

On the predictability of common risk factors in the US and UK interest rate swap markets: Evidence from non-linear and linear models

Ilias Lekkos
Research Department
Eurobank Ergasias, Greece

Costas Milas
Department of Economics
Keele University, UK

and

Theodore Panagiotidis*
Department of Economics
Loughborough University, UK

October 2005

Abstract

This paper explores the ability of common risk factors to predict the dynamics of US and UK interest rate swap spreads within a linear and a non-linear framework. We reject linearity for the US and UK swap spreads in favour of a regime-switching smooth transition vector autoregressive (STVAR) model, where the switching between regimes is controlled by the slope of the US term structure of interest rates. The first regime is characterised by a "flat" term structure of US interest rates, while the alternative is characterised by an "upward" sloping US term structure. We compare the ability of the STVAR model to predict swap spreads with that of a non-linear nearest-neighbours model as well as that of linear AR and VAR models. We find some evidence that the nearest-neighbours and STVAR models predict better than the linear AR and VAR models. However, the evidence is not overwhelming as it is sensitive to swap spread maturity. We also find that within the non-linear class of models, the nearest-neighbours model predicts better than the STVAR model US swap spreads in periods of increasing risk conditions and UK swap spreads in periods of decreasing risk conditions.

Keywords: Interest rate swap spreads, term structure of interest rates, regime switching, smooth transition models, nearest-neighbours, forecasting.

JEL classification: C51, C52, C53, E43.

* Corresponding author: Dr Theodore Panagiotidis, Department of Economics Loughborough University, Leicestershire LE11 3TU, Email: t.panagiotidis@lboro.ac.uk

1. Introduction

Determination of the risk factors that affect the dynamics of the spread between fixed-for-floating interest rate swaps and the underlying government bond yields is important for both market participants and policy makers. Relative movements in interest rate swap spreads are nowadays widely used by policy makers as indicators for credit and liquidity conditions in the economy. Kocic and Quintos (2001) point out that practitioners have also interest in identifying the risk factors driving the evolution of swap spreads in order to perform “rich-cheap” analysis and construct relative value trades. A clear understanding of the dynamics of the risk factors in swap markets will allow market participants to construct more accurate swap pricing models and policy makers to extract more accurate information on credit and liquidity conditions in the economy. Previous work (discussed below) has already examined whether different proxies for liquidity and credit risk as well as proxies for market structure can account for the variability of interest rate swap spreads.

The current paper contributes to the existing literature by focusing on the interlinkages between the international interest rate swap markets. Previous work by Smith, Smithson and Wakeman (1988) showed that under the assumption of no default and liquidity risk, the fixed rate of an interest rate swap can be considered as the yield of an identical maturity that trades at par. Subsequent contributions have shown that swap spreads represent a reward for the investors above government bond yields for bearing either liquidity risk in the interbank market, e.g. Grinblatt (1995), or both liquidity and default risk in swap markets, e.g. Duffie and Singleton (1997). Brown, Harlow and Smith (1994) have argued that swap spreads can also be used to cover hedging costs for swap market deals. Sorensen and Bollier (1994) argue that the swap spreads reflect the price of a series of European options to default implicitly held by the counterparty that is in-the-money during the initial stages of the swap contract. Lang, Litzenberger and Liu (1998) and Fehle (2000) examine how the swap market structure can affect spreads through the supply and demand for swaps.

The theory of swap-pricing models has been accompanied by substantial empirical work. Sun, Sundaresan and Wang (1993) look at the relationship between swap rates and Treasury yields as well as yields on interbank par bonds. They find that although swap rates are highly correlated with treasury yields, the swap rates are significantly higher than treasury yields, irrespective of the shape of the treasury yield curve. Minton (1997) looks at the relationship between swap rates and Eurodollar futures rates as well as yields on portfolios of non-callable corporate bonds. Minton's (1997) work identifies the shape of the term structure of default-free interest rates, the level of interest rates and the volatility of short-term interest rates as additional factors affecting swap spreads. Brown, Harlow and Smith (1994) model swap spreads as a function of the difference between Eurodollar LIBOR rates and the corresponding maturity Treasury bill rates (TED) and various measures for credit risk and hedging costs for swap market dealers. They find that while all of these factors are significant, their explanatory power is low. Eom, Subrahmanyam and Uno (2001) report that the slope and curvature of the default-free interest rates and the corporate bond yields are significant factors in the determination of the Japanese swap spreads, while factors like the TED spreads and short-term interest rates play only a minor role.

The above studies estimate single linear swap spread regressions. Duffie and Singleton (1997) and Lekkos and Milas (2001) extend the analysis of swap spreads to a multivariate vector autoregression (VAR) framework. Duffie and Singleton (1997) find that default risk is

significant in affecting longer maturity swap spreads. Lekkos and Milas (2001) examine the ability of factors such as the level, volatility and slope of the zero-coupon government bond yield curve, the TED spread and the corporate bond spread to describe the term structure of the US and UK swap spreads. They find that the slope of the term structure has a significant countercyclical effect across maturities, whereas the TED and corporate spreads play a smaller role and their significance varies across maturities.

More recently, Lekkos and Milas (2004) and In (2005) examine in detail the issue of international linkages between interest rate swap markets.¹ These links can be due to common variations in the business cycles across economies.² Lekkos and Milas (2004) employ non-linear smooth transition vector autoregressive (STVAR) models to show that the slope of the US term structure affects significantly swap spread dynamics in the UK. Similar findings are reported by In (2005) who employs multivariate VAR-EGARCH models to show that the slope of the US term structure has a significant effect on the Japanese and U.K. swap markets.

The papers discussed above have looked at the availability of risk factors to explain the dynamics of swap spreads within sample. To the best of our knowledge, no previous study has attempted to forecast swap spreads out-of-sample. The current paper explores the out-of-sample ability of a number of risk factors to predict the dynamics of the term structure of swap spreads. The risk factors we employ are: estimates of the corporate bond spreads of the two countries, the interest rate differentials between the US and UK government bonds, and the slopes of the term structures of zero-coupon government bonds of the two countries. The interest rate differentials are used to provide evidence of arbitrage trades between the two markets. The corporate bond spreads are used as proxies for credit risk. The slopes of the term structure can be used to test whether the option to default is priced in swap markets; increases in the long-term interest rates imply that, during the first stages of the swap contract, the fixed rate will be higher than the expected short-term LIBOR. Therefore, the fixed-rate payer will be exposed to the possibility of default of the floating-rate payer during the later stages of the contract. This exposure is priced in a higher swap spread.

