The Rodney L. White Center for Financial Research

An International Examination of Affine Term Structure Models and the Expectations Hypothesis

Huarong Tang
Yihong Xia

26-05

The Wharton School
University of Pennsylvania

An International Examination of Affine Term Structure Models and the Expectations Hypothesis*

Huarong Tang[†] **and Yihong Xia**[‡]

March 29, 2005

Abstract

We examine the yield curve behavior and the relative performance of affine term structure models using government bond yield data from Canada, Germany, Japan, UK, and US. We find strong predictability of forward rates for excess bond returns and reject the expectations hypothesis across all five countries. A three-factor model is sufficient to capture movements in the yield curve of Canada, Japan, UK, and US, but may not be enough for Germany. An exhaustive comparison among affine term structure models with no more than three factors reveals that the three-factor essential affine model ($A_1(3)E$) with only one factor affecting the volatility of the short rate but with all three factors affecting the price of risk performs best in all five countries. Simulations provide inconclusive evidence on whether this best affine model can successfully generate the rich yield curve behavior observed in the data.

*We thank Andy Abel, Michael Brennan, Qiang Dai, Chris Geczy, Rich Kihlstrom, and Allan White for helpful discussions. Xia acknowledges financial support from the Weiss Center for International Financial Research and the Rodney L. White Center for Financial Research at the Wharton School.

[†]Ph.D. candidate, Finance Department, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6367. Email: tangh@wharton.upenn.edu.

[‡]Assistant Professor, Finance Department, The Wharton School, University of Pennsylvania; 2300 Steinberg Hall-Dietrich Hall; Philadelphia, PA 19104-6367. Phone: (215) 898-3004. Fax: (215) 898-6200. E-mail: yxia@wharton.upenn.edu.

1 Introduction

Despite the large theoretical literature on dynamic term structure models and extensive empirical investigations of their performance using U.S. data,[1] there is as yet virtually no systematic study of the behavior of the yield curve or the performance of popular dynamic term structure models in other major government bond markets.[2] This paper extends the empirical examination of dynamic term structure models to Canada (CA), Germany (GM), Japan (JP), and the United Kingdom (UK), in addition to the United States (US).

Although empirical regularities in the US term structures have been documented very extensively in the literature, the behavior of foreign term structures has not been well studied because of the lack of foreign data. We use a newly constructed data set for the one- to ten-year constant maturity zero coupon nominal government bond yield data for CA, GM, JP, UK and US to study the term structure behavior in these five countries and to examine how well affine term structure models account for these empirical regularities. This is the first paper that carries out international comparisons of term structure models and such comparisons are bound to be informative because of the wide differences in monetary policies, government bond markets, interest and inflation processes across these countries during the sample period of 1985 to 2002. For example, while CA, US and UK have experienced high inflation and high nominal rates in 1980s, Germany and Japan has enjoyed low rates throughout the period. While Japan has witnessed its interest rate gradually decline to zero and has concerned about deflation, Germany has weathered shocks from the unification and the introduction of Euro. In light of these differences, an important issue is whether empirical regularities found for the U.S. term structure carry over to foreign markets and, in turn, how affine term structure models perform in these markets.

Before making a formal comparison of affine term structure models, we first examine the information content of foreign term structures and check whether stylized findings about US yield curve behavior, such as the failure of the expectations hypothesis and the predictive power of forward rates for excess future bond returns, extend to the other four OECD countries. The resulting empirical findings in turn provide benchmarks against which the best affine term structure model is later examined.

Under the expectations hypothesis with a constant risk premium the coefficient from the regression of the change of the $\tau-$period zero-coupon yield on the term spread, which is defined as the difference between the τ-period bond yield and the one-year bond yield, should be unity for all bond maturities τ.

[1] See Dai and Singleton (2003) for a comprehensive survey.

[2] There are a few exceptions. For example, Brennan and Xia (2003) estimates a three-factor Gaussian model where the three factors are the real interest rate, the expected rate of inflation and the volatility of the pricing kernel.

Many studies, notably Fama (1984), Stambaugh (1988), and Campbell and Shiller (1991), have found that the expectations hypothesis is violated in the US yield data. We test the expectations hypothesis in all five countries using two regression specifications. In the first specification, we test whether the current maturity-adjusted term spread forecasts changes in the $\tau-$period zero-coupon bond yield with a theoretical predictive coefficient of one. In the second specification, we test whether the current term spread forecasts average short rate changes with a coefficient of one. These two specifications are theoretically equivalent under the null of expectations hypothesis, but may lead to different empirical results if the expectations hypothesis is violated in the data.

Although there is variation across countries, the slope estimates from the first regression are often negative and become increasingly so as the bond maturity increases. However, the yields used in the estimation are highly persistent and the point estimates of the coefficients are usually not significantly different from the theoretical value of unity when the standard errors are corrected for autocorrelation. In contrast, the slope estimates from the second regression are all positive and in most cases are close to unity, which is consistent with the prediction of the expectations hypothesis. The slope estimates also slightly increase with the bond maturity τ. Our simulation results suggest that the paradoxical results from the two specifications are consistent with a simple model in which both the interest rate and the term premium follow AR(1) processes.

Finally, we test whether forward rates forecast future excess bond returns. Predictability of bond excess returns by forward rates provides another piece of evidence that the expectations hypothesis is violated and that the bond market risk premia are time varying. We find that forward rates have predictive power in all five countries. The effect is especially strong in GM, JP and US, and it is mainly concentrated on bonds with maturities less than six years.

These results suggest that empirical regularities in term structures are broadly similar despite the wide differences in government bond markets, and they provide overwhelming out-of-sample evidence that bond market risk premia are time-varying so that models that allow for a time-varying risk premium are required to fit the yield data well. Important dynamic term structure models include the complete affine term structure models of Duffie and Kan (1996) and Dai and Singleton (2001), the essential affine term structure model of Dai and Singleton (2002) and Duffee (2002), the affine model with regime shifts of Bansal and Zhou (2002), and the Gaussian-quadratic model of Ahn, Dittmar, and Gallant (2002). In this paper, we focus on the comparison of the relative performance of models in the affine term structure family which include both the complete and the essential affine models.

Preliminary analysis shows that in all five countries the first three principal components of yields explain over 99% of the total yield curve variation and the components can be broadly characterized as "level," "slope," and "curvature" ,[3] but each country has its unique features. In particular, there is evidence of more than three factors in the German yield curve. It is possible that German re-unification and the introduction of the Euro, which both occurred during our sample period, caused structural breaks in the German bond market so that a model with four factors or with regime shifts such as that used by Bansal *et. al.* (2003) is likely to provide a better fit.

We estimate all three-factor essentially affine models using a Kalman filter/Quasi Maximum Likelihood (QML) approach, and then compare all nested models using Wald test as well as likelihood function values. Non-nested models, as well as nested models that are not eliminated in the first step, are also estimated using the same approach and then compared using the Schwarz criterion and the in- and the out-of-sample pricing errors. We find that the essential three-factor model, $A_1(3)E$, in which only one factor affects the instantaneous volatility of the short rate but all three factors affect the price of risk, performs best in all five countries. The best model has very small *average* in- and out-of- sample prediction errors, although it is still associated with *occasional* large mis-pricing. The consistent superior performance of $A_1(3)E$ across all five term structures suggest that this model is flexible enough to capture both universal stylized facts and country-specific empirical regularities. This provides preliminary evidence that implications from US term structure data may be robust enough to extend to foreign markets.

The model-implied risk free rate tracks the one-month Treasury bill rates quite well for UK, CA and JP, but there is a noticeable gap between the two for Germany around 1998-99 when the Euro was introduced, and for the US in 2000 when the Internet Bubble burst. The time series of the estimated price of risk in the bond market shows great variation during the sample. This partly reflects the fact that the parameters governing the prices of risks are estimated very imprecisely, but despite the difficulty in pinning down parameters, the estimated price of risk varies with important monetary events and stock market performance in a plausible fashion.

Finally, we examine whether the best model, $A_1(3)E$, can generate the stylized regression results discussed above. We find that estimation results from simulated data for all three regressions are generally different from those obtained from the actual data in both magnitude and patterns across different maturities, but the differences are statistically insignificant (at the 5% level). Thus, the coefficient estimates observed in the data could conceivably be generated by an $A_1(3)E$ model. This conclusion must be tempered by the

[3]These labels were first proposed by Litterman and Scheinkman (1991) for the U.S. yield data.

observation that the difficulty in obtaining precise estimates of the affine term structure model parameters, especially those governing risk premia, translates into low power in refuting the model.

The rest of the paper is organized as follows. In Section 2, we give a brief discussion and summary statistics of the yield curve data and provide a principal components analysis of the yield curve. In section 3, we provide evidence on the violation of the expectations hypothesis in all five countries under three regression specifications. Section 4 introduces the affine term structure models considered in the paper. Section 5 contains details of the estimation procedure, the estimation results, and the comparison of all affine term structure models with up to three factors. Section 6 presents simulation evidence on whether the "best" model (in the sense of best fitting the yield curve) within the affine term structure family can successfully generate the stylized empirical findings documented in Section 3. Section 7 summarizes and concludes the paper.

2 Term Structure Data

The basic data are estimated constant-maturity zero-coupon government bond yields for Canada (CA), Germany (GM), Japan (JP), the United Kingdom (UK), and the United States (US). The raw data consist of bond prices, coupon rates, and coupon, issue, and redemption dates for all available government bonds outstanding on the second of each month from Datastream. A cubic spline was fitted to these coupon-bond yields with maturities up to forty years for each country each month.[4] Zero-coupon yields for maturities of six months and from one to ten years are estimated at the beginning of each month. Yields for maturities of six months, one, two, three, five, seven, and ten years (or nine years for JP and GM) are used in the in-sample analysis and estimation, while yields for maturities of four, six, and eight years are reserved for out-of-sample analysis. The data start in January 1983 and end in May 2002 for CA, GM and JP, but end in February 2003 for UK and in January 2003 for US. The maximum number of observations is 233 for CA, GM and JP, 241 for the US, and 242 for the UK. However, there are only 203 (193) observations for the ten-year bond in GM (JP).

Summary statistics for the bond yield data of each country are reported in Table 1. In all countries, average yields are generally increasing, while their standard deviations are generally decreasing with maturity.[5] For example, the sample mean increases from about 6.4% for the six-month bond to 7.6%

[4] See Brennan and Xia (2004) for more details of the raw data and the estimation of the zero-coupon bond yields. Brennan and Xia (2004) created a slightly different zero-coupon yield dataset for their study of the co-variation of risk premium in the bond and the foreign exchange markets by fitting a cubic spline to government coupon bonds with maturities *less than 20 years*.

[5] The sample mean decreases at the long end for GM and JP, most likely caused by missing observations from the ten-year

for the ten-year bond in US, while the sample volatility decreases with the bond maturity from 2.29% to 1.03% in US. As compared with the widely-used McCulloch-Kwon (1993) U.S. yield data, which has monthly observations from January 1953 to December 1998, our U.S. data yield higher sample means but smaller sample volatilities for all maturities. This is especially true at the long end: our ten-year yield has a sample volatility of only 1.03% as compared with 2.7% of the McCulloch-Kwon data.

All yields in all countries are highly persistent. The first order autocorrelation is over 0.96 in CA, GM, UK and US, and is above 0.99 in JP. There is also substantial cross-correlation (not reported here for brevity) between yields of different maturities. The cross-correlation structure is quite similar for CA, UK and US, where the correlations range from 0.8 between the two most distant maturities (six months and ten years) to 0.99 between adjacent maturities. The cross-correlations for Japan are all greater than 0.94. On the other hand, cross-correlations are much lower in Germany where the correlation is only about 0.5 between short and long maturity bond yields. High cross-correlations suggest that a small number of common factors drive the co-movement of the bond yields across different maturities. This is confirmed by the finding that the first three principal components explain a high proportion of the total variation of the eleven yields on bonds with maturities between six months and ten years. In CA, UK and US, the first component explains about 95%, the second around 3-4%, and the third about 0.2-1.8% of the total variation. In JP, the first component accounts for over 98% of the total variation, the second about 1.12%, and the third 0.24%. In contrast, the first component of the German bond yields explains only about 82.8% of the total variation, while the second and the third components contribute 15.6% and 1.1%, respectively. The first three components together explain over 99.9% of the total variation in CA, JP, UK and US, and about 99.6% of the total variation in GM.

The weights of the first three principal components on the eleven bond yields, which are plotted in Figure 1, show broadly similar patterns across all five countries.[6] The weights of the first principal component are similar across yields so that this component can be interpreted as a level component. The second component has negative weights on the three shortest bonds in CA, GM, JP and UK, and positive weights on the long bonds, and the magnitude of the negative and positive weights is similar, implying that the second component is close to the difference between the long and the short bond yields and can be interpreted as a slope component. In US, however, the pattern of the weights of the second component is more like an inverted and skewed "U" shape, so that it can be interpreted as a combination of a slope in the short end and a curvature in the long end. In all five countries, the third component has weights of the

bond.

[6] Litterman and Scheinkman (1991) first carried out the principal component analysis for the US yield data.

same sign on the short and long bonds but of the opposite sign on the bonds with intermediate maturities. In addition, the weights on the short and the long bonds in CA, UK and US are almost the same, so the weights are close to the calculation of a second difference in bond yields so that the component corresponds to the curvature of the yield curve. The magnitudes of the weights on the short and the long bonds in the other two countries, however, are very skewed, so that the component can be interpreted approximately as a combination of a curvature at the short end and another curvature at the long end. Therefore, although the properties of the first three components can be broadly characterized as "level," "slope," and "curvature" in all five countries, they show some interesting country-specific patterns.

3 The Expectations Hypothesis Re-visited

Before comparing the affine term structure models, we first examine whether the stylized empirical findings for US yields, the failure of the expectations hypothesis and the predictive power of forward rates for excess future bond returns, extend to the other four OECD countries. This examination not only provides evidence on the robustness of the findings from the US data, but also generates empirical results that can be used as benchmarks for the best affine term structure model to match.

The violation of the expectations hypothesis in the U.S. yield data is well documented and there is compelling empirical evidence that the expected excess returns on the U.S. Treasury bonds exhibit predictable variation over time.[7] In contrast, results for other countries are mixed. On the one hand, Jorion and Mishkin (1991) found no, or much weaker, evidence against the expectations hypothesis for UK, Germany, and Switzerland, and Hardouvelis (1994) found strong support *for* the expectations hypothesis in all G7 countries except the US. On the other hand, Bekaert, Wei and Xing (2002) find strong statistical evidence against the expectations hypothesis of the term structure in US, UK and GM.

While this is not the first paper extending the analysis of US term structure behavior to other countries, it is still worthwhile to revisit the problem for three reasons. First, earlier studies have only used a subset of term structures, mostly with maturities less than five years. Since information content in the short and long end of the term structure is likely to be different, it is important to re-examine the question by using the entire term structure. Second, existing results are mixed and are all based on data before 1992. As international capital market integration evolves over time,[8] it is interesting to check whether recent data give more or less consistent results across different countries. Finally, earlier studies, which

[7] A long list of empirical studies include Fama (1984), Fama and Bliss (1987), Stambaugh (1988), Campbell and Shiller (1991), Bekaert, Hodrick and Marshall (1997), and Cochrane and Piazzesi (2002), to name a few.

[8] See, for example, Bekaert and Harvey (1995) and Brennan and Xia (2004).

were before the development of affine term structure models, could not link the term structure behavior to dynamic models in the international setting, so this paper provides the first international comparison of affine models with the empirical regularities in the term structure as benchmarks.

Following Campbell and Shiller (1991), we define the expectations hypothesis (EH) of the term structure as the proposition that the continuously compounded long term yield is equal to the average of the expected future short term (one-period) interest rates and a constant term premium. Formally,

$$y_t^\tau = \alpha_\tau + \frac{1}{\tau}\sum_{k=0}^{\tau-1} \mathrm{E}_t\left[y_{t+12k}^1\right], \tag{1}$$

where y_t^τ is the $\tau-$year zero-coupon yield at month t, y_{t+12k}^1 is the one-year interest rate at month $t+12k$, and α_τ is the constant term premium. The EH implies that the expected excess return on long term bonds is constant over time.

Campbell and Shiller (1991) propose two approaches to test the expectations hypothesis. The first is based on the observation that equation (1) holds for both the $\tau-$year bond at month t and the $(\tau-\Delta)$-year bond at month $t+12\Delta$, where Δ is measured in years. Taking the difference between them implies that the maturity weighted term spread $\frac{y_t^\tau - y_t^1}{\tau-1}$ predicts the change of the $\tau-$period zero-coupon yield from month t to month $t+12\Delta$, $y_{t+12\Delta}^{\tau-\Delta} - y_t^\tau$. Since our data are monthly but the zero-coupon bond maturities vary from one to ten years ($\Delta = 1$), our first approach is thus based on the following regression:

$$y_{t+12}^{\tau-1} - y_t^\tau = a_0 + a_1 \frac{y_t^\tau - y_t^1}{\tau - 1} + \epsilon, \quad \tau = 2, \cdots, 10. \tag{2}$$

Under the expectations hypothesis, a_1 is uniformly one for all bond maturities τ.

