Short Rate Nonlinearities and Regime Switches

Andrew Ang*
Columbia University

Geert Bekaert†
Columbia University and NBER

First Version: January 10, 2000
This Version: December 19, 2000

JEL Classification: C22, C32, E43, G12
Keywords: short rate, term spread, drift, volatility, regime-switching

*Columbia Business School, 805 Uris Hall, 3022 Broadway, New York, NY 10027. Email: aa610@columbia.edu, ph 212 854 9154

†Columbia Business School, 802 Uris Hall, 3022 Broadway, New York, NY 10027. Email: gb241@columbia.edu, ph 212 854 9156. Geert Bekaert acknowledges financial support from an NSF grant. The authors would like to thank Michael Brandt, Michael Johannes, James Hamilton, Matt Pritsker, Ken Singleton, Liuren Wu, and seminar participants at Stanford University and the January 1999 meeting of the Econometric Society for helpful comments. The paper greatly benefitted from the insightful comments of an anonymous referee. A much shorter first draft of this paper originally appeared as Section 6 of the working paper "Regime Switches in Interest Rates."

Abstract

Using non-parametric estimation methods, various authors have shown distinct non-linearities in the drift and volatility function of the US short rate, which are inconsistent with standard affine term structure models. We document how a regime-switching model with state dependent transition probabilities between regimes can replicate the patterns found by the non-parametric studies. To do so, we use data from the UK and Germany in addition to US data and include term spreads in some of our models. We also examine the drift and volatility function of the term spread.

1 Introduction

Recent years have seen a proliferation of work on the stochastic properties of interest rates. In a number of influential articles, Aït-Sahalia (1996a, 1996b), Stanton (1997), Conley et al. (1997), and Boudoukh et al. (1999) use non-parametric techniques to show distinct non-linearities in the drift function of the short rate. The non-parametric studies show mean reversion at high rates to be much stronger than in normal ranges of the short rate, where the mean-reversion is close to zero. The conditional volatility estimated by these authors also takes on a convex shape which increases with the level of the short rate. This finding is very important for the modeling of the term structure and related derivatives pricing, since many models typically use linear drift and volatility models (such as the Duffie-Kan (1996) affine class of term structure models). Nevertheless, these findings are still somewhat controversial and this line of research has been criticized from a number of directions.

First, non-parametric models in general are over-parameterized and may have poor small sample properties. In fact, with a Vasicek (1977) model, Pritzker (1999) finds that 2755 years of data are required to obtain the accuracy of the kernel estimator implied by the asymptotic estimation using Aït-Sahalia (1996a)'s sample length of 22 years. Worse, Chapman and Pearson (2000) show that Aït-Sahalia's findings may be entirely spurious, primarily because of the lack of data at the extremes of the interest rate ranges (where the non-linearities are found). They show that in small samples, an affine parameterization of the short rate conditional mean may produce the non-linearities in non-parametric estimations. However, Jones (2000) uses a Bayesian setting to show that rejections of linear drifts may be driven by implicit prior beliefs that contain a non-trivial amount of information about the shape of the drift function. Under the Jeffreys prior, a non-informative prior robust to reparameterization, the non-linear results disappear. Downing (1999) shows that while finite sample bias of the non-parametric kernel estimators may account for apparent non-linearities in conditional means, he strongly rejects the null hypothesis of linear volatilities.

Second, the lack of a parametric model is of course problematic in terms of further modeling of the term structure. In fact, these articles generally ignore information in term spreads.[1] This is surprising since the inclusion of term spreads both from an econometric perspective (they are known to Granger-cause short rates) and from a modeling perspective (they are closely linked to short rate dynamics in most term structure models) would help identification tremendously. Some progress has been made here. Ahn and Gao (1999) provide an interesting non-linear term structure model that captures some of the dynamics found by Aït-Sahalia (1996a). Unlike

[1] With the exception of Boudoukh, Richardson, Stanton and Whitelaw (1999).

the other extant non-affine term structure models,[2] in Ahn and Gao's model the factors which drive the short rate are the same factors which completely determine the dynamics of the entire yield curve, as in the affine term structure models. Ahn and Gao's short rate model has a quadratic conditional mean and a cubic conditional variance. Their model matches the non-parametric shapes of the conditional variance very accurately, but captures less adequately the sharp downward slope of the non-parametric conditional mean at high interest rates.

In this article, we contribute to this debate in a number of ways. First, we provide an alternative, parametric model that can match the non-linear patterns detected before. Our model is a regime-switching model which Gray (1996) and Ang and Bekaert (1998) have shown to forecast interest rates well. In such a model there is an unobserved state (a regime variable) that follows a Markov chain and governs the switching between two potentially linear processes (see Hamilton (1989)). Second, we explicitly incorporate information in term spreads, which not only leads to more efficient estimation but also allows us to derive drift functions for the term spread. Moreover, we show that the probability of the transition of one regime to another regime depends on the spread, and that the short rate and spread Granger-cause each other.

Third, we use information from short rates and spreads in three different countries. Because of the extremely high persistence of short rates, using information from other countries is a much more effective way to increase the sample size than lengthening the sample itself. The various experiences of different countries are particularly helpful to give us more observations of the distribution of interest rates near very low and very large short rates. In particular, in addition to interest rates from the US we use data from Germany and the UK.

Fourth, most term structure models treat the stochastic volatility of the short rate as a function of the short rate itself, but only a few papers allow for time-variation in the coefficients of the volatility function. While Ball and Torous (1999) model stochastic interest rate volatility using an exponential GARCH model, we let the short rate volatility depend only on the prevailing regime. Our model also generates stochastic volatility, in a different form: conditional on the regime, the interest rate is homoskedastic, but stochastic volatility is generated by the switching of regimes.[3] As the probability of being in a particular regime conditional on past information varies through time, the conditional moments of the short rate also vary through time. However, more rapid mean reversion presumably occurs during much more volatile periods, and hence al-

[2] See Beaglehole and Tenney (1992), and in particular Constantinides (1992).

[3] Bekaert, Hodrick and Marshall (2000) combine regime-switching volatility with volatility depending on the level of the interest rates within the regime. In a three-state model applied to interest rates from seven countries they find significant within-regime heteroskedasticity. Gray (1996) combines GARCH and regime-switching variances in a univariate model.

lowing for a link between volatility and mean dynamics may improve identification. Our model accomplishes this.

Finally, regime-switching models can accomodate unit root regimes and still remain covariance stationary (See Ang and Bekaert (1988) and Holst et al. (1994)). This makes regime-switching models ideal models to capture the non-linearities of short rates. In particular, they can capture the unit root (or near-unit root) behavior at normal levels of the short rate by having a unit root regime. A second regime can have a higher conditional mean with much higher mean reversion and higher volatility. This enables higher short rates to be associated with higher mean reversion and higher conditional volatility.

Our paper is related to Brandt (1999) who finds that a regime-switching model is a good auxiliary model to estimate continuous-time short rate dynamics with non-linear drifts and volatilities in an Efficient Method of Moments setting. Using this process he cannot reject the Aït-Sahalia (1996a) and Conley et al. (1997) specifications of the short rate. In contrast, we work directly with regime-switching processes and investigate their implied drift and volatility functions. We also use term spread and international cross-sectional data.

This paper is organized as follows. Section 2 sets out the various regime-switching models we estimate. Section 3 briefly describes the data and provides parameter estimates and some simple hypotheses tests for the various models. We show how the parameter estimates reveal an economically intuitive model that could be the result of an interaction between fundamental shocks to inflation or real rates and monetary policy action. Section 4 provides the main empirical results graphing and discussing the drift and volatility functions implied by the main models. We find that when we allow the probability of transitioning from one regime to another to depend on the short rate or spread, we obtain plots of the conditional means and volatilities of short rate dynamics which closely mimic the non-parametric estimations. Section 5 concludes.