We assess the ability of the risk factors above to predict swap rates by employing two types of non-linear models: a smooth transition vector autoregressive (STVAR) model and a nearest-neighbours (NN) model. A STVAR model is a regime switching model where the transition from one regime to the other occurs in a smooth way. The switching between regimes is controlled by an observed state variable. This feature of the STVAR model, that the transition from one regime to the other is a function of the underlying variables, allows us to test the ability of the different economic variables to best describe the non-linear dynamics of the term structure swap spreads.³ In particular, we find that the slope of the US term structure of interest rates best describes the transition between the two regimes in both the US and UK swap spreads across maturities. The second type of non-linear models we examine is a nearest-neighbours

¹ Lekkos and Milas (2001) have provided some preliminary evidence on the impact of US factors on UK swap markets and Eom, Subrahmanyam and Uno (2001) on the links between US and Japanese swap markets. Fehle (2000) examines the impact that US swap spreads have on British, German, French, Japanese, Spanish and Dutch swap markets.

² Lumsdaine and Prasad (1997) show that business cycles in each economy are not independent; instead they are affected, in different degrees, by a "world business cycle". Harvey (1991) finds that the correlation between the world and US business cycles is 87%.

³ Among other studies, Ang and Bekaert (2002), Bekaert, Hodrick and Marshall (2001), Hamilton (1988) and Gray (1996) use Hamilton's (1989) Markov regime switching (MRS) model to explain the dynamics of short and long-term interest rates. In contrast to STVAR models, MRS models assume that an unobserved Markov state variable drives the switching between regimes.

(NN) model. Contrary to the STAR model, which relies on global information in order to predict swap spreads, the NN model is a non-parametric local information model that uses a number of nearest neighbours to compute a weighted average estimate of swap spreads.⁴ The ability of the non-linear STVAR and NN models to predict swap spreads is compared with that of linear AR and VAR models. There is some evidence (although not overwhelming) of forecasting superiority of the non-linear models over the linear ones. Further, the NN model does better than the STVAR model in predicting US swap spreads during periods of increasing risk conditions. For the UK, the NN model forecasts swap spreads better than the STVAR model during periods of decreasing risk conditions.

The structure of the paper is organised as follows. The next section describes the data. Sections 3 and 4 discuss the STVAR and NN models, respectively. Section 5 reports the forecasting results. Finally, section 6 concludes.

2. The data

The data set consists of weekly observations from June 1991 to June 2001. We proxy the slope of the term structure of interest rates (denoted by *USslope* and *UKslope*, respectively) with the difference between the yields of the 10-year default-free zero-coupon bonds and the 3-month T-Bill rates. The US and UK zero-coupon yields are provided by the Bank of England. They are estimated by fitting a set of cubic splines to the prices of observed coupon-paying government bonds. The quality of the fit is controlled by a penalty function that restricts the curvature of the implied forward rates (see Anderson and Sleath (1999)). Zero-coupon yields are also used to estimate the difference between the 3-year, 7-year and 10-year US and UK interest rates, denoted by *dif_3*, *dif_7* and *dif_10*, respectively. The US corporate spreads (denoted by *UScorp*) are estimated as the difference between Moody's AAA corporate bond yield index and the yields of the 10-year Treasury bonds. The UK corporate spread (denoted by *UKcorp*) is estimated as the difference between the corporate bond yield index provided by Datastream and the 10-year UK government bond yield. Finally, the US and UK swap spreads (denoted by *USsp_i* and *UKsp_i*, respectively, with $i = 3, 7$ and 10 years) are estimated as the difference between the bootstrapped zero-coupon swap rates and the corresponding maturity default-free zero-coupon rates.

Figure 1 plots the US and UK swap spreads across maturities, whereas Table 1 reports the descriptive statistics for the US and UK swap spreads and the relevant risk factors. In both markets, swap spreads increase, on average, with maturity. Same maturity US and UK swap spreads are roughly equal but UK swap spreads are more volatile. The UK slope is also more volatile compared to the US slope. This can be explained by the fact that UK rates are, on average, higher than US rates over the sample period. Finally, the mean spread between US corporate and US treasury yields is 119 basis points and the corresponding UK corporate spread is 92 basis points.

⁴ Recent applications of NN models in finance include e.g. Diebold and Nason (1990), Meese and Rose (1990), Gençay (1999), Jaditz and Riddick (2000), and Pérez-Rodríguez, Torra and Andrada-Félix (2005).

3. The Smooth Transition Vector Autoregressive (STVAR) model

3.1 The theoretical STVAR model

We define a vector of state variables; one for each maturity we examine. For each maturity, this vector contains the relevant swap spreads as well as the US and UK term structure slopes, the difference between US and UK interest rates and the US and UK corporate spreads. We focus on the 3-year, 7-year and 10-year maturity swap spreads. For each of these maturities the vector of state variables is given by:

$$\mathbf{y}_t = [USslope, UKslope, dif_i, UScorp, UKcorp, USsp_i, UKsp_i]' \quad (1)$$

where $i = 3, 7$ and 10 years. The corresponding STVAR model can be specified as:

$$\mathbf{y}_t = \left(\boldsymbol{\mu}_1 + \sum_{j=1}^p \boldsymbol{\Phi}_{1,j} \mathbf{y}_{t-j} \right) (1 - G(s_t)) + \left(\boldsymbol{\mu}_2 + \sum_{j=1}^p \boldsymbol{\Phi}_{2,j} \mathbf{y}_{t-j} \right) G(s_t) + \boldsymbol{\varepsilon}_t, \quad (2)$$

where \mathbf{y}_t is the $(k \times 1)$ time series vector defined above, $\boldsymbol{\Phi}_{1,j}$ and $\boldsymbol{\Phi}_{2,j}$, $j = 1, \dots, p$, are $(k \times k)$ matrices, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are $(k \times 1)$ vectors, and $\boldsymbol{\varepsilon}_t \sim iid(0, \boldsymbol{\Sigma})$. $G(s_t)$ is the transition function that controls the regime switching dynamics of \mathbf{y}_t . The STVAR model is a regime switching model where the transition between the two alternative regimes is controlled by a transition function $G(\cdot)$ which is continuous and bounded between 0 and 1. Values of zero by the transition function identify the one regime and values of 1 identify the alternative and the transition between the two regimes occurs in a smooth way, i.e. the model does not allow jumps from one regime to the other. The regime that occurs at any time t is not probabilistic. Instead, it is determined but the transition variable s_t and the functional form of the transition function $G(s_t)$. We focus our attention on the ‘logistic’ function:

$$G(s_t; \gamma, c) = \{1 + \exp[-\gamma(s_t - c) / \sigma(s_t)]\}^{-1}, \gamma > 0, \quad (3)$$

where $\sigma(s_t)$ is the sample standard deviation of s_t . Model (3) allows for asymmetric adjustment to positive and negative deviations of s_t relative to c . The parameter c is the threshold between the two regimes, in the sense that $G(s_t)$ changes monotonically from 0 to 1 as s_t increases, and takes the value of $G(s_t) = 0.5$ at $s_t = c$. The parameter γ determines the smoothness of the change in the value of the logistic function and thus the speed of the transition from one regime to the other. When $\gamma \rightarrow 0$, the ‘logistic’ function equals a constant (i.e. 0.5), and when $\gamma \rightarrow +\infty$, the transition from $G(s_t) = 0$ to $G(s_t) = 1$ is almost instantaneous at $s_t = c$.