The second approach is motivated by subtracting the short rate from both sides of equation (1) and re-arranging, so that the average expected future change in the one-year rate equals the current term spread:

$$\frac{1}{\tau}\sum_{k=0}^{\tau-1} \mathrm{E}_t\left[y_{t+12k}^1\right] - y_t^1 = -\alpha_\tau + y_t^\tau - y_t^1$$

which can be equivalently re-written as

$$\frac{1}{\tau}\sum_{k=1}^{\tau-1}\left[y_{t+12k}^1 - y_t^1\right] = b_0 + b_1\left[y_t^\tau - y_t^1\right] + \epsilon, \quad \tau = 2, \cdots, 10. \tag{3}$$

Under the expectations hypothesis, the slope coefficient b_1 is again one for all bond maturities τ and the intercept $b_0 = -\alpha_\tau$ is a maturity-specific constant.

The estimates of a_1 together with the OLS and the Newey-West adjusted standard errors are reported in Table 2. The adjusted regression R^2 is less than 1.4% for CA, GM, UK and US so that in these four countries the term spread has virtually no predictive power for future yield changes. The R^2 is, however, much larger for JP, where it increases from around -0.3% for $\tau = 2$ to 9.5% for $\tau = 6$ and then decreases to 5.2% for $\tau = 9$. The estimates reported in Table 2 are inconsistent with the expectation hypothesis: the slope coefficient is mostly negative across the five countries and the nine bond maturities, and it becomes more negative as the bond maturity increases. Since both the dependent variable and the regressors are highly persistent, the Newey-West standard errors are typically twice as large as the OLS standard errors, and the slope estimates are all insignificant according to the Newey-West $t-$ratios (except for bonds with maturities of five, six, seven, and eight years in Japan). The slope estimates are only significantly different from the theoretical value of unity in GM and JP, suggesting that there is significant statistical evidence against the expectations hypothesis only in these two countries.

The OLS estimate of the slope coefficient $\hat{b}_1$ in equation (3) is reported in Table 3. In stark contrast to the results in Table 2, there is a very strong relation between average future short rate changes and the current yield spread, and the expectations hypothesis is not rejected in most cases. In Canada, for example, $\hat{b}_1$ ranges from 0.83 to 1.14, which is significantly different from zero but not from one, and the regression $\bar{R}^2$ ranges from 10% to 62%. In general, both the $\bar{R}^2$ and the slope estimate increase with τ, which measures the distance of the spread $y_t^\tau - y_t^1$. Interestingly, while Japan has the strongest predictive relation in Table 2, it has the weakest in Table 3, where the slope coefficient is not significantly different from zero for $\forall \tau \leq 6$, and is less than 0.5 and significantly different from one but not from zero for $\tau = 4, 5, 6$. The violation of the expectations hypothesis is also observed for $\tau = 6, 7, 8, 9$ in Germany and US, where the slope estimate is significantly greater than one.

In summary, while there are country-specific patterns in the results, the broad picture is similar across these five countries in that estimates from first regression specification clearly contradict the expectations hypothesis while those from the second one are mostly supportive of the hypothesis. The tension between these two sets of results is consistent with the paradox documented in Campbell and Shiller (1991) using only US yield data and a different sample period. Thus, the slope of the term structure gives no, or a wrong, forecast of the short-term change in the long term bond yield, but it gives a strong forecast which is consistent with the expectations hypothesis for long term changes in the one-year bond rate.

Bekaert, Hodrick and Marshall (1997) demonstrate via Monte Carlo simulation that both the point estimates and the OLS standard errors in these two regression-based tests of the expectations hypothesis are severely biased in small samples under the null hypothesis. They argue that results from the two regressions would provide more consistent rejections of the expectations hypothesis if small sample biases were taken into account. We examine the small sample bias in the parameter estimates in two ways. First, we derive the bias-adjusted parameter estimates as well as bias-adjusted standard errors using the bias adjustment formula given in Stambaugh (1997) and the estimation approach proposed in Amihud and Hurvich (2004). Second, we estimate the parameters and the standard errors using the bootstrap approach. We find that both the bias adjustment in the first approach and the difference between the bootstrap and the OLS estimates are negligible in the two regressions.

Two explanations of the paradox have been offered. The first, which was offered by Campbell and Shiller (1991), argues that current long term yields do move in the direction stipulated by the expectations hypothesis, but that they under-react to current short rates and/or over-react to future short rates. This explanation relies on the failure of the rational expectations hypothesis but assumes a constant risk premium. The second explanation, which was argued in Fama and Bliss (1987), assumes that expectations are rational, but that the risk (or expected term) premium are time varying. A stochastic but slow-moving risk premium would allow the expectations hypothesis to hold approximately in the short run but imply its failure in the long run.

To provide direct evidence of time variation in bond market risk premium, we regress excess bond returns on lagged forward rates in the spirit of Fama and Bliss (1987), Cochrane and Piazzesi (2002), and Bansal *et. al.* (2003), all of whom use the Fama-Bliss zero-coupon constant maturity US bond yield data with maturities from one to five years. The expectations hypothesis, that long yields are the average of expected future short yields plus a constant term premium, implies that excess returns should not be predictable, so any predictability of excess bond returns by forward rates provides additional evidence against the expectations hypothesis.

Let r^{τ}_{t+12} denote the excess log holding period return from buying a τ-period bond at time t and then selling it one year later, then

$$r^{\tau}_{t+12} \equiv \ln P_{t+12}(\tau-1) - \ln P_t(\tau) - y^1_t, \quad \forall \tau \geq 2.$$

Then define the time t log forward rate between time $t+i-1$ and $t+i$ as the rate from holding an i-year and shorting an $(i-1)$-year zero-coupon bond at t:

$$f_t(0,1) = -\ln P_t(1) = y^1_t, \quad \text{and} \quad f_t\,(i-1,i) \equiv \ln P_t(i-1) - \ln P_t(i), \quad \forall i \geq 2.$$

Ten $f_t(i, i+1)$'s for $i = 0, \cdots, 9$ can be calculated from the yield data.

We examine a multiple regression of $r_{t+12}(\tau)$ ($\tau = 2, \cdots, 10$) on several forward rates.[9] To avoid problems caused by high cross-correlation between adjacent forward rates, we restrict the regressors to be the one year spot rate ($f(0,1)$) and two- and eight-year forward rates ($f(2,3)$ and $f(8,9)$):[10]

$$r^{\tau}_{t+12} = c_0 + c_1 f_t(0,1) + c_2 f_t(2,3) + c_3 f_t(8,9) + e_t, \quad \tau = 2, \ldots, 10. \tag{4}$$

The results reported in Table 4 reveal significant joint forecasting power in the three forward rates, but the magnitude varies from country to country with CA the weakest and JP the strongest. While $\bar{R}^2$ increases with τ in CA (2.6% to 13%) and US (31% to 34%), it first increases and then decreases (hump-shaped) with τ in GM (27.5% to 33.4% and then to 25.9%), JP (28.6% to 36.3% and then to 22.5%) and UK (13.7% to 17.6% and then to 8.9%). The individual Newey-West adjusted $t-$ratios are mostly significant at the 5% level except for those in CA. Interestingly, similar to the finding in Cochrane and Piazzesi (2002), the point estimates of the coefficients exhibit an inverted "V-"shape in JP and UK where $\hat{c}_1 < 0$, $\hat{c}_2 > 0$ and $\hat{c}_3 < 0$ for all τ, but it only has such a pattern in CA, GM and US for bonds with $\tau \leq 6$. This suggests that information content is different at different maturities of the term structure in different countries, but the support for a time varying bond risk premium is largely consistent along the whole term structure and across all countries.

Since both the excess bond returns and the forward rates are highly persistent, we also examine univariate and multivariate cointegrating regressions (results not reported for brevity). The predictive power of the forward rates remains in all five countries, but now only for selective bond maturities, most of which are less than six years. The log likelihood decreases with τ, indicating a weaker predictive relation for long maturity bonds. The strong predictive power comes mainly from near- or medium- term forward rates, $f(i, i+1)$, with $i \leq 4$. Although the significance of $\hat{c}_i$ ($i = 1, 2, 3$) is weaker in the cointegrating regression, the cointegrating and the OLS regressions yield broadly consistent results in support of the forecasting power of forward rates and in rejecting the static expectations hypothesis.

A challenge faced by term structure models is then whether they can successfully explain the qualitative and quantitative behavior of the time-varying excess bond returns and forecastable term structures. Dai and Singleton (2002) show that the U.S. term structure forecastabilities are consistent with a three-factor (essentially) affine model, while Brandt and Chapman (2002) find that the Gaussian-quadratic model is

[9] A univariate regression finds predictive power of forward rates in four out of five countries, but the $\bar{R}^2$'s are much lower than those from a multi-variate regression reported in Table 4 below.

[10] Similar results are obtained by using any other forward rates $f(i, i+1)$ with $0 < i \leq 4$ to replace $f(2,3)$ and $f(i, i+1)$ with $i \geq 6$ to replace $f(8,9)$. The results are very different, however, when we include all forward rates in the regression, indicating unreliable point estimates when highly correlated neighboring forward rates are used simultaneously.

better at matching the second moments of the term structure. We will limit our analysis to the affine term structure family and examine whether the affine model, which provides the best fit of the yield curve, can generate the above-documented stylized facts of the yield data in all five countries (CA, GM, JP, UK and US).

4 Affine Term-Structure Models

The no-arbitrage condition implies that the time-t price of a default-free zero-coupon bond that matures at T, $P(t,T)$, is given by

$$P(t,T) = \mathrm{E}_t^Q \left[e^{-\int_t^T r(s)ds} \right], \tag{5}$$

where r is the (instantaneous) risk free rate and E^Q denotes the expectation under the risk-neutral (equivalent martingale) measure Q.

A N-factor affine term structure model (ATSM) is obtained by assuming that the instantaneous short rate $r(t)$ is an affine function of an N-vector of unobservable state variables $X(t) = (X_1(t), X_2(t), \cdots, X_N(t))$:

$$r(t) = \delta_0 + \delta_1' X(t), \tag{6}$$

where the scalar δ_0 and the $N \times 1$ vector δ_1 are constants, and the state variables $X(t)$ follow an "affine diffusion" under the risk neutral measure Q:

$$dX(t) = \tilde{K} \left(\tilde{\theta} - X(t) \right) dt + \Sigma \sqrt{S(t)} dW^Q(t), \tag{7}$$

where $\tilde{K}$ and Σ are $N \times N$ matrices of constants, and $\tilde{\theta}$ is an N-vector of constants. The $N \times N$ matrix $S(t)$ is diagonal and its i^{th} diagonal element is given by

$$S(t)_{ii} = \alpha_i + \beta_i' X(t), \tag{8}$$

where α_i is a scalar and β_i, the i^{th} column of the matrix $\beta \equiv [\beta_1, \cdots, \beta_N]$, is an N-vector. Finally, W^Q is an $N-$dimensional independent standard Brownian motion under the risk neutral measure Q. Note that both the drifts, $\tilde{K}\left(\tilde{\theta} - X(t)\right)$, and the conditional variance, $\Sigma S(t) \Sigma'$, of the state variable process (7) are affine functions of the state variables $X(t)$ under the measure Q.

Duffie and Kan (1996) show that under these assumptions the bond price $P(t,T)$ in equation (5) can be expressed as an exponential affine function of the underlying state variables, $X(t)$:

$$P(t,T) = \exp\left(A(\tau) - B(\tau)' X(t)\right) \tag{9}$$

where $\tau = T - t$, and the parameters $A(\tau)$ and $B(\tau)$ satisfy the following ordinary differential equations:

$$\begin{aligned} \frac{dA(\tau)}{d\tau} &= -\tilde{\theta}'\tilde{K}'B(\tau) + \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma' B(\tau)\right]_i^2 \alpha_i - \delta_0, &(10)\\ \frac{dB(\tau)}{d\tau} &= -\tilde{K}'B(\tau) - \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma' B(\tau)\right]_i^2 \beta_i + \delta_1, &(11)\end{aligned}$$

with initial conditions $A(0) = 0$ and $B(0) = 0$.

While equation (7) specifies the state variable process under the risk neutral measure Q, the time series of bond prices is governed by the physical measure. To estimate the model using observable bond prices and yields, it is necessary to specify the dynamics of the price of risk which links the dynamics of the state variables $X(t)$ under the risk-neutral measure with the dynamics under the physical measure. To this end, write the process of the pricing kernel, M, as

$$\frac{dM}{M} = -rdt - \Lambda_t' dW,$$

where the $N-$vector Λ is the price of risk associated with the innovations dW.

To ensure that the drifts and conditional variances of $X(t)$ remain affine functions of $X(t)$ under the physical measure, Λ is assumed to be a special affine function of the state variables, as suggested in Duffee (2002):

$$\Lambda_t = \sqrt{S(t)}\left(\lambda_1 + S(t)^{-1} I \lambda_2 X(t)\right), \tag{12}$$

where λ_1 is a constant N-vector, and λ_2 is a constant $N \times N$ matrix. The indicator matrix I is diagonal with the i^{th} element given by

$$I_{ii} = \begin{cases} 1 & \text{if } \inf\left(\alpha_i + \beta_i' X(t)\right) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Thus, the dynamics of $X(t)$ under the physical measure are simply

$$dX(t) \quad = K\left(\theta - X(t)\right)dt + \Sigma\sqrt{S(t)}dW(t). \tag{13}$$

Matrix K and vector θ are related to their measure-Q counterparts as follows:

$$\begin{aligned} K &= \tilde{K} - \Sigma\Phi - \Sigma I^{-}\lambda_2, \\ \theta &= K^{-1}\left(\tilde{K}\tilde{\theta} + \Sigma\Psi\right), \end{aligned}$$

where Φ is a N by N matrix with the i^{th} row defined by $\lambda_1(i)\,\beta(i)'$, and Ψ is a N-vector with the i^{th} element given by $\lambda_1(i)\,\alpha(i)$.

The specification of the price of risk in equation (12) defines the "essential" affine model of Duffee (2002) and Dai and Singleton (2002). When $\lambda_2 = 0$, the model reduces to the so-called "complete" affine model examined by Dai and Singleton (2001). Under "complete" affine term structure models, both the drift and the variance of the state variables, X, and the pricing kernel, M, are affine functions of X under P and Q. Under "essential" affine term structure models, the drift and the variance of $X(t)$ are still affine functions of $X(t)$, but the pricing kernel variance $\Lambda_t\Lambda_t'$ is no longer affine in $X(t)$. The pricing kernel variance, however, does not affect the pricing of zero-coupon bonds in equation (5). Therefore, the "essential" model with $\lambda_2 \neq 0$ retains the tractability of the "complete" affine model, but has the benefit of allowing the price of risk to vary independently of the short rate volatility.

The benefit of this added flexibility can be illustrated easily in a simple one-factor model. For example, if the short rate follows a Vasicek process with constant drift and volatility, then the price of risk Λ is also constant under the complete affine term structure model. If the short rate follows a square root process (CIR model) with a constant drift but a volatility that is proportional to the square root of the short rate, then Λ is also proportional to the square root of the short rate. This close link between short rate volatility and the price of risk severely limits the model specification. To obtain models that have a constant short rate volatility but a time-varying price of risk or vice versa, we need to extend the "complete" affine model to the "essential" affine family. More flexibility is achieved if multiple factors drive the dynamics of the short rate.[11]

5 Estimation and Model Comparison

To estimate the ATSM, additional constraints must be imposed on the parameters to ensure that the model is admissible and identifiable. For example, restrictions must be imposed to ensure that $S_{ii}(t) \geq 0$. Dai

[11] Brennan and Xia (2003) estimate an essential affine three-factor Gaussian model across the five countries considered in this paper. There model which is a special case of the models considered in this paper assumes that r and Λ follow correlated Markov Gaussian processes.

and Singleton (2001) (DS) introduced the canonical form $\mathcal{A}_m(N)$ to denote an $N-$factor model with the first m factors entering the matrix S, which determines the volatility of the state variables and in turn affects the short rate volatility. Up to N factors can affect the price of risk under the more general essential affine term structure framework. We use $\mathcal{A}_m(N)C$ to denote the complete ATSM and $\mathcal{A}_m(N)E$ to denote the essential ATSM. When $m = N$, the essential and the complete affine models coincide since the indicator matrix in equation (12), I, is zero in this case. DS provide sufficient parameter restrictions and normalizations to guarantee admissibility and identification of the affine model, and these are reproduced in Appendix A for completeness.

We start from a three-factor essential affine model, $\mathcal{A}_m(3)E$, with the number of factors affecting the short rate volatility $m = 0, 1, 2, 3$ for each country, and then examine its two sets of nested models. The first set of nested models of $\mathcal{A}_m(3)E$ ($m = 0, 1, 2$) is its complete model counterparts, $\mathcal{A}_m(3)C$, under the constraint $\lambda_2 = 0$ with λ_2 defined in (12). The second set of nested models contains essential models with fewer factors $\mathcal{A}_m(N)E$ ($m \leq N < 3$). When $m = 3$, there are no nested models. The details of the parameter constraints under the nested models are also contained in Appendix A.

While the Wald test is the correct test for the first set of nested models, it is invalid for the second set of nested models because parameters associated with excluded factors such as the elements in the third row or third column of the K matrix only exist under the alternative of three-factor models and are not identified under the null of one- or two- factor models. Andrews and Ploberger (1994) and Hansen (1996) point out that the presence of such nuisance parameters invalidates classical Wald, LM or LR tests. A suitable test in the current setting is however still unavailable in the literature. Andrews and Ploberger (1994) explore optimal testing but do not discuss methods to obtain critical values in practice. Hansen (1996) only considers regression models with additive nonlinearity, and it is not obvious how to extend his method to nonlinear latent-variable maximum likelihood estimation settings. Therefore, we use the Wald test to formally test the first set of nested models but only treat the Wald test results together with likelihood function value comparisons as suggestive examinations of the dimension of the factor model (the second set of nested models).