2 Models

We consider univariate short rate models and bivariate models which contain the short rate and spread.[4] In each regime the short rate (and spread) process is a linear function, and the regime variable itself follows a Markov chain with possibly time-varying transition probabilities. In Section 2.1 we describe the univariate short rate models and in Section 2.2 we describe the bivariate term spread models.

[4] Hamilton (1988), Lewis (1991), Evans and Lewis (1995), Sola and Driffill (1994), Gray (1996), Evans (1999), Ang and Bekaert (1998), Bansal and Zhou (1999), Veronesi and Yared (1999), and Bekaert, Hodrick and Marshall (2000) all examine empirical models of regime switches in interest rates.

We estimate the models by maximum likelihood using the recursive algorithm developed by Hamilton (1989).[5] We estimate the models for US data, and also estimate the models jointly over three countries (US, UK and Germany) using the cross-sectional approach of Bekaert, Hodrick and Marshall (2000). For the three-country cross-sectional estimation we assume the innovations and regimes are independent across countries. Although pair-wise interest rate correlations across countries are non-zero, Monte Carlo results in Bekaert, Hodrick and Marshall (2000) suggest that the assumption of independence is not rejected by the data.

2.1 Short Rate Univariate Model

The dynamics of the short rate r_t in the univariate model are given by:

$$r_t = \mu(s_t) + \rho(s_t) r_{t-1} + \sigma(s_t)\epsilon_t \tag{1}$$

where the IID errors $\epsilon_t \sim N(0,1)$.[6] Equivalently we can write the model as a function of Δr_t:

$$\Delta r_t = \mu(s_t) - (1 - \rho(s_t)) r_{t-1} + \sigma(s_t)\epsilon_t$$

Conditional on the regime s_t the short rate is a Vasicek (1977) model and has drift $\mu(s_t) - (1 - \rho(s_t)) r_{t-1}$ and volatility $\sigma(s_t)$.

The regime variable s_t is either 1 or 2 and following Diebold et al. (1994) has transition probabilities:

$$p(s_t = j | s_{t-1} = j; \mathcal{I}_{t-1}) = \frac{e^{a_j + b_j r_{t-1}}}{1 + e^{a_j + b_j r_{t-1}}},\ j = 1, 2. \tag{2}$$

where $\mathcal{I}_{t-1}$ is the information set. In this model $I_{t-1} = \{r_{t-1}, r_{t-2}, \ldots, r_0\}$. We consider two cases. The first case has constant transition probabilities, ($b_j = 0$) where we denote $p(s_t = 1 | s_{t-1} = 1; \mathcal{I}_{t-1}) = P$ and $p(s_t = 2 | s_{t-1} = 2; \mathcal{I}_{t-1}) = Q$. The second case has time-varying transition probabilities ($b_j \neq 0$).

This model is similar to the model in Bekaert, Hodrick and Marshall (2000) with the exception that Bekaert, Hodrick and Marshall allow for within-regime heteroskedasticity. By keeping our within-regime processes linear, the non-linearities are entirely driven by regime-switching, not by other features of the model.

[5] See also Hamilton (1994) and Gray (1996).

[6] We will denote the regime variable by subscripts, so $\mu(s_t = i) = \mu_i$, $\rho(s_t = i) = \rho_i$ and $\sigma(s_t = i) = \sigma_i$.

2.2 Term Spread Bivariate Model

The term spread model is a switching bivariate first-order VAR of the short rate r_t and spread z_t:

$$Y_t = \mu(s_t) + A(s_t)Y_{t-1} + \Sigma^{\frac{1}{2}}(s_t)\epsilon_t \tag{3}$$

where $Y_t = (r_t\ z_t)'$, and $\epsilon_t \sim N(0, I)$. We estimate the Cholesky decomposition $R(s_t)$ of $\Sigma(s_t)$ where $\Sigma(s_t) = R(s_t)R(s_t)'$.

The transition probabilities are logistic functions of both lagged short rates and spreads:

$$p(s_t = j|s_{t-1} = j; \mathcal{I}_{t-1}) = \frac{\exp(a_j + b_j r_{t-1} + c_j z_{t-1})}{1 + \exp(a_j + b_j r_{t-1} + c_j z_{t-1})} \quad j = 1, 2. \tag{4}$$

where $\mathcal{I}_{t-1} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_0\}$. As with the univariate case, we consider the case of constant transition probabilities ($b_j = c_j = 0$) and time-varying transition probabilities ($b_j \neq 0$, $c_j \neq 0$).

2.3 Drift and Volatility Functions

In a simple AR(1) specification of the short rate $r_t = \mu + \rho r_{t-1} + \sigma\epsilon_t$, the conditional drift of Δr_t is given by $\mu - (1-\rho)r_{t-1}$ and the conditional volatility is given by σ. In a regime-switching model, such as our short rate univariate model, the conditional drifts and volatilities will be functions of the ex-ante probability of being in a regime $p(s_t|\mathcal{I}_{t-1})$. In this model, the conditional drift $f(r_{t-1}; \mathcal{I}_{t-1})$ is given by:

$$f(r_{t-1}; \mathcal{I}_{t-1}) = \mathrm{E}(\Delta r_t|\mathcal{I}_{t-1}) = \sum_{i=1}^{2} p_{i,t-1}(\mu_i - (1-\rho_i)r_{t-1}) \tag{5}$$

where $p_{i,t-1} = p(s_t = i|\mathcal{I}_{t-1})$ is the ex-ante probability and subscripts on μ_i and ρ_i denote the regime. Equation (5) shows the conditional drift is a weighted average of the drifts conditional on the regime.

In contrast, the conditional volatility in the regime-switching short rate model is not simply an average of the regime-dependent volatilities, as pointed out by Gray (1996). The conditional volatility $g(r_{t-1}; \mathcal{I}_{t-1})$ is given by:

$$\begin{aligned} g(r_{t-1}; \mathcal{I}_{t-1})^2 &= \mathrm{E}(r_t^2|\mathcal{I}_{t-1}) - [\mathrm{E}(r_t|\mathcal{I}_{t-1})]^2 \\ &= \left[\sum_{i=1}^{2} p_{i,t-1}[(\mu_i + \rho_i r_{t-1})^2 + \sigma_i^2]\right] - \left[\sum_{i=1}^{2} p_{i,t-1}(\mu_i + \rho_i r_{t-1})\right]^2 \\ &= \left[\sum_{i=1}^{2} p_{i,t-1}\sigma_i^2\right] + p_{1,t-1}p_{2,t-1}[(\mu_1 - \mu_2) + (\rho_1 - \rho_2)r_{t-1}]^2 \end{aligned} \tag{6}$$

The first term is the average of the regime-dependent variances, and the second term adds a jump component to the variance due to the switching effect of moving from one regime to another.

Note that the conditional mean $f(r_{t-1}; \mathcal{I}_{t-1})$ and volatility $g(r_{t-1}; \mathcal{I}_{t-1})$ are functions of both the lagged short rate and the information set at $t-1$ through the ex-ante probabilities $p_{i,t-1}$. Through time, the ex-ante probabilities are updated through a ratio of likelihoods which also depends on the transition probabilities (which can be functions of $\mathcal{I}_{t-1}$ as in our time-varying transition probability models). The ex-ante probability can be written as:

$$p_{i,t-1} = p(s_t = i|\mathcal{I}_{t-1}) = \sum_{j=1}^{2} p(s_t = i|s_{t-1} = j; \mathcal{I}_{t-1}) p(s_{t-1} = j|\mathcal{I}_{t-1}) \tag{7}$$

where the first term in the sum is the transition probability which can be state-dependent. For the univariate model the second term may be decomposed by Bayes' Rule as:

$$\begin{aligned} p(s_{t-1} = j|\mathcal{I}_{t-1}) &= \frac{\phi(r_{t-1}, s_{t-1} = j|\mathcal{I}_{t-2})}{\phi(r_{t-1}|\mathcal{I}_{t-2})} \\ &= \frac{\phi(r_{t-1}|s_{t-1} = j; \mathcal{I}_{t-2}) p(s_{t-1} = j|\mathcal{I}_{t-2})}{\sum_{m=1}^{2} \phi(r_{t-1}|s_{t-1} = m, \mathcal{I}_{t-2}) p(s_{t-1} = m|\mathcal{I}_{t-2})} \end{aligned}$$

where $\phi(r_{t-1}|s_{t-1} = j; \mathcal{I}_{t-2})$ is the conditional density of r_{t-1} given $s_{t-1} = j$. As $p_{i,t-1}$ varies through time, the regime-switching models produce both stochastic means and stochastic volatilities.