3.2 Linearity testing in a STVAR model

Testing for linearity in the STVAR model (2) using the ‘logistic’ transition model (3) is equivalent to testing the null hypothesis $H_0: \gamma = 0$ against the alternative $H_1: \gamma > 0$. To do this, define $w_t = (y_{1t-1}, \dots, y_{1t-p}, y_{2t-1}, \dots, y_{2t-p}, \dots, y_{kt-1}, \dots, y_{kt-p})$ and assume that the transition variable

(denoted by s_t) is known. Following Luukkonen, Saikkonen and Teräsvirta (1988), linearity testing equation by equation is based on a first-order Taylor approximation of the transition function around $\gamma = 0$. We first estimate $y_{it} = \beta_{i0} + \sum_{j=1}^{pk} \beta_{ij} w_{jt} + \varepsilon_{it}$ and then use the estimated residuals e_{it} to run the following regression: $e_{it} = \alpha_{i0} + \sum_{j=1}^{pk} \alpha_{ij} w_{jt} + \sum_{j=1}^{pk} \delta_i s_t w_{jt} + \eta_{it}$. Denote the estimated residuals by v_{it} . A Lagrange Multiplier (*LM*) test can be constructed as: $LM = T(SSR_0 - SSR_1) / SSR_0$, where $SSR_0 = \sum e_{it}^2$ and $SSR_1 = \sum v_{it}^2$. Under the null hypothesis of linearity the *LM* statistic is distributed as a $\chi^2(pk)$. In small samples, the χ^2 test may be heavily oversized. Therefore, it is preferable to use the equivalent *F* version of the *LM* test statistic, which is given by $F = [(SSR_0 - SSR_1) / pk] / [SSR_1 / (T - (2pk + 1))]$. It is well known that neglected heteroskedasticity may lead to spurious rejection of linearity. To tackle this problem, we use Wooldridge's (1990, 1991) heteroskedasticity-robust versions of the tests. These tests can be used without having to specify the exact form of heteroskedasticity (see Granger and Teräsvirta, 1993). To compute a heteroskedasticity robust version of the *LM* test statistic reported above, first we estimate $y_{it} = \beta_{i0} + \sum_{j=1}^{pk} \beta_{ij} w_{jt} + \varepsilon_{it}$ and save the estimated residuals e_{it} . We then regress the auxiliary regressors $s_t w_{jt}$ on w_{jt} and save the residuals r_{jt} . Finally, we regress 1 on $e_{it} r_{jt}$. The explained sum of squares from this last regression is the heteroskedasticity robust *LM* test statistic.

Both the χ^2 and *F* versions of the *LM* statistic are equation specific tests for linearity. To test the null hypothesis $H_0: \gamma = 0$ in all equations simultaneously, we need a system-wide test. Following Weise (1999), define $\Omega_0 = \sum e_t e_t' / T$ and $\Omega_1 = \sum v_t v_t' / T$ as the estimated variance-covariance residual matrices from the restricted and the unrestricted estimated equations, respectively. The appropriate log-likelihood system-wide test statistic is given by $LR = T\{\log|\Omega_0| - \log|\Omega_1|\}$, which, under the null hypothesis of linearity is asymptotically distributed as $\chi^2(pk^2)$.

3.3 Empirical STVAR models

3.3.1 Linear VAR models and linearity testing

We begin by estimating a benchmark linear VAR (one for each maturity) over “rolling” fixed-length windows of data, where the first data window runs from June 1991 until December 1998, and each successive data window is constructed by shifting the preceding window ahead by one week.⁵ Bearing in mind that a high-order VAR may cause over-fitting and make it more difficult to get converging estimates for the non-linear models, we use $p = 2$ lags in models (1). For each data window, we test for non-linearities and select the best candidate for the transition variable s_t . To account for the presence of autocorrelation and heteroskedasticity in our models of weekly swap spreads, we use the Generalised Method of Moments (GMM; see Hansen, 1982) estimation technique, which is robust to heteroskedasticity and autocorrelation of unknown form.

⁵ After estimating our models for each “rolling” data window, we forecast one week ahead as discussed in more detail in section 5 below.

We use all lagged variables in (1), as possible transition candidates s_t . To save space, Table 2 reports (for the first data window from June 1991 to December 1998) equation specific *LM* tests and system-wide *LR* linearity tests for the different transition variable candidates only in the case of the 3-year US and UK swap spread equations.⁶ The common approach is to select the appropriate transition variable associated with the smallest *p*-value. The *LM* tests identify the first lag of the slope of the US term structure of interest rates as the most appropriate transition variable. The *LR* system tests indicate that all VAR equations react in a non-linear way not only to $USslope_{t-1}$, but also to all other lagged variables in the system. However, the *LR* tests are less informative than the *LM* tests as the corresponding *p*-values are almost always equal to zero. Therefore, we proceed by using $USslope_{t-1}$ as the transition variable across all equations.⁷ The intuitive reason for this choice is related to the ability of the slope of the term structure to predict economic expansions and recessions; in particular, steep slopes tend to precede periods of economic expansion, whereas flat or negative slopes tend to indicate recessions (for more details see e.g. the recent survey by Stock and Watson, 2003, and references therein). Our empirical choice is also consistent with the findings of Harvey (1991) who found that the US term structure is able to forecast real economic growth in the UK whereas Ang and Bekaert (2002) provided evidence that the US slope Granger-causes the UK term structure.

3.3.2 Estimation of STVAR models and regime identification

As for the linear models, we estimate STVAR models for each “rolling” data window. In order to estimate the STVAR models, we follow Granger and Teräsvirta (1993) and Teräsvirta (1994) in scaling the ‘logistic’ function (3) by dividing it by the standard deviation of the transition variable $\sigma(s_t)$, so that γ becomes a scale-free parameter. Doing so, avoids slow convergence or overestimation associated with estimates of γ . Following from the above scaling, we set $\gamma = 1$ as a starting value and the sample mean of s_t as a starting value for the threshold c . At the same time, the estimates of the linear VAR equations for the $USsp_i$ and $UKsp_i$, ($i = 3, 7$, and 10) provide the starting values for the parameters in the STVAR model (2).

