposition in the risk-neutral measure. Under the risk-neutral measure,

$$dX(t) = \widetilde{\kappa} \left[\widetilde{\vartheta} - X(t)\right] dt + S_t d\widetilde{W}(t) \tag{9}$$

where parameter vector $\widetilde{\vartheta}$ and parameter matrix $\widetilde{\kappa}$ are just aggregation from their single-country risk-neutral counterparts.

Before we proceed to the specification analysis of the joint term structure model, it is helpful to define two terms. First, a two-country joint term structure model is *decomposable* if it can be decomposed in physical measure into two single-country affine models. It can be shown that all affine joint term structure models are decomposable as long as the dynamics of the common factors do not depend on the dynamics of the country-specific factor. The dynamics of the country-specific factors may or may not depend on the dynamics of the common factors. Moreover, they are also

[12] Admissibility restrictions in Dai and Singleton (2000) and Aït-Sahalia and Kimmel (2002) also apply to our two-country affine term structure model. Additional restrictions on the structure of the affine two-country term structure are discussed in this section and in Appendix 1.

decomposable in the risk-neutral measure. We restrict our analysis to decomposable models, which has the important advantage of reducing the dimensionality of the joint model by one.

Second, a two-country joint term structure model is *symmetric* if the submodels for domestic and foreign countries have the same structure. We restrict our analysis to symmetric models. In these cases, each country has the same number of local and foreign factors, although the extent to which they are affected by these factors can be different. While the symmetry assumption simplifies the classification analysis, it can be easily relaxed and the methodology can be extended to non-symmetric models. In the next section, we provide specification analysis of three-factor decomposable symmetric joint term structure models for two countries.

3 Specification Analysis of Three-Factor Joint Term Structure Models

In this section, we demonstrate the structure of three-factor affine joint term structure models. We build upon the single-country specification analysis by Dai and Singleton (2000) described in the previous section. We classify all admissible models into subfamilies and within each subfamily derive the maximal model that nests existing models. In Appendix 1, we go on to classify symmetric four-factor joint term structure models. The main idea is to treat the two-country model as a single model where some factor(s) are common and the others are country-specific. We then study possible specifications for common factors. Dropping trivial cases, we come up with specifications for all interesting models.

3.1 $\mathbf{A}_0(3)$ *model*

For model $\mathbf{A}_0(3)$, the stochastic differential equation (8) takes the form

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} -X_{1t} \\ -X_{2t} \\ -X_{3t} \end{bmatrix} dt + d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{bmatrix}. \tag{10}$$

Note that for the general case when entries in the lower triangular submatrix of κ are non-zero, only X_3 can be an individual country factor in (10). If we assume a symmetric nature for dependence between single country term structures then we would have either one or three common factors. The non-trivial case is that of one common factor. Without loss of generality we can assume that X_1 is the common factor, X_2 is the domestic individual factor and X_3 is the foreign individual factor. In other words, domestic country is affected by X_1 and X_2, while foreign country is affected by X_1 and X_3. In this case, we need to add the additional restriction that $\kappa_{32} = 0$. Note that κ_{21} can be non-zero. Thus for each country we have an $\mathbf{A}_0(2)$ model with one common factor. Specifically, for the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t)$, and

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 \\ \kappa_{21} & \kappa_{22} \end{bmatrix} \begin{bmatrix} -X_{1t} \\ -X_{2t} \end{bmatrix} dt + d\begin{bmatrix} W_{1t} \\ W_{2t} \end{bmatrix},$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_1^* X_1(t) + \delta_3^* X_3(t)$, and the dynamics of factors are

$$d\begin{bmatrix} X_{1t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 \\ \kappa_{31} & \kappa_{33} \end{bmatrix} \begin{bmatrix} -X_{1t} \\ -X_{3t} \end{bmatrix} dt + d\begin{bmatrix} W_{1t} \\ W_{3t} \end{bmatrix}.$$

3.2 $\mathbf{A}_1(3)$ *model*

For model $\mathbf{A}_1(3)$, equation (8) becomes

$$\begin{aligned} d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} &= \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{2t} \\ -X_{3t} \end{bmatrix} dt \\ &+ \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_{1t}} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_{1t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{bmatrix}. \end{aligned} \tag{11}$$

The only symmetric case is when X_1 is the common factor. To ensure that X_2 and X_3 are single country factors we have to add extra restrictions: $\kappa_{23} = \kappa_{32} = 0$. For each country we have an $\mathbf{A}_1(2)$ model. Factor X_1 enters both drift and volatility terms in (11)), and no single country factor affects volatility directly. For the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t)$, and

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 \\ \kappa_{21} & \kappa_{22} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{2t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 \\ 0 & \sqrt{1+\beta_{21}X_{1t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \end{bmatrix},$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_1^* X_1(t) + \delta_3^* X_3(t)$, and the dynamics of factors are

$$d\begin{bmatrix} X_{1t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 \\ \kappa_{31} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{3t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 \\ 0 & \sqrt{1+\beta_{31}X_{1t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{3t} \end{bmatrix}.$$

3.3 $\mathbf{A}_2(3)$ *model*

Similarly, $\mathbf{A}_2(3)$ has the following dynamics:

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ -X_{3t} \end{bmatrix} dt \tag{12}$$

$$
+\begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_{1t}+\beta_{32}X_{2t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{bmatrix}
$$

In this subfamily, the only symmetric case is that in which X_3 is the common factor. This implies that $\kappa_{12} = \kappa_{21} = \kappa_{32} = \kappa_{33} = 0$ and $\beta_{31} = \beta_{32} = 0$. For the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_3 X_3(t)$, and

$$
d\begin{bmatrix} X_{1t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{3t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 \\ 0 & 1 \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{3t} \end{bmatrix},
$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_2^* X_2(t) + \delta_3^* X_3(t)$, and the dynamics of factors are

$$
d\begin{bmatrix} X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{22} & 0 \\ 0 & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{3t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{2t}} & 0 \\ 0 & 1 \end{bmatrix} d\begin{bmatrix} W_{2t} \\ W_{3t} \end{bmatrix}.
$$

In this case, the common factor only affects correlation through the drift term of (12)) while both single-country factors affect volatility as well as correlation. For each country we have a particular case of the $\mathbf{A}_1(2)$ model.

3.4 $\mathbf{A}_3(3)$ *model*

Finally, for the $\mathbf{A}_3(3)$ equation, (8) simplifies to

$$
d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ \vartheta_3 - X_{3t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 \\ 0 & 0 & \sqrt{X_{3t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{bmatrix}. \tag{13}
$$

Without loss of generality, X_1 is the common factor. This implies that $\kappa_{12} = \kappa_{13} = \kappa_{23} = \kappa_{32} = 0$. For the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t)$, and

$$
d\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 \\ \kappa_{21} & \kappa_{22} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 \\ 0 & \sqrt{X_{2t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \end{bmatrix},
$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_1^* X_1(t) + \delta_3^* X_3(t)$, and the dynamics of factors are

$$
d\begin{bmatrix} X_{1t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 \\ \kappa_{31} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_3 - X_{3t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 \\ 0 & \sqrt{X_{3t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{3t} \end{bmatrix}.
$$

Factor X_1 enters both drift and volatility terms in (13). Additionally, for each country the country-specific factor also affects both drift and volatility.

This completes the specification analysis for three-factor model. In Appendix 1, we report the results for the classification of the symmetric four-factor two-common-factor models. In the next section, we will discuss estimation and ranking methodologies.

4 Estimation and Ranking Methodologies

After classifying the models, the next task is to find out which model best characterizes the joint term structure of interest rates of dollars and euros. We approach this task in three steps. We first examine the principal components of the individual country term structures and the joint term structure. We then estimate the models using the approximate maximum likelihood method. Finally, we examine the in-sample and out-of-sample performances of different models. We now describe the empirical methodologies underlying these three steps.

4.1 Principal Component Analysis

Principal component analysis has been used before on bond markets in multiple contexts. Litterman and Sheinkman (1991) characterize the common factors that affect US bond markets. Wadhwa (1999) conducts principal component analysis of the implied volatilities in the swaptions market. Heidari and Wu (2003) include both interest rates and interest rate options and examine whether a common finite-dimensional system spans both types of instruments.

Following Litterman and Sheinkman (1991) and Heidari and Wu (2003), we conduct principal analysis to examine the factors underlying the US and Euro interest rate markets. As Heidari and Wu (2003) explain, although principal component analysis is traditionally performed on excess returns of assets, it can also be used directly on interest rates.

We perform principal component analysis of the Euribor and LIBOR rates to identify the common factors underlying the two yield curves. We call these the Euribor and LIBOR factors. We also form a portfolio of Euribor and LIBOR interest rates, perform principal component analysis of the joint term structure, and create factors underlying the joint term structure. We call these factors the common factors. We use principal component analysis to gauge how many factors are needed to characterize the Euribor and LIBOR term structures. As we will see in the results section, we need two or three factors to characterize the domestic term structures and four factors to characterize the joint term structure.

We then regress the LIBOR factors and Euribor factors on the common factors. We do this to examine how the common factors are related to the Euribor and LIBOR factors. Finally, we assess the number of common factors needed in the joint term structure model. In order to do that, we perform principal component analysis on the residuals of LIBOR after it is regressed on the first one or two principal components of the joint model. We do the same for Euribor. That way, we can examine the number of factors after the common factors are taken out.

The principal component analysis provides guidance for us on the total number of factors and the numbers of common vs local factors needed in explaining the LIBOR and Euribor term structure. However, one can say that the structural factors estimated in a formal term structure model could be different from that of the linear factors found in a principal component analysis. To assess this possibility, Heidari and Wu (2003) conduct a simulation analysis.