To find the drift and volatility function of Δr_{t-1} as a function of r_{t-1} only, we must integrate the effect of the path $\{r_{t-2}, r_{t-3}, \ldots, r_0\}$ out of the functions $f(r_{t-1}; \mathcal{I}_{t-1})$ and $g(r_{t-1}, \mathcal{I}_{t-1})$. The conditional drift as a function only of r_{t-1}, denoted by $f(r_{t-1})$, is given by:

$$f(r_{t-1}) = \mathrm{E}(\Delta r_t|r_{t-1}) = \sum_{i=1}^{2} p_i(\mu_i - (1-\rho_i) r_{t-1}) \tag{8}$$

where p_i is the stable probability $p_i = p(s_t = i|r_{t-1})$. The stable probability will be a function of r_{t-1} and can be evaluated numerically. The conditional volatility as a function only of r_{t-1}, denoted by $g(r_{t-1})$, is given by:

$$\begin{aligned} g(r_{t-1})^2 &= \mathrm{E}(r_t^2|r_{t-1}) - [\mathrm{E}(r_t|r_{t-1})]^2 \\ &= \left[\sum_{i=1}^{2} p_i \sigma_i^2\right] + p_1 p_2 [(\mu_1 - \mu_2) + (\rho_1 - \rho_2) r_{t-1}]^2 \end{aligned} \tag{9}$$

The above two formulae show how the drift and volatility of Δr_t in a regime-switching model can potentially be non-linear functions of r_{t-1}. Both equations (8) and (9) are functions of the regime-dependent parameters, and the stable probability p_i can depend on r_{t-1}. In particular, for the conditional drift in equation (8) the drift will be a weighted average of the drift in each regime, with the weights being the stable probabilities p_i. As p_i changes across the lagged short rate, the drift will also change.

We integrate out the path numerically as follows. We simulate out the system recording r_{t-1} and r_t. We divide the observations into bins of r_{t-1} of 50 basis points. Within each bin we calculate the values

$$\mathrm{E}(\Delta r_t | r_{t-1})$$

which is an empirical estimate of the drift $f(r_{t-1})$ and

$$\sqrt{\mathrm{E}[(\Delta r_t - \mathrm{E}(\Delta r_t | r_{t-1}))^2 | r_{t-1}]}$$

which is an estimate of the volatility $g(r_{t-1})$. We will plot the average drift and volatilities within each bin at the mid-point of the bin to obtain appropriate drift and volatilities for the regime-switching models as a function of r_{t-1}. We also record the regime realizations to estimate $\mathrm{E}(s_t | r_{t-1})$, the average regime.[7] We require over 20 million observations in the simulation to accurately pin down the the drifts and volatilities at very low and high interest rate levels.

A similar analysis applies for the bivariate model, except in addition to integrating out s_t, the spread z_t must also be integrated out to obtain drift and volatilities for the short rate as a function only of r_{t-1}.

3 Data and Estimation

3.1 Data

Our empirical work uses monthly observations on 3 month short rates and 5 year long rates of zero coupon bonds from the US, Germany and Great Britain from January 1972 to August 1996. The data is an updated set of the Jorion and Mishkin (1991) data series.[8] Table (1) reports central moments and the first three autocorrelations for short rates and spreads for each country and the correlations between these variables. We note that short rates for Germany and short rates and spreads for the UK do not show excess kurtosis. Short rates are very persistent, with the UK showing the least persistence. Spreads are also highly autocorrelated but less so than short rates. Spreads are on average lower in the UK and Germany than in the US but they are

[7] We can equivalently calculate $p_1 = \mathrm{E}(s_t = 1 | r_{t-1})$ and then use equations (8) and (9) to obtain the drift and volatility. This still involves simulating a large number of interest rate paths. The procedure taken here is easily adapted to integrating out z_t in addition to the interest rate paths in the bivariate systems.

[8] See Bekaert, Hodrick and Marshall (2000) for further details.

more variable. The correlations between short rates across countries range between 0.44 for the US and German rates and 0.67 for the US and UK short rate.

3.2 Estimation Results

We first discuss the parameter estimates for the various models in more detail in Sections 3.2.1 and 3.2.2 and then provide an economic interpretation of the main empirical patterns in Section 3.2.3.

3.2.1 Univariate Short Rate Models

The univariate short rate model coefficients are given in Tables (2), (3) and (4) for the joint, individual and US estimations, respectively. In Table (2) for the time-varying transition probability estimation, the first regime has a near-unit root ($\rho_1 = 0.9896$) with a constant within-regime mean of 0.0502 and within-regime volatility of 0.2180. In contrast, the second regime is much more mean-reverting ($\rho_2 = 0.9315$), with a higher constant ($\mu_2 = 0.7284$) and higher regime-dependent volatility ($\sigma_2 = 1.0441$). Although the state-dependent transition probabilities make these numbers harder to interpret, the within-regime means ($\mu_i/(1 - \rho_i)$) do not change very much relative to the constant transition probability model. They are 4.83% for the first and 10.63% for the second regime.

Figure (1) shows the transition probabilities from the joint univariate short rate model with time-varying probabilities. The transition probabilities are significantly affected by the lagged short rate. In particular, in the first regime b_1 is significantly negative, so as interest rates rise in the first regime there is an increased probability of transitioning to the second regime. In the second regime, as short rates rise, there is an increased chance of remaining in the higher volatile regime ($b_2 > 0$). A likelihood ratio test rejects the null hypothesis of constant transition probabilities (p-value = 0.0000). The other parameters for the constant transition probability model are very similar to the ones obtained in the time-varying transition probability model.

In Table (3) we show results for the joint constant probability model estimated allowing for different parameters across countries. The parmeter patterns across regimes are very similar, but we still strongly reject the overall equality of the parameters across countries. Individual Wald tests for the different parameters suggest that the rejection is likely driven by the different volatility levels across countries. Nevertheless, given the qualitative analogy between the parameters across countries, we will continue to focus most of our attention on the joint model with equal parameters. Comparing the standard errors between Table (2) and (3), the efficiency gain in imposing equality of parameters is clear. This efficiency gain is even more apparent for the time-varying probability model, as is illustrated for the US univariate model reported in Table (4). The parameter patterns are again similar, but the transition probability parameters are estimated very imprecisely whereas in the joint model we obtain significance for both parameters capturing the state dependence of the transition probabilities. We also note that the first regime in the US has a true unit root in the time-varying transition probability model, but is also extremely persistent ($\rho_1 = 0.9980$) in the constant transition probability model. Given the parameter differences we uncover across countries, we will also report drift and volatility functions for the US estimation.

3.2.2 Bivariate Term Spread Models

The bivariate term spread model coefficients are listed in Tables (5) and (6). The results share many characteristics with the univariate results. In addition to the constant in the short rate equation being lower (higher) in the first (second) regime, the constant in the spread equation is positive in the first regime and negative in the second. The high short rate regime is associated with negative term spreads. The persistence of the system in each regime can be seen by looking at the moduli of the eigenvalues for the companion matrices A_i. For the time-varying probabilty system estimated over the US, UK and Germany in Table (5), the moduli of the eigenvalues for A_1 are 0.9870 and 0.9423 and for A_2 they are 0.9145 (repeated), showing the first regime to be more persistent. In the first regime, both the short rate and spread Granger-cause each other (significant p-values for $A_1[1,2]$ and $A_1[2,1]$), but in the second regime we cannot reject the hypothesis that lagged short rates and spreads do not affect each other.