The regime identification of our models is reported in Figure 2, which plots (for the first data window from June 1991 to December 1998) the value of the transition function estimated for the 3-year swap spreads system over time. The periods from June 1991 to December 1991 and from January 1995 to December 1998 are classified into the first regime, while the periods from June 1992 to June 1993 and from March 1994 to August 1994 are classified into the second regime. Notice, however, that the economic interpretation of the two regimes is not straightforward because they do not always coincide with periods of economic expansion and recession. Recall that the first regime that corresponds to a flat term structure should identify periods of economic recession, while the alternative regime should identify periods of economic expansion.

Our regime classification captures the recession that ended in December 1991 and the subsequent recovery of the US economy. Nevertheless, the years from 1995 to 1998, which are classified into the first regime, were periods of significant economic expansion. This period is identified with the first regime because the US slope (see figure 2) started disinvesting by the end of 1994 and continued to move downwards between 1995 and 1998 despite the fact that

⁶ Detailed Tables with linearity tests for the remaining US and UK swap maturities are available on request. As for the 3-year swap maturities, these select $USslope_{t-1}$ as the appropriate transition variable. See the discussion below.

⁷ Linearity test results for each successive data window are consistent with those reported here in the sense that $USslope_{t-1}$ is chosen as the most appropriate transition variable.

these were periods of robust economic growth. Bearing in mind that the relationship between changes in the term structure and subsequent changes in economic activity is probabilistic and that our sample does not contain a significant number of changes from expansion to recession (and vice versa), we cannot explore further the reasons for this apparent broken link between the term structure and economic activity. That said, the ability of our model to classify correctly the recession ending in 1991 is consistent with Estrella, Rodrigues and Schich, (2000) who find that models using the US slope are stable in predicting recessions but become unstable when predicting output growth.

4. Nearest-neighbours (NN) model

We only give a very brief summary of the nearest-neighbours model; for a detailed discussion see e.g. the recent papers by Gençay (1999) and Jaditz and Riddick (2000). In order to estimate y_t conditional on its history $(y_{t-1}, \dots, y_{t-n})$, convert the time series process $\{y_t\}_{t=1}^T$ into n past history components of the form $y_t^n = (y_t, y_{t-1}, \dots, y_{t-n+1})$. The idea here is to take the most recent history available and then retrieve the k nearest neighbours by searching over the set of all n histories. That is, in order to estimate y_t conditional on the information available at $t-1$, compute the distance between the vector $y_{t-1}^n = (y_{t-1}, y_{t-2}, \dots, y_{t-n})$ and its k nearest neighbours to derive the estimator $\sum_{i=1}^k \lambda_{ti} y_i$, where λ_{ti} are the k nearest-neighbour weights. These are calculated using the sup norm $\|y\| = \max_i |y_i|$.⁸ The optimal number of nearest-neighbours is determined by the minimum Mean Squared Prediction Error (MSPE) achieved by regressing y_t on all possible nearest neighbours. Relying on optimally chosen local information as opposed to the global information used by the rest of the models employed in our paper, may prevent overfitting (see e.g. Gençay, 1999). It should also be pointed out that the optimal number of nearest-neighbours changes as we estimate nearest-neighbours models for each “rolling” data window.

5. Forecasting analysis

5.1 Some theoretical issues on forecasting

In order to assess the usefulness of the non-linear models, we carry out our forecasting exercise over “rolling” fixed-length windows of data, where the first data window runs from June 1991 until December 1998, and each successive data window is constructed by shifting the preceding window ahead by one week. Therefore, we re-estimate our models for each data window and then produce out-of-sample forecasts for the US and UK swap spreads over $h = 1$ week ahead.

We compute out-of-sample forecasts from STVAR models, univariate NN (NN) models, linear VAR models and univariate autoregressive (AR) swap spread models.⁹ Forecasting performance is evaluated using the Mean Squared Prediction Error (MSPE) and the Mean Absolute Prediction Error (MAPE) criteria. Further, in order to see whether the non-linear models outperform the AR and VAR models, we employ the Diebold and Mariano (1995) test.

⁸ Alternatively, one can use Euclidean distances. We use the sup norm because it is computationally less intensive. Jaditz and Riddick (2000) point out that the sup norm is not worse than the Euclidean one.

⁹ We use two lags for all AR swap spread equations except for the *USsp_7* and *USsp_10* equations where three lags are used; lags are selected based on the Akaike Information Criterion.

This is computed by weighting the forecast loss differentials between two competing models equally, where the loss differential for observation t is given by $d_t \equiv [g(e_{it|t-h}) - g(e_{jt|t-h})]$, where $g(\cdot)$ is a general function of forecast errors (e.g. MSPE or MAPE). The null hypothesis of equal accuracy of the forecasts of two competing models can be expressed in terms of their corresponding loss functions, $E[g(e_{it|t-h})] = E[g(e_{jt|t-h})]$, or equivalently in terms of their loss differential, $E[d_t] = 0$. Let $\bar{d} = \frac{1}{P} \sum_{t=R+h}^{R+P+h-1} d_t$ denote the sample mean loss differential over t observations, such that there are P out-of-sample point forecasts and R observations have been used for estimation. The Diebold-Mariano test statistic follows asymptotically the standard normal distribution:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi\hat{f}_d(0)}{P}}} \xrightarrow{d} N(0,1), \quad (4)$$

where $N(\cdot)$ is the normal distribution and $\hat{f}_d(0)$ is a consistent estimate of the spectral density of the loss differential at frequency 0.

To counteract the tendency of the DM test statistic to reject the null too often when it is true in cases where the forecast errors are not bivariate normal, Harvey, Leybourne and Newbold (1997) propose a modified Diebold-Mariano test statistic:

$$DM^* = \left[\frac{P+1-2h+P^{-1}h(h-1)}{P} \right]^{1/2} DM \xrightarrow{d} t_{(P-1)}, \quad (5)$$

where DM is the original Diebold and Mariano (1995) test statistic for h -steps ahead forecasts and $t_{(P-1)}$ refers to Student's t distribution with $P-1$ degrees of freedom.