In the simulation, Heidari and Wu estimate a three-factor Gaussian affine model on LIBOR using the quasi-maximum likelihood method and an extended Kalman filter. They use these extracted factors to produce a simulated series of LIBOR rates. They then conduct principal component analysis on the simulated interest rates. By design, the simulated series of LIBOR rates are governed by three dynamic factors in a non-linear function. When Heidari and Wu (2003) conduct principal component analysis on the LIBOR series, they show that exactly three principal components explain 100% of the series. Heidari and Wu's results echo Singleton and Umantsev (2003)'s findings that LIBOR rates are approximately linear in terms of the state variables in an affine model. Hence, principal component analysis is useful in identifying the number of factors in an affine framework.

4.2 Approximate Maximum Likelihood Estimation

After we identify the number of factors, we proceed to estimate the models. The best estimation method for affine term structure model should be maximum likelihood estimation, given its consistency and asymptotic efficiency. However, except for the case of a multi-factor Gaussian model, the transition density of an affine model generally has no closed-form. In these cases, maximum likelihood estimation is infeasible and alternative estimation methods have to be used.

Most papers use the quasi-maximum likelihood estimation (e.g. Han and Hammond (2003), Dewachter and Maes (2001), Brennan and Xia (2006), Tang and Xia (2006)) because of its ease of application. As Aït-Sahalia and Kimmel (2002) point out, two assumptions are needed in quasi-maximum likelihood estimation. First, the density of the state vector conditional on the previous observation is assumed to follow a multivariate Gaussian distribution. Second, the mean vector and covariance matrix of the state vector are assumed to be proportional to the length of time between observations. As Aït-Sahalia and Kimmel (2002) point out, both these assumptions are unlikely to hold. Only some affine yield models have a Gaussian transition density, and even in those cases, the assumptions of quasi-maximum likelihood estimation regarding the means and variances of the transition density are not accurate.

Another popular method is the simulation-based efficient method of moments of Galant and Tauchen (1996). Dai and Singleton (2000) use this in their affine term structure estimation. The efficient method of moments is efficient as the number of moment conditions goes to infinity with the number of data observations. However, Duffee and Stanton (2002) find that this method performs poorly in a small sample in the context of affine term structure model. Other potential estimation methods include simulated maximum likelihood estimation of Brandt and Santa-Clara (2002), and Durham and Gallant (2002), or the empirical characteristic function method of Singleton (2001) and Jiang and Knight (2002). These methods are computationally intensive for a scalar diffusion and especially difficult for multivariate diffusions.

Following Aït-Sahalia (1996, 2002a, 2002b) and Aït-Sahalia and Kimmel (2002), we estimate the joint term structure models using the approximate maximum likelihood method. The approximate likelihood method provides extremely fast and accurate estimations for affine models (see the comparison in Jensen and Polsen (2002) and Egorov, Li and Xu (2003)) when the data are sampled daily as in our case. The disadvantage of this method is that it requires preliminary work in obtaining a closed-form formula for the approximate likelihood through linear expansions. Fortunately, Aït-Sahalia and Kimmel (2002) derive this analytic formula for two-factor and three-factor models. Since the relationship between the state vector and bond yields is affine, as in equation (5), we can derive the transition function of the bond yields from the transition function of the state vector through a change of variables and multiplication by a Jacobian. Let $p_X(\Delta t, x|x_o;\theta)$ denote the transition function, that is the conditional density of $X(t+\Delta t) = x$ given $X(t) = x_o$. Let $p_Y(\Delta, y|y_o;\theta)$ also denote the transition function of the vector of yields $Y(t+\Delta t) = y$ given $Y(t) = y_o$. In our case, with daily data, Δt is the inverse of the number of trading days in a year ($\Delta t \approx 1/250$). We obtain latent factors X by inverting a system of equations (5) by taking enough yields. To guarantee invertability of state vector X the rank of this system should be equal to the number factors N. This system can be written in matrix form as $Y = \Gamma_o(\theta) + \Gamma'(\theta)X$. It follows that $X = \Gamma'^{-1}(\theta)(Y - \Gamma_o(\theta))$. Hence,

$$p_Y(\Delta t, y|y_o;\theta) \equiv \Gamma'^{-1}(\theta) p_X(\Delta t, \Gamma'^{-1}(\theta)(y - \Gamma_o(\theta))|\Gamma'^{-1}(\theta)(y_o - \Gamma_o(\theta));\theta). \tag{14}$$

Noting that the yields vector follows a Markov process and applying Bayes rule, the log-likelihood function for discrete data on the yield vector y_t sampled at dates $t_0, t_1, ..., t_n$ is obtained:

$$L_n(\theta) \equiv n^{-1} \sum_{i=1}^{n} l_Y\left(t_i - t_{i-1}, y_{t_i}|y_{t_{i-1}};\theta\right), \tag{15}$$

where $l_Y = \ln p_Y$.

To estimate this likelihood function, we need to derive a closed-form approximation for l_Y and for the log-likelihood function of the discretely sampled vector of yields. Aït-Sahalia and Kimmel (2002) use the highly accurate linear expansion method described in Aït-Sahalia (1999, 2002) to derive the analytical formula in two- and three-factor affine term structure models. In our classification, we break down the joint term structure model into two-country models with a lower dimension of factors. This makes it possible to apply Aït-Sahalia and Kimmel (2002)'s results for three-factor models toward our four-factor models of joint term structure.[13]

In summary, here is an overview of the approximate maximum likelihood estimation procedure. Given an initial value of parameter vector θ we can can estimate $\Gamma_o(\theta)$ and $\Gamma(\theta)$. Affine structure implies a system of ordinary differential equations for $\Gamma_o(\theta)$ and $\Gamma(\theta)$. It also provides us linear transformation from the observed yields $Y(t_i)$ to the latent factors (or state variables) $X(t_i)$ for $i = 0, 1, 2, ..., n$. Close form approximation of Aït-Sahalia provides transition density

[13] For five-factor joint term structure models with three common factors, or for four-factor models with four common factors, we need closed-form formulas for approximate likelihood of four-factor affine models. We use Mathematica code to obtain these formulas, which are available from authors upon request.

$p_Y\left(t_i - t_{i-1}, y_{t_i}|y_{t_{i-1}};\theta\right)$ for $i = 1, 2, ..., n$ and thus likelihood function $L_n(\theta)$. In the end we maximize the likelihood function $L_n(\theta)$. Thus, the only role the affine structure plays in the estimation method is to simplify the transformation from observed yields to state variables.

This procedure can be extended for the case when we want to use more yields than the number of factors in our model. In this situation, we usually assume that some yields are observed with an error and make assumptions on the structure of these errors. The simplest assumption is to assume that all errors are independent and normally distributed with unknown standard deviation and zero mean. The log likelihood function $L_n(\theta)$ in that case should be augmented by additional term accounting for these errors.

4.3 In-Sample and Out-of-Sample Tests

To assess the in-sample and out-of-sample goodness of fit of the models, we adopt an omnibus nonparametric specification test for continuous time models derived from Hong, Li and Zhao (2006). This test is based upon the transition density capturing the full dynamics of a continuous time process. The basic idea is the following: if a model is correctly specified, then the probability integral transform of data via the model transition density should be i.i.d. U[0,1]. This probability integral transform can be called the "generalized residuals" of the continuous time model. We test this i.i.d. U[0,1] hypothesis for the model generalized residuals by comparing the kernel estimator of the joint density of the generalized residuals with the product of two U[0,1] densities.

Dai and Singleton (2000) point out that it has been challenging to formally compare the relative goodness of fit of different affine models, given that these models have non-nested specifications with different estimation methods. Our non-parametric approach allows comparison of the performance across different non-nested models via a metric measuring the distance of the model generalized residuals from i.i.d. U[0,1]. As the transition density can capture the full dynamics of $\{X_t\}$, the omnibus test has power against any model misspecification. In addition, this test significantly improves the size and power performance of the marginal density-based test.

Suppose we have a random sample of interest rates $\{r_{\tau\Delta}\}_{\tau=1}^{L}$ of size L, where Δ is the time interval at which the data are observed or recorded. For a given continuous-time interest rate model, there is a model-implied transition density of

$$\frac{\partial}{\partial r} P\left(r_{\tau\Delta} \leq r | I_{(\tau-1)\Delta}, \theta\right) = p(r, \tau\Delta | I_{(\tau-1)\Delta}, \theta), \qquad 0 < r < \infty,$$

where θ is an unknown finite-dimensional parameter vector, $I_{(\tau-1)\Delta} = \{r_{(\tau-1)\Delta}, r_{(\tau-2)\Delta}, ..., r_{\Delta}\}$ is the information set available at time $(\tau - 1)\,\Delta$. We divide the whole sample into two sub-samples: an estimation sample $\{r_{\tau\Delta}\}_{\tau=1}^{R}$ of size R, which is used to estimate model parameters, and a forecast sample $\{r_{\tau\Delta}\}_{\tau=R+1}^{L}$ of size $n = L - R$, which is used to evaluate density forecast.[14] We can then

[14] One can also use rolling estimation or recursive estimation. We expect that our test procedures are applicable to these different estimation methods under suitable regularity conditions.

define the probability integral transform of the data in the forecast sample with respect to the model-implied transition density:

$$Z_\tau(\theta) \equiv \int_{-\infty}^{r_{\tau\Delta}} p(r, \tau\Delta | I_{(\tau-1)\Delta}, \theta) dr, \qquad \tau = R+1, ..., L. \tag{16}$$

If the continuous-time model is correctly specified in the sense that there exists some θ_0 such that the model-implied transition density $p(r, \tau\Delta | I_{(\tau-1)\Delta}, \theta_0)$ coincides with the true transition density of interest rates, then the transformed sequence $\{Z_\tau(\theta_0)\}$ is i.i.d. $U[0,1]$. Intuitively, the $U[0,1]$ distribution indicates proper specification of the stationary distribution of $r_{\tau\Delta}$, and the i.i.d. property characterizes the correct specification of its dynamic structure. If $\{Z_\tau(\theta)\}$ is not i.i.d. $U[0,1]$ for all $\theta \in \Theta$, then $p(r, \tau\Delta | I_{(\tau-1)\Delta}, \theta)$ is not optimal and there exists room for further improvement. Thus density forecast evaluation boils down to testing whether $\{Z_\tau(\theta)\}$, which is often referred to as the "generalized residuals" of the model-implied transition density $p(r, \tau\Delta | I_{(\tau-1)\Delta}, \theta)$, follows i.i.d. $U\left[0,1\right]$.