Table (5) also shows the conditional covariances to be larger in the second regime. Pfann, Schotman and Tschernig (1996) show that the correlations between the short rate and spread change as the level of the short rate changes. The implied correlations conditional on the regime of the short rate and spread are -0.4375 (-0.7673) in regime 1 (2) for the joint bivariate model with time-varying probabilities. (For comparison the unconditional correlation from the data is -0.5838.) A Wald test rejects the hypothesis that the correlations are equal across regimes with a p-value of 0.0000.

Finally, the coefficients in the transition probabilities are significant at the 1% level in the first regime. As both short rates and spreads increase, the probability of staying in the first regime decreases (positive a_1 and b_1). In the second regime, only the coefficient on the lagged short rate is border-line significant (p-value of 0.0542) and the p-value of the second coefficient is slightly over 0.11. However, the positive point estimates indicate that as both short rates and spreads increase the probability of staying in the second regime increases. A likelihood ratio test rejects the null hypothesis of constant transition probabilities with a p-value of 0.0000.

The US bivariate term spread model parameters are presented in Table (6). The results are qualitatively similar to the joint estimation results. The time-varying probability model failed to converge due to insufficient data, especially at higher interest rates. Instead, we report a restricted model where the transition probabilities depend only on the spread. A likelihood ratio test for coefficients on the lagged spread to be equal to zero in the transition probabilities is rejected with a p-value of 0.0167.

3.2.3 Interpretation of the Results

The estimation results are characterized by one regime producing unit root, or near-unit root behavior with lower conditional volatility, and a second regime which is more mean-reverting with higher conditional volatility. This result is shared across the univariate short rate and bivariate term spread models.

Economists such as Mankiw and Miron (1986) argue that the smoothing actions of the US Fed make the short rate behave like a random walk, and the first regime corresponds to "normal" periods. When extraordinary shocks occur, interest rates switch to the second regime and are driven up, volatility becomes higher and interest rates become more mean-reverting. In the second regime, policy makers switch from interest-rate smoothing to inflation fighting. This economic underpinning of the estimation results is particularly plausible for the US. One of the marked high inflation rate episodes was 1979-1982, and the high interest rates were clearly partially caused by high inflation. This logic could also motivate the dependence of the transition probability on the level of the short rate. Since short rates reflect expected inflation, they signal to the authorities that a high inflation regime is likely to occur.

To show the plausibility of this economic rationale for the regime switching model, Figure (2) shows the smoothed probabilities (the probability of the high interest rate regime) for the three countries over the sample period together with short rates and inflation rates. The model used is the spread model jointly estimated over the US, UK and Germany, because the cross-sectional estimation allows more power to pin down the coefficients of the non-linear regime-switching models, particularly the coefficients of the transition probabilities.

For the US, the major inflationary periods around 1975 and 1979-1982 indeed coincide with high probabilities for the high-interest rate regime. The fit is not perfect, as was also pointed out by Evans and Lewis (1995), and the correlation between inflation and smoothed probabilities is 0.54 in the US. At 0.50, this correlation is very similar in Germany. Apart from the mid-seventies and early eighties, Germany also witnessed a relatively high-inflation period after the unification with East Germany in the early nineties. All of these periods coincide with high probabilities for the high-interest rate regime. The worst fit is for the UK. The regime switching model switches regimes so often that it is difficult to distinguish high-interest rate episodes. This being said, inflation is also more variable and higher in the UK than in the US or Germany, and the correlation between inflation and the high interest regime smoothed probability remains solidly positive at 0.36.

4 Implied Nonlinearities of Regime-Switching Models

We organize our discussion into four sub-sections. The first subsection reviews the results reported in the previous literature. Section 4.2 discusses our results for the drift and volatility functions for the short rate from our models estimated using data from all countries. Section 4.3 reports the drifts and volatility functions for the spread. Finally, in Section 4.4 we check robustness by graphing the drift and volatility functions for the models estimated using US data only.

4.1 Literature Review

There is now a large literature documenting empirical non-linearities in interest rates. Aït-Sahalia (1996a) parametrically specifies the drift and volatility functions for US 7-day Eurodollar spot rate changes using non-linear functions. Aït-Sahalia finds a highly non-linear drift with strong mean-reversion at very low and high interest rates but the drift is essentially zero in the middle region. The volatility function assumes a J-shape so the spot rate is more volatile outside the middle region, with the highest volatility occurring at very high interest rates. We reproduce his findings in Figure (3).[9] Aït-Sahalia uses the following parameterizations for the drift and volatility functions:

$$
\begin{aligned}
f^A(r_{t-1}) &= \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + \alpha_3 / r_{t-1} \\
g^A(r_{t-1})) &= \sqrt{\beta_0 + \beta_1 r_{t-1} + \beta_2 r_{t-1}^{\beta_3}} \qquad (10)
\end{aligned}
$$

The inverse α_3/r_{t-1} in the conditional mean imparts an asymptote of the short rate drift at zero. Aït-Sahalia estimates β_3 as approximately 2, which imparts a quadratic shape to the conditional variance in Figure (3).

[9] In Figure (3) the Figure 4c of Aït-Sahalia (1996a) shows the estimated diffusion of Δr_t, which is the conditional variance.

Conley et al. (1997)'s drift estimations on overnight Fed funds rate changes look very similar to Aït-Sahalia's plots, but without the strong mean reversion at high interest rates. In their formulation, stationarity at high interest rates is induced by increasing volatility. Stanton (1997)'s non-parametrically estimated drift on daily 3 month T-bill rate changes is zero until high interest rates where the drift becomes very negative. Stanton's non-parametrically estimated volatility looks very similar to Aït-Sahalia's, with volatility increasing at higher levels of interest rates. However, Stanton finds the volatility to be mostly convex over most short rate levels, except at very high short rates where the conditional volatility becomes concave, and the lowest volatility happens at the lowest rates. Stanton's pictures differ somewhat from Aït-Sahalia (1996a) as he does not impose a parametric form onto the drift and volatility functions. Aït-Sahalia (1996a)'s parameterizations may not be sufficiently rich enough to capture the non-parametric shapes which Stanton finds. We reproduce Stanton's figures in Figure (4).

These findings suggest that interest rates exhibit strong non-linear drifts, with the drift being zero over much of the support of the data, but strongly mean-reverting at low or high interest rates. The volatility of interest rates generally increases with the level of the interest rate with the lowest volatility appearing in the low to middle range of the support.[10] In the next section we show that the drift and volatility functions implied by the regime-switching models can mimic these features.

We note that regime-switching models may not be the only models which can reproduce the non-parametric estimations of the drift and volatility. If multiple factors drive the term structure, then conditioning only on the short rate may induce apparent non-linearities in the drift and volatility functions because these now depend on the entire history of the short rate process. However, certain classes of multi-factor models can be ruled out. For example, in the Duffie-Kan (1996) affine class the short rate can always be written as a linear function of factors. In this case, the drift will be a linear function of a single state variable, after integrating out other state variables. In addition, in multi-factor CIR models the factors must be positively correlated to ensure admissibility. Single-factor models without regime-switching may also capture some non-linearities. For example, the constant elasticity of volatility model of Chan, Karolyi, Longstaff and Sanders (1992) is able to capture the convex shape of the conditional volatility by specifying the conditional variance to be a cubic function. However, these models still parameterize the conditional drift to be affine. Ahn and Gao (1999) use a quadratic drift, but this model does not satisfactorily capture the highly mean-reverting drift at high interest rates levels. Regime-switching models can reproduce the non-parametric estimations of both

[10] Johannes (1999) shows that the high mean-reversion at high interest rates survives the addition of jumps in the short rate process, but introducing jumps causes the conditional volatility to appear more linear.

the conditional mean and conditional volatility.