Recently, van Dijk and Franses (2003) argue that the uniform weighting scheme employed by the DM and DM^* tests may be unsatisfactory for frequently encountered situations in which some observations are more important than others. For example, in a swap spread forecasting exercise, large positive swap spread observations generally signal periods of increasing risk conditions in the economy.

van Dijk and Franses (2003) modify the test statistic by weighting more heavily the loss differentials for observations that are deemed to be of greater substantive interest. In their approach, the weighted mean loss differential is given by $\bar{d}_w = \frac{1}{P} \sum_{t=R+h}^{R+P+h-1} w(\omega_t) d_t$, where $w(\omega_t)$ is the information set available at time t . Letting y_t be the variable to be forecast, two particular cases van Dijk and Franses (2003) study are:

$$w_{LT}(\omega_t) = 1 - \Phi(y_t), \quad (6)$$

where $\Phi(\cdot)$ is the cumulative distribution function of y_t , to focus on the left tail of the distribution of y_t , and:

$$w_{RT}(\omega_t) = \Phi(y_t), \quad (7)$$

to focus on the right tail of the distribution of y_t . A necessary condition for the associated test statistic to have an asymptotic standard normal distribution under the null hypothesis of equal forecast accuracy is that the weight function $w(\omega_t)$ be a twice continuously differentiable mapping to the $[0,1]$ interval. The weighted DM statistic is computed as:

$$W - DM = \frac{\bar{d}_w}{\sqrt{\frac{2\pi\hat{f}_{dw}(0)}{P}}} \quad (8)$$

where $\hat{f}_{dw}(0)$ is a consistent estimate of the spectral density of the loss differential at frequency 0. The weighted DM^* test statistic is given by:

$$W - DM^* = \left[\frac{P+1-2h+P^{-1}h(h-1)}{P} \right]^{1/2} W - DM \quad (9)$$

Once again following Harvey et al. (1997), van Dijk and Franses propose using the Student's t distribution with $P - 1$ degrees of freedom to obtain critical values for the $W-DM^*$ test. In our forecasting exercise, the Left-tailed $W-DM^*$ statistic focuses on the ability of the competing models to forecast small swap spread values which is interpreted as evidence of decreasing risk conditions in the economy. On the other hand, the Right-tailed $W-DM^*$ statistic focuses on the ability to forecast large spread values which is interpreted as evidence of periods of increasing risk conditions in the economy.

5.2 Empirical results

The results of our forecasting exercise are reported in Tables 3 and 4. We report the MSPE criteria for the different US and UK swap spread models (results using MAPE criteria led to very similar conclusions and are available on request). The statistical significance of the forecasting performance of the non-linear STVAR and NN models relative to the linear VAR and AR models is examined using the modified DM^* , Left-tailed $W-DM^*$ and Right-tailed $W-DM^*$ criteria. For both Tables 3 and 4, the top entry in $[.]$ contains the p -values for the modified DM^* statistic against the one-sided alternative that the MSPE of the competing model is lower. The middle entry in $[.]$ contains the p -values for the modified Left-tailed $W-DM^*$ statistic whereas the bottom entry in $[.]$ contains the p -values for the modified Right-tailed $W-DM^*$.

From Table 3, the NN model produces the lower MSPE for two out of three US swap spread maturities. In particular, our results suggest forecasting superiority of the NN model over the AR, VAR and STVAR models for short and long US swap spread maturities (i.e. $USsp_3$ and $USsp_10$). However, the ability of the NN model to predict small spread values at the long end (i.e. $USsp_10$) is not better than that of the VAR model (i.e. p -value for the Left-tailed $W-DM^*$ equals 0.589) or that of the STVAR model (i.e. p -value for the Left-tailed $W-DM^*$ equals 0.144). The STVAR model does not beat the VAR model at any maturity, but it outperforms the AR model at the short maturity (i.e. $USsp_3$). For the 7-year US swap spread, the non-linear NN and STVAR models do not outperform the linear models. On the other hand, for the 7-year US swap spread, the NN model seems to outperform the STVAR model in predicting

large swap spreads (i.e. p -value for the Right-tailed W - DM^* equals 0.041).

The NN model produces the lower MSPE for two out of three UK swap spread maturities (see Table 4). In statistical terms, however, the NN model does not beat the AR model at any maturity. On the other hand, the NN model outperforms the linear VAR model at all maturities. Further, it outperforms the STVAR model at the 3-year and 7-year swap spread maturities. The NN model predicts small swap spread values better than the STVAR model at all maturities. The STVAR model outperforms the VAR model for short and long maturities. For the medium-to-long maturity swap spreads (i.e. the $UKsp_7$ swap spread), the STVAR model is able to predict better than the VAR model only small swap spread values (i.e. p -value for the Left-tailed W - DM^* equals 0.004).

Overall, our forecasting exercise for the US and UK swap spreads suggests some forecasting superiority of the non-linear models against the linear ones. This is more the case for the NN non-linear model, which seems to occasionally predict better than the rest of the non-linear and linear models. Another interesting finding of our forecasting exercise is that for all maturities, the NN model is able to predict better than the STVAR model US swap spreads during periods of increasing risk conditions and UK swap spreads during periods of decreasing risk conditions. The ability of the STVAR models to do better than linear AR and VAR models appears sensitive to swap spread maturity. Previous studies also find some superiority of NN relative to other models. Gençay (1999) finds that NN models outperform other linear and non-linear models for a number of exchange rates, whereas Pérez-Rodríguez, Torra and Andrada-Félix (2005) find some evidence that NN models outperform linear AR and STAR models in predicting Spanish stock returns. On the other hand, Teräsvirta, van Dijk and Medeiros (2005) use a number of G7 macroeconomic time series to show that STAR models forecast better compared to linear models whereas the forecasting performance of NN models is more mixed.

6. Conclusions

This paper explores the ability of common risk factors to predict US and UK interest rate swap spreads within a linear and a non-linear framework. We reject linearity in favour of a regime-switching STVAR model for the US and UK swap spread dynamics and we show that the switching between regimes is controlled by the slope of the US term structure of interest rates. The first regime is characterised by a "flat" term structure of US interest rates, while the alternative is characterised by an "upward" sloping US term structure. In economic terms, the two regimes do not always coincide with periods of economic expansion and recession. This can be interpreted as evidence of a break in the relationship between US real output growth and the slope of the US term structure. We do not explicitly test for a break but possible reasons include the shift in US monetary policy from reactive to proactive as a response to the Asian and Russian financial crises and the Long-Term Capital Management (LTCM) collapse in 1998.

The ability of the non-linear STVAR model to predict swap spreads is compared with that of a non-linear NN model as well as that of linear AR and VAR models. We find evidence that the NN and STVAR models predict better than the linear AR and VAR models. However, the evidence is not overwhelming as it is sensitive to swap spread maturity. Within the non-linear class of models, the NN model predicts US swap spreads better than the STVAR model during periods of increasing risk conditions. Further, the NN model forecasts UK swap spreads better than the STVAR model during periods of decreasing risk conditions. Encouraged by these findings, it seems promising to compare the forecasting performance of a broader variety of non-linear swap spread models in future research.