We measure the distance between a forecast density model and the true transition density by comparing a kernel estimator $\hat{g}_j(z_1, z_2)$ for the joint density of $\{Z_\tau, Z_{\tau-j}\}$ and unity, the product of two $U[0,1]$ densities, where j is a lag order. The detail of this kernel estimator is reported in Appendix 2. Simulation studies in Hong and Li (2005) show that the tests perform well in small samples even for highly persistent financial data.

Hong and Li (2005) propose an in-sample specification test that uses a quadratic form between $\hat{g}_j(z_1, z_2)$ and 1, the product of two $U[0,1]$ densities. This test is extended to the out-of-sample context in Hong, Li and Zhao (2006) as

$$\hat{Q}(j) \equiv \left[(n-j)h \int_0^1 \int_0^1 \left[\hat{g}_j(z_1, z_2) - 1\right]^2 dz_1 dz_2 - h\Psi_h^0\right] / V_0^{1/2}, \quad j = 1, 2, ..., \tag{17}$$

where j is a prespecified lag order, the nonstochastic centering and scaling factors are

$$\Psi_h^0 \equiv \left[(h^{-1} - 2)\int_{-1}^1 k^2(u)du + 2\int_0^1 \int_{-1}^b k_b^2(u)dudb\right]^2 - 1, \tag{18}$$

$$V_0 \equiv 2\left[\int_{-1}^1 \left[\int_{-1}^1 k(u+v)k(v)dv\right]^2 du\right]^2, \tag{19}$$

and $k_b(\cdot) \equiv k(\cdot)/\int_{-1}^b k(v)dv$. Note that the modification of the kernel $k(\cdot)$ in the boundary regions affects the centering constant Ψ_h^0, although not the asymptotic variance V_0.

Under suitable regularity conditions, $\hat{Q}(j) \to N(0,1)$ in distribution when the continuous-time model is correctly specified. In a simulation experiment mimicking the dynamics of U.S. interest rates via the Vasicek model, Hong and Li (2005) find that the in-sample version of $\hat{Q}(j)$ has good sizes for $n \geq 250$ (i.e., about one year of daily data). This is a substantial improvement over other nonparametric tests (see Aït-Sahalia 1996 and Pritsker 1998).

With various choices for lag order j, $\hat{Q}(j)$ can reveal useful information regarding which lag order significantly departs from i.i.d. $U[0,1]$. This is analogous to the use of the sample autocorrelation function in the linear time series context. If a large set of $\{\hat{Q}(j)\}$ is considered, then some of them will probably be significant even if the null is true, due to statistical sampling variation. In fact, on average one out of twenty will be significant at the 5% level under the null. On the other hand, the choice of lag order j is expected to have a significant impact on the power of $\hat{Q}(j)$. Moreover, when comparing two different models, it is desirable to use a single portmanteau test statistic. For this purpose, we consider the following portmanteau evaluation statistic

$$\hat{W}(p) = \frac{1}{\sqrt{p}} \sum_{j=1}^{p} \hat{Q}(j). \tag{20}$$

Like many time series test statistics, we still have to choose the lag truncation order p. The power of $\hat{W}(p)$ is still affected by the choice of p, but not as much as the power of $\hat{Q}(j)$ is affected by the choice of individual lag order j. We can show that for any p, $\hat{W}(p) \to N(0,1)$ in distribution when the continuous-time model is correctly specified. Intuitively, when the forecast model is correctly specified, we have $cov[\hat{Q}(i), \hat{Q}(j)] \to 0$ in probability for $i \neq j$ as $n \to \infty$. That is, $\hat{Q}(i)$ and $\hat{Q}(j)$ are asymptotically independent whenever $i \neq j$. Thus, the portmanteau test statistic $\hat{W}(p)$ is a normalized sum of approximately i.i.d. N(0,1) random variables, and so is asymptotically N(0,1). This test may be viewed as a generalization of the popular Box-Pierce-Ljung type autocorrelation test from a linear time series context to a continuous-time context with an out-of-sample setting.

Under model misspecification, we can show that as $n \to \infty, \hat{Q}(j) \to \infty$ in probability whenever $\{Z_\tau, Z_{\tau-j}\}$ are not independent or $U[0,1]$. As long as model misspecification occurs such that there exists some lag order $j \in \{1, ..., p\}$ at which $\hat{Q}(j) \to \infty$, we have $\hat{W}(p) \to \infty$ in probability. Therefore, the portmanteau test statistic $\hat{W}(p)$ can be used as an omnibus procedure to evaluate the out-of-sample density forecast performance of a continuous-time model.

5 Data Description and Empirical Results

We apply the econometric methodology described above on the euro interest rates, i.e. Euribor and the US interest rates, i.e. LIBOR. We use daily data from July 1, 1999 til June 30, 2003. It consists of 1) Euribor with maturities of 1, 3, 6 and 9 months, 2) LIBOR with maturities of 1, 3 and 6 months, 3) Euro-Euribor interest rate swap with maturities of 1 to 10 years, and 4) USD-LIBOR interest rate swap with maturities of 1 to 10 years. We use daily exchange rate data provided by GTIS.

As our empirical inquiry is specific to dollars and euros, we are limited to the relatively short time sample after the Euro was launched. The euro was launched on January 1, 1999 for accounting purposes and electronic fund transfers, although the euro notes and coins were not issued as legal tender until 2002.[15] To minimize any noise due to the initial adoption of the euro, we start our data series on July 1, 1999.

Both Euribor and LIBOR, denoted here for simplicity as $B(\tau)$ are simply compounded interest rates, related to continuously compounded spot rates $Y(\tau)$ by

$$B(\tau) = \frac{1}{\tau}\left(e^{\tau Y(\tau)} - 1\right), \tag{21}$$

where the maturities τ follow the actual-over-360-day-counting convention for both currencies, starting two business days forward. The swap rates for both LIBOR and Euribor have payment intervals of six months and are related to the zero prices (discount factors) by

$$SWAP(\tau) = 2\frac{1 - B(\tau)}{\sum_{i=1}^{2\tau} B(i/2)}. \tag{22}$$

Using these relationships we recover LIBOR and Euribor zero coupon yields for maturities starting from 1 year till 10 years using swap rates. Table 1 presents the summary statistics for the levels and changes in these zero coupon bond yields over the sample period. Panel A shows the means, medians, standard deviations, skewness and kurtosis for the levels and changes of LIBOR of different maturities. Panel B shows the same statistics for Euribor. Panel C shows the correlation between the levels and changes of LIBOR and Euribor of different maturities. Figure 1 presents these zero coupon bond yields graphically.

We first perform principal component analysis in order to understand how many factors we need to capture most of the behavior of LIBOR and Euribor. We consider the principal components of the joint term structure using both LIBOR and Euribor of all maturities. We then consider the first three principal components for LIBOR and Euribor separately.

Table 2 reports the variations of LIBOR, Euribor and the joint LIBOR-Euribor portfolio that are explained by the first seven principal components. Although most of the variation in LIBOR, Euribor and joint term structure seems to be captured by the first principal component, it is commonly accepted that a single factor is not enough for modeling even single-country term structure for interest rates. More factors would typically be desirable, but computation time imposes a constraint. As a result, usually three factors (and occasionally two factors) are used to explain the term structure in a single country.

Table 2 also shows that the first three principal components capture 99.97% variation in LIBOR and 99.93% in Euribor. In the case of the joint term structure, the first two principal components explain 98.28% of the variation; three principal components explain 99.68%; and four principal components explain 99.84%. Only when the number of principal components goes up to four is

[15]The euro is the monetary unit of the European Union. Countries that use the euro include twelve of the fifteen EU member states. Greece, the 12 th member, did not adopt the Euro until 2001. All the others adopted the euro in 1999. The euro replaced the national currencies previously used in these countries. Before adopting the EU, these countries had to meet economic criteria set by the EU on inflation levels, budget deficits, and currency stability.

the variation captured comparable to the 99.83% for the first two principal components of LIBOR. Thus, in order to capture variation of up to the first three factors in LIBOR or Euribor, we need at least four term structure factors. Obviously, explaining 0.16% more (from 99.68% to 99.84%) of the variation in the joint term structure only represents weak suggestive evidence that we need four factors instead of three. To examine this issue more carefully, we have to examine the identities of these factors.

Table 3 reports the common factors in the joint term structure models. We regress the first three principal components of LIBOR (i.e. LIBOR factors) on the first six principal components of the joint term structure (i.e. common factors). We then regress the first three principal components of Euribor (i.e. Euribor factors) on the first six common factors. The results of these regressions are summarized in Table 3. Panel A reports the results from regressing LIBOR and Euribor factors on six common factors. As seen in the first row, when the first LIBOR factor is regressed upon the first six common factors, the coefficient on the first common factors is 0.905. This suggests that the first common factor is most associated with the first LIBOR factor. Similarly, we find that the second common factor is most associated with the second Euribor factor with a coefficient of 0.874. The third common factor is most associated with the second LIBOR factor, with a coefficient of 0.868. The fourth common factor is most associated with the second Euribor factor, with a coefficient of -0.831. This suggests that we need to go up to a four-factor joint term structure model in order to capture the variation embedded in the second Euribor factor in a two-factor domestic term structure model.