4.2 Short Rate Drift and Volatilities from the Joint Estimation

4.2.1 Drift Function

Figure (5) reports the short rate drift function for the four models we estimate: the univariate short rate model with constant and time-varying transition probabilities, and the bivariate short rate-term spread models with constant and time-varying probabilities. Figure (5) shows very clearly that the feature of the model which drives the shape of the drift function is the time-varying logistic transition probability. The models with constant transition probabilities produce a drift function that is nearly linear, except at the very edges where some slight curvature is present which is due to sampling error.[11] This kind of function could be well approximated by a simple linear autoregressive model. The drift functions from the models with the time-varying transition probabilities on the other hand closely resemble the drift presented in Stanton (1997).

The drift function is at first very flat (but downward sloping) until an interest rate of about 10% and then turns steeply negative at higher interest rates. Stanton's drift starts turning negative around 14%. For most of the range, the function looks similar to Aït-Sahalia's (1996a) but the drift there starts turning negative at only 18%. We do not generate the non-linearity he produces at very low short rates, but that is exactly the range of interest rates where the biases documented by Chapman and Pearson (2000) play a large role. Aït-Sahalia's increasing drift at low interest rates may also be due to the α_3/r_{t-1} function he imposes in the conditional mean (see equation (10)). Despite some small differences in the details, which are perhaps due to the use of different data sets, what is remarkable here is that the state-dependent probability model can reproduce the shape of the non-parametrically estimated drift functions.

What drives the drift function in the regime-switching models? The drift function for the regime-switching process is a weighted average of the linear drift functions in each regime, with the weights determined by the different amount of time spent in each regime at different interest rate levels. The non-linearity is induced by this weighting. When transition probabilities are constant, there is little scope for non-linearity and the drift function retains a fairly linear shape. However, in the state-dependent model, as interest rates increase a much faster transition into the stronger mean-reverting regime occurs and much more time is spent in the second regime at higher short rates. This non-linear weighting is illustrated in Figure (6). The dotted line

[11] At very low or very high interest rates, because these are in the extremes of the interest rate distribution, not many observations are present: for example, in the constant probability bivariate spread model, the proportion of observations lying in the range 18-20% is simulated to be only 0.0018.

represents the linear drift in the first regime, which features a near-unit root and hence very flat drift function. The dashed line represents the linear drift in the second much more mean -reverting regime, where the drift slopes downwards much more steeply. At low interest rates, the first regime totally dominates and the drift function is flat, but when interest rates reach 10%, it becomes more and more likely that the interest rates are drawn from the second regime and the function curves steeply downward. The non-parametric result found before can now potentially be given an economic explanation. It is well known that monetary authorities smooth interest rates, as documented by Mankiw and Miron (1986)). This explains the near-unit root behavior of short rates in the normal range giving rise to an almost flat drift function. It takes large shocks to bring interest rates outside this range and once there both policy actions and mean reversion in fundamentals (inflation for example) bring about rapid mean reversion.

4.2.2 Volatility Function

Figure 7 plots the conditional volatility functions of the short rate for the same four models considered in Figure (5). The constant probability universe and bivariate models yield similar shapes for the conditional volatility and bear a strong resemblance to Aït-Sahalia (1996a)'s J-shaped estimations (see Figure (3)). The time-varying probability models produce a drawn-out S-shape volatility. The volatility first is rather flat at very low interest rates (0-4%) and then increases strongly with interest rates, and finally flattens out around 15%. This shape is very similar to what Stanton (1997) finds (see Figure (4)). As in the case of the drift functions, the time-varying probabilities allow more non-linearities than the constant transition probabilities.

What drives the volatility functions in the regime-switching models? Although the conditional volatility is a complex function of the model parameters (see Section 2), a major component is a simple weighting of the two volatility parameters in the two regimes, which are assumed constant. Hence, the expected regime also plays a critical role in driving the volatility function. This explains why the range of volatilities reached in the constant probability model is much lower than in the time-varying probability model, since the weighting of the two regimes does not vary very much with the interest rate. Figure (8) explicitly shows the link between the shape of the volatility function and $\mathrm{E}(s_t)$ the expected value of the regime (between 1, the low-volatility regime and 2, the high-volatility regime), which was also recorded during our simulations. The model used in Figure (8) is the univariate short rate model with time-varying probabilities. The "Average Regime" function has the exact same shape as the volatility function does.

Figure (8) shows that we are more likely to be in the first low-volatility regime at low short rates (0-8%) while at high short rates (above 14%) we are much more likely to be in the second high-volatility state. This pattern is expected from the state-dependent transition probabilities: recall Figure (1) which plots the transition probabilities as a function of the lagged short rate for the same model as Figure (8). At low interest rates, the probability of remaining in the low-volatility regime is high and as the short rate increases, the probability of transitioning to the second regime increases steeply. At high interest rates, the probability of remaining in the second regime is high, while at low interest rates if we are in the second regime the probability of transitioning to the first regime increases.

Note that the left column of Figure (7) shows that the constant transition probability models show a J-shape, which is naturally shared by the expected regime as a function of r_{t-1} (not shown). There is not much differentiation of the regimes across short rates in these models. The minima of the volatility curves coincide with the lowest values of $\mathrm{E}(s_t)$, where the model is most likely to be in the low-volatility regime. The flatness of the curves derives from a large and a relatively high stable probability of being in this regime. This means that the mixing of the two different regimes does not result in noticeably different modes in the distribution. For example, for the univariate short rate model, the regime-dependent unconditional volatility in regime 2 is 2.8344 (versus 0.6523 in regime 1), and the stable probability of being in regime 2 is 0.3419.[12] This high-volatility distribution tends to flatten the effect of the more concentrated first distribution. In contrast, the state dependence of the time-varying probability models enables the expected regime to be more differentiated across the support of the short rate.

4.3 Spread Drift and Volatility Functions from the Joint Estimation

The bivariate regime-switching models permit investigating the drifts and volatilities of term spreads. These are presented in Figure (9). In the constant probability model, the drift function for the spread is linear and downward sloping, reflecting the higher mean reversion found for spreads. The time-varying probability model generates a drift function that is very similar in shape to the one from the constant probability model, with only very slight, almost imperceptible non-linearities. This is not so surprising since the spread dynamics do not differ very much across regimes (See Table (5)).

However, spreads are much more variable in the second regime and the transition probabilities depend on the spread in the time-varying transition probability model. Hence, we would expect more divergence between the two models for the volatility function. Figure (9) confirms this. For the constant probability model, the volatility function has a U-shape. For normal

[12] The regime-dependent unconditional variance is $\sigma_i^2/(1-\rho_i^2)$, and the stable probability is $(1-Q)/(2-P-Q)$.

levels of the spread (slightly negative to slightly over 1%), volatility is low, but high and low spreads are both associated with higher spread volatility. In other words, high-volatility short rate regimes can coincide both with upward and downward sloping yield curves. In the time-varying transition probability model, the spread volatility curve is no longer symmetric and its range is larger. In particular, volatility decreases until spreads are above the normal range and then increases rapidly, but spread volatility at unusually large negative spreads is higher than at unusually large positive spreads.

The mechanism for this result lies in the state-dependence of the transition probabilities. Negative term spreads typically coincide with very high short rates and hence one is likely in the high volatility regime. As spreads tend to their normal range, short rates drop and the probability of switching into the normal regime increases. This is only partially counteracted by the fact that negative spreads generally decrease the probability of staying in the high variance regime, an effect that weakens at higher, less negative spreads. At normal (positive) levels of the spread, the economy is in the first unit-root regime, and volatility is low. When spreads increase, the probability of staying in this low-volatility regime decreases (note the negative c_1 coefficient in equation (4)). Since the probability of switching into the high variance regime increases the volatility function slopes upward again. The curve is steeper here, since the state-dependence in the first regime is stronger (see Table (5)).