The estimation is carried out using bond yields with maturities of six months, one, two, three, five, seven, and ten years. The ten-year bond yield in GM and JP is plagued with many missing observations, so it is replaced by the nine-year bond yield for these two countries. Bond yields with maturities of four, six, and eight years are saved for out-of-sample analysis.

Equation (9) implies that the bond yield y_t is a linear function of the factors $X(t)$

$$y_t(\tau_j) \equiv -\frac{\ln P_{j,t}}{\tau_j} = -A(\tau_j) + B(\tau_j)' X(t) + \epsilon_t(\tau_j), \quad j = 1, \cdots, 7, \tag{14}$$

where τ_j is the j^{th} bond's time-to-maturity, $A(\tau_j)$ and $B(\tau_j)$ are defined by the ODE (10-11), while ϵ is added to reflect the possibility of model misspecification and measurement errors in the bond yield data. The complete and the essential affine term structure models assume the same state variable dynamics under the risk neutral measure Q and have the same bond pricing formula. They, however, make different assumptions for the price of risk, which lead to different state variable dynamics under the physical measure P. Therefore, equation (14) relates observable bond yields $y_t(\tau_j)$ to different expressions of $A(\tau_j)$, $B(\tau_j)$ and $X(t)$ under P for the complete and affine models.

Equation (14) is the basis of our empirical studies of ATSM. When ϵ is assumed to be zero, the N unobservable factors, X, can be inverted from any N bond yields. Since there is no a priori information governing the choice of bonds, we choose to use all seven bonds by assuming $\epsilon(\tau_j) \neq 0 \; \forall j$. In this case, the state variables $X(t)$ are treated as latent variables in the estimation.

Exact maximum likelihood estimation (MLE) is generally preferred for the estimation of the ATSM, because it has the property of consistency and asymptotic efficiency. When $X(t)$ is treated as a vector of latent variables, a Kalman filter algorithm is used together with the MLE. In the case of $A_m(N)E$ models, however, the exact MLE is feasible only when $m = 0$, i.e., only in the case of multi-factor Gaussian models. In all other nonlinear models, a quasi-maximum likelihood (QML) estimation has to be used. The exact Kalman-filter is also replaced by an approximated version in which the first order Taylor expansion is applied to the dynamics of the vector $X(t)$ in equation (13). Another widely used method is the Efficient Method of Moments (EMM) technique combined with a semi-nonparametric (SNP) auxiliary model. The EMM/SNP is more flexible and can be used to estimate non-affine term structure models such as the quadratic model proposed by Ahn, Dittmar, and Gallant (2002). Duffee and Stanton (2002) find, however, that the EMM/SNP estimation method yields quite poor small sample results within the affine term structure model family. On the other hand, the approximate Kalman filter/QML approach, despite its theoretical inconsistency, produces the best results. Therefore, we use the approximate Kalman filter/QML approach to estimate the ATSM.

The observable implications of a given model consist of a set of transition equations and a set of observation equations. The set of observation equations is given by (14). The set of transition equations for the unobservable state variables $X(t)$ is a discretized version of the dynamics of $X(t)$ in equation

(13). To simplify the estimation, $\epsilon(\tau_j)$ is assumed to have variance proportional to the inverse of the bond maturity, $\sigma^2_{\epsilon(\tau_j)} = \frac{\sigma^2_e}{\tau_j}$, and to have zero correlation both with one another and with innovations in the transition equations. When the dynamics of $X(t)$ are nonlinear, the transition equation is approximated by the conditional mean and the conditional variance of $X(t)$, both of which are derived in de Jong (2000) for the complete affine model and in Duffee (2002) for the essential affine model. For completeness, the conditional mean and variance together with the details of the estimation procedure are summarized in Appendix B.

There are three steps in the estimation. First, four non-nested $A_m(3)E$ models with $m = 0, 1, 2, 3$ are estimated for each country, and the two sets of nested models are examined under each three-factor ($N = 3$) model. Both the Wald test result[12] and the comparison of likelihood function values suggest that three-factor models dominate their one-factor ($N = 1$) or two-factor ($N = 2$) counterparts in all cases and for all countries. The Wald test, however, fails to reject the nested model $A_2(3)C$ in all countries. In this case, the same estimation approach is used in the second step to estimate the term structure model $A_2(3)C$. Results from these two steps are omitted for brevity. In summary, we find that complete affine models except for $A_2(3)C$ are all easily rejected in favor of essential models, and that models with fewer than three factors are inferior to three-factor models in all five countries.

The four non-nested $A_m(3)E$ ($m = 0, 1, 2, 3$) models together with $A_2(3)C$ are compared in the last step under three different metrics. The first metric is the Schwarz Criterion:

$$BIC = \text{loglikelihood} - \frac{1}{2} N_\theta \ln(T)$$

where N_θ is the number of parameters and T is the number of observations. The Schwarz Criterion compares the log likelihood function value with a penalty for over-parameterization. The higher is BIC, the better is the model. The second metric is the in-sample absolute pricing error, PE, averaged across all the seven bonds and over all the observations. The absolute pricing error at each time t is defined as

$$PE_t = \frac{1}{J} \sum_{j=1}^{J} |\hat{y}_t(\tau_j) - y_t(\tau_j)|$$

where $J = 7$ is the number of bonds used in the estimation, $\hat{y}_t(\tau_j)$ is the predicted, and $y_t(\tau_j)$ is the

[12] As we discussed earlier, the Wald test is not valid in the test of the dimension of a factor model due to the presence of nuisance parameters, so the results are only suggestive.

observed yield for maturity τ_j at t. The time-series average absolute pricing error

$$PE = \frac{1}{T}\sum_{t=1}^{T} PE_t$$

is then used to compare different models. The third metric is the out-of-sample absolute pricing error, PEO, averaged across the three yields not used in the estimations.

Table 5 reports the values of BIC, PE and PEO for the five models. The best value is reported in bold face. First, all three metrics choose the same best model $A_1(3)E$ in JP, UK and US. In CA, $A_1(3)E$ performs the best according to the BIC criterion and is the second best according to both the PE and the PEO criteria. In GM, $A_1(3)E$ is the best model according to both PE and PEO while it is the second-to-best model based on BIC. Taking into account of all three metrics, $A_1(3)E$ seems to be the best model for all five countries. In light of the wide difference in the monetary policies and government bond markets, this finding is striking and suggests that the $A_1(3)E$ model is flexible enough to fit many different term structure behaviors in different markets.

Table 6 reports the parameter estimates for $A_1(3)E$ for the five countries. Due to the large number of parameters (24) and the short sample period (only 233-242 monthly observations), it is not surprising to find that most of the parameter estimates have large asymptotic standard errors and are thus not statistically significant.[13]

The estimate of the measurement error volatility σ_e is statistically significant but economically small in all five countries; it varies from about 8-9 basis points in CA and UK to about 29 basis points in GM. This implies a pricing error of about 11 (2-3) basis points for the six-month (ten-year) bond in CA and about 41 (9) basis points for the six-month (ten-year) bond in GM. The estimate of δ_0, the constant term in the short rate process (6), is positive and significant in GM, JP, UK and US, but it is close to zero and insignificant in CA. The coefficients δ_i ($i = 1, 2, 3$) are all small. In four out of the five countries, at least one δ_i is significantly different from zero at the 5% level. The long run mean of the first state variable, θ_1, is only statistically significantly different from zero in CA, GM and UK. The estimates for β_{12} and β_{13} are very imprecise with large standard errors, but they are close to the numbers reported by Duffee (2002).

The mean-reversion parameter for the first factor, K_{11}, is generally small in magnitude (from 0.02 in

[13] We estimate the maximum number of possible parameters allowed under the identification and admissibility. Both Duffee (2002) and Dai and Singleton (2002) found similarly large standard errors associated with their parameter estimates under the most general parameter specifications, but they then re-estimate a preferred model by exogenously setting certain parameters to zero.

UK to 0.15 in GM, but 0.43 in U.S.) but highly significant in all five countries, indicating that the first state variable X_1 is very persistent. The mean-reversion parameter K_{22} for the second state variable is only significant in CA, where the estimate of 0.52 is much larger than its first-factor counterpart. The estimates for the other four countries, although still positive, are smaller in magnitude and statistically insignificant. The estimates of the third-factor mean-reverting parameter, K_{33}, ranging from 0.25 to 1.6, are only significant in US and UK. The off-diagonal elements in the matrix K are generally statistically insignificant except for K_{23} in US. The point estimates suggest that X_2 and X_3 are generally more transient as compared to X_1.

Estimates of the elements of the first price of risk vector, λ_1, are mostly negative across all five countries. In CA, GM, UK and US, at least one element in λ_1 is significantly different from zero, but all three elements of λ_1 in JP are insignificant. Finally, the estimates of the matrix λ_2 are generally quite imprecise. The estimates for the element $\lambda_{2,21}$ are often around 60 with even larger standard errors. There is no statistically significant estimate out of the six elements of λ_2 in GM, JP, and UK, but there is one significant estimate in CA and two in US.

Table 7 reports summary statistics on the yield prediction errors. The mean and the standard deviation of the absolute prediction errors are reported in the first two rows while the maximum and the minimum of the raw prediction errors are reported in the last two rows. The left panel reports results for bonds used in the estimation, while the right panel contains statistics for bonds not used in the estimation.

The average absolute prediction error generally declines with the yield maturity. For example, it ranges from 40 basis points for the six-month yield to about 27 basis points for the ten-year yield in Canada. Similar magnitude of the average absolute mispricing are observed in GM, UK and US. Japan's mispricing at around 20 basis points across all maturities is the smallest. The standard deviation of the prediction error also declines with maturity, which is around 30 basis points at the short end and is about 18 basis points at the long end. Judging from these two statistics, the model does a good job in fitting the term structure across all five countries. The maximum and the minimum prediction errors, however, paint a different picture. They indicate very poor fit of the model for at least one point of the sample. For example, the maximum absolute prediction error for the six-month bond ranges from over 140 basis points in the US to about 275 basis points in Germany. Even at the long end, the largest prediction error is still at around 100 basis points. The small average prediction error together with occasionally large error confirms the observation that the predicted and the actual yield curves are almost on top of one another for most periods, but there are a few points of large deviations during the sample period.

To study the behavior of the factors extracted by the Kalman filter algorithm, we first examine their correspondence with the first three principal components "level," "slope," and "curvature." The correlations between one of the factors and one of the principal components are reported in Table 8. There is always one extracted factor, X_1 for CA, GM, JP, and US and X_2 for GM, that is highly correlated with the "level" component, and the correlation coefficient is above 0.9 in CA, JP, UK and US and is around 0.87 in GM. The correspondence between one of our extracted factors, X_3 for CA, GM, UK, and US and X_2 for JP, with the "slope" component is weaker, ranging from only 0.45 in JP to about 0.81 in CA. The correspondence between one of our extracted factors, X_2 for CA, GM, and US, X_2 for JP, and X_1 for UK, and the "curvature" component is the weakest, with the highest correlation at about 0.7 in the US and the lowest at 0.4 in UK. Therefore, the behavior of the three factors, although all from the $A_1(3)E$ model, differs from country to country.

In addition to the above analysis, we also examine how well the model-implied instantaneous risk free rate, $\hat{r}(t)$,

$$\hat{r}(t) \equiv \hat{\delta}_0 + \hat{\delta}_1' \hat{X}(t), \tag{15}$$

tracks the one-month Treasury bill rate, TB. The sample mean (sample volatility) of $\hat{r}(t)$ for CA, GM, JP, UK and US is, respectively, 7.11%, 5.38%, 3.55%, 8.13% and 6.31% (2.63%, 1.90%, 2.36%, 3.19%, and 2.42%), which compares with the sample mean (sample volatility) of the corresponding one-month Treasury Bill rates at 7.10%, 4.95%, 3.21%, 8.30%, and 5.46% (2.98%, 1.78%, 2.63%, 3.10%, and 2.04%). The time series of both $\hat{r}(t)$ and TB of the five countries are plotted in Figure 2. Since the one-month Treasury bill rate[14] is not used in the estimation, this plot provides an out-of-sample examination of the goodness-of-fit of the best model at the very short end of the yield curve.

Figure 2 shows that $A_1(3)E$ works the best in UK, where the fitted short rates and the one-month Treasury bill rates are virtually coincident. The model also works well in CA and JP. In CA, $\hat{r}$ is close to TB after 1987, but $\hat{r}$ is substantially lower than TB around 1986. In JP, $\hat{r}$ is close to TB before 1994, but the estimated short rate at around 1% is higher than the T-bill rate, which is virtually zero in late 1990s. The model does not seem to provide a good fit for the US T-bill rates. Although the time series dynamics of $\hat{r}$ closely tracks that of the TB, the implied short rates are higher than TB in most periods. The implied short rate runs up to a much higher level in 1999 and then it drops to a lower level than TB

[14] The one-month Treasury bill rates for the U.S. are from CRSP, the rates for CA and UK are from Datastream, and the rates for GM is from Bloomberg. The Japan T-bill rates from March 1993 to December 2002 are from Datastream, and those from January 1983 to February 1993 are provided by Kent Daniel.

in 2002. The model works well in GM before 1997, but $\hat{r}$ has a large dip in 1997-98 and then a run-up in 1999 while the T-bill rates has much smaller swings during this period. It is possible that the replacement of Mark-denominated bonds by Euro-denominated bonds caused structural break or instability in the yield curve, leading to the poor performance of any model without regime shifts. This figure suggests that while $A_1 3(E)$ is the best model among the family of affine term structure models and provides a reasonably good fit for the term structure, alternative non-affine models may work better for some countries than the others.

Although the parameters governing the price of risk, λ_1 and λ_2, are estimated with large errors, it is still instructive to examine the time series of the price of risk. The estimated vector of market price of risk is given by

$$\hat{\Lambda}_t = \sqrt{\hat{S}(t)} \left(\hat{\lambda}_1 + \hat{S}(t)^{-1} I \hat{\lambda}_2 \hat{X}(t) \right), \tag{16}$$

where the i^{th} element, $\hat{\Lambda}(i)$ ($i = 1, 2, 3$), of the 3×1 vector $\hat{\Lambda}(t)$ gives the price of risk associated with the i^{th} state variable at time t.

Figure 3 contains the plot of $\hat{\Lambda}(i)$ ($i = 1, 2, 3$) for the five countries. The one standard deviation bound is calculated but not plotted in the figure. In general, $\hat{\Lambda}(1)$ is smaller in magnitude with tight one standard deviation bounds, and it is significantly different from zero in most periods. The other two prices of risk, $\Lambda(2)$ and $\Lambda(3)$, however, are associated with large one standard deviation bounds so that they are generally statistically insignificant despite their large magnitude. This reflects the difficulty in obtaining a good estimate of the market price of risk in general affine term structure models.[15] While $\hat{\Lambda}(1)$ increases over time in CA, GM and JP, it declines during the same sample period in UK and US. There is a huge spike in Germany's $\hat{\Lambda}(2)$ at the end of 1998, which is accompanied by a plunge of $\hat{\Lambda}(3)$ at the same time. A similar pattern is observed in the US during the period of 2000-2003, when a large decline in $\hat{\Lambda}(2)$ coincides with a run-up in $\hat{\Lambda}(3)$. Note that all $\hat{\Lambda}(i)$'s ($i = 1, 2, 3$) peaked (in the absolute value) around the stock market crash in year 2000, suggesting some co-movement of prices of risk in the international capital markets.

Figure 4 plots,

$$\hat{\eta} \equiv \sqrt{\hat{\Lambda}\hat{\Lambda}'},$$

which is the estimated maximum Sharpe ratio in the bond market. As in Figure 3, the precision of $\hat{\Lambda}$ is

[15] Dai and Singleton (2002) and Duffie (2002), for example, also report large standard errors for their price of risk estimates in general affine term structure models.

very low. Nonetheless, we detect some interesting patterns in the plot.

As compared to the risk free rate in Figure 2, the market price of risk displays much more volatility and contains some spikes in the plot. The sample mean (standard deviation) of $\hat{\eta}$ in CA, GM, JP, UK and US are respectively, 0.88 (0.57), 0.98 (0.41), 2.32 (1.12), 0.98 (0.55), and 2.02 (1.13). In CA, $\hat{\eta}$ swings from almost zero in early 1983 and late 1980s to almost three around 1985-86. It steadily increases during the period of 1989-1992 before it declines in early 1990s. The GM $\hat{\eta}$ is stable and lies mostly below one before 1991 (the unification), but it exhibits more variation afterwards and even shot up to almost three when the Euro was introduced in 1998-99. On the one hand, this is consistent with the previous observation that the three-factor affine models may be a poor for German yield data which can then lead to unreasonable parameter and factor estimates. On the other hand, it may rationally reflect the higher compensation of risk required by investors during this period, which brings a new monetary regime and a lot of uncertainty in the implementation of the new currency. The price of risk for Japanese bond market also seems to contain a structural break. Before 1991, $\hat{\eta}$ has long swings around the mean of 1.4, but it then steadily increases over time after 1991 to the peak of 5 in 1996 before it declines toward the mean of 2.3. The UK total market price of risk $\hat{\eta}$ fluctuates around its mean of 0.98 with a range from zero to two except for the spike of 3 occurred in 2000. There are two spikes in the US $\hat{\eta}$, with one in September 1985 and the other in September 2000. While the first run-up in $\hat{\Lambda}$ happens right after the Volcker monetary experiment during 1979-1982, the second run-up coincides with the burst of the Internet bubble in year 2000. Therefore, the price of risk in the bond market is affected by both country-specific monetary policies and common international factors.