Panel B reports the R-square from regressing LIBOR and Euribor factors on the first five factors estimated in the joint term structure model. The first common factor explains 99.245% of variation of the first LIBOR factor and almost none of the variation of the second and third LIBOR factors. On the other hand, the first common factor explains only 86.267% of variation of the first Euribor factor. If we go up to the first three factors in a joint model, we can explain most of the variations in the first LIBOR factor (99.99%), second LIBOR factor (96.5%), and first Euribor factor (99.97%). However, we are not able to fully capture the variations in the second Euribor factor (only 83.37%). Only when we go up to the the fourth joint component will we be able to explain 99.96% of the second Euribor component. If the goal is to capture the first two factors of each domestic model, then we need a four-factor joint term structure model. The first three factors do not seem to be enough. Furthermore, if the goal is to capture the first three factors of each domestic model, then we would need a five-factor or even a six-factor model to capture the variation. Hence, Table 2 shows that at least a four-factor model is needed to capture the joint dynamics.

Panel C reports the results from regressing the log of exchange rates on the first six joint term structure common factors. The R-square is 0.59%, implying that the exchange rate is largely orthogonal to the common factors from the joint term structure. This confirms the results in Inci and Lu (2004) and Han and Hammond (2003) that exchange rate is a separate factor.

In summary, from Table 3 we see that the exchange rate factor only has a minimal relationship to the dynamics of the yield curve factors. In contrast, the movement in LIBOR affects the yield curve movement in Euribor and vice versa. Hence, the rest of the empirical work will focus on establishing the local and common factors driving the two yield curves.

In Table 4, we turn our attention to the number of local factors versus common factors in a joint term structure model. Each column presents the result of a principal component analysis and reports the variation explained by the first seven principal components. To set the benchmark, column 2 reports the result of a principal component analysis conducted on LIBOR only (same as in table 1). Column 3 contains the results of the principal component analysis on the residuals of LIBOR after its regression on the first common factor (i.e. first principal component in joint term structure). Column 4 contains the results of the principal component analysis of residuals for LIBOR after the regression on the first two common factors. These columns suggest the number of local factors needed to explain the variation in LIBOR after a common factor is incorporated. After the first common factor, one local LIBOR factor would only capture 60.75% of the variation. We still need two local factors to capture most of the variations. After two common factors, however, one local LIBOR factor would explain about 90.56% of the variation. Therefore, the best combination here is to have two common factors and a local LIBOR factor. Columns 5, 6, 7 present the results for Euribor. After 2 common factors, 88.23% would be explained by the first local factor. Again, the best combination here is to have two common factors and a local Euribor factor.

Overall, Table 3 and Table 4 show that the best way to explain the two yield curves is to have four factors in total: one LIBOR local factor, one Euribor local factor, and two common factors. If we have only three factors, then the best case scenario is to have a local LIBOR factor, a local Euribor factor and a common factor. In this case, while the first LIBOR factors and Euribor factors can be explained by the common factor reasonably well, at 99.25% and 86.27% respectively, the second and third LIBOR and Euribor factors cannot be explained well by the local factors. Only 60.76% of the LIBOR variation and about 85% of Euribor variation can be explained, which is far from satisfactory. Motivated by the above evidence, in our estimation we consider four-factor models in the paper. The results in Tables 3 and 4 show that a model with two common components and one country-specific component in each country can explain over 99.9% of first principal components of LIBOR and Euribor as well as 91% of the residual variation in the LIBOR and 88.2% of the residual variation in Euribor.

Another motivation for consideration of models with more than three factors comes from recent papers by Cochrane and Piazzesi (2002, 2004) and Dai, Singleton and Yang (2004). This literature addresses bond risk premia, and, in particular, it shows that up to five factors might be needed for affine models to forecast bond risk premia in US government bonds. Hence, a four-factor joint term structure model would be more preferable to a three-factor model.

Since the increasing number of factors significantly increases the number of parameters to be estimated, this poses a substantial challenge for the term structure literature. This makes it crucial to adopt approximate maximum likelihood estimation, a fast and efficient method, as described in the methodology section. We conduct the approximate maximum likelihood estimation for models $\mathbf{A}_0(4), \mathbf{A}_1(4), \mathbf{A}_{2,1}(4), \mathbf{A}_{2,2}(4), \mathbf{A}_3(4)$ and $\mathbf{A}_4(4)$, and examine the in-sample and out-of-sample performance of these models.

Table 5 reports the non-parametric Pormanteau statistics for both in-sample and out-of sample performance. We show the results for four factor completely affine models of joint term structure of LIBOR and Euribor. We choose the first half of the sample (from July 1, 1999 to June 30, 2001) as the estimation sample and the second half (from July 1, 2001 to June 30, 2003) as the forecast sample.

Table 5 also shows the non-parametric Pormanteau statistics $W(p)$ for $p = 5, 10$, and 20, where p represents the lag truncation order. The use of $W(p)$ with various lag orders can reveal which lag order significantly departs from iid. U[0,1]. As a robustness check, we separately report results when we limit our sample to only two-year and five-year zero coupon bonds for LIBOR and Euribor.

The $W(p)$ statistics have a standard normal distribution and the results show that all models are rejected. One caveat about this non-parametric test is that it is extremely powerful. As an in-sample statistic, it tests whether the entire distribution is captured by the model. As an out-of-sample forecast statistic, it tests whether the forecast of the entire distribution is accurate. As a result, this test often rejects a model which may match the moments very well and which may not be rejected in other specification tests. We therefore use this specification test as a way to compare and rank different models rather than a way to reject models. In so comparing models, out-of-sample performance is especially important. This is because introducing more complicated models with larger numbers of estimated parameters creates a potential danger of overfitting noise in the data. Thus the model that performs better out-of-sample might be more effective in capturing the underlying data generating process.

Our results show that the $\mathbf{A}_1(4)$ model is the best model to capture joint term structure both in-sample and out-of-sample. The next best sets of models in terms of performance are $\mathbf{A}_0(4)$ and $\mathbf{A}_{2,2}(4)$. The $\mathbf{A}_{2,1}(4)$ and $\mathbf{A}_3(4)$ models perform significantly worse, and $\mathbf{A}_4(4)$ has the worst performance by far both in-sample and out-of-sample.

We can interpret our results in light of the tradeoff between flexibility in variance and factor correlations as pointed out in Dai and Singleton (2000). Recall that in $\mathbf{A}_m(N)$, m is the number of state variables that affect the instantaneous variance of $X(t)$. The Gaussian model ($m = 0$) implies that none of the state variables affect the variance. At the other extreme, when $m = 4$, all state variables affect the variances. On the surface, it might look as though $\mathbf{A}_4(4)$ would be the most flexible specification and should have the best empirical performance. However, because parameter restrictions are imposed to ensure admissibility, models with $m = 4$ will have zero correlation across the state variables. Hence, as Dai and Singleton point out, in moving from $m = 0$ to $m = 4$, there is a tradeoff between the flexibility in specifying the conditional variance and that in allowing for conditional correlation among the factors.

Our findings can be interpreted in light of this tradeoff. We find that $\mathbf{A}_1(4)$ is the best model. In the $\mathbf{A}_1(4)$ model, both countries' term structures are $\mathbf{A}_1(3)$. So, both countries have richer volatility and correlation structures while providing substantial freedom in correlation of factors. In $\mathbf{A}_1(4)$, the volatility is described by a single common factor while correlation is driven by two common factors and country specific factors. This provides maximum flexibility in fitting correlation of factors among models that have non-Gaussian structure (i.e. among models $\mathbf{A}_1(4)$, $\mathbf{A}_{2,1}(4)$, $\mathbf{A}_{2,2}(4)$,$\mathbf{A}_3(4)$, $\mathbf{A}_4(4)$.)

Model $\mathbf{A}_0(4)$ allows for maximal flexibility in fitting factor correlation, and thus the correlation of interest rates. This results in $\mathbf{A}_0(4)$ having an overall good performance. But in this case, none of the state variables can affect the variance. Hence, the variance is restricted to be homoskedastic and hence the model performs worse than $\mathbf{A}_1(4)$.

Models $\mathbf{A}_{2,1}(4)$, $\mathbf{A}_{2,2}(4)$ and $\mathbf{A}_3(4)$ all have two factors driving volatility. They all assume a restricted $\mathbf{A}_2(3)$ for each country. In model $\mathbf{A}_{2,2}(4)$, volatility in each country is driven by the country-specific factor, while both common factors are Gaussian. In model $\mathbf{A}_{2,1}(4)$, for each country the volatility is driven by both common factors. In model $\mathbf{A}_3(4)$, for each country the volatility is driven by one common factor and the country-specific factor. This in turn gives less freedom to capture correlation among factors. Among these three models, $\mathbf{A}_{2,1}(4)$ fits the data better than $\mathbf{A}_3(4)$ and $\mathbf{A}_{2,2}(4)$. But overall, our estimation suggests that these models perform worse than $\mathbf{A}_1(4)$ both in-sample and out-of-sample.

Finally, for model $\mathbf{A}_4(4)$, each country's term structure is described by a restricted $\mathbf{A}_3(3)$ model. Although these models allow for time varying volatility and correlated factors, they impose zero conditional correlations with positive unconditional correlation. This is clearly counterfactual. This results in $\mathbf{A}_4(4)$ having the worst fit of all six models both in-sample and out-of-sample.