4.4 Drift and Volatility Functions for the US

Given the difficulty in estimating regime-switching models in general, and the lack of extreme interest rate data in the US, we feel that the joint estimation is the most reliable guide towards the true shape of the drift and volatility functions. For completeness, Figures (10) and (11) plot drift and volatility functions for the estimations that use only US data. Because of the superiority of the time-varying transition probability model, we only show results for that model.

Figure (10) shows results for the univariate model. Even for US data, the drift function takes the Stanton (1997) form, with a flat part and then a steeply downward sloping part, for interest rates beyond 9 to 10%. The volatility curve also looks like what Stanton finds, where the low and high interest range portions are now virtually entirely flat. The results for the bivariate model, reported in Figure (11), are less encouraging, but have to be interpreted with caution. Remember that we were not successful in obtaining convergence using the short rate in the transition probability function so the state dependence here is of a different, less complete form than in the full model we discussed before. As a consequence, the drift function is linear and downward sloping and the volatility function U-shaped. For spreads, we do get functions that resemble the ones reported in the joint estimation. In particular, the drift function is near linear, and the volatility function has an asymmetric U-shape, with two distinctions relative to our previous findings. First, the asymmetry in volatility between low and high spreads is much more striking. This is primarily due to the fact that the identification of the negative spread part of the function for US data is limited to the 1979-1982 monetary targeting period. Second, there is an almost flat portion at very low and high spreads, which is due to the expected regime converging to remain close to 2 at these values.

5 Conclusion

An economically intuitive regime-switching model replicates the non-linear patterns in the drift and volatility functions of short rates found by non-parametric studies. The critical feature of the model that generates the required non-linearities is the state dependence of the transition probabilities. In our univariate short rate regime-switching models these probabilities depend logistically on the lagged level of the short rate. In our bivariate model using the short rate and term spread, these are logistic functions of lagged short rates and spreads. At a detailed level, our empirical findings are remarkably close to the non-parametric findings of Stanton (1997). Although the shapes we found are generally similar to Aït-Sahalia (1996a)'s findings, there are some differences, especially at very low interest rates. Since there are few data points at these ranges, we suspect that these differences may be generated by the semi-non-parametric nature of Aït-Sahalia (1996a)'s method, which imposes a parametric form on the drift and volatility function.

Our results have several implications. First, for the term structure literature it is important to build models that embed these non-linearities. It is unlikely for example that affine models can ever generate the non-linearities documented in both the drift and volatility functions. This paper has shown that models with regime switches, or perhaps jumps are rich enough to mimic the non-parametrically estimated short rate drift and volatility. Although some progress has been made in this area, much more remains to be done. Naik and Lee (1994) and Veronesi and Yared (1999) develop a continuous-time regime-switching term structure model, but allow switching only in the conditional mean. Bansal and Zhou (1999) and Evans (1999) have recently developed discrete-time term structure models accommodating regime switches, but they do not allow for state-dependent transition probabilities and their solution technique makes some strong assumptions. Consequently, the extant models are unlikely to generate the required non-linearities.

Second, our results have implications for the vast macro-economic literature on the effects of policy shocks on the economy and asset prices (see for example, Gali (1992)). Identifying policy shocks from linear VAR's seems very inappropriate given the dynamics of interest rates illustrated here. Since the regimes may well be caused by changes in monetary policy operating procedures, policy analysis should take these regime dynamics into account. Ang and Bekaert (1998) specifically consider how dynamic impulse responses may differ between a simple VAR and a regime-switching VAR.

Finally, although the non-linear patterns in itself are of much interest, eventually we should attempt to understand what economic forces drive them. Are they induced by shifts in expected inflation or low-frequency changes in real rates? Evans and Lewis (1995) document that regime changes in inflation do not perfectly coincide with regime changes in interest rates, and on-going work by Bekaert and Marshall (2000) attempts to trace regime changes to changes in real rate or inflation regimes.

References

[1] Ahn, D. H., and B. Gao, 2000, “A Parametric Nonlinear Model of Term Structure Dynamics,” *Review of Financial Studies*, 12, 4, 721-762.

[2] Aït-Sahalia, Y., 1996a, “Testing Continuous-Time Models of the Spot Interest Rate”, *Review of Financial Studies*, 9, 2, 385-426.

[3] Aït-Sahalia, Y., 1996b, “Nonparametric Pricing of Interest Rate Derivative Securities”, *Econometrica*, 64, 3, 527-560.

[4] Ang, A., and G. Bekaert, 1998, “Regime Switches in Interest Rates”, working paper.

[5] Ball, C. A., and W. N. Torous, 1999, “The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence”, *Journal of Finance*, 54, 6, 2339-2359.

[6] Bansal, R., and H. Zhou, 1999, “Term Structure of Interest Rates with Regime Shifts”, working paper.

[7] Beaglehole, D., and M. Tenney, 1992, “A Nonlinear Equilibrium Model of the Term Structure of Interest Rates: Corrections and Additions”, *Journal of Financial Economics*, 32, 345-354.

[8] Bekaert, G., and D. A. Marshall, 2000, “Inflation and Real Rate Regimes in Interest Rates”, work in progress, Federal Reserve Bank of Chicago.

[9] Bekaert, G., R. J. Hodrick, and D. A. Marshall, 2000, “ “Peso Problem” Explanations for Term Structure Anomalies”, working paper.

[10] Boudoukh, J., M. Richardson, R. Stanton, and R. Whitelaw, 1999, “A Multifactor, Nonlinear, Continuous-Time Model of Interest Rate Volatility”, NBER working paper 7213.

[11] Brandt, M., 1999, “Regime Switching or Nonlinear Diffusions? An Efficient Method of Moments Analysis,” working paper.

[12] Chan, K. C., A. Karolyi, F. A. Longstaff, and A. B. Sanders, 1992, “An Empirical Comparison of Alternative Models of the Short-Term Interest Rate,” *Journal of Finance*, 47, 1209-1228.

[13] Chapman, D. A., and N. D. Pearson, 2000, “Is the Short Rate Drift Actually Nonlinear?”, *Journal of Finance*, 55, 1, 355-388.

[14] Conley, T. G., L. P. Hansen, E. G. J. Luttmer, J. A. Scheinkman, 1997, “Short-Term Interest Rates as Subordinated Diffusions”, *Review of Financial Studies*, 10, 3, 525-577.

[15] Constantinides, G., 1992, “A Theory of the Nominal Structure of Interest Rates”, *Review of Financial Studies*, 5, 531-552.

[16] Diebold, F. X., J. H. Lee, and G. C. Weinbach, 1994, “Regime switching with time-varying transition probabilities”, in Hargreaves, C., ed., *Time Series Analysis and Cointegration*, Oxford University Press.

[17] Downing, C., 1999, “Nonparametric Estimation of Multifactor Continuous Time Interest Rate Models,” working paper.

[18] Duffie, D., and R. Kan, 1996, “A Yield-Factor Model of Interest Rates”, *Mathematical Finance*, 6, 379-406.

[19] Evans, M. D. D., 1999, “Regime Shifts, Risk and the Term Structure”, working paper.

[20] Evans, M. D., and K. K. Lewis, 1995, “Do Expected Shifts in Inflation Affect Estimates of the Long-Run Fisher Relation?” *Journal of Finance*, 50, 1, 225-253.

[21] Gali, J., 1992, "How Well Does the IS-LM Model Fit Postwar US Data?", *Quarterly Journal of Economics*, 107, 709-738.

[22] Gray, S. F., 1996, "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process", *Journal of Financial Economics*, 42, 27-62.

[23] 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.