6 Term Structure Forecasts: Some Simulation Evidence

Although the $A_1(3)E$ model provides a reasonable fit to the data and leads to reasonable estimates of the instantaneous short rate and market prices of risk, the ultimate challenge for the model is to explain the qualitative and quantitative behavior of term structure predictability that was documented in Section 3. In this section, we provide simulation evidence on the small-sample behavior of the term structure forecastability.[16]

We simulate 2000 time series of yields for maturities of one to ten years using parameter estimates from the best model, $A_1(3)E$, for each country. The length of each time series corresponds to the number

[16]Dai and Singleton (2002), Duffee (2002), and Brandt and Chapman (2002) all examine whether essential affine models can accommodate various stylized facts in the US yield curve.

of observations in the actual historical data and varies between 233 and 242 months. The details of the simulation are relegated to Appendix C.

The regressions (2), (3), and (4) are carried out on each simulated dataset. First, the annual yield change of a bond of maturity τ years is regressed on the yield spread between the τ-year and the one-year yield as in equation (2). Second, the average of future short rate changes is regressed on the yield spread as in equation (3). Finally, excess bond returns are regressed on the forward rates as in equation (4). The slope coefficient estimates $\hat{a}_1$, $\hat{b}_1$, and $\hat{c}_i(i = 1, 2, 3)$, which correspond to the historical estimates reported in Tables 2, 3 and 4, are recorded. The means and standard deviations of the estimates $\hat{a}_1$, $\hat{b}_1$, and $\hat{c}_i(i = 1, 2, 3)$ from the 2000 simulations are then interpreted as the simulated point estimates and the simulated standard errors.

Figure 5 plots the simulated slope estimates of $\hat{a}_1$ together with its two-standard-error bounds. For comparison, the estimates obtained from the historical data which were reported in Table 2 are also plotted in the figure. In contrast to the observation that the sample estimates from Table 2 are mostly negative and become more so when the bond maturity τ increases, the simulated estimates based on $A_1(3)E$ and its MLE parameters are mostly positive and show an upward trend with τ in CA, GM, and US.[17] The estimates $\hat{a}_1$ in JP are close to zero and flat in τ while the estimates in UK are decreasing with τ. Interestingly, although the best models across the five countries are all $A_1(3)E$, they generate different small-sample patterns in the slope coefficient estimates. The slope coefficients are estimated with such large error in the small sample that the historical estimates from section 2 fall safely within the two standard error bounds except for those corresponding to long maturities in CA and short maturities in UK. Therefore, the pattern of slope coefficients we found in the historical data could conceivably have been generated by the $A_1(3)E$ model.

Figure 6 plots the slope estimates of $\hat{b}_1$ from the simulated data together with its two-standard-error bounds. The estimates obtained from the historical data which were reported in Table 3 are also plotted in the figure. Consistent with the historical estimates reported in Table 3, the simulated slope estimates increase with maturity τ in CA, GM, JP, and US. In UK, however, the simulated slope estimates decline with τ, which is opposite to the pattern observed in the historical estimates. While the historical estimates all fall inside the two standard error bounds of the simulated estimates in CA, GM, JP and US, UK's sample estimates at both the short and the long ends fall outside the bounds.

[17] In light of the findings of Dai and Singleton (2002), who show that only $A_0(3)E$ model has a reasonable chance in accommodating the coefficient patterns in Table 2 using US yield data, it is not surprising that our best models across five countries, all of which are $A_1(3)E$, fail to generate a similar pattern as that observed in Table 2.

Figure 7 plots the simulated coefficient estimates $\hat{c}_1$, $\hat{c}_2$ and $\hat{c}_3$ together with their two-standard-error bounds from regression (4). The results are plotted only for bonds with three years to maturity ($\tau = 3$), since the results for other maturities are broadly similar. Corresponding coefficient estimates based on historical data and reported in Table 5 are also plotted. The two-standard-error bounds are tighter in GM and UK as compared to those in CA, JP and US. The simulated coefficient estimates are very close to the historical estimates in CA, and they are not statistically different from the historical ones in GM and US. The coefficient estimates from the simulated data, however, are statistically different from their historical counterparts in JP and UK. Although the simulated estimates of the coefficients exhibit an inverted "V-"shape in four of the five countries, this pattern disappears as τ increases.

The results from Figures 5 to 7 taken together imply that the best essential affine models $A_1(3)E$ whose parameter estimates were reported in Table 6 could conceivably generate the stylized empirical results that were documented in Section 3. However, it is also possible that the difficulty in obtaining precise estimates of the affine term structure model parameters, especially those governing risk premia, translates into very low power in refuting the model. Therefore, it is premature to conclude that an affine term structure model is successful in generating the rich yield curve behaviors and strong term structure forecastability observed in the data.

7 Conclusion

In contrast to the large and growing empirical dynamic term structure literature examining the US yield curve, similar analysis for other developed bond markets is still lacking. In this paper, we extend analysis of the yield curve and the performance comparison of dynamic affine term structure models to five developed government bond markets: Canada, Germany, Japan, United Kingdom and the United States.

Although the first three principal components of yields can be broadly characterized as "level," "slope," and "curvature," each country has unique yield curve features. Using three frequently used regression specifications, we first provide extensive empirical evidence that the bond risk premium is time varying in all five countries.

Since there is no prior evidence on which affine term structure model fits the yield curve best in countries other than US, this paper examines all possible models within the family of affine dynamic term structure models with one to three factors. The models are estimated using the Kalman filter/QML approach, and the Wald test is carried out to compare nested models. The Schwartz criteria and the in- and out-of- sample prediction errors are used to compare non-nested models. The three-factor essentially affine model, $A_1(3)E$, in which only one factor affects the instantaneous volatility of short rates but all three factors affect the price of risk, appears to be the best model in all five countries. This finding is consistent with results in Dai and Singleton (2002) and Duffee (2002) on US data. The $A_1 3(E)$ model provides very good average fit of the yield curve in all five countries, but has occasional large mis-pricing during the sample.

Finally, we examine whether the best affine model can explain qualitative and quantitative aspects of yield curve behavior such as the forecast power of forward rates for expected bond excess returns. Our simulation evidence shows that, while point estimates of regression coefficients exhibit patterns different from the point estimates obtained from historical data, they are not statistically different from them owing to the large estimation errors that arise in the small sample regression studies. This lack of power renders an inconclusive answer to the question of whether the best affine term structure model can successfully generate the observed stylized empirical findings in the data. This paper has restricted itself to the analysis of affine term structure models. Further work is required to assess the performance of quadratic term structure models and term structures models with regime shifts on international data.

To conclude the paper, we note that while there are country-specific empirical regularities, qualitative results and their implications are strikingly consistent across all five countries. It is thus likely that common international factors and economic forces as well as country specific monetary policies are important driving forces behind these government term structures.

Appendix

A. Parameter Restrictions and Nested Hypothesis Tests

This appendix contains information on the parameter constraints in the five canonical affine term structure models, which are explicitly estimated in this paper, to ensure admissibility and identification. In all three-factor models, $\Sigma = I_{3\times 3}$ is always an identity matrix. The parameter δ_0 is unconstrained in all five models. Other parameter constraints are summarized in Table A1.

Under each essential affine model $A_m(N)E$, there are several nested models. The model $A_0(3)E$ nests three alternative models: (1) $A_0(2)E$ with the parameter constraints $H_0 : \lambda_{1,3} = \lambda_{2,31} = \lambda_{2,32} = \lambda_{2,33} = \lambda_{2,13} = \lambda_{2,23} = \delta_1(3) = 0$; (2) $A_0(1)E$ with parameter constraints $H_0 : \lambda_{1,3} = \lambda_{2,31} = \lambda_{2,32} = \lambda_{2,33} = \lambda_{2,13} = \lambda_{2,23} = \delta_1(3) = 0$, and $\lambda 1_2 = \lambda 2_{12} = \lambda 2_{21} = \lambda 2_{22} = \delta_1(2) = 0$; and (3) $A_0(3)C$ with parameter constraints $H_0 : \lambda_2 = 0$.

The model $A_1(3)E$ also nests three models: (1) $A_1(2)E$ with parameter constraints $H_0 : \lambda_{1,3} = \lambda_{2,31} = \lambda_{2,32} = \lambda_{2,33} = \lambda_{2,23} = \delta_1(3) = 0$; (2) $A_1(1)C$ with parameter constraints $H_0 : \lambda_{1,3} = \lambda_{2,31} = \lambda_{2,32} = \lambda_{2,33} = \lambda_{2,23} = \delta_1(3) = 0$, and $\lambda_{1,2} = \lambda_{2,21} = \lambda_{2,22} = \delta_1(2) = 0$; and (3) $A_1(3)C$ with $H_0 : \lambda_2 = 0$.

The model $A_2(3)E$ nests two models: (1) $A_2(2)C$ with parameter constraints $H_0 : \lambda_{1,3} = \lambda_{2,31} = \lambda_{2,32} = \lambda_{2,33} = \delta_1(3) = 0$; and (2) $A_2(3)C$ with $H_0 : \lambda_2 = 0$.

Finally, the model $A_2(3)C$ nests $A_2(2)C$ with $H_0 : \lambda 1_3 = \delta_1(3) = 0$ while there is no nested models under $A_3(3)C$.

B. Estimation Procedure

de Jong (2000) and Duffee (2002) show that the conditional mean and the conditional variance of the state variables X are given by

$$\mathrm{E}\left[X_s|X_t\right] = \left(I - e^{-K(s-t)}\right)\theta + e^{-K(s-t)}X_t, \tag{B1}$$

$$\mathrm{Var}\left[X_s|X_t\right] = Lb_0L^{'} + \sum_{i=1}^{N}\left(\sum_{l=1}^{N} Lb_lL^{'}L^{-1}_{l,i}\right)X_{t,i}, \tag{B2}$$

where the matrix L is from the diagonalization of K in $K = LDL^{-1}$ with D as the diagonal matrix. The diagonal elements of D are denoted $d_1, \cdots, d_N$.

Before defining b_0 and b_l ($l = 1, \cdots, N$), we need to introduce several new variables:

$$\begin{aligned}
\alpha^* &= \alpha, \quad \theta^* = L^{-1}\theta, \quad \Sigma^* = L^{-1}\Sigma, \quad \beta^* = \beta L, \\
G_0 &= \Sigma^* diag\left(\alpha^*\right)\left(\Sigma^*\right)', \quad G_i = \Sigma^* diag\left(\beta_i^*\right)\left(\Sigma^*\right)' \quad (i = 1, \cdots, N),
\end{aligned}$$

where β_i^* is the i^{th} row of the matrix β^* and G_0 and G_i are $N \times N$ matrices.

The variance of X_s conditional on X_t depends on $s - t$. Define $F_0(s-t)$, $F_i(s-t)$ ($i = 1, \cdots, N$), and $H_i(s-t)$ ($i = 1, \cdots, N$) as time-varying $N \times N$ matrices with typical element (j, k) given by

$$\begin{aligned}
F_0(s-t)^{(j,k)} &\equiv \left(d_j + d_k\right)^{-1} G_0^{(j,k)} \left(1 - e^{-(d_j+d_k)(s-t)}\right), \\
F_i(s-t)^{(j,k)} &\equiv \left(d_j + d_k\right)^{-1} G_i^{(j,k)} \left(1 - e^{-(d_j+d_k)(s-t)}\right), \\
H_i(s-t)^{(j,k)} &\equiv \left(d_j + d_k - d_i\right)^{-1} G_i^{(j,k)} \left(e^{-d_i(s-t)} - e^{-(d_j+d_k)(s-t)}\right),
\end{aligned}$$

then we have

$$b_0 = F_0(s-t) + \sum_{i=1}^{N} \theta_i^* \left[F_i - H_i\right], \quad \text{and} \quad b_l = H_l.$$

In the Kalman filter/QML approach, the transition equation for the state variables X is the discretized version equation (13) with the conditional mean and conditional variance given in (B1-B2); and the observation equation is given in (14). The observation error ϵ is assumed to be independent of the innovations to the state variables dW. In addition, the variance-covariance of the observation error $V \equiv \text{Var}\left(\epsilon\right)$ is assumed to be a $J \times J$ diagonal matrix with the j^{th} diagonal element given by $V_{jj} = \frac{\sigma^2}{\tau_j}$. This implies that the pricing error of each bond due to possibly measurement error or model mis-specification is independent of the bond maturity. The details of the estimation procedure are as follows:

- The selection of initial parameter values.

  1. Duffee (2002)'s estimated optimal values (from his web site) were used as the initial values for each model.
  2. The two standard-deviation boundaries of the converged estimates were calculated from the above initial values.
  3. Set the converged estimate as the first set of initial parameter values, and draw other initial parameter values randomly within the two standard deviation boundaries until 50 converged outputs were obtained.

- The evaluation of the estimation.

1. The minimization process was carried out using the simplex method (Fortran routine DUMPOL).
2. A specific parameter set is discarded if
   (a) The program fails to converge.
   (b) The converged estimate $\hat{\theta}$ is not valid (see definition below).
   (c) $\hat{\theta} \pm \triangle\theta$ is invalid, where $\triangle\theta = \min\{\text{abs}(\theta) \cdot 10^{-5}, 10^{-9}\}$ if $\theta \neq 0$; and $\triangle\theta = 10^{-9}$ otherwise.
   (d) The eigenvalue of the information matrix is complex.
3. An estimate $\hat{\theta}$ is invalid if
   (a) It does not satisfy the requirements of the canonical form (see Appendix A).
   (b) The numerical solutions of ODE in equations (10-11) cannot be achieved with the desired precision within the maximum 100,000 steps.
   (c) Positive-definite matrices fail to be positive definite at $\hat{\theta}$.

- The treatment of missing observations.

1. The system matrices from the observation equation are allowed to be different across time in the Kalman filter algorithm.
2. If bond yields of certain maturities are missing at a specific date, then only available bond yields of other maturities are used in the updating. Therefore, the number of observation equations may be different at different time, leading to different system matrices.

C. Simulation

In the simulation, the MLE parameter estimates are treated as the "true" parameter values. The state variable $X(t)$ is generated by assuming that the discretized process for $X(t)$ is a stochastic difference equation

$$\Delta X(t+h) = K(\theta - X(t))\,h + \Sigma\sqrt{S(t)}\varepsilon(t+h)\sqrt{h}, \tag{C1}$$

where $h = \frac{1}{360}$ (one day) is the time interval between two observation, and $\varepsilon(t+h)$ has an $N \times 1$ standard normal distribution. The terms $K(\theta - X(t))$ and $\Sigma\sqrt{S(t)}$ are exactly the same as the drift and volatility terms in equation (13).

To generate state variables X at the monthly frequency, the sequence of values $\{X(t+h), X(t+2h), \cdots, X(t+30h)\}$ is simulated from equation (C1) by starting from $X(t)$ at the beginning of the month and then saving $X(t+30h)$ as the observation at the end of the month. Using the long run mean θ as the initial value $X(0)$, a sample of 466 monthly observations of state variables are generated. The first 233 observations are discarded and the last 233 observations are used to construct a sample of bond yields with maturities of one to ten years according to equation (14).

Each set of simulated bond yield data consists of 233 observations for CA, GM and JP, and of 242 and 241 observations for UK and US. The regressions (2), (3), and (4) are carried out and the slope coefficient estimates $\hat{a}_1$, $\hat{b}_1$, and $\hat{c}_i(i=1,2,3)$ are recorded in each of the 2000 simulated datasets. The mean and standard deviation of the 2000 estimates $\hat{a}_1$, $\hat{b}_1$, and $\hat{c}_i(i=1,2,3)$ are then calculated as the point estimate and the standard error of the coefficients.

Table 1

Summary Statistics of Zero-Coupon Constant Maturity Government Bond Yields

This table reports summary statistics for the estimated zero-coupon constant maturity government bond yields with maturities of six months and one to ten years. The sample starts from January 1983 and ends in May 2002 for CA, GM, and JP, but ends in February 2003 for UK and in January 2003 for US.