References

- Anderson, N. and J. Sleath (1999). New estimates of the UK real and nominal yield curves, Bank of England Quarterly Bulletin, 39, 384-392.
- Ang, A. and G. Bekaert (2002). Regime switches in interest rates. Journal of Business and Economic Statistics 20, 163-182.
- Bekaert, G., R.J. Hodrick and D.A. Marshall (2001). "Peso Problem" explanations for term structure anomalies. Journal of Monetary Economics, 48, 241-270.
- Brown, K., W.V. Harlow and D.J. Smith (1994). An empirical analysis of interest rates swap spreads. Journal of Fixed Income 3, 61-78.
- Diebold, F.X. and R.S. Mariano (1995). Comparing predictive accuracy. Journal of Business and Economic Statistics, 13, 253-263.
- Diebold, F.X. and J.A. Nason (1990). Nonparametric exchange rate prediction? Journal of International Economics, 28, 315-322.
- Duffie, D. and M. Huang (1996). Swap rates and credit quality. Journal of Finance, 51, 921-949.
- Duffie D. and K.J. Singleton (1997). An econometric model of the term structure of interest-rate swap yields. Journal of Finance, 52, 1287-1321.
- Eom, Y.H., M.G. Subrahmanyam and J. Uno (2001). Credit risk and the yen interest rate swap market. Working Paper 042, Stern Business School, New York.
- Estrella, A., A.P. Rodrigues and S. Schich (2000). How stable is the predictive power of the yield curve? Evidence from Germany and the United States. Federal Reserve Bank of New York, Staff Report 113.
- Fehle, F. (2000). The components of interest rate swap spreads: Theory and international evidence. Unpublished manuscript, University of South Carolina.
- Franses, P.H. and D. van Dijk (2000). Non-linear time series models in empirical finance. Cambridge University Press, Cambridge.
- Gençay, R. (1999). Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules. Journal of International Economics, 47, 91-107.
- Granger, C.W.J., and T. Teräsvirta (1993). Modelling nonlinear economic relationships. Oxford University Press, Oxford.
- Gray, S.F. (1996). Modeling the conditional distribution of interest rates as a regime switching process. Journal of Financial Economics, 42, 42-62.

- Grinblatt, M. (1995). An analytic solution for interest-rate swap spreads. Unpublished manuscript, Anderson Graduate School of Management, University of California, Los Angeles.
- Hamilton, J.D. (1988). Rational expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control*, 12, 385 - 423.
- Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57, 357-384.
- Hansen, L.P. (1982). Large sample properties of Generalised Method of Moments estimators. *Econometrica*, 50, 1029-1054.
- Harvey, C.R. (1991). The term structure and world economic growth. *Journal of Fixed Income*, June, pp. 7-19.
- Harvey, D., S. Leybourne and P. Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13, 281-291.
- In, F. (2005). Volatility spillovers across international swap markets: the U.S., Japan, and the U.K. *Journal of International Money and Finance* (forthcoming).
- Jaditz, T. and L.A. Riddick (2000). Time-series near-neighbor regression. *Studies in Nonlinear Dynamics and Econometrics*, 4, 35-44.
- Kocic, A., and C. Quintos (2001). Identifying relative value through the forecasting of swap spreads. Lehman Brothers, Fixed Income Derivatives Research.
- Lang, L.H.P., R.H. Lintzenberger, and A.L. Liu (1998). Determinants of interest rate swap spreads. *Journal of Banking and Finance*, 22, 1507-1532.
- Lekkos, I. and C. Milas (2004). Common risk factors in the US and UK interest rate swap markets: Evidence from a non-linear vector autoregression approach. *Journal of Futures Markets*, 24, 221-250.
- Lekkos, I. and C. Milas (2001). Identifying the factors that affect interest rate swap spreads: some evidence from the US and the UK. *Journal of Futures Markets*, 21, 737-768.
- Lumsadaine, R.L. and E.S. Prasad (1997). Identifying the common component in international economic fluctuations, National Bureau of Economic Research, Working Paper 5984.
- Luukkonen, R., P. Saikkonen and T. Teräsvirta (1988). Testing linearity against smooth transition autoregressive models. *Biometrika*, 75, 491-499.
- Meese, R.A. and A.K. Rose (1990). Non-linear, nonparametric, nonessential exchange rate estimation. *American Economic Review*, 80, 192-196.
- Minton, B.A. (1997). An empirical examination of basic valuation models for plain vanilla

- U.S. interest rate swaps. *Journal of Financial Economics*, 44, 251-277.
- Pérez-Rodríguez, J., S. Torra and J. Andrada-Félix (2005). STAR and ANN models: forecasting performance on the Spanish “Ibex-35” stock index. *Journal of Empirical Finance*, 12, 490-509.
- Smith, C.W., C.W. Smithson and L.M. Wakeman (1988). The market for interest rate swaps. *Financial Management*, 17, 34-44.
- Sorensen, E.H. and T.F. Bollier (1994). Pricing swap default risk. *Financial Analysts Journal*, 50, 23-33.
- Stock, J.H. and M.W. Watson (2003). Forecasting output and inflation: the role of asset prices. *Journal of Economic Literature* 41, 788-829.
- Sun, T.S., Sundaresan, S., Wang, C. (1993). Interest rate swaps - An empirical investigation. *Journal of Financial Economics* 34, 77-99.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89, 208-218.
- Teräsvirta, T., D. van Dijk, and M.C. Medeiros (2005). Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting*, 21, 755-774.
- van Dijk, D. and P.H. Franses (2003). Selecting a nonlinear time series model using weighted tests of equal forecast accuracy. *Oxford Bulletin of Economics and Statistics*, 65, 727-744.
- van Dijk, D., T. Teräsvirta, and P.H. Franses (2002). Smooth transition autoregressive models – a survey of recent developments. *Econometric Reviews*, 21, 1-47.
- Weise, C.L. (1999). The asymmetric effects of monetary policy: a nonlinear vector autoregression approach. *Journal of Money, Credit, and Banking*, 31, 85-108.
- Wooldridge, J.M. (1990). A unified approach to robust, regression-based specification tests. *Econometric Theory*, 6, 17-43.
- Wooldridge, J.M. (1991). On the application of robust, regression-based diagnostics to models of conditional means and conditional variances. *Journal of Econometrics*, 47, 5-46.