[24] Hamilton, J. D., 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, 57, 357-384.

[25] Hamilton, J. D., 1994, *Time Series Analysis*, Princeton University Press.

[26] Holst, U., G. Lindgren, J. Holst and M. Thuvesholmen, 1994, "Recursive Estimation in Switching Autoregressions with a Markov Regime", *Journal of Time Series Analysis*, 15, 5, 489- 506.

[27] Johannes, M. S., 1999, "Jumps in Interest Rates: A Nonparametric Approach," working paper.

[28] Jones, C. S., 2000, "Nonlinear Mean Reversion in the Short-Term Interest Rate," working paper.

[29] Jorion, P., and F. Mishkin, 1991, "A Multicountry Comparison of Term-Structure Forecasts at Long Horizons", *Journal of Financial Economics*, 29, 59-80.

[30] Lewis, K. K., 1991, "Was There a "Peso Problem" in the U.S. Term Structure of Interest Rates: 1979-1982?", *International Economic Review*, 32, 159-173.

[31] Mankiw, N. G., and J. A. Miron, 1986, "The Changing Behavior of the Term Structure of Interest Rates", *Quarterly Journal of Economics*, 101, 211-228.

[32] Naik, V., and M. H. Lee, 1994, "The Yield Curve and Bond Option Prices with Discrete Shifts in Economic Regimes", working paper.

[33] Pfann, G. A., P. C. Schotman, and R. Tschernig, 1996, "Nonlinear interest rate dynamics and implications for the term sturcture", *Journal of Econometrics*, 74, 149-176.

[34] Pritsker, M., 1998, "Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models", *Review of Financial Studies*, 11, 3, 449-487.

[35] Sola, M., and J. Driffill, 1994, "Testing the Term Structure of Interest Rates Using a Vector Autoregressive Model with Regime Switching", *Journal of Economic Dynamics and Control*, 18, 601-628.

[36] Stanton, R., 1997, "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk", *Journal of Finance*, 52, 5, 1973-2002.

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

[38] Veronesi, P., and F. Yared, 1999, "Short and Long Horizon Term and Inflation Risk Premia in the US Term Structure: Evidence from an Integrated Model for Nominal and Real Bond Prices Under Regime Shifts," working paper.

Table 1: Summary Statistics of Data

	US short rate	US spread	GER short rate	GER spread	UK short rate	UK spread
Central Moments						
Means	7.1104	1.2443	6.6350	0.6349	10.1008	0.2456
Stdev	2.8206	1.3776	2.6650	1.7285	3.0524	1.6802
Skew	0.9668	-0.7596	0.7878	-0.6988	-0.0076	-0.3503
Kurt	3.8689	3.7987	2.8478	3.5518	2.1834	2.7805
Autocorrelations						
ρ_1	0.9743	0.8689	0.9836	0.9670	0.9690	0.9380
ρ_2	0.9411	0.7698	0.9552	0.9204	0.9318	0.8866
ρ_3	0.9113	0.6982	0.9212	0.8708	0.8891	0.8351

Cross-correlations

	US short rate	US spread	GER short rate	GER spread	UK short rate
US spread	-0.5838				
GER short rate	0.4372	-0.3663			
GER spread	-0.2824	0.3350	-0.8861		
UK short rate	0.6706	-0.4289	0.4590	-0.3066	
UK spread	-0.3101	0.3216	-0.3245	0.3372	-0.7891

Sample period January 1972 to September 1996. Short rates are 3 month short rates, spreads are the difference between 5 year zero coupon bond long rates and the short rate. The i-th autocorrelation is denoted by ρ_i.

Table 2: Univariate Short Rate Model: Joint Estimation

Constant Probs				Time-Varying Probs			
Parameter	Estimate	Std Error	p value	Parameter	Estimate	Std Error	p value
μ_1	0.0482	0.0274	0.0789	μ_1	0.0502	0.0274	0.0670
μ_2	0.7216	0.2188	0.0010	μ_2	0.7284	0.2276	0.0014
ρ_1	0.9892	0.0004	0.0000	ρ_1	0.9896	0.0004	0.0000
ρ_2	0.9320	0.0055	0.0000	ρ_2	0.9315	0.0056	0.0000
σ_1	0.2246	0.0122	0.0000	σ_1	0.2180	0.0111	0.0000
σ_2	1.0270	0.0500	0.0000	σ_2	1.0441	0.0513	0.0000
P	0.9128	0.0210	0.0000	a_1	4.0530	0.5385	0.0000
				b_1	-0.2700	0.0655	0.0000
Q	0.8322	0.0411	0.0000	a_2	-0.0124	0.0916	0.8920
				b_2	0.1182	0.0281	0.0000

Model given by: $r_t = \mu(s_t) + \rho(s_t)r_{t-1} + \sigma(s_t)\epsilon_t$, $\epsilon_t \sim$ IID $N(0,1)$. The transition probabilities are given by $p(s_t = 1|s_{t-1} = 1) = P$ and $p(s_t = 2|s_{t-1} = 2) = Q$ in the case of the constant probability model, and $p(s_t = i|s_{t-1} = i; \mathcal{I}_{t-1}) = \frac{e^{a_i + b_i r_{t-1}}}{1+e^{a_i+b_i r_{t-1}}}$, $i = 1, 2$. for the time-varying probability model. Subscripts in the Table denote the regime. The models are estimated cross-sectionally across the US, Germany and the UK. A likelihood ratio test for $b_i = 0$ yields a p-value of 0.0000.

Table 3: Univariate Short Rate Model: Different Parameters Across Countries

	US results		GER results		UK results		Wald Test
	Param	SE	Param	SE	Param	SE	p-value
μ_1	0.0426	0.0789	-0.0283	0.0135	-0.0106	0.0492	0.6579
μ_2	0.6847	0.5754	0.4157	0.2339	1.2308	0.4363	0.2544
ρ_1	0.9980	0.0129	1.0000	0.0000	0.9926	0.0050	0.9897
ρ_2	0.9265	0.0529	0.9536	0.0267	0.8956	0.0396	0.8006
σ_1	0.2849	0.0162	0.1677	0.0104	0.2019	0.0161	0.0000
σ_2	1.2552	0.1178	0.7114	0.0567	1.1435	0.0818	0.0000
P	0.9782	0.0112	0.9407	0.0237	0.6882	0.0524	0.8078
Q	0.9216	0.0371	0.8780	0.0514	0.5630	0.0812	0.7830

Model given by: $r_t = \mu(s_t) + \rho(s_t)r_{t-1} + \sigma(s_t)\epsilon_t$, $\epsilon_t \sim$ IID $N(0,1)$. The transition probabilities are given by $p(s_t = 1|s_{t-1} = 1) = P$ and $p(s_t = 2|s_{t-1} = 2) = Q$. Each country has different parameters. Subscripts in the Table denote the regime. The models are estimated cross-sectionally across the US, Germany and the UK. The last column lists a p-value of a Wald Test of parameter equality across the US, Germany and UK. A likelihood ratio test for parameters being the same across the US, Germany and UK yields a p-value of 0.0000.