In summary, our tables examine the feature of an ideal affine model that captures the joint term structures in increasing specificity. Tables 2, 3, and 4 suggest that this term structure model should consist of four factors. In particular, Table 4 suggests that two of these four factors should be common factors and the other two should be local factors. Finally, Table 5 shows that this model should have an $\mathbf{A}_1(4)$ factor structure. That is, in this 4-factor model, one common factor should drive volatility, while the other common factor and the local factors should not drive volatility.

6 Conclusion

International bond market positions in euros and dollars create exposure to different interest rate risks and exchange rate risk for banks and investors. In order to manage these risks, one needs an appropriate model for the joint term structure of interest rate for dollars and euros. In this paper, we provide a thorough analysis of multi-factor joint affine term structure models for dollars and euros.

Our paper systematically examines, decomposes and classifies joint term structure models with up to four local and common factors. We then provide new evidence on the joint term structure using daily data in LIBOR and Euribor from July 1999 to June 2003.

Our methodology builds on, and extends, three main streams of research. First, we take advantage of recent advances in domestic term structure specification analysis. Specifically, we show that the system for classifying domestic affine model structures can be adapted to provide classification of joint affine models in two countries. Second, we extend the applications of estimation techniques recently developed by Aït-Sahalia (1996, 2002a, and 2002b) and Aït-Sahalia and Kimmel (2002). The absence of a closed-form solution for transition density of affine models makes maximum likelihood estimation infeasible. We show that Aït-Sahalia's approximate maximum likelihood techniques are applicable in estimating joint term structure models. Third, we utilize the latest advances in nonparametric tests. Recently developed by Hong and Li (2005) and Hong, Li and Zhao (2006), these nonparametric tests provide powerful alternatives to standard out-of-sample tests. We apply these nonparametric tests to identify the factor model with the best out-of-sample properties.

We find that a four-factor model with two common factors and two local factors best describes the term structures of Euribor and LIBOR. Our analysis of principal components for daily data for LIBOR and Euribor motivates us to study a new set of four-factor models. We conclude that the best model is one with two common factors and two local factors. This model has an $\mathbf{A}_1(4)$ factor structure where one common factor drives volatility, while the other common factor and the local factors do not drive volatility. This model seems to provide the best tradeoff in terms of flexibility in modeling correlation and volatility. This opens up the possibility of improving previously developed models and explaining some remaining puzzles about joint dynamics of interest rates.

We plan to extend our current study in a few different directions. We first intend to study the implications of our proposed models for the forward premium puzzle. Motivated by our principal component analysis that exchange rate plays a secondary role in affecting domestic term structure, the current study only focuses on capturing the joint dynamics of the interest rates. In future extensions, we can study exchange rate dynamics more specifically and examine the forward premium puzzle. We also plan to develop a methodology to characterize the risk involved in an international bond portfolio. The joint term structure model allows us to characterize in detail the Value-at-Risk (VaR) and other risk characteristics of these portfolios. We can then study the out-of-sample performance of joint term structure models in forecasting the density and VaR of international bond portfolios. In addition, our methodology can be extended to other affine models like those developed by Duffee (2002) and Cheridito, Filipović, and Kimmel (2006). Finally, we plan to combine modeling joint term structure of interest rates with the dynamics of interest rate derivatives like individual countries' caps and floors and possibly cross-country derivatives.

REFERENCES

Ahn, D., 2004, Common Factors and Local Factors: Implications for Term Structures and Exchange Rates, *Journal of Financial and Quantitative Analysis* 39, 69-101.

Aït-Sahalia, Y., 1996, Testing Continuous-Time Models of the Spot Interest Rate, *Review of Financial Studies* 9, 385-426.

Aït-Sahalia, Y., 2002a, Maximum-Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approach, *Econometrica* 70, 223-262.

Aït-Sahalia, Y., 2002b, Closed-Form Likelihood Expansions for Multivariate Diffusions, Working paper, Princeton University.

Aït-Sahalia, Y., and R. Kimmel, 2002, Estimating Affine Multifactor Term Structure Models Using Closed-Form Likelihood Expansions, Working paper, Princeton University.

Backus, D.K., S. Foresi, and C.I. Telmer, 2001, Affine Term Structure Models and the Forward Premium Anomaly, *Journal of Finance* 56, 279-304.

Bakshi, G., and Z. Chen, 1997, Equilibrium Valuation of Foreign Exchange Claims, *Journal of Finance* 52, 799-826.

Bansal, R., 1997, An Exploration of the Forward Premium Puzzle in Currency Markets, *Review of Financial Studies* 10, 369-403.

Brandt, M., and P. Santa-Clara, 2002, Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets, *Journal of Financial Economics* 63, 161-210.

Brennan, M., and Y. Xia, 2006, International Capital Markets and Foreign Exchange Risk, forthcoming, *Review of Financial Studies*

Cheridito, P., D. Filipović, and R. Kimmel, 2006, Market Price of Risk Specifications for Affine Models: Theory and Evidence, *Journal of Financial Economics*, forthcoming.

Cochrane J. H., and Piazzesi, M. 2004, Bond Risk Premia, NBER Working paper No. 9178.

Cox, J.C., J.E. Ingersoll and S.A. Ross, 1985, A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385-407.

Dai, Q., and K. Singleton, 2000, Specification Analysis of Affine Term Structure Models, *Journal of Finance 55,* 1943-1978.

Dai, Q., and K. Singleton, 2002, Term Structure Dynamics in Theory and Reality, *Review of Financial Studies* 16, 631-678.

Dai, Q., K. Singleton, and W. Yang, 2004, Predictability of Bond Risk Premia and Affine Term Structure Models, Stanford University working paper.

Dewachter, H., and K. Maes, 2001, An Admissible Affine Model for Joint Term Strucutre Dynamics of Interest Rates, KU Leuven working paper.

Diebold, F.X., J. Hahn and A.S. Tay**,** 1999, Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns of Foreign Exchange, *Review of Economics*

and Statistics 81, 661-673.
Duffee, G., 2002, Term Premia and Interest Rate Forecasts in Affine Models, *Journal of Finance* 57, 405-443.
Duffie, D., and R. Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6, 379-406.
Durham, G. and R. Gallant, 2002, Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes, *Journal of Business and Economic Statistics* 20, 297-316.
Egorov, A., H. Li, and Y. Xu, 2003, Maximum Likelihood Estimation of Time-Inhomogeneous Diffusions, *Journal of Econometrics* 114, 107-139.
European Central Bank, 2004, The Euro Bond Market Study, December, Frankfurt, Germany.
Filatov, Victor S. and Peter Rappoport, 1992, Is Complete Hedging Optimal for International Bond Portfolio? *Financial Analysts Journal*, July-August, 37-47.
Gallant, A.R. and G. Tauchen, 1996, Which Moments to Match?, *Econometric Theory* 12, 657-681.
Gallant, A.R. and G. Tauchen, 2001, Efficient Method of Moments, UNC Working paper.
Glen, J. and P. Jorion, 1993, Currency Hedging for International Portfolios, *Journal of Finance* 48, 1865-86.
Han B., and P. Hammond, 2003, Affine Models of Joint Dynamics of Exchange Rates and Interest Rates, Working paper.
Hansen, L.P. and J.A. Scheinkman, 1995, Back to the Future: Generating Moment Implications for Continuous Time Markov Processes, *Econometrica* 63, 767-804.
Heidari, M. and L. Wu, 2003, Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates?, *Journal of Fixed Income*, 75-84.
Hodrick R., 1987, The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets (Harwood Academic Publishers, New York).
Hodrick R., and M. Vassalou, 2002, Do we need Multi-country Models to Explain Exchange Rate and Interest Rate and Bond Return Dynamics?, *Journal of Economic Dynamics and Control* 26, 1275-1299.
Hong, Y., and H. Li, 2005, Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates, *Review of Financial Studies* 18, 37-84.
Hong, Y., H. Li, and F. Zhao, 2006, Can the Randon Walk Model be Beaten in Out-of-Sample Density Forecasts: Evidence from the Intraday Exchange Rates, working paper, Cornell University.
Inci, Ahmet Can and Biao Lu, 2004, Exchange Rates and Interest Rates: Can Term Structure Models Explain Currency Movements? *Journal of Economic Dynamics and Control*, Vol. 28, pp. 1595-1624.
Jensen, B., and R. Poulsen, 2002, Transition Densities of Diffusion Processes: Numerical Comparison of Approximation Techniques, *Journal of Derivatives* 9, 1-15.

Jiang, G. and J. Knight, 2002, Estimation of Continuous-Time Processes via the Empirical Characteristic Function, *Journal of Business and Economic Statistics* 20, 198-212.

Levich, Richard M., 2001, International Financial Markets: Prices and Policies, 2nd Edition.

Levich, Richard, and Lee Thomas, 1993, The Merits For Active Currency Risk Management: Evidence from International Bond Portfolio, *Financial Analysts Journal*, 9-10, 63-70.

Litterman, R., and J. Scheinkman, 1991, Common Factors Affecting Bond Returns, *Journal of Fixed Income* 1, 54-61.

Meese, R. and K. Rogoff, 1983, Empirical Exchange Rate Models of the Seventies: Do They Fit Out-of-Sample?, *Journal of International Economics* 14, 3-24.

Mosburger, G. and P. Schneider, 2005, Modelling International Bond Markets with Affine Term Structure Models, University of Vienna working paper.

Pedersen, A. R., 1995, A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations, *Scandinavian Journal of Statistics*, 22, 55-71.

Piazzesi, M. 2002, Affine Term Structure Models, Working paper, University of Chicago.