1. Canada

	Bond Yield Maturities										
Securities	0.5	1	2	3	4	5	6	7	8	9	10
mean (% per year)	7.25	7.37	7.59	7.78	7.95	8.09	8.18	8.25	8.30	8.34	8.39
Std. Dev. (% per year)	2.54	2.46	2.34	2.27	2.23	2.21	2.19	2.15	2.11	2.08	2.07
No. of Observation	233	233	233	233	233	233	233	233	233	233	233
Autocorrelation	0.977	0.979	0.982	0.984	0.986	0.987	0.988	0.988	0.989	0.989	0.989

2. Germany

	Bond Yield Maturities										
Securities	0.5	1	2	3	4	5	6	7	8	9	10
mean (% per year)	5.50	5.62	5.78	5.85	5.92	6.03	6.18	6.31	6.41	6.47	6.35
Std. Dev. (% per year)	1.84	1.62	1.48	1.46	1.44	1.39	1.33	1.27	1.22	1.17	1.10
No. of Observation	233	233	233	233	233	233	233	233	233	233	203
Autocorrelation	0.968	0.976	0.980	0.985	0.988	0.988	0.988	0.987	0.985	0.982	0.975

3. Japan

	Bond Yield Maturities										
Securities	0.5	1	2	3	4	5	6	7	8	9	10
mean (% per year)	3.34	3.28	3.31	3.50	3.71	3.90	4.05	4.16	4.27	4.37	3.94
Std. Dev. (% per year)	2.49	2.48	2.45	2.39	2.33	2.25	2.17	2.10	2.06	2.04	1.78
No. of Observation	233	233	233	233	233	233	233	233	233	233	193
Autocorrelation	0.993	0.995	0.994	0.993	0.993	0.993	0.992	0.992	0.993	0.993	0.990

4. United Kingdom

	Bond Yield Maturities										
Securities	0.5	1	2	3	4	5	6	7	8	9	10
mean (% per year)	7.65	8.06	8.12	8.18	8.23	8.27	8.29	8.31	8.31	8.30	8.27
Std. Dev. (% per year)	2.72	2.67	2.44	2.34	2.30	2.29	2.30	2.30	2.30	2.29	2.28
No. of Observation	197	241	242	242	242	242	242	242	242	242	242
Autocorrelation	0.988	0.986	0.985	0.984	0.985	0.985	0.986	0.987	0.988	0.988	0.989

5. United States

	Bond Yield Maturities										
Securities	0.5	1	2	3	4	5	6	7	8	9	10
mean (% per year)	6.38	6.45	6.60	6.75	6.91	7.06	7.21	7.35	7.47	7.56	7.60
Std. Dev. (% per year)	2.29	2.24	2.20	2.18	2.14	2.05	1.91	1.73	1.54	1.31	1.03
No. of Observation	241	241	241	241	241	241	241	241	241	241	241
Autocorrelation	0.986	0.986	0.987	0.987	0.987	0.986	0.985	0.983	0.980	0.976	0.966

Table 2

The Regression of Change in Long-Term Yields on the Yield Spread

This table reports the OLS slope coefficient from equation (2):

$$y_{t+12}^{\tau-1} - y_t^{\tau} = a_0 + a_1 \frac{y_t^{\tau} - y_t^{1}}{\tau - 1} + \epsilon, \quad \tau = 2, \cdots, 10,$$

where the yield change of a τ-year bond from month t to $t+12$ is regressed on the yield spread between the $\tau-$period bond and the one-year bond at month t. The OLS standard error is in parentheses and the Newey-West adjusted standard error is in the brackets. The adjusted R^2 is in percentage. Regression results in GM and JP for the ten-year bonds are omitted due to missing data.

Country	Bond Maturity τ								
	2	3	4	5	6	7	8	9	10
CA	0.669	0.384	0.180	0.081	0.081	-0.027	-0.316	-0.725	-1.140
	(0.33)	(0.32)	(0.31)	(0.33)	(0.36)	(0.40)	(0.45)	(0.51)	(0.56)
	[0.70]	[0.69]	[0.69]	[0.73]	[0.80]	[0.90]	[1.00]	[1.11]	[1.21]
$\bar{R}^2$	1.4	0.2	-0.3	-0.4	-0.4	-0.5	-0.2	0.5	1.4
GM	-0.041	-0.391	-0.481	-0.425	-0.435	-0.543	-0.679	-0.871	
	(0.13)	(0.16)	(0.17)	(0.19)	(0.22)	(0.25)	(0.29)	(0.33)	
	[0.25]	[0.28]	[0.33]	[0.36]	[0.40]	[0.47]	[0.56]	[0.65]	
$\bar{R}^2$	-0.4	2.4	3.0	1.7	1.4	1.7	2.0	2.6	
JP	0.225	-0.662	-1.717	-2.268	-2.451	-2.362	-2.177	-1.982	
	(0.33)	(0.49)	(0.50)	(0.49)	(0.50)	(0.52)	(0.54)	(0.55)	
	[0.64]	[1.07]	[1.19]	[1.20]	[1.21]	[1.25]	[1.30]	[1.32]	
$\bar{R}^2$	-0.3	0.4	4.7	8.5	9.5	8.1	6.5	5.2	
UK	0.009	-0.033	-0.041	-0.053	-0.086	-0.150	-0.248	-0.380	-0.545
	(0.22)	(0.25)	(0.29)	(0.32)	(0.35)	(0.37)	(0.40)	(0.42)	(0.44)
	[0.43]	[0.49]	[0.56]	[0.64]	[0.72]	[0.80]	[0.87]	[0.94]	[1.01]
$\bar{R}^2$	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.3	-0.1	0.2
US	-0.031	-0.260	-0.414	-0.406	-0.235	-0.222	-0.278	-0.238	-0.364
	(0.28)	(0.29)	(0.32)	(0.38)	(0.47)	(0.54)	(0.56)	(0.52)	(0.46)
	[0.62]	[0.64]	[0.70]	[0.82]	[0.96]	[1.02]	[1.05]	[1.07]	[0.97]
$\bar{R}^2$	-0.4	-0.1	0.3	0.1	-0.3	-0.4	-0.3	-0.4	-0.2

Table 3

The Regression of Short Rate Change on the Yield Spread

This table reports the OLS slope coefficient from equation (3):

$$\frac{1}{\tau}\sum_{k=1}^{\tau-1}\left[y_{t+12k}^{1}-y_{t}^{1}\right]=b_{0}+b_{1}\left[y_{t}^{\tau}-y_{t}^{1}\right]+\epsilon,\quad \tau=2,\cdots,10,$$

where the average future change of the one-year bond is regressed on the yield spread between the $\tau-$period bond and the one-year bond. The OLS standard error is in parentheses and the Newey-West adjusted standard error is in the brackets. The adjusted R^2 is in percentage. Regression results in GM and JP for the ten-year bonds are omitted due to missing data.

Country	Bond Maturity τ								
	2	3	4	5	6	7	8	9	10
CA	0.834	0.832	0.896	0.976	1.000	1.043	1.087	1.128	1.144
	(0.17)	(0.12)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)
	[0.35]	[0.24]	[0.18]	[0.14]	[0.11]	[0.11]	[0.11]	[0.10]	[0.11]
$\bar{R}^2$	9.9	18.2	31.2	46.4	56.3	60.3	61.6	65.3	61.8
GM	0.479	0.526	0.677	1.298	1.401	1.453	1.504	1.530	
	(0.07)	(0.07)	(0.08)	(0.10)	(0.09)	(0.09)	(0.08)	(0.07)	
	[0.13]	[0.13]	[0.16]	[0.18]	[0.16]	[0.16]	[0.14]	[0.13]	
$\bar{R}^2$	19.3	21.6	28.1	46.0	57.2	62.6	69.6	78.3	
JP	0.612	0.690	0.426	0.371	0.443	0.622	0.942	1.467	
	(0.17)	(0.20)	(0.18)	(0.16)	(0.15)	(0.15)	(0.16)	(0.15)	
	[0.32]	[0.41]	[0.39]	[0.33]	[0.30]	[0.31]	[0.38]	[0.30]	
$\bar{R}^2$	5.3	5.1	2.2	2.3	4.4	9.0	17.8	41.7	
UK	0.505	0.637	0.755	0.872	0.981	1.041	1.125	1.162	1.250
	(0.11)	(0.10)	(0.10)	(0.09)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)
	[0.22]	[0.16]	[0.17]	[0.16]	[0.16]	[0.15]	[0.14]	[0.13]	[0.13]
$\bar{R}^2$	8.1	15.2	22.9	32.2	42.6	51.2	60.2	67.1	73.3
US	0.485	0.715	0.786	1.143	1.414	1.387	1.369	1.299	1.114
	(0.14)	(0.12)	(0.14)	(0.13)	(0.11)	(0.09)	(0.07)	(0.06)	(0.04)
	[0.31]	[0.26]	[0.27]	[0.24]	[0.16]	[0.15]	[0.14]	[0.10]	[0.06]
$\bar{R}^2$	4.5	13.1	13.4	29.3	47.3	57.4	68.8	77.9	82.5

Table 4

The Regression of Bond Excess Returns on the Forward Rates

This table reports the OLS regression coefficients from equation (3):

$$r^{\tau}_{t+12} = c_0 + c_1 f_t(0,1) + c_2 f_t(2,3) + c_3 f_t(8,9) + e_t, \quad \tau = 2, \cdots, 10.$$

The Newey-West adjusted $t-$ratios are reported in the brackets. The adjusted R^2 is in percentage. Regression results in GM and JP for the ten-year bond are omitted due to missing data.

		Bond Maturity τ								
Country	Coefficient	2	3	4	5	6	7	8	9	10
CA	c_0	-0.007	-0.015	-0.024	-0.036	-0.049	-0.062	-0.076	-0.090	-0.102
		[0.71]	[0.90]	[1.12]	[1.37]	[1.60]	[1.76]	[1.89]	[1.98]	[2.01]
	c_1	-0.112	-0.317	-0.560	-0.771	-0.898	-1.071	-1.343	-1.706	-2.136
		[0.54]	[0.82]	[1.03]	[1.14]	[1.13]	[1.18]	[1.32]	[1.52]	[1.73]
	c_2	0.163	0.463	0.802	1.044	1.053	0.997	0.991	1.108	1.422
		[0.89]	[1.39]	[1.76]	[1.85]	[1.59]	[1.31]	[1.13]	[1.11]	[1.26]
	c_3	0.097	0.154	0.216	0.354	0.639	1.022	1.458	1.866	2.135
		[0.68]	[0.59]	[0.60]	[0.78]	[1.18]	[1.62]	[2.02]	[2.29]	[2.36]
	$\bar{R}^2$	2.6	4.6	6.8	8.1	8.2	8.7	10.1	11.7	13.0
GM	c_0	-0.013	-0.032	-0.037	-0.033	-0.033	-0.047	-0.079	-0.129	
		[1.55]	[1.81]	[1.41]	[1.01]	[0.88]	[1.05]	[1.52]	[2.21]	
	c_1	-0.229	-0.455	-0.677	-0.836	-0.964	-1.084	-1.206	-1.369	
		[2.32]	[2.38]	[2.51]	[2.62]	[2.71]	[2.70]	[2.67]	[2.73]	
	c_2	0.452	0.966	1.263	1.261	1.202	1.172	1.085	0.875	
		[4.16]	[4.31]	[3.89]	[3.19]	[2.61]	[2.18]	[1.77]	[1.30]	
	c_3	0.038	0.086	0.131	0.280	0.523	0.897	1.544	2.576	
		[0.21]	[0.22]	[0.21]	[0.36]	[0.56]	[0.81]	[1.21]	[1.82]	
	$\bar{R}^2$	27.5	33.4	32.0	26.7	23.2	21.7	22.4	25.9	

Table 4 (continue)

Country	Coefficient	Bond Maturity τ								
		2	3	4	5	6	7	8	9	10
JP	c_0	0.003	0.014	0.021	0.025	0.029	0.028	0.021	0.011	
		[1.07]	[2.30]	[2.27]	[2.05]	[1.88]	[1.50]	[0.99]	[0.48]	
	c_1	-0.775	-1.903	-3.117	-4.150	-4.825	-5.220	-5.554	-5.807	
		[3.60]	[4.57]	[5.33]	[5.67]	[5.57]	[5.40]	[5.39]	[5.33]	
	c_2	1.339	3.162	4.920	6.291	7.109	7.444	7.593	7.536	
		[5.75]	[6.82]	[7.49]	[7.70]	[7.57]	[7.40]	[7.32]	[6.88]	
	c_3	-0.503	-1.192	-1.700	-2.023	-2.197	-2.106	-1.783	-1.303	
		[4.11]	[4.88]	[4.68]	[4.31]	[3.97]	[3.44]	[2.72]	[1.87]	
	$\bar{R}^2$	28.6	34.8	36.3	34.6	30.5	26.1	23.5	22.5	
UK	c_0	0.000	-0.009	-0.020	-0.030	-0.039	-0.045	-0.047	-0.047	-0.043
		[0.02]	[0.84]	[1.23]	[1.44]	[1.53]	[1.52]	[1.43]	[1.28]	[1.08]
	c_1	-0.439	-0.779	-1.058	-1.293	-1.485	-1.635	-1.747	-1.827	-1.883
		[2.77]	[2.82]	[2.82]	[2.76]	[2.65]	[2.50]	[2.34]	[2.18]	[2.04]
	c_2	0.974	1.980	2.779	3.327	3.623	3.689	3.551	3.236	2.772
		[3.69]	[4.23]	[4.39]	[4.32]	[4.06]	[3.68]	[3.22]	[2.69]	[2.14]
	c_3	-0.501	-1.022	-1.380	-1.534	-1.500	-1.309	-0.991	-0.572	-0.072
		[2.52]	[2.92]	[2.94]	[2.69]	[2.27]	[1.75]	[1.20]	[0.63]	[0.07]
	$\bar{R}^2$	13.7	16.6	17.6	17.2	15.9	14.0	12.1	10.3	8.9
US	c_0	-0.087	-0.160	-0.228	-0.299	-0.377	-0.456	-0.524	-0.571	-0.583
		[5.06]	[5.19]	[5.24]	[5.41]	[5.85]	[6.28]	[6.44]	[6.31]	[5.85]
	c_1	-0.632	-1.286	-1.922	-2.462	-2.855	-3.267	-3.775	-4.438	-5.293
		[3.36]	[3.98]	[4.43]	[4.60]	[4.56]	[4.58]	[4.73]	[5.24]	[6.40]
	c_2	1.200	2.310	3.322	4.201	4.941	5.698	6.457	7.156	7.707
		[5.50]	[6.14]	[6.52]	[6.64]	[6.66]	[6.80]	[6.93]	[7.14]	[7.52]
	c_3	0.605	1.139	1.669	2.263	2.946	3.617	4.213	4.708	5.033
		[4.54]	[5.09]	[5.47]	[5.86]	[6.44]	[6.92]	[7.10]	[7.16]	[7.18]
	$\bar{R}^2$	31.0	32.4	31.9	30.6	30.6	32.1	33.3	33.6	33.7

Table 5

The Schwartz Criterion, the In-, and out-of- Sample Prediction Errors for Non-nested Model Comparisons

This table reports the Schwarz Criterion BIC, the in-sample pricing error PE, and the out-of-sample pricing error PEO for all models that cannot be rejected in the first two steps of the estimation. The values of BIC are given in thousands while the values of PE and PEO are given in basis points. The value of the best model of each country is in bold face.

		Affine Term Structure Models				
Country	Criterion	$A_0(3)E$	$A_1(3)E$	$A_2(3)E$	$A_2(3)C$	$A_3(3)C$
CA	BIC	10.024	**10.093**	9.662	9.565	9.690
CA	PE	33.9	31.5	40.2	32.1	**31.1**
CA	PEO	31.7	30.1	37.7	31.0	**29.7**
GM	BIC	**9.215**	9.197	8.957	9.108	9.084
GM	PE	23.6	**22.3**	24.7	23.4	24.6
GM	PEO	20.0	**18.7**	19.3	19.1	19.2
JP	BIC	9.331	**9.375**	9.336	9.333	9.272
JP	PE	22.2	**21.1**	22.1	22.1	22.1
JP	PEO	22.0	**21.3**	21.9	21.8	22.2
UK	BIC	10.164	**10.209**	10.096	10.064	10.151
UK	PE	30.7	**29.3**	31.0	31.5	30.3
UK	PEO	28.7	**27.9**	29.3	29.4	28.6
US	BIC	9.642	**9.658**	8.376	8.421	8.750
US	PE	32.4	**29.1**	34.9	34.8	35.4
US	PEO	32.8	**29.6**	35.4	34.5	34.2

Table 6

Parameter Estimates for the Best of Model of Each Country

This table reports the parameter estimates for $A_1(3)E$, the best model of each country. The parameters are estimated using Kalman filter/QML approach. The asymptotic standard error is given in the parenthesis. For the model of $A_1(3)E$, $\alpha_1 = \beta_{12} = \beta_{13} = 0$, $\alpha_2 = \alpha_3 = \beta_{11} = 1$, and the first row of λ_2 is all zero. For other parameter constraints, please refer to Appendix A. The parameter with * denotes significance at 5%.