Table 1
Descriptive statistics

	Mean	Max	Min	Std. Dev	Skewness	Kurtosis
<i>USsp_3</i>	0.353	0.958	0.014	0.193	0.779	2.872
<i>USsp_7</i>	0.433	1.141	0.088	0.237	1.167	3.313
<i>USsp_10</i>	0.440	1.227	0.093	0.249	1.255	3.792
<i>UKsp_3</i>	0.333	1.010	0.002	0.217	0.634	2.358
<i>UKsp_7</i>	0.413	1.198	0.002	0.277	0.775	2.306
<i>UKsp_10</i>	0.466	1.280	0.003	0.302	0.910	2.635
<i>USslope</i>	1.581	4.141	0.690	1.182	0.382	2.139
<i>UKslope</i>	0.463	3.976	-2.669	1.619	0.196	1.704
<i>Dif_3</i>	-1.100	0.699	-5.294	1.097	-0.996	3.986
<i>Dif_7</i>	-0.813	1.087	-3.064	0.965	0.106	1.916
<i>Dif_10</i>	-0.591	1.323	-2.157	0.991	0.338	1.781
<i>UScorp</i>	1.196	2.150	0.630	0.346	0.759	2.804
<i>UKcorp</i>	0.920	2.068	0.015	0.418	0.257	2.422

Notes: The Table reports descriptive statistics for the swap spreads and the risk factors defined in section 2. The sample refers to weekly data from June 1991 to June 2001.

Table 2

Linearity tests for: $\mathbf{y}_t = [USslope, UKslope, dif_3, UScorp, UKcorp, USsp_3, UKsp_3]'$

Transition variable	<i>USsp_3</i> equation	<i>UKsp_3</i> equation	System test
<i>USslope</i> _{t-1}	0.001	0.029	0.000
<i>USslope</i> _{t-2}	0.006	0.039	0.000
<i>UKslope</i> _{t-1}	0.037	0.497	0.000
<i>UKslope</i> _{t-2}	0.093	0.585	0.002
<i>Dif_3</i> _{t-1}	0.173	0.164	0.000
<i>Dif_3</i> _{t-2}	0.155	0.107	0.000
<i>UScorp</i> _{t-1}	0.382	0.092	0.000
<i>UScorp</i> _{t-2}	0.401	0.054	0.001
<i>UKcorp</i> _{t-1}	0.023	0.105	0.001
<i>UKcorp</i> _{t-2}	0.030	0.206	0.000
<i>USsp_3</i> _{t-1}	0.021	0.108	0.000
<i>USsp_3</i> _{t-2}	0.037	0.266	0.000
<i>UKsp_3</i> _{t-1}	0.288	0.197	0.000
<i>UKsp_3</i> _{t-2}	0.334	0.088	0.000

Notes: The Table reports *p*-values of equation specific Lagrange Multiplier *F* statistics and system-wide *LR* test statistics for the *USsp_3* and *UKsp_3* equations. These are based on Wooldridge's (1990, 1991) heteroskedasticity-robust versions of the tests. The null hypothesis is linearity. The alternative hypothesis is the STVAR model.

Table 3

Mean Squared Prediction Error (MSPE) for the US spreads using modified DM^* , Left-tailed $W-DM^*$ and Right-tailed $W-DM^*$

MSPE				p -values in [.]				
STVAR	NN	VAR	AR	AR vs. STVAR	AR vs. NN	VAR vs. STVAR	VAR vs. NN	STVAR vs. NN
<i>USsp_3</i>								
0.160	0.049	0.135	0.166	[0.001] [0.011] [0.000]	[0.000] [0.000] [0.000]	[0.999] [0.999] [0.999]	[0.000] [0.000] [0.000]	[0.000] [0.000] [0.000]
<i>USsp_7</i>								
0.095	0.090	0.089	0.084	[0.997] [0.995] [0.996]	[0.893] [0.974] [0.678]	[0.935] [0.996] [0.654]	[0.566] [0.970] [0.136]	[0.118] [0.515] [0.041]
<i>USsp_10</i>								
0.105	0.077	0.095	0.085	[1.000] [0.996] [1.000]	[0.048] [0.010] [0.001]	[0.999] [0.990] [0.999]	[0.000] [0.589] [0.000]	[0.000] [0.144] [0.000]

Notes: MSPE for rolling window one step ahead out-of sample forecasts from 1999:1 to 2001:26. The top entry in [.] contains the p -values for the modified DM^* statistic of Harvey, Leybourne and Newbold (1997) against the one-sided alternative that the MSPE of the competing model is lower. The middle entry in [.] contains the p -values for the modified Left-tailed $W-DM^*$ statistic of van Dijk and Franses (2003). The bottom entry in [.] contains the p -values for the modified Right-tailed $W-DM^*$ statistic of van Dijk and Franses (2003).

Table 4
Mean Squared Prediction Error (MSPE) for the UK spreads using modified DM^* , Left-tailed $W-DM^*$ and Right-tailed $W-DM^*$

MSPE				p -values in [.]				
STVAR	NN	VAR	AR	AR vs. STVAR	AR vs. NN	VAR vs. STVAR	VAR vs. NN	STVAR vs. NN
<i>UKsp 3</i>								
0.053	0.049	0.068	0.048	[0.998] [0.999] [0.878]	[0.752] [0.990] [0.400]	[0.000] [0.000] [0.000]	[0.000] [0.000] [0.000]	[0.018] [0.021] [0.110]
<i>UKsp 7</i>								
0.075	0.050	0.070	0.051	[1.000] [0.999] [0.999]	[0.421] [0.384] [0.463]	[0.376] [0.004] [0.999]	[0.000] [0.000] [0.001]	[0.000] [0.000] [0.000]
<i>UKsp 10</i>								
0.053	0.052	0.067	0.053	[0.549] [0.923] [0.225]	[0.320] [0.369] [0.305]	[0.000] [0.000] [0.000]	[0.000] [0.000] [0.016]	[0.362] [0.043] [0.722]

Notes: MSPE for rolling window one step ahead out-of sample forecasts from 1999:1 to 2001:26. The top entry in [.] contains the p -values for the modified DM^* statistic of Harvey, Leybourne and Newbold (1997) against the one-sided alternative that the MSPE of the competing model is lower. The middle entry in [.] contains the p -values for the modified Left-tailed $W-DM^*$ statistic of van Dijk and Franses (2003). The bottom entry in [.] contains the p -values for the modified Right-tailed $W-DM^*$ statistic of van Dijk and Franses (2003).