Pritsker, M. 1998, Nonparametric Density Estimation and Tests of Continuous-Time Interest Rate Models, *Review of Financial Studies* 11, 449-87.

Singleton, K., 2001, Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function, *Journal of Econometrics 102, 111-141.*

Singleton, K., and L. Umantsev, 2003, Pricing Coupon-Bond Options and Swaptions in Affine Term Structure Models, *Mathematical Finance 12, 427-446*.

Tang, H. and Y. Xia, 2006, An International Exmaination of Affine Term Structure Models and the Expectation Hypothesis, Wharton School working paper.

Tauchen, G., 1997, New Minimum Chi-Square Methods in Empirical Finance, in: D. Kreps and K.Wallis, eds., *Advances in Econometrics: Seventh World Congress* (Cambridge University Press, Cambridge).

Tauchen, G., 2002, Comment on "Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes" by Durham and Gallant, *Journal of Business and Economic Statistics* 20, 331-332.

Vasicek, O., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* 5, 177-188.

Wadhwa, P., 1999, An Empirical Analysis of the Common Factors Governing U.S. Dollar-LIBOR Implied Volatility Movements, *Journal of Fixed Income* 9, 61-68.

7 Appendix 1: Specification Analysis of Four-Factor Joint Term Structure Models

7.1 $\mathbf{A}_0(4)$ *model*

For model $\mathbf{A}_0(4)$ stochastic differential equation (8) takes the form

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ \kappa_{41} & \kappa_{42} & \kappa_{43} & \kappa_{44} \end{bmatrix} \begin{bmatrix} -X_{1t} \\ -X_{2t} \\ -X_{3t} \\ -X_{4t} \end{bmatrix} dt + d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \\ W_{4t} \end{bmatrix}. \tag{23}$$

Note that for the general case when all entries in the lower triangular submatrix of κ are non-zero, only X_4 can be an individual country factor since it enters the equation (23) by itself. All other factors enter in at least one equation for another factor in (23). If we assume symmetry, then we would have either zero, or two, or four common factors. The non-trivial case is that of two common factors. Without loss of generality we can assume that X_1 and X_2 are the common factors while X_3 is the domestic individual factor and X_4 is the foreign individual factor. In that case we need to add additional restriction that $\kappa_{43} = 0$. Thus for each country we have an $\mathbf{A}_0(3)$ model with two common factors. Specifically, for the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t) + \delta_3 X_3(t)$, and

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} -X_{1t} \\ -X_{2t} \\ -X_{3t} \end{bmatrix} dt + d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{bmatrix},$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_1^* X_1(t) + \delta_2^* X_2(t) + \delta_4^* X_4(t)$, and the dynamics of factors are

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{41} & \kappa_{42} & \kappa_{44} \end{bmatrix} \begin{bmatrix} -X_{1t} \\ -X_{2t} \\ -X_{4t} \end{bmatrix} dt + d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{4t} \end{bmatrix}.$$

7.2 $\mathbf{A}_1(4)$ *model*

For model $\mathbf{A}_1(4)$, equation (8) becomes

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & \kappa_{24} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & \kappa_{34} \\ \kappa_{41} & \kappa_{42} & \kappa_{43} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{2t} \\ -X_{3t} \\ -X_{4t} \end{bmatrix} dt \tag{24}$$

$$+\begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_{1t}} & 0 & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_{1t}} & 0 \\ 0 & 0 & 0 & \sqrt{1+\beta_{41}X_{1t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \\ W_{4t} \end{bmatrix}.$$

Assumption of symmetry of the model relative to domestic vs. foreign country implies that X_1 should be a common factor. Without loss of generality X_2 is the second common factor. To ensure that X_3 and X_4 are single country factors we have to add extra restrictions: $\kappa_{23} = \kappa_{24} = \kappa_{34} = \kappa_{43} = 0$. For each country we have an $\mathbf{A}_1(3)$ model with two common factors X_1 and X_2. Factor X_1 enters into both the drift and volatility terms in (24) while factor X_2 enters only the diffusion term in (24). Moreover, no single country factor affects volatility directly since the equation for X_1 in the system (24) does not depend on any other factor. For the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t) + \delta_3 X_3(t)$, and

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{2t} \\ -X_{3t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_{1t}} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_{1t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{bmatrix},$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_1^* X_1(t) + \delta_2^* X_2(t) + \delta_4^* X_4(t)$, and the dynamics of factors are

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{41} & \kappa_{42} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{2t} \\ -X_{4t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_{1t}} & 0 \\ 0 & 0 & \sqrt{1+\beta_{41}X_{1t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{4t} \end{bmatrix}.$$

7.3 $\mathbf{A}_2(4)$ *model*

Similarly, $\mathbf{A}_2(4)$ has the following dynamics:

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & \kappa_{34} \\ \kappa_{41} & \kappa_{42} & \kappa_{43} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ -X_{3t} \\ -X_{4t} \end{bmatrix} dt \tag{25}$$

$$+\begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_{1t}+\beta_{32}X_{2t}} & 0 \\ 0 & 0 & 0 & \sqrt{1+\beta_{41}X_{1t}+\beta_{42}X_{2t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \\ W_{4t} \end{bmatrix}$$

Now we have two possibilities for non-trivial symmetric choice of common factors. The first one is the $\mathbf{A}_{2,1}(4)$ model. This is the case when $\kappa_{34} = \kappa_{43} = 0$, and X_1 and X_2 are the two common factors. In this model both common factors affect both the volatility and drift terms of (25), and no single-country factor affects the volatility term. For each country we have an $\mathbf{A}_2(3)$ model. For the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t) + \delta_3 X_3(t)$, and

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ -X_{3t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_{1t}+\beta_{32}X_{2t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{bmatrix},$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_1^* X_1(t) + \delta_2^* X_2(t) + \delta_4^* X_4(t)$, and the dynamics of factors are

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{41} & \kappa_{42} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ -X_{4t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 \\ 0 & 0 & \sqrt{1+\beta_{41}X_{1t}+\beta_{42}X_{2t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{4t} \end{bmatrix}.$$

The second model is $\mathbf{A}_{2,2}(4)$. This is the case when X_3 and X_4 are the two common factors, while X_1 and X_2 are the single-country factors in (25). This leads to restrictions $\kappa_{12} = \kappa_{21} = \kappa_{31} = \kappa_{32} = \kappa_{41} = \kappa_{42} = 0$ and $\beta_{31} = \beta_{32} = \beta_{41} = \beta_{42} = 0$. We further impose, without loss of generality, $\kappa_{34} = 0$. In this case the common factors affect only correlation structure (i.e. the drift term). For the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_3 X_3(t) + \delta_4 X_4(t)$, and

$$d\begin{bmatrix} X_{1t} \\ X_{3t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ 0 & \kappa_{33} & 0 \\ 0 & \kappa_{43} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ -X_{3t} \\ -X_{4t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{3t} \\ W_{4t} \end{bmatrix},$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_2^* X_2(t) + \delta_3^* X_3(t) + \delta_4^* X_4(t)$, and the dynamics of factors is

$$d\begin{bmatrix} X_{2t} \\ X_{3t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{22} & 0 & 0 \\ 0 & \kappa_{33} & 0 \\ 0 & \kappa_{43} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_2 - X_{2t} \\ -X_{3t} \\ -X_{4t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{2t}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} d\begin{bmatrix} W_{2t} \\ W_{3t} \\ W_{4t} \end{bmatrix}.$$

Thus for both the domestic and the foreign country we have a special case of $\mathbf{A}_1(3)$ model. Common factors X_3 and X_4 follow a bivariate Gaussian process while each country specific factor is a univariate square root process. This case turns out to be a particular case of $\mathbf{A}_1(4)$ model if the additional restriction is imposed that $\vartheta_1 = \vartheta_2$.

7.4 $\mathbf{A}_3(4)$ *model*

For the $\mathbf{A}_3(4)$ model, equation (8) specializes to

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ \kappa_{41} & \kappa_{42} & \kappa_{43} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ \vartheta_3 - X_{3t} \\ -X_{4t} \end{bmatrix} dt \tag{26}$$

$$
+\begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 & 0 \\ 0 & 0 & \sqrt{X_{3t}} & 0 \\ 0 & 0 & 0 & \sqrt{1+\beta_{41}X_{1t}+\beta_{42}X_{2t}+\beta_{43}X_{3t}} \end{bmatrix} d \begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \\ W_{4t} \end{bmatrix}
$$

Without loss of generality, non-trivial symmetric common factors are X_1 and X_4. This implies that $\beta_{42} = \beta_{43} = 0$, and $\kappa_{12} = \kappa_{13} = \kappa_{23} = \kappa_{32} = \kappa_{42} = \kappa_{43} = 0$. For the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t) + \delta_4 X_4(t)$, and

$$
d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{41} & 0 & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ -X_{4t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 \\ 0 & 0 & \sqrt{1+\beta_{41}X_{1t}} \end{bmatrix} d \begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{4t} \end{bmatrix},
$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_1^* X_1(t) + \delta_2^* X_2(t) + \delta_4^* X_4(t)$, and the dynamics of factors are

$$
d\begin{bmatrix} X_{1t} \\ X_{3t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{31} & \kappa_{33} & 0 \\ \kappa_{41} & 0 & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_3 - X_{3t} \\ -X_{4t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{X_{3t}} & 0 \\ 0 & 0 & \sqrt{1+\beta_{41}X_{1t}} \end{bmatrix} d \begin{bmatrix} W_{1t} \\ W_{3t} \\ W_{4t} \end{bmatrix}.
$$