	Model: $A_1(3)E$								
Constants	CA			GM			JP		
δ_0	0.0066	(0.0074)		0.0432^*	(0.0120)		0.1203^*	(0.0310)	
σ_e	0.0008^*	(0.00002)		0.0022^*	(0.00005)		0.0021^*	(0.00004)	
Parameters	Index i			Index i			Index i		
	1	2	3	1	2	3	1	2	3
δ_i	0.0042^*	0.0000	0.0073^*	0.0004	$2.11E-9$	0.0117^*	-0.0005^*	0.0001	0.0002
	(0.0006)	(0.0002)	(0.0014)	(0.0003)	(0.0005)	(0.0050)	(0.0002)	(0.0009)	(0.0045)
θ_i	21.1482^*	0	0	39.8416^*	0	00	93.9499	0	0
	(4.0049)			(20.1873)			(95.8792)		
K_{1i}	0.0478^*	0	0	0.1506^*	0	0	0.0356^*	0	0
	(0.0051)			(0.0119)			(0.0063)		
K_{2i}	-0.0000	0.5161^*	3.2337	-1.6912	0.0594	4.6844	-0.2429	0.0059	6.2579
	(0.0567)	(0.2199)	(2.7879)	(6.8021)	(0.7840)	(21.3211)	(11.9537)	(21.1830)	(56.2555)
K_{3i}	-0.0000	0.0060	0.2506	-0.0040	-0.1138	1.5207	-0.4206	-0.0962	1.5579
	(0.0484)	(0.0284)	(0.2486)	(0.0997)	(0.4636)	(0.7876)	(4.0006)	(5.4523)	(21.1966)
β_{1i}	1	10.5953	0.1047	1	9.5914	0.0207	1	10.0959	1.3734
		(12.8186)	(0.0709)		(77.9973)	(0.0432)		(62.9907)	(28.5664)
λ_{1i}	-0.0227^*	-6.1130	-2.8957^*	-0.0350^*	-6.0213	-1.2088	0.0058	-6.3325	0.1713
	(0.0051)	(4.7322)	(1.2893)	(0.0114)	(26.0404)	(1.5701)	(0.0039)	(39.0317)	(23.7682)
$\lambda_{2,2i}$	64.8963	-0.5323^*	-0.9770	57.7593	0.0773	5.2376	64.0284	-0.4179	2.6956
	(100.2326)	(0.2487)	(1.9672)	675.6679	(0.4379)	(22.7773)	(62.0354)	(8.2873)	(23.5923)
$\lambda_{2,3i}$	0.4187	-0.0794	-0.0413	0.0379	0.0130	-0.5497	-0.1506	-0.0718	0.0438
	(0.3296)	(0.0528)	(0.2758)	(0.0911)	(0.0607)	(0.4120)	(35.1388)	(1.5160)	(8.2992)
LL		10158.49			9262.16			9440.1633	

Table 6 (continued)

	Model: $A_1(3)E$					
Constants	UK			US		
δ_0	0.1232^*	(0.0143)		0.1466^*	(0.0218)	
σ_e	0.0009^*	(0.00002)		0.0019^*	(0.00004)	
Parameters	Index i			Index i		
	1	2	3	1	2	3
δ_i	-0.0005	0.0012	0.0035	-0.0039^*	0.0002	0.0000
	(0.0004)	(0.0011)	(0.0030)	(0.0012)	(0.0011)	(0.0005)
θ_i	29.5491^*	0	0	14.1498	0	0
	(5.8932)			(9.5182)		
K_{1i}	0.0180^*	0	0	0.4254^*	0	0
	(0.0069)			(0.0509)		
K_{2i}	-0.5616	0.1006	2.6879	-1.0532	0.0267	1.6321^*
	(0.5035)	(0.1921)	(3.3785)	(4.5621)	(0.0365)	(0.2026)
K_{3i}	-0.0018	-0.0402	0.8187^*	-1.0557	0.0073	0.4453^*
	(0.1065)	(0.0647)	(0.2661)	(4.4834)	(0.0479)	(0.1871)
β_{1i}	1	10.9528	0.2079	1	14.2853	1.0597
		(20.6811)	(0.1963)		(121.8975)	(9.4008)
λ_{1i}	0.0482^*	-5.8385	-0.4892	-0.1094^*	-4.2300	-3.8555
	(0.0130)	(8.5791)	(1.9740)	(0.0476)	(20.2047)	(19.4405)
$\lambda_{2,2i}$	64.4729	-0.0408	2.7388	59.5679	0.3477^*	1.9063
	(129.9537)	(0.2841)	(3.1974)	(754.1299)	(0.1639)	(1.6473)
$\lambda_{2,3i}$	0.2383	0.0251	0.1729	4.8784	-0.1316	-0.8961^*
	(0.5087)	(0.0440)	(0.3403)	(59.5906)	(0.1250)	(0.1429)
LL		10274.12			9701.28	

Table 7

Summary Statistics of Prediction Errors from Model $A_1(3)E$

This table reports the summary statistics of the prediction errors, defined as the difference between the actual and the predicted yields at each month across all maturities, from the best model $A_1(3)E$. The mean and the standard deviation are for the absolute prediction errors while the maximum and the minimum are for the raw (before taking absolute value) prediction errors. All numbers are in basis points.

	In Sample							Out of Sample			
Yield Maturity	0.5	1	2	3	5	7	10	4	6	8	9
Canada											
mean(ABS)	40.2	37.1	33.0	30.7	28.6	26.9	24.4	29.6	27.6	26.1	25.2
std(ABS)	37.6	34.8	30.1	26.5	22.6	20.0	18.4	24.0	21.2	19.6	19.0
Max	250.5	221.2	171.0	148.5	120.0	90.7	94.8	134.1	104.0	82.7	82.7
Min	-140.3	-123.7	-105.2	-97.6	-99.5	-106.0	-103.4	-92.7	-98.4	-111.3	-112.3
Germany											
mean(ABS)	34.3	23.4	23.6	20.4	18.7	17.0	18.1	19.3	17.7	17.2	20.1
std(ABS)	35.7	26.6	20.6	15.2	14.1	13.9	14.0	14.1	13.9	13.3	16.2
Max	352.7	264.5	146.5	73.5	102.8	104.6	105.1	93.2	105.0	104.8	104.7
Min	-121.0	-107.6	-106.6	-87.5	-62.2	-54.7	-71.8	-70.7	-58.5	-49.7	-92.0
Japan											
mean(ABS)	23.6	18.4	21.4	20.7	22.2	21.8	19.3	21.3	22.5	19.7	21.9
std(ABS)	21.3	18.9	20.1	19.2	18.7	17.6	16.9	18.4	18.6	16.8	18.3
Max	108.4	107.3	116.8	100.4	90.5	95.4	91.6	87.2	97.2	94.8	88.9
Min	-133.2	-102.5	-96.8	-104.4	-92.2	-81.1	-83.4	-100.3	-84.4	-82.3	-84.0
UK											
mean(ABS)	29.0	32.0	31.5	30.3	29.0	28.0	25.8	29.6	28.6	27.3	26.5
std(ABS)	33.6	32.5	29.9	28.3	26.0	24.3	22.1	27.0	25.1	23.5	22.8
Max	177.1	170.0	201.9	205.4	186.5	167.2	146.0	197.5	176.0	159.9	153.1
Min	-170.9	-187.2	-156.8	-133.0	-137.1	-145.2	-131.8	-127.3	-143.5	-142.9	-138.0
US											
mean(ABS)	35.4	32.1	29.5	28.4	28.1	28.5	22.1	28.1	27.9	28.6	26.1
std(ABS)	27.9	25.4	21.9	20.9	21.6	20.5	16.7	21.6	20.6	21.1	19.2
Max	104.8	104.5	99.8	85.4	86.4	97.5	54.9	81.2	87.2	108.1	93.7
Min	-140.6	-130.7	-119.3	-109.8	-97.9	-102.2	-91.8	-102.3	-100.6	-102.7	-100.5

Table 8

Correlations between Estimated Three Factors and the Level, Slope and Curvature Components

This table reports the correlation between the extracted three factors from the Kalman filter estimation and the factors of "level," "slope," and "curvature" from the principal component analysis. "SV" stands for the state variable with which the correlation is calculated.

	Level		Slope		Curvature	
	SV	Correlation	SV	Correlation	SV	Correlation
CA	X_1	0.9357	X_3	0.8135	X_2	0.5847
GM	X_1	0.8741	X_3	0.7700	X_2	0.5072
JP	X_1	0.9619	X_2	0.4561	X_3	0.7908
UK	X_2	0.9894	X_3	0.7994	X_1	0.4148
US	X_1	0.9013	X_3	0.5310	X_2	0.6893

Table A1

Parameter Restrictions for Three-Factor Affine Models

This table contains the parameter constraints to ensure admissibility and identification for all three-factor models considered in the paper.

Model	$A_0(3)E$			Model: $A_1(3)E$			$A_2(3)E$		
	Index i			Index i			Index i		
	1	2	3	1	2	3	1	2	3
δ_1	$\delta_1(1) \geq 0$	$\delta_1(2) \geq 0$	$\delta_1(3) \geq 0$	$\delta_1(1)$	$\delta_1(2) \geq 0$	$\delta_1(3) \geq 0$	$\delta_1(1)$	$\delta_1(2)$	$\delta_1(3) \geq 0$
α	1	1	1	0	1	1	0	0	1
β_{1i}	0	0	0	1	$\beta_{12} \geq 0$	$\beta_{13} \geq 0$	1	0	$\beta_{13} \geq 0$
β_{2i}	0	0	0	0	0	0	0	1	$\beta_{23} \geq 0$
β_{3i}	0	0	0	0	0	0	0	0	0
θ	0	0	0	$\theta_1 \geq 0$	0	0	$\theta_1 \geq 0$	$\theta_2 \geq 0$	0
K_{1i}	$K_{11} > 0$	0	0	$K_{11} > 0$	0	0	$K_{11} > 0$	$K_{12} \leq 0$	0
K_{2i}	K_{21}	$K_{22} > 0$	0	$K_{21} \leq 0$	K_{22}	K_{23}	$K_{21} \leq 0$	$K_{22} > 0$	0
K_{3i}	K_{31}	K_{32}	$K_{33} > 0$	$K_{31} \leq 0$	K_{32}	K_{33}	$K_{31} \leq 0$	$K_{32} \leq 0$	K_{33}
		Eigenvalues of $K > 0$			Eigenvalues of $K > 0$			Eigenvalues of $K > 0$	
					$K_1\theta > 0$			$K_1\theta > 0$ and $K_2\theta > 0$	
I^-_{1i}	1	0	0	0	0	0	0	0	0
I^-_{2i}	0	1	0	0	1	0	0	0	0
I^-_{3i}	0	0	1	0	0	1	0	0	1
$\lambda_{1,i}$	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$
$\lambda_{2,1i}$	$\lambda_{2,11}$	$\lambda_{2,12}$	$\lambda_{2,13}$	0	0	0	0	0	0
$\lambda_{2,2i}$	$\lambda_{2,21}$	$\lambda_{2,22}$	$\lambda_{2,23}$	$\lambda_{2,21}$	$\lambda_{2,22}$	$\lambda_{2,23}$	0	0	0
$\lambda_{2,3i}$	$\lambda_{2,31}$	$\lambda_{2,32}$	$\lambda_{2,33}$	$\lambda_{2,31}$	$\lambda_{2,32}$	$\lambda_{2,33}$	$\lambda_{2,31}$	$\lambda_{2,32}$	$\lambda_{2,33}$

Table A1 (continued)

Model	$A_2(3)C$			Model: $A_3(3)C$		
	Index i			Index i		
	1	2	3	1	2	3
δ_1	$\delta_1(1)$	$\delta_1(2)$	$\delta_1(3) \geq 0$	$\delta_1(1)$	$\delta_1(2)$	$\delta_1(3)$
α	0	0	1	0	0	0
β_{1i}	1	0	$\beta_{13} \geq 0$	1	0	0
β_{2i}	0	1	$\beta_{23} \geq 0$	0	1	0
β_{3i}	0	0	0	0	0	1
θ	$\theta_1 \geq 0$	$\theta_2 \geq 0$	0	$\theta_1 \geq 0$	$\theta_2 \geq 0$	$\theta_3 \geq 0$
K_{1i}	$K_{11} > 0$	$K_{12} \leq 0$	0	$K_{11} > 0$	$K_{12} \leq 0$	$K_{13} \leq 0$
K_{2i}	$K_{21} \leq 0$	$K_{22} > 0$	0	$K_{21} \leq 0$	$K_{22} > 0$	$K_{23} \leq 0$
K_{3i}	$K_{31} \leq 0$	$K_{32} \leq 0$	K_{33}	$K_{31} \leq 0$	$K_{32} \leq 0$	$K_{33} > 0$
	Eigenvalues of $K > 0$			Eigenvalues of $K > 0$		
	$K_1\theta > 0$ and $K_2\theta > 0$			$K_i\theta > 0$ ($i = 1, \cdots, 3$)		
I^-_{1i}	0	0	0	0	0	0
I^-_{2i}	0	0	0	0	0	0
I^-_{3i}	0	0	0	0	0	0
$\lambda_{1,i}$	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$
$\lambda_{2,1i}$	0	0	0	0	0	0
$\lambda_{2,2i}$	0	0	0	0	0	0
$\lambda_{2,3i}$	0	0	0	0	0	0

Figure 1
Weights of the First Three Principal Components on the Eleven Bonds

The figure plots the weights of the first three principal components on the eleven zero coupon bond yields with maturities of six months and one to ten years .

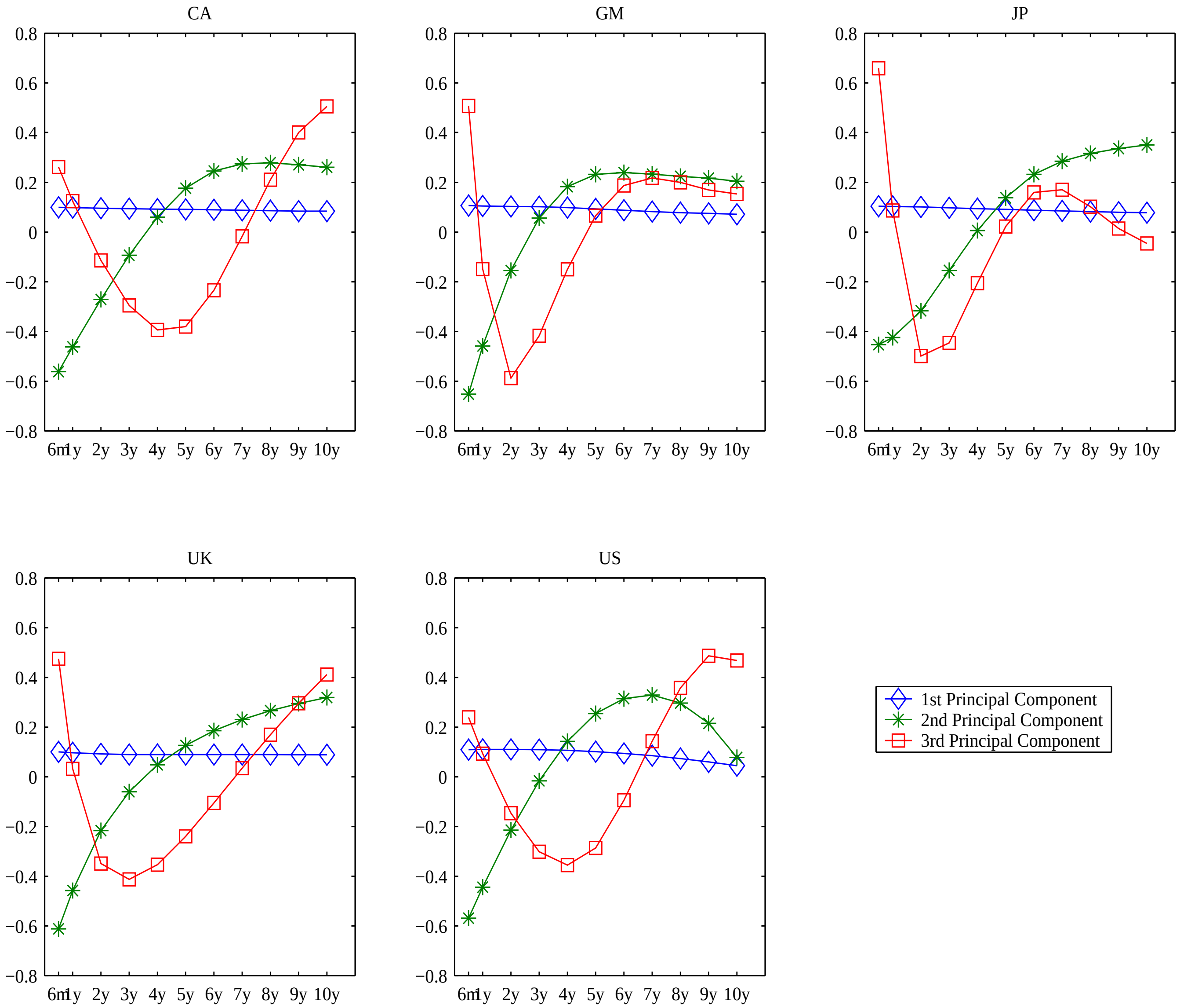


Figure 2

Model Implied Instantaneous Interest Rates and the One-month Treasury Bill Rates

The figure plots the estimated short rate from the best model in each country against the one-month Treasury bill rate, which is not used in the estimation. The sample is from January 1983 to December 2002. The solid line stands for the implied short rate while the dashed line is for the one-month Treasury bill rate.