Figure 1: 3, 7, and 10-year US and UK swap spreads

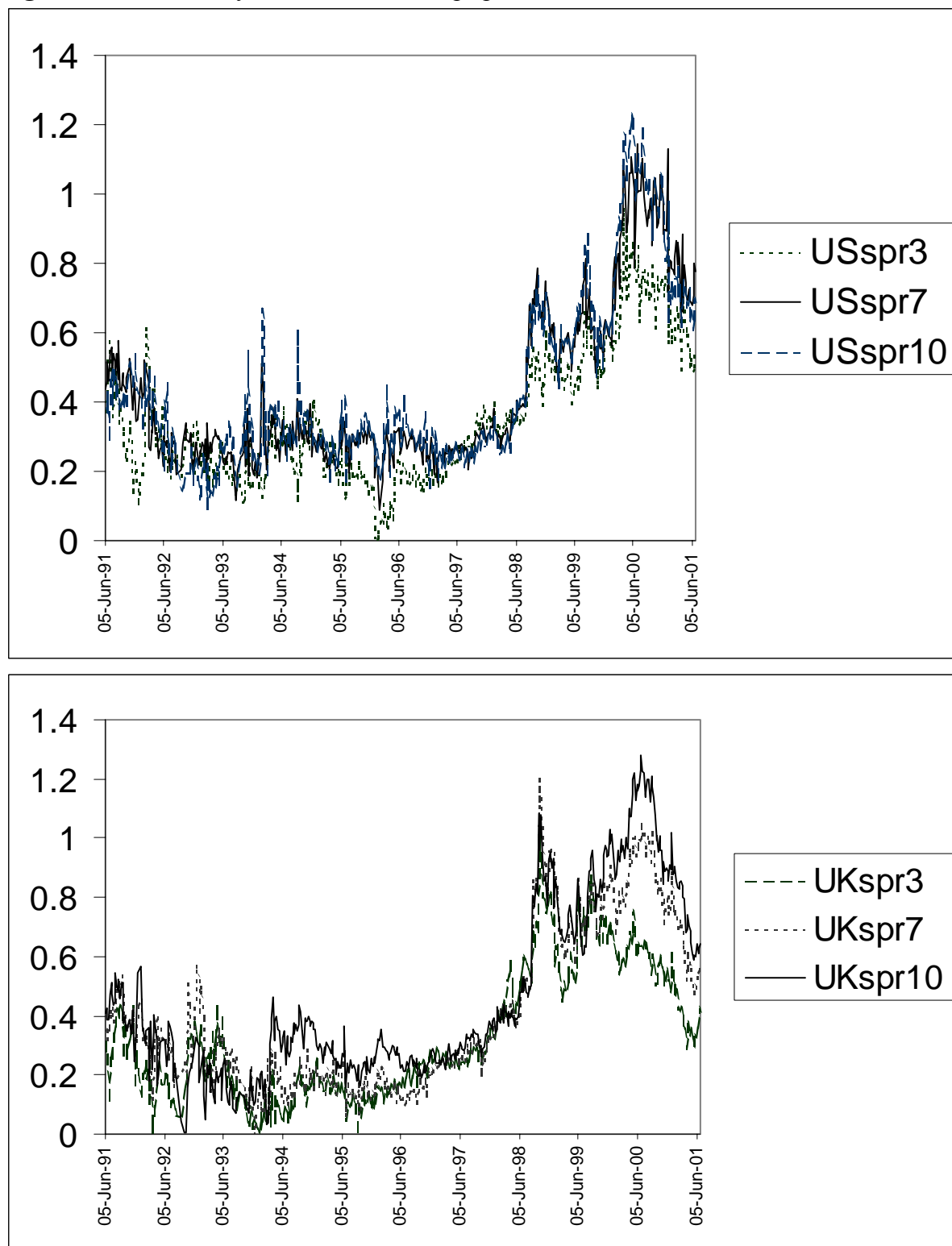
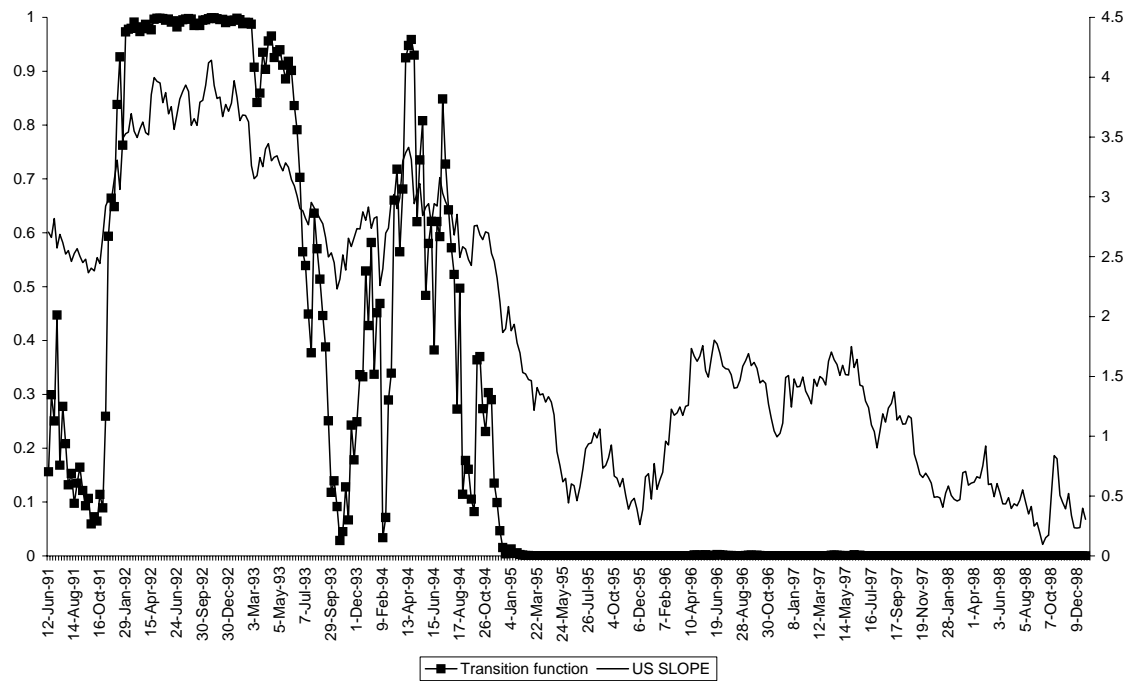


Figure 2: Regime classification and the slope of the US term structure of interest rates



Notes:

The figure plots the slope of the US term structure (solid line, right-hand axis, in percent) and the transition function for the 3-year swap spreads system (line with blocks, left-hand axis) over time. The transition function is estimated as:

$G(USslope_{t-1}; \gamma, c) = \{1 + \exp[-10.528(USslope_{t-1} - 2.837) / \sigma(USslope_{t-1})]\}^{-1}$. Values of the transition function close to zero identify a period with the first regime while values of the transition function close to one identify a period with the second regime.