7.5 $\mathbf{A}_4(4)$ *model*

Finally, for the $\mathbf{A}_4(4)$ model, equation (8) simplifies to

$$
d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & \kappa_{14} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & \kappa_{24} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & \kappa_{34} \\ \kappa_{41} & \kappa_{42} & \kappa_{43} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ \vartheta_3 - X_{3t} \\ \vartheta_4 - X_{4t} \end{bmatrix} dt \tag{27}
$$

$$
+\begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 & 0 \\ 0 & 0 & \sqrt{X_{3t}} & 0 \\ 0 & 0 & 0 & \sqrt{X_{4t}} \end{bmatrix} d \begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \\ W_{4t} \end{bmatrix}.
$$

Without loss of generality, non-trivial symmetric common factors are X_1 and X_2. This implies that $\kappa_{13} = \kappa_{14} = \kappa_{23} = \kappa_{24} = \kappa_{34} = \kappa_{43} = 0$. For the domestic country we have $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t) + \delta_4 X_4(t)$, and

$$
d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ \vartheta_3 - X_{3t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 \\ 0 & 0 & \sqrt{X_{3t}} \end{bmatrix} d \begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{bmatrix},
$$

while for the foreign country the spot rate is $r^*(t) = \delta_0^* + \delta_1^* X_1(t) + \delta_2^* X_2(t) + \delta_4^* X_4(t)$, and the dynamics of factors are

$$d\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{4t} \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{41} & \kappa_{42} & \kappa_{44} \end{bmatrix} \begin{bmatrix} \vartheta_1 - X_{1t} \\ \vartheta_2 - X_{2t} \\ \vartheta_4 - X_{4t} \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_{1t}} & 0 & 0 \\ 0 & \sqrt{X_{2t}} & 0 \\ 0 & 0 & \sqrt{X_{4t}} \end{bmatrix} d\begin{bmatrix} W_{1t} \\ W_{2t} \\ W_{4t} \end{bmatrix}.$$

For each country we get a particular case of $\mathbf{A}_3(3)$ model.

8 Appendix 2: Kernel Estimators for Non-parametric In-Sample and Out-of-sample Tests

Our kernel estimator of the joint density is, for any integer $j > 0$,

$$\hat{g}_j(z_1, z_2) \equiv (n-j)^{-1} \sum_{\tau=R+j+1}^{L} K_h\left(z_1, \hat{Z}_\tau\right) K_h\left(z_2, \hat{Z}_{\tau-j}\right), \qquad 0 \le z_1, z_2 \le 1, \tag{28}$$

where $\hat{Z}_\tau = Z_\tau(\hat{\theta}_R)$, $\hat{\theta}_R$ is any $\sqrt{R}$-consistent estimator for θ_0, and $K_h(z_1, z_2)$ is a boundary-modified kernel function.

The boundary-modified kernel function is defined as follows. For $x \in [0,1]$, we define

$$K_h(x,y) \equiv \begin{cases} h^{-1}k\left(\frac{x-y}{h}\right) / \int_{-(x/h)}^{1} k(u)du, & \text{if } x \in [0,h), \\ h^{-1}k\left(\frac{x-y}{h}\right), & \text{if } x \in [h, 1-h], \\ h^{-1}k\left(\frac{x-y}{h}\right) / \int_{-1}^{(1-x)/h} k(u)du, & \text{if } x \in (1-h, 1], \end{cases} \tag{29}$$

where $k(\cdot)$ is a prespecified symmetric probability density, and $h \equiv h(n)$ is a bandwidth such that $h \to 0, nh \to \infty$ as $n \to \infty$. Throughout our empirical analysis, we use the quartic kernel

$$k(u) = \frac{15}{16}(1-u^2)^2 \mathbf{1}(|u| \le 1), \tag{30}$$

where $\mathbf{1}(\cdot)$ is the indicator function. In practice, the choice of bandwidth h is more important than the choice of the kernel $k(u)$. Like Scott (1992), we choose $h = \hat{S}_Z n^{-\frac{1}{6}}$, where $\hat{S}_Z$ is the sample standard deviation of $\{\hat{Z}_\tau\}_{\tau=R+1}^{L}$. This simple bandwidth rule attains the optimal rate for bivariate kernel density estimation.

The modified kernel in (25) can automatically deal with the boundary bias problem associated with standard kernel estimation. As is well known (e.g., Härdle 1990, pp. 130-133), a standard kernel density estimator gives biased estimates near the boundaries of data, because a standard kernel provides an asymmetric coverage of the data in the boundary regions. In contrast, the weighting functions in the denominators of $K_h(x,y)$ for $x \in [0,h) \cup (1-h, 1]$ account for the asymmetric coverage and ensure that the estimator (29) is asymptotically unbiased uniformly over the entire support $[0,1]$ for the generalized residuals. The modified-kernel in (30) has several advantages over some existing alternative solutions to the boundary bias problem in the literature. One alternative is to simply ignore the data in the boundary regions and only use the data in the interior region. Such a trimming procedure is simple, but in the present context, it would lead to the loss of a significant amount of information. For a nearly uniformly distributed transformed sequence $\{Z_\tau\}$, the data in the boundary region is still about 10% when the sample size $n = 5000$ and the bandwidth $h = \hat{S}_Z n^{-\frac{1}{6}}$, where $\hat{S}_Z$ is the sample standard deviation of $\{\hat{Z}_\tau\}_{t=R+1}^{L}$. For a financial time series such as interest rates, one may be particularly interested in the tail distribution of the underlying process, which is exactly contained in (and only in) the boundary regions. Alternatively, we can also use the so-called jackknife kernel to eliminate the boundary bias, as in Chapman and Pearson (2000) and Diebold, Hahn and Tay (1999). In the present context, the jackknife kernel, however, has the undesirable property that it may generate negative density estimates in the boundary regions. It also induces a relatively large variance for the kernel estimates in the boundary regions, adversely affecting the power of the test in finite samples. In contrast, our modified kernel always produces nonnegative density estimates with a smaller variance in the boundary regions.

One advantage of this approach is that since there is no serial dependence in $\{Z_\tau\}$ under correct model specification, nonparametric joint density estimators and related test statistics are expected to perform well in finite samples. This is appealing because there exists persistent dependence in interest rate time series data. Another advantage is that there is no asymptotic bias for nonparametric density estimators under the null hypothesis of correct model specification either, because the conditional density of Z_τ given $\{Z_{\tau-1}, Z_{\tau-2}, ...\}$ is uniform (i.e. a constant). Moreover, our test can be applied to time-inhomogeneous continuous-time processes, because $\{Z_\tau\}$ is always i.i.d. $U[0,1]$ under correct model specification.[16]

[16] Egorov, Li and Xu (2002) extend Aït-Sahalia's (2002) Hermite expansion approach to obtain accurate closed-form approximation for the transition density of time-inhomogeneous diffusion models.

Table 1
Summary Statistics

This table reports the summary statistics of the level and change series of 1-month, 6-month, 2-year, 5-year and 10-year zero coupon bond yields. The sample is from July 1, 1999 to June 30, 2003. Panel A shows the summary statistics for LIBOR. Panel B shows the summary statistics for Euribor. Panel C shows the correlations between levels of LIBOR and Euribor (upper table), and correlations between changes in LIBOR and Euribor (lower table)

Panel A: LIBOR series

	mean	median	std dev	skewness	kurtosis
Levels					
1-month LIBOR	3.7980	3.7500	2.0759	0.0554	1.3036
6-month LIBOR	3.8950	3.7200	2.1608	0.0839	1.3104
2-year LIBOR swap	4.6016	4.6050	1.9382	-0.0450	1.5888
5-year LIBOR swap	5.3947	5.4550	1.4539	-0.2392	1.8970
10-year LIBOR swap	5.9291	5.9650	1.1090	-0.2606	2.0225
Changes in					
1-month LIBOR	-0.0039	0	0.0513	1.3412	138.8800
6-month LIBOR	-0.0043	0	0.0412	-2.7039	28.6190
2-year LIBOR swap	-0.0044	-0.0050	0.0643	0.1589	5.3032
5-year LIBOR swap	-0.0036	0	0.0691	0.2979	4.6131
10-year LIBOR swap	-0.0028	0	0.0684	0.2593	5.8087

Panel B: Euribor series

	mean	median	std dev	skewness	kurtosis
Levels					
1-month Euribor	3.6501	3.3820	0.7956	0.1994	1.8049
6-month Euribor	3.7280	3.5380	0.8095	0.0895	2.0770
2-year Euribor swap	4.1152	4.2600	0.8384	-0.4313	2.5734
5-year Euribor swap	4.6898	4.8300	0.6868	-0.7307	2.9384
10-year Euribor swap	5.2237	5.3450	0.5487	-0.8148	2.9866
Changes in					
1-month Euribor	-0.0005	-0.0010	0.0331	-0.7725	72.2850
6-month Euribor	-0.0007	-0.0010	0.0266	-0.6282	19.0240
2-year Euribor swap	-0.0011	0	0.0437	0.5543	4.6783
5-year Euribor swap	-0.0011	0	0.0467	0.5008	4.2775
10-year Euribor swap	-0.0010	0	0.0425	0.4178	4.1020

Table 1 Summary Statistics (continued)

Panel C: Correlations

Levels	6-month LIBOR	2-year LIBOR	5-year LIBOR	6-month Euribor	2-year Euribor	5-year Euribor
6-month LIBOR	1	0.979	0.941	0.697	0.786	0.802
2-year LIBOR		1	0.989	0.684	0.831	0.870
5-year LIBOR			1	0.672	0.844	0.898
6-month Euribor				1	0.922	0.837
2-year Euribor					1	0.979

Changes In	6-month LIBOR	2-year LIBOR	5-year LIBOR	6-month Euribor	2-year Euribor	5-year Euribor
6-month LIBOR	1	0.395	0.295	0.380	0.279	0.179
2-year LIBOR		1	0.897	0.262	0.699	0.650
5-year LIBOR			1	0.218	0.678	0.686
6-month Euribor				1	0.363	0.200
2-year Euribor					1	0.922

Table 2
Principal Component Analysis for LIBOR, Euribor and Joint Term Structure Models

Table 2 reports the variations of LIBOR, Euribor and joint LIBOR-Euribor term structure models that are explained by the first seven principal components. The estimation is done using daily LIBOR and Euribor from July 1, 1999 till June 30, 2003. Column 1 lists the number of principal components. Column 2 reports the LIBOR term structure model variation that is explained by the corresponding principal components. Column 3 reports the Euribor term structure model variation that is explained by the corresponding number of principal components. Column 4 reports the joint term structure model variation that is explained by the corresponding number of principal components. The results for principal components ranging from 1 to 7 are first reported, followed by the results from the first two up to first seven principal components.