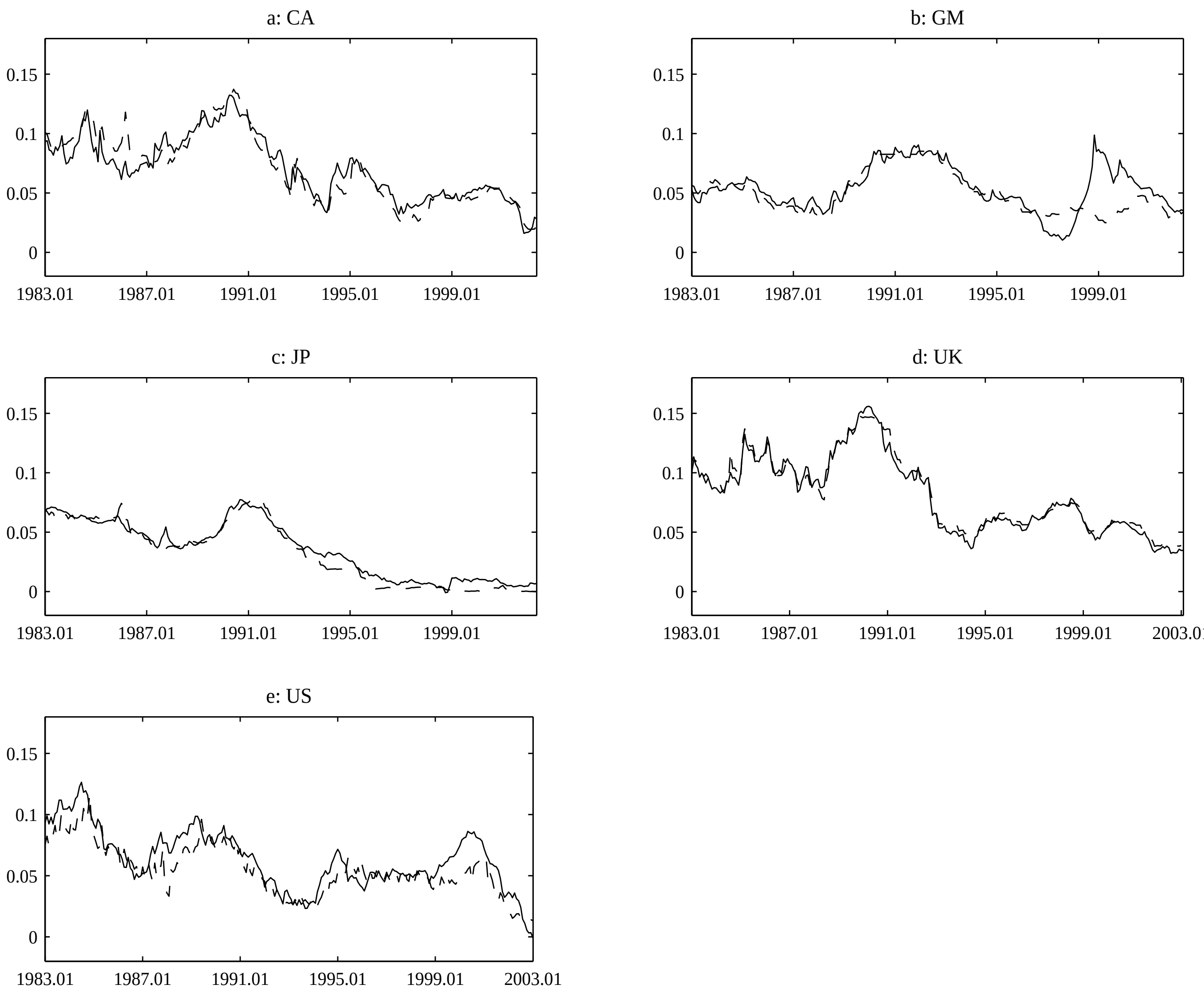


Figure 3
Model Implied Price of Risk

The figure plots the estimated market price of risk derived from the MLE estimated parameters and state variables from the best model of each country,

$$\hat{\Lambda}_t = \sqrt{\hat{S}\,(t)}\left(\hat{\lambda}_1 + \hat{S}(t)^{-1} I \hat{\lambda}_2 \hat{X}\,(t)\right).$$

The i^{th} element, $\hat{\Lambda}(i)$ $(i = 1, 2, 3)$, of the 3×1 vector $\hat{\Lambda}(t)$ is the price of risk associated with the i^{th} state variable at time t. The sample is from January 1983 to December 2002.

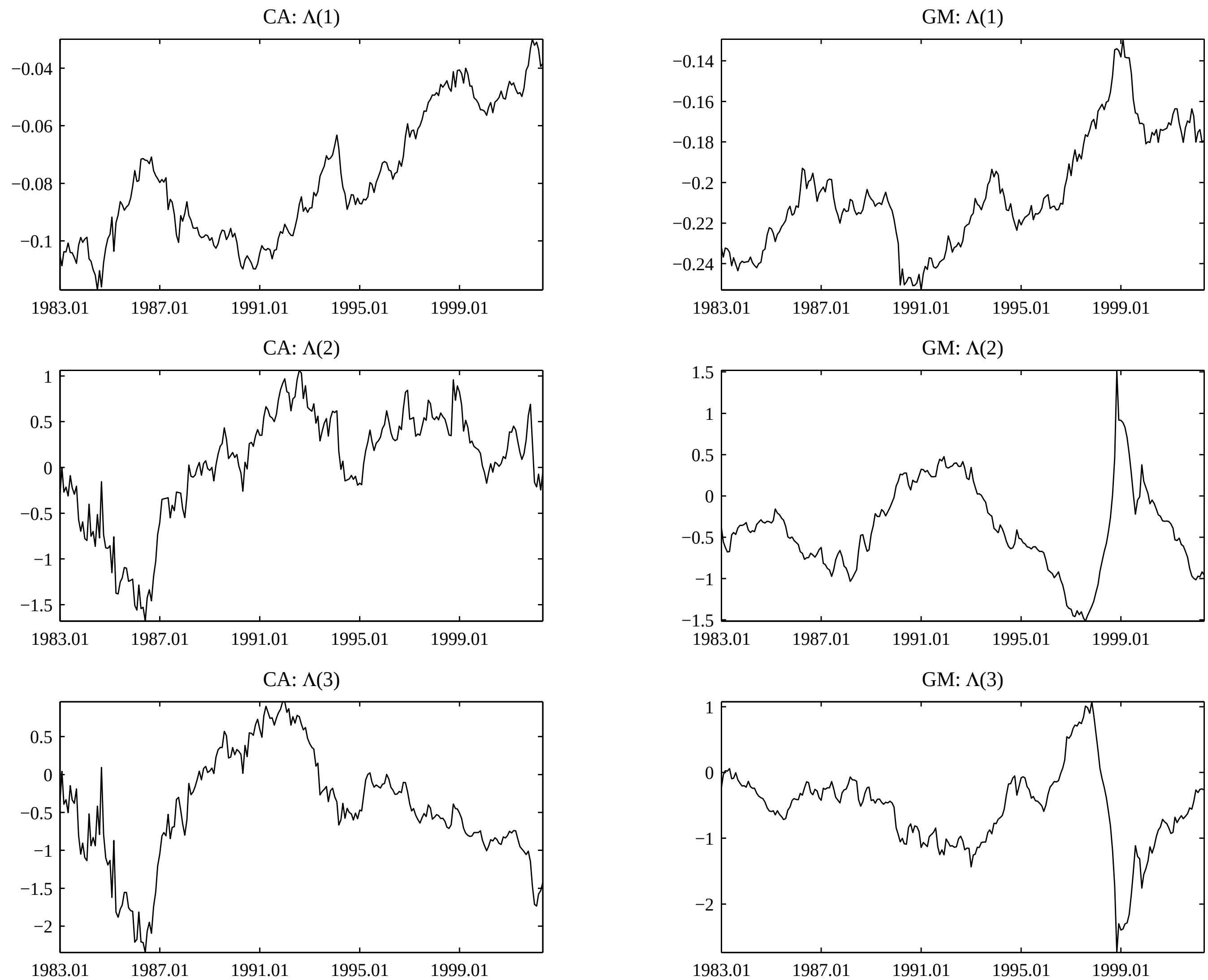


Figure 3 (continued)

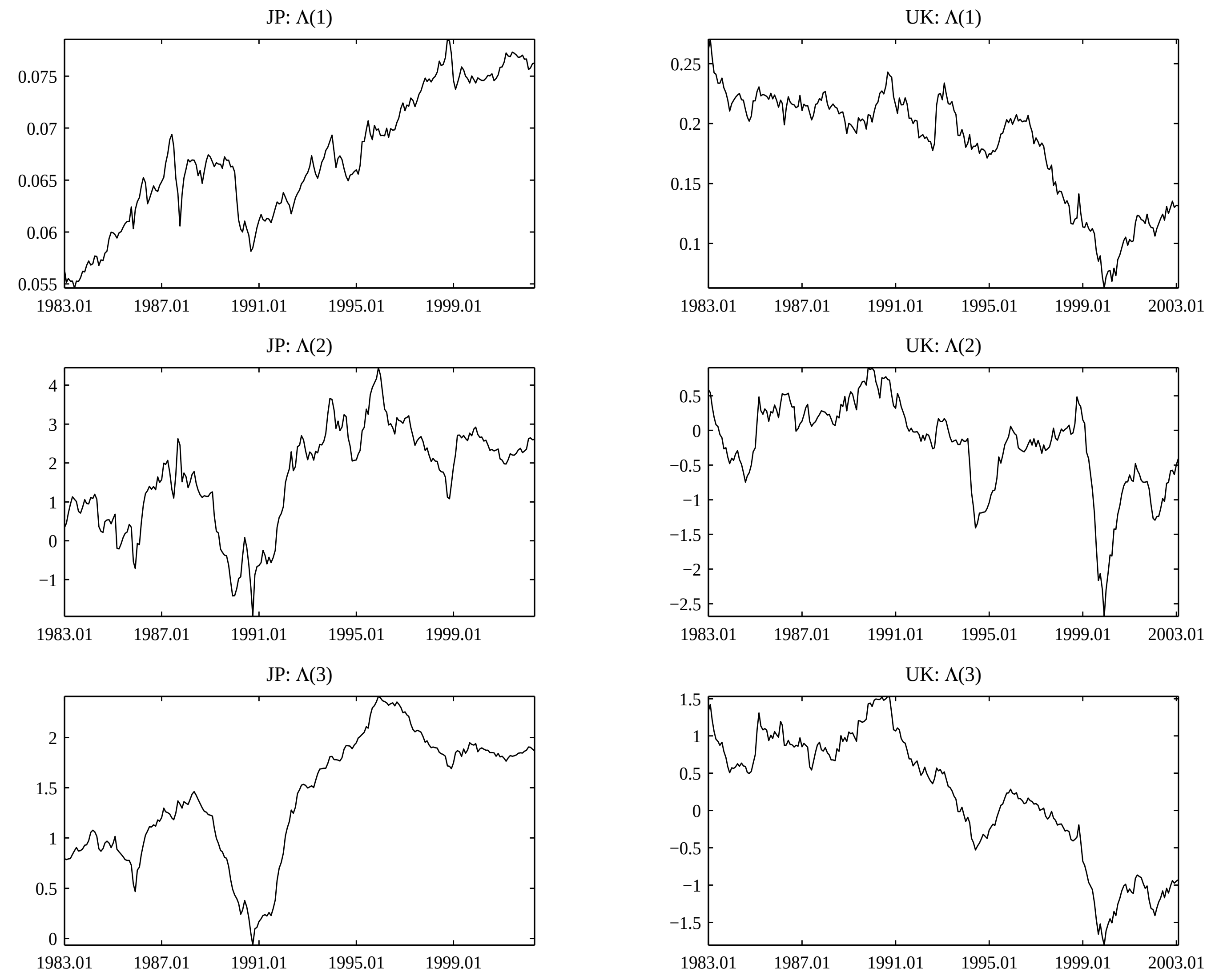


Figure 3 (continued)

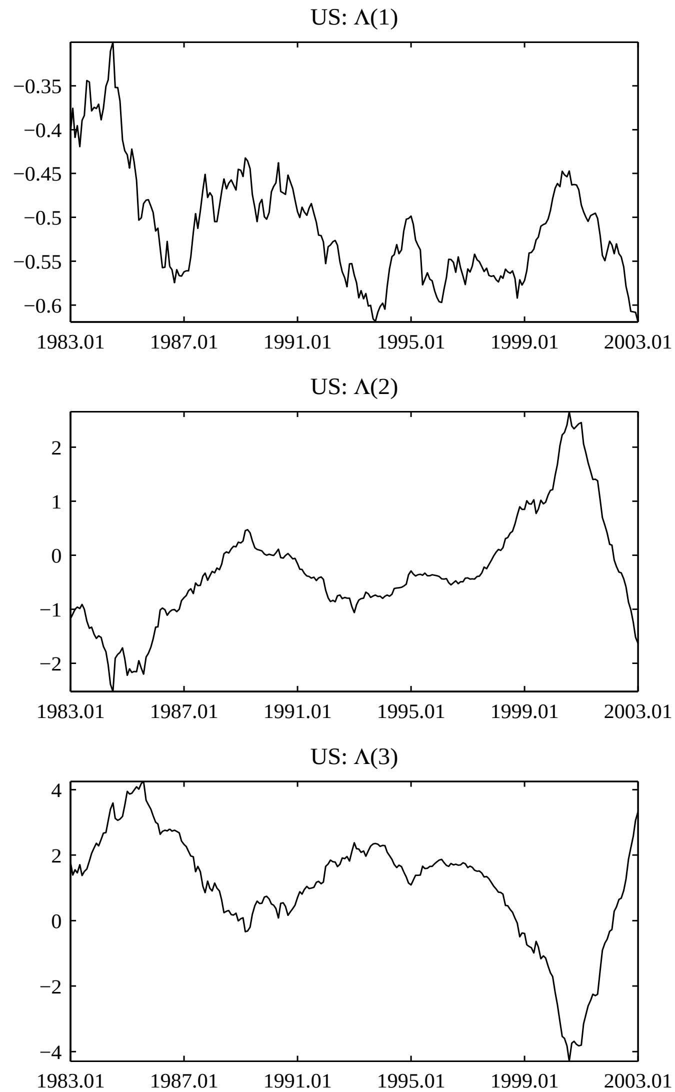


Figure 4

Model Estimated Total Market Price of Risk

The figure plots the estimated total market price of risk, derived from the Kalman filter/QML estimated parameters and state variables from the best model of each country, $\hat{\eta} \equiv \sqrt{\hat{\Lambda}\hat{\Lambda}'}$. The sample is from January 1983 to December 2002.

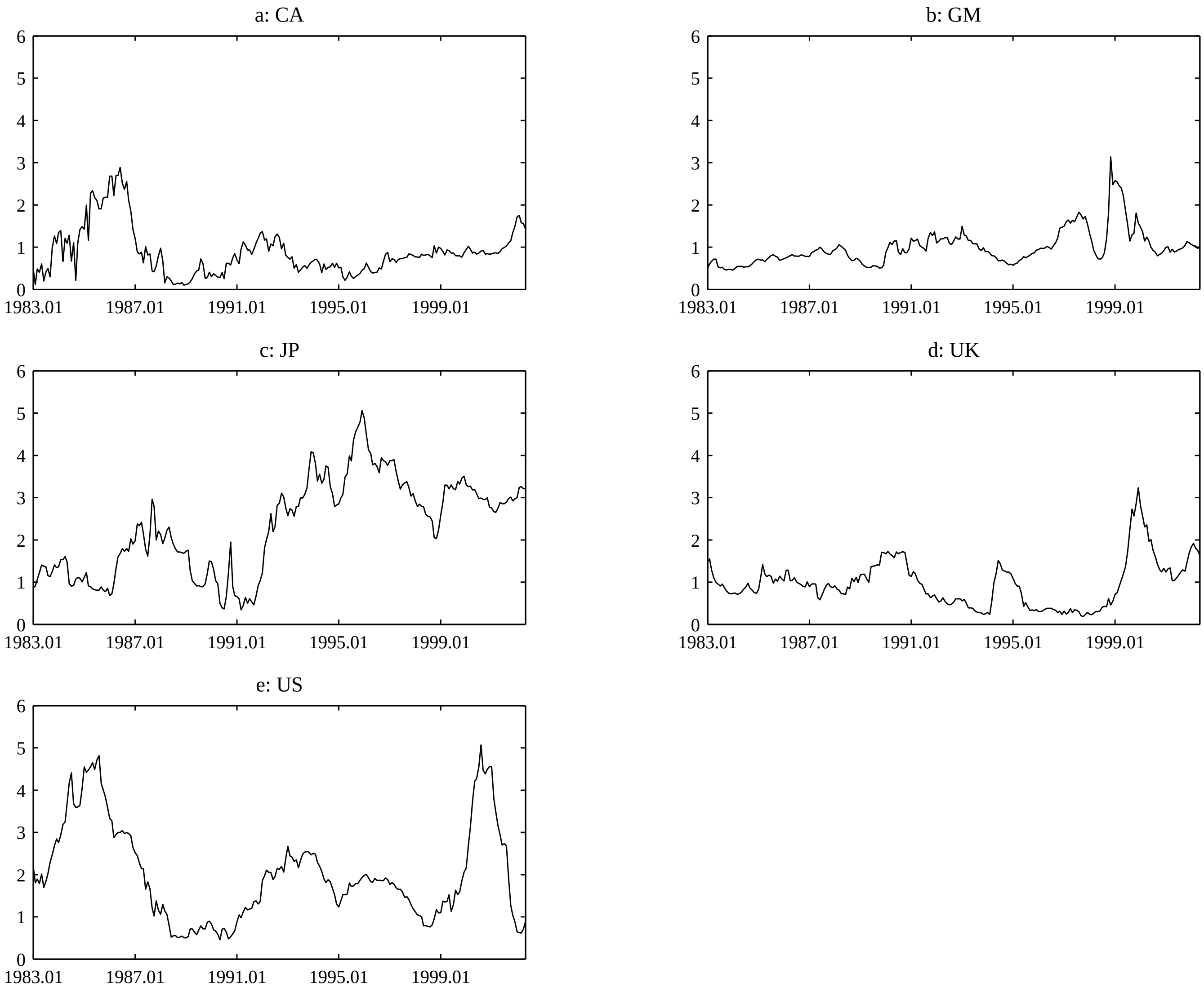


Figure 5

Simulated Slope Estimates of Regressing the Long-Term Yield Change on the Yield Spread

The figure plots the average slope estimates, $\hat{a}_1$, from the regression,

$$y_{t+12}^{\tau-1} - y_t^{\tau} = a_0 + a_1 \frac{y_t^{\tau} - y_t^{1}}{\tau - 1} + \epsilon, \quad \tau = 2, \cdots, 10,$$

against the bond maturity τ using 2000 sets of simulated yields. The two standard error bounds around $\hat{a}_1$ and the sample estimates from Table 2 are also plotted. The simulated yields contain 233 monthly observations for CA, GM, and JP, and 242 and 241 observations for UK and US, the same as the sample size of the actual data in the respective five countries.

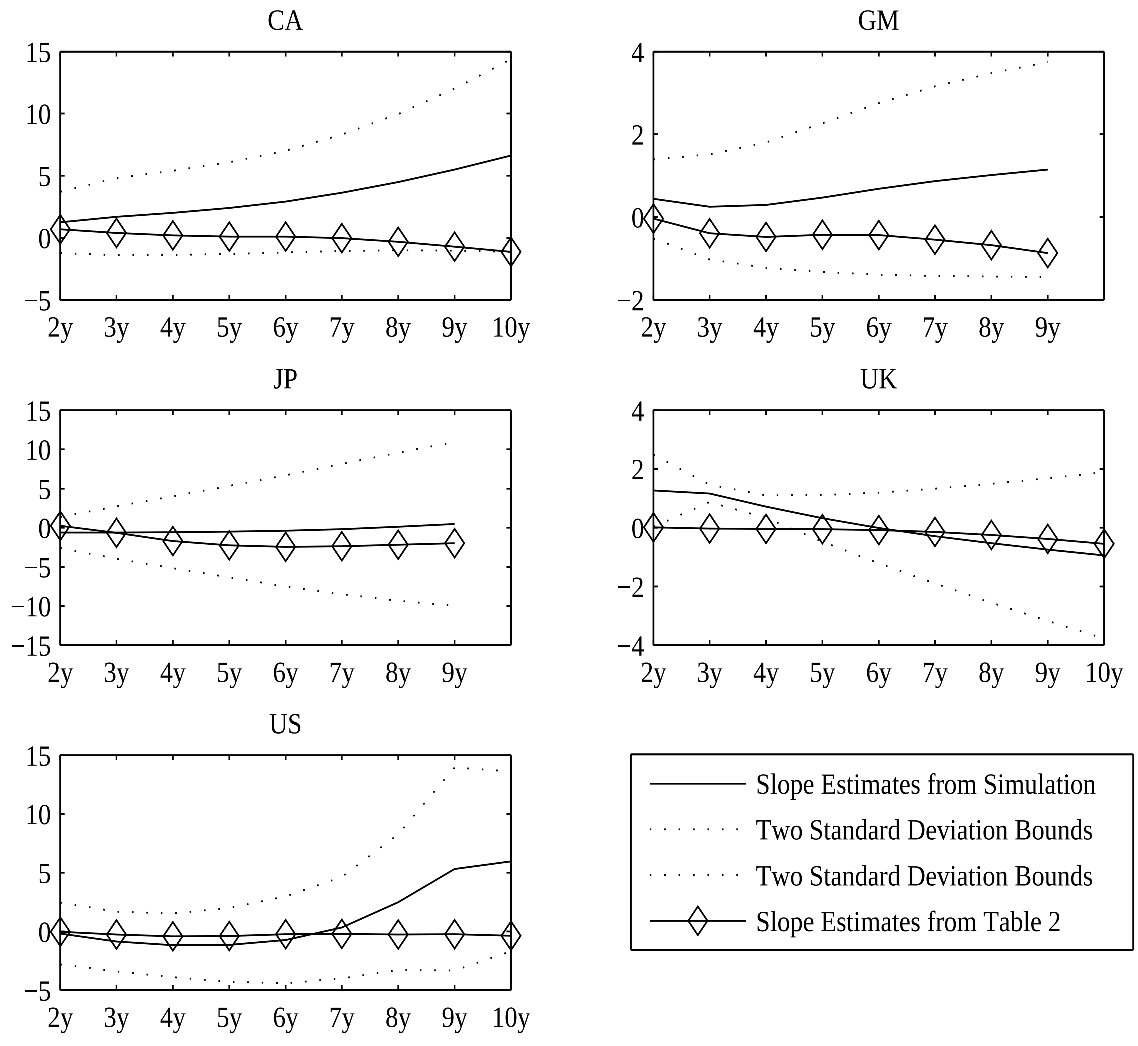


Figure 6

Simulated Slope Estimates of Regressing the Short Rate Change on the Yield Spread

The figure plots the average slope estimates $\hat{b}_1$ from the regression

$$\frac{1}{\tau}\sum_{k=1}^{\tau-1}\left[y_{t+12k}^{1}-y_t^1\right]=b_0+b_1\left[y_t^{\tau}-y_t^1\right]+\epsilon,\quad \tau=2,\cdots,10,$$

against the bond maturity τ using 2000 sets of simulated yields. The two standard error bounds around $\hat{b}_1$ and the sample estimates from Table 3 are also plotted. The simulated yields contain 233 monthly observations for CA, GM, and JP, and 242 and 241 observations for UK and US, the same as the sample size of the actual data in the respective five countries.

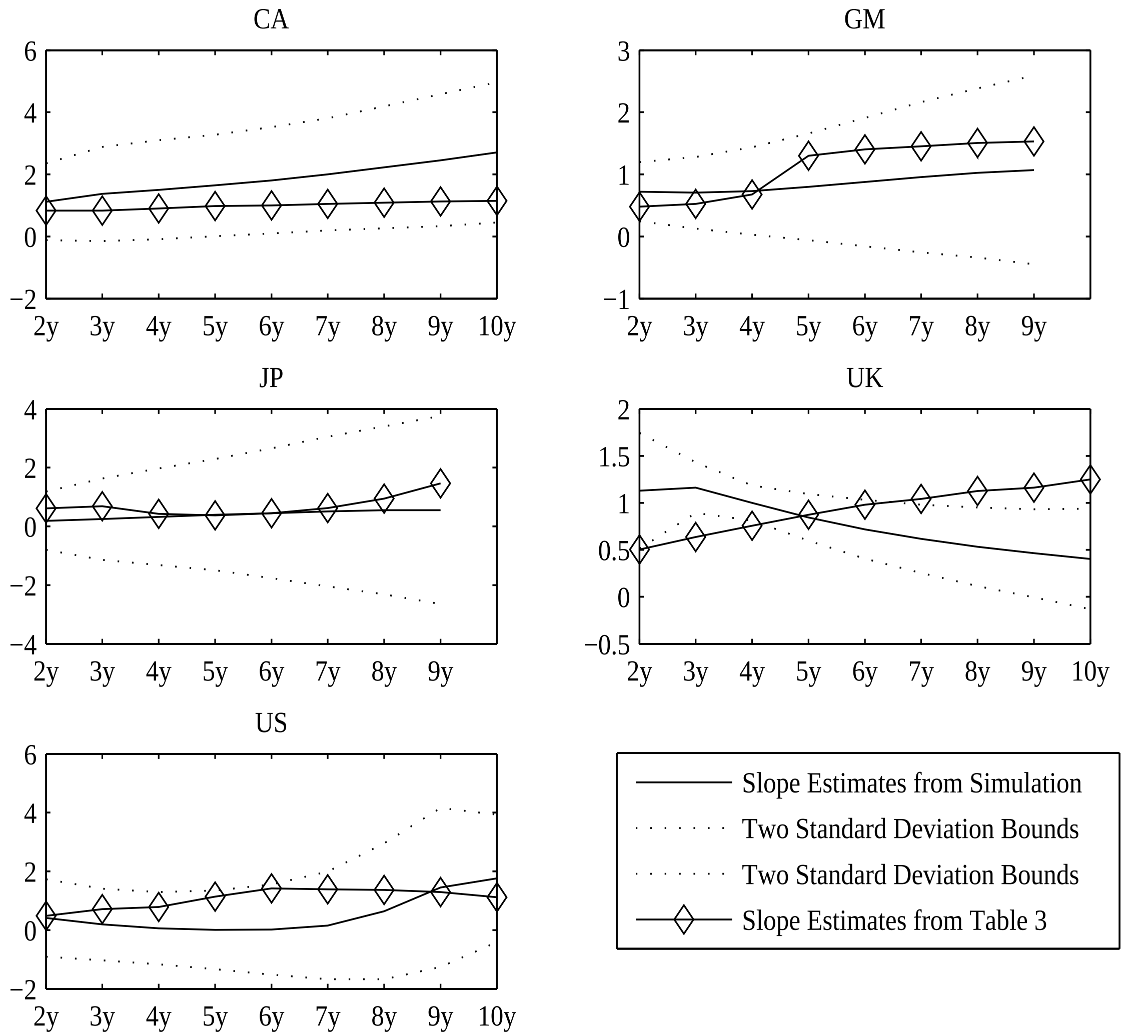


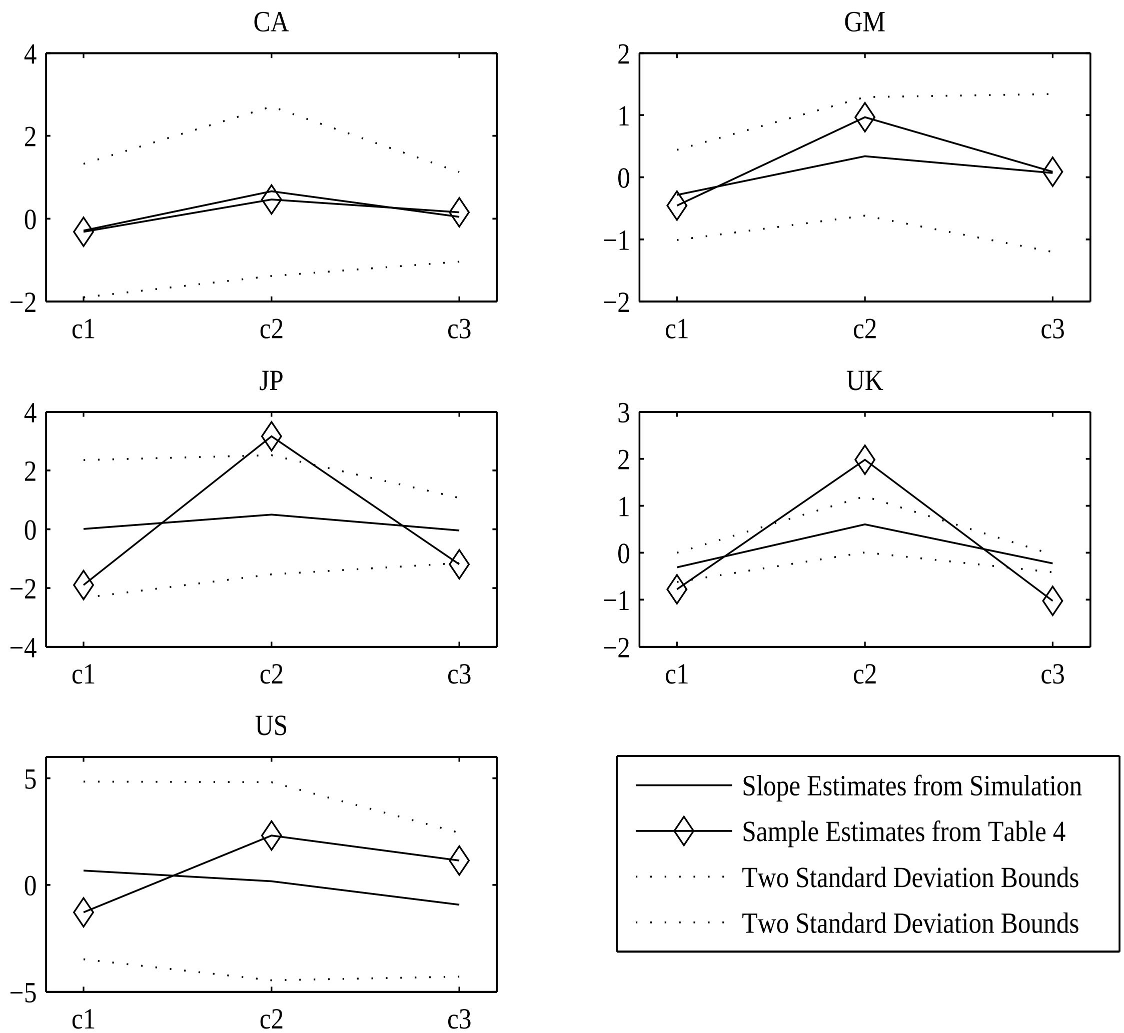
Figure 7

Simulated Slope Estimates of Regressing the Excess Bond Returns on the Forward Rates

The figure plots the average slope estimates, $\hat{c}_1$, $\hat{c}_2$ and $\hat{c}_3$, from the regression

$$r^{\tau}_{t+12} = c_0 + c_1 f_t(0,1) + c_2 f_t(2,3) + c_3 f_t(8,9) + e_t, \quad \tau = 3.$$

for bond yields with maturity $\tau = 3$ using 2000 sets of simulated yields. The two standard error bounds around $\hat{c}_1$, $\hat{c}_2$ and $\hat{c}_3$ and the sample estimates with the corresponding maturity ($\tau = 3$) from Table 4 are also plotted. The simulated yields contain 233 monthly observations for CA, GM, and JP, and 242 and 241 observations for UK and US, the same as the sample size of the actual data in the respective five countries.



References

Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant, 2002, Quadratic term structure models: theory and evidence, *Review of Financial Studies* 15, 243-288.

Andrews, D. W. K., and W. Ploberger, 1994, Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica* 62, 1383-1414.

Bekaert, Geert, and Campbell Harvey, 1995, Time-Varying World Market Integration, *Journal of Finance* 50, 403-444.

Bekaert, Geert, and Robert J. Hodrick, and David A. Marshall, 1997, On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates, *Journal of Financial Economics* 44, 309-348.

Bekaert, Geert, Min Wei, and Yuhang Xing, 2002, Uncovered interest rate parity and the term structure, working paper, Columbia University.

Bansal, Ravi, George Tauchen, and Hao Zhuo, 2003, Regime-shifts, risk premiums in the term structure, and the business cycle, working paper, Duke University.

Brandt, Michael W., and David A. Chapman, 2002, Comparing multifactor models of the term structure, working paper, Duke University.

Brandt, Michael W., and Pedro 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, Michael J., and Yihong Xia, 2003, International Capital Markets and Foreign Exchange Risk, working paper, University of Pennsylvania.

Campbell, John Y., and Robert J. Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economics Studies* 58, 495-514.

Cochrane, John H., and Monika Piazzesi, 2002, Bond risk premia, working paper, UCLA.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *The Journal of Finance* 55, 1943-1978.

Dai, Qiang, and Kenneth J. Singleton, 2002, Expectation puzzles, time-varying risk premia, and dynamic models of the term structure, *Journal of Financial Economics* 63, 415-441.

Dai, Qiang, and Kenneth J. Singleton, 2003, Term structure dynamics in theory and reality, *Review of Financial Studies* 16, 631-678.

Dai, Qiang, Kenneth J. Singleton, and Wei Yang, 2003, Regime shifts in a dynamic term structure model of U.S. Treasury bond yields, 2003, working paper, New York University.

de Jong, Frank, 2000, Time-series and cross-section information in affine term structure models, *Journal of Business and Economics Statistics* 18, 300-314.

Driessen, Joost, Bertrand Melenberg, and Theo Nijman, 2000, Common factors in international bond returns, working paper, Tilburg University.

Duffee, Gregory R., 2002, Term premia and interest rate forecasts in affine models, *The Journal of Finance* 57, 405-443.

Duffee, Gregory R., and Richard H. Stanton, 2002, Estimation of dynamic term structure models, working paper, UC Berkeley.

Duffie, Darrel, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379-406.

Fama, Eugene F., 1984, The information in the term structure of interest rates, *Journal of Financial Economics* 13, 509-528.

Fama, Eugene F., and Robert T. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680-692.

Hardouvelis, Gikas A., 1994, The term structure spread and future changes in long and short rates in the G7 countries: is there a puzzle? *Journal of Monetary Economics* 33, 255-284.

Hansen, Bruce E., 1996, Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* 64, 413-430.

Jorion, Philippe, and Frederic Mishkin, 1991, A multicountry comparison of term-structure forecasts at long horizons, *Journal of Financial Economics* 29, 59-80.

Litterman, Robert, and Jose Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* 1, 54-61.

McCulloch, John H., and H. Kwon, 1993, U.S. term structure data: 1947-1991, working paper, Ohio State University.

Piazzesi, Monika, 2003, Affine term structure models, working paper, UCLA.

Stambaugh, Robert F., 1988, The information in the forward rates: implications for models of the term structure, *Journal of Financial Economics* 21, 41-69.

The Rodney L. White Center for Financial Research
The Wharton School
University of Pennsylvania
3254 Steinberg Hall-Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104-6367

(215) 898-7616
(215) 573-8084 Fax
http://finance.wharton.upenn.edu/~rlwctr

The Rodney L. White Center for Financial Research is one of the oldest financial research centers in the country. It was founded in 1969 through a grant from Oppenheimer & Company in honor of its late partner, Rodney L. White. The Center receives support from its endowment and from annual contributions from its Members.

The Center sponsors a wide range of financial research. It publishes a working paper series and a reprint series. It holds an annual seminar, which for the last several years has focused on household financial decision making.

The Members of the Center gain the opportunity to participate in innovative research to break new ground in the field of finance. Through their membership, they also gain access to the Wharton School's faculty and enjoy other special benefits.

Members of the Center
2004 – 2005

Directing Members

Aronson + Johnson + Ortiz, LP
Geewax, Terker & Company
Goldman, Sachs & Co.
Hirtle, Callaghan & Co.
Morgan Stanley
Merrill Lynch
The Nasdaq Educational Foundation
The New York Stock Exchange, Inc.
Twin Capital

Founding Members

Ford Motor Company Fund
Merrill Lynch, Pierce, Fenner & Smith, Inc.
Oppenheimer & Company
Philadelphia National Bank
Salomon Brothers
Weiss, Peck and Greer