Number of	Percentage of variations explained		
Principal components	LIBOR	Euribor	Joint
1	98.456%	96.368%	94.738%
2	1.377%	3.263%	3.540%
3	0.137%	0.301%	1.401%
4	0.018%	0.046%	0.162%
5	0.006%	0.012%	0.086%
6	0.002%	0.005%	0.049%
7	0.001%	0.001%	0.001%
First two	99.833%	99.631%	98.278%
First three	99.970%	99.932%	99.679%
First four	99.988%	99.978%	99.841%
First five	99.994%	99.990%	99.927%
First six	99.996%	99.995%	99.976%
First seven	99.998%	99.996%	99.984%

Table 3
Common Factors in the Term Structure Models

Table 3 reports the results from regressing the LIBOR and Euribor factors on to the common factors. LIBOR and Euribor factors are principal components from LIBOR and Euribor term structure models, while common factors are principal components from the joint term structure model. Panel A reports the results from regressing LIBOR and Euribor factors on six common factors. Panel B reports the R-square from regressing LIBOR and Euribor factors on the first six common factors. Panel C reports the results from regressing log of exchange rates on six common factors. In panel A, the first column lists the dependent variables, i.e. the first three LIBOR or Euribor factors. The next six columns reports the regression coefficients on the first six common factors. The last column reports the R2 of the regressions. In panel B, the first column lists the dependent variables, i.e. the first three LIBOR and the first three Euribor factors. The next column reports the R2 when the dependent variable is regressed upon the first factor of the joint term structure model. The next 4 columns report the result when the regressions are run on the first two factors, first three factors, first four factors, and first five factors, respectively. In panel C, the first column is the dependent variable, i.e. log of exchange rates. The next six columns report the coefficients on the first six common factors. The last column reports the R2 of the regression.The estimation is done using daily LIBOR and Euribor data from July 1, 1999 till June 30, 2003.

Panel A: Results from regressing LIBOR and Euribor factors on all six common factors

	Regression coefficients on common factor						
	1st factor	2nd factor	3rd factor	4th factor	5th factor	6th factor	R^2
LIBOR 1st factor	0.905	0.406	0.055	0.107	-0.349	-0.019	0.99999
LIBOR 2nd	-0.001	0.008	0.868	-0.483	-0.021	-0.052	0.99992
LIBOR 3rd	0.001	-0.085	0.002	0.022	-0.974	0.128	0.99933
Euribor 1st	-0.425	0.874	0.097	0.196	-0.077	-0.031	0.99999
Euribor 2nd	0.017	0.246	-0.478	-0.831	-0.032	0.041	0.99980
Euribor 3rd	-0.004	-0.046	-0.077	0.031	-0.138	-0.932	0.98548

Table 3 Panel B: R^2 from regressing LIBOR and Euribor factors on the first five common factors

	R^2				
	Only 1st factor	First 2 factors	First 3 factors	First 4 factors	First 5 factors
LIBOR 1st factor	0.99245	0.99920	0.99997	0.99999	0.99999
LIBOR 2nd	0.00006	0.00029	0.96525	0.99977	0.99980
LIBOR 3rd	0.00167	0.23574	0.23581	0.23652	0.99198
Euribor 1st	0.86267	0.99900	0.99966	0.99997	0.99999
Euribor 2nd	0.03921	0.35896	0.83370	0.99955	0.99969
Euribor 3rd	0.02197	0.14200	0.27543	0.27801	0.30464

Panel C: Results from regressing log exchange rates on all six common factors

	Regression coefficients on common factor						
	1st factor	2nd factor	3rd factor	4th factor	5th factor	6th factor	R^2
Log Exchange Rates	0.00003	0.00000	0.00006	0.00025	-0.00025	-0.00047	0.00599

Table 4
Principal Component Analysis for LIBOR and Euribor Models after taking out 1 or 2 common factors

Table 4 reports the variations of LIBOR and Euribor residuals that are explained by the first seven principal components. The estimation is done using daily LIBOR and Euribor data from July 1, 1999 till June 30, 2003. Column 1 lists the number of principal components. Column 2 reports the LIBOR term structure model variation that is explained by the corresponding number of principal components. Column 3 reports the variations explained for the residuals of LIBOR after it is regressed on the first common factor. Column 4 reports the variations explained for the residuals of LIBOR after it is regressed on the first two common factors. Column 5 reports the Euribor term structure model variation that is explained by the corresponding number of principal components. Column 6 reports the variations explained for the residuals of Euribor after it is regressed on the first common factors. Column 7 reports the variations explained for the residuals of Euribor after it is regressed on the first two common factors.

	Percentage of variations explained					
Principal Component	LIBOR only	LIBOR after 1 common factor	LIBOR after 2 common factors	Euribor only	Euribor after 1 common factor	Euribor after 2 common factors
1	0.98456	0.60756	0.90956	0.96368	0.85074	0.88229
2	0.01377	0.33526	0.06909	0.03263	0.13176	0.09397
3	0.00137	0.04449	0.01182	0.00301	0.01409	0.01453
4	0.00018	0.00753	0.00409	0.00046	0.00212	0.00468
5	0.00006	0.00265	0.00181	0.00012	0.00070	0.00181
6	0.00002	0.00101	0.00140	0.00004	0.00027	0.00110
7	0.00001	0.00054	0.00082	0.00002	0.00013	0.00060

Table 5
Non-parametric Pormanteau statistics for the In-sample and Out-of-sample Performance of Affine Models

Table 5 reports the non-parametric pormanteau statistics W(p) defined in equation (19) in the text for in-sample and out-of-sample performances of six specifications of four-factor affine models. p represents the lag truncation order and equals 6, 10 and 20 in our case. We separately report results when our sample is limited to only 2-year and 10-year zero coupon bonds for LIBOR and Euribor, respectively. The estimation is done using daily LIBOR and Euribor data from July 1, 1999 till June 30, 2003. W(p) has a standard normal distribution.

			In-Sample			Out-of-Sample		
Model	Maturity	Int. Rate	W(6)	W(10)	W(20)	W(6)	W(10)	W(20)
$\mathbf{A}_0(4)$	Combined		65.22	76.19	82.04	98.22	103.52	101.35
	2y	LIBOR	33.62	33.40	41.72	69.79	73.86	73.08
	10y	LIBOR	34.01	38.54	37.15	66.97	76.22	74.63
	2y	Euribor	39.29	44.73	42.37	60.01	82.03	70.93
	10y	Euribor	45.20	66.32	65.13	73.00	77.18	70.01
$\mathbf{A}_1(4)$	Combined		57.03	59.28	66.63	67.22	73.76	86.13
	2y	LIBOR	20.10	24.87	27.62	28.19	33.01	35.77
	10y	LIBOR	26.03	30.02	38.15	25.90	29.62	30.17
	2y	Euribor	31.01	33.81	39.31	34.03	37.98	39.00
	10y	Euribor	27.88	27.91	38.44	40.55	48.44	48.74
$\mathbf{A}_{2,1}(4)$	Combined		103.21	108.08	111.37	167.49	183.03	201.72
	2y	LIBOR	50.07	53.77	55.92	92.29	97.32	101.24
	10y	LIBOR	52.34	56.13	58.11	93.99	96.11	98.13
	2y	Euribor	51.88	54.44	56.02	95.43	102.02	103.55
	10y	Euribor	52.90	56.01	58.30	92.22	99.71	102.70
$\mathbf{A}_{2,2}(4)$	Combined		65.23	70.03	102.64	97.03	102.04	103.75
	2y	LIBOR	31.55	67.32	87.25	61.01	63.28	67.21
	10y	LIBOR	48.53	42.03	38.09	29.01	35.55	40.38
	2y	Euribor	50.01	53.37	40.22	67.72	77.73	100.90
	10y	Euribor	54.34	68.11	56.13	70.12	90.01	93.11
$\mathbf{A}_3(4)$	Combined		110.71	121.01	123.55	227.73	244.32	290.11
	2y	LIBOR	53.20	55.12	58.11	87.12	91.16	93.33
	10y	LIBOR	49.14	50.97	56.00	88.11	92.20	94.70
	2y	Euribor	52.19	54.12	56.79	88.03	93.81	94.91
	10y	Euribor	49.93	56.12	68.33	89.54	93.72	94.92
$\mathbf{A}_4(4)$	Combined		242.90	351.99	490.82	335.03	480.72	504.33
	2y	LIBOR	91.33	133.17	140.05	118.20	165.66	168.91
	10y	LIBOR	92.13	140.01	127.78	138.04	149.14	159.14
	2y	Euribor	123.45	135.55	142.21	135.51	177.82	182.90
	10y	Euribor	135.15	152.79	190.13	154.12	163.88	180.01

Figure 1. Term Structures from July 1, 1999 till June 30, 2003