Table 4: Univariate Short Rate Model: US Estimation

Constant Probs				Time-Varying Probs			
Parameter	Estimate	Std Error	p value	Parameter	Estimate	Std Error	p value
μ_1	0.0426	0.0789	0.5135	μ_1	0.0364	0.0206	0.0776
μ_2	0.6847	0.5754	0.2341	μ_2	0.6291	0.6275	0.3161
ρ_1	0.9980	0.0005	0.0000	ρ_1	1.0000	0.0000	0.0604
ρ_2	0.9265	0.0142	0.0000	ρ_2	0.9306	0.0149	0.0000
σ_1	0.2849	0.0162	0.0000	σ_1	0.2863	0.0175	0.0000
σ_2	1.2552	0.1178	0.0000	σ_2	1.2869	0.1325	0.0000
P	0.9782	0.0112	0.0000	a_1	7.4678	2.6077	0.0042
				b_1	-0.5590	0.2945	0.0577
Q	0.9216	0.0370	0.0000	a_2	-7.1864	5.0917	0.1581
				b_2	1.1202	0.7292	0.1245

Model given by: $r_t = \mu(s_t) + \rho(s_t)r_{t-1} + \sigma(s_t)\epsilon_t$, $\epsilon_t \sim$ IID $N(0,1)$. The transition probabilities are given by $p(s_t = 1|s_{t-1} = 1) = P$ and $p(s_t = 2|s_{t-1} = 2) = Q$ in the case of the constant probability model, and $p(s_t = i|s_{t-1} = i; \mathcal{I}_{t-1}) = \frac{e^{a_i + b_i r_{t-1}}}{1+e^{a_i+b_i r_{t-1}}}$, $i = 1, 2$. for the time-varying probability model. Subscripts in the Table denote the regime. The models are estimated on US data. A likelihood ratio test for $b_i = 0$ yields a p-value of 0.0001.

Table 5: Bivariate Spread Model: Joint Estimation

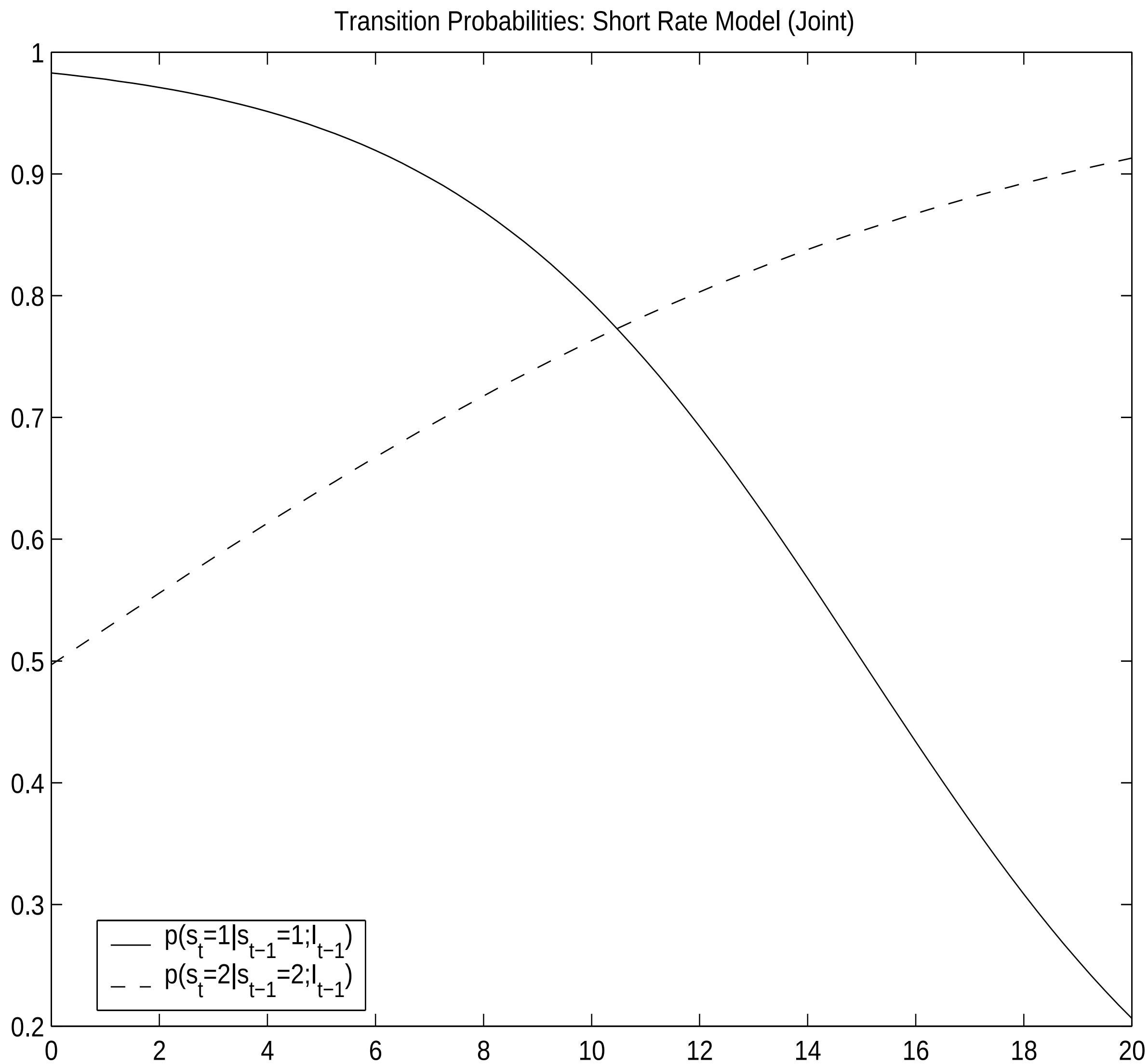
Constant Probs				Time-Varying Probs			
Parameter	Estimate	Std Error	p value	Parameter	Estimate	Std Error	p value
μ_{11}	-0.0985	0.0603	0.1024	μ_{11}	-0.0908	0.0620	0.1429
μ_{12}	0.2507	0.0700	0.0003	μ_{12}	0.2498	0.0709	0.0004
μ_{21}	0.8142	0.3013	0.0069	μ_{21}	0.9232	0.3395	0.0065
μ_{22}	-0.1649	0.2631	0.5308	μ_{22}	-0.2419	0.2877	0.4004
$A_1[1,1]$	1.0048	0.0074	0.0000	$A_1[1,1]$	1.0040	0.0077	0.0000
$A_1[1,2]$	0.0430	0.0136	0.0015	$A_1[1,2]$	0.0427	0.0136	0.0016
$A_1[2,1]$	-0.0255	0.0086	0.0032	$A_1[2,1]$	-0.0246	0.0088	0.0050
$A_1[2,2]$	0.9271	0.0156	0.0000	$A_1[2,2]$	0.9253	0.0156	0.0000
$A_2[1,1]$	0.9243	0.0281	0.0000	$A_2[1,1]$	0.9156	0.0309	0.0000
$A_2[1,2]$	-0.0286	0.0438	0.5141	$A_2[1,2]$	-0.0375	0.0473	0.4281
$A_2[2,1]$	0.0178	0.0247	0.4708	$A_2[2,1]$	0.0234	0.0265	0.3777
$A_2[2,2]$	0.9072	0.0385	0.0000	$A_2[2,2]$	0.9124	0.0411	0.0000
$R_1[1,1]$	0.2313	0.0119	0.0000	$R_1[1,1]$	0.2332	0.0137	0.0000
$R_1[1,2]$	-0.1110	0.0153	0.0000	$R_1[1,2]$	-0.1223	0.0166	0.0000
$R_1[2,2]$	0.2504	0.0111	0.0000	$R_1[2,2]$	0.2514	0.0096	0.0000
$R_2[1,1]$	1.0318	0.0529	0.0000	$R_2[1,1]$	1.0614	0.0531	0.0000
$R_2[1,2]$	-0.6974	0.0491	0.0000	$R_2[1,2]$	-0.7130	0.0508	0.0000
$R_2[2,2]$	0.5789	0.0274	0.0000	$R_2[2,2]$	0.5960	0.0289	0.0000
P	0.9017	0.0182	0.0000	a_1	7.1210	1.2215	0.0000
				b_1	-0.6146	0.1283	0.0000
				c_1	-0.5885	0.2152	0.0062
Q	0.8068	0.0408	0.0000	a_2	-1.1749	1.1249	0.2963
				b_2	0.2074	0.1077	0.0542
				c_2	0.2417	0.1541	0.1169

Model is $Y_t = A(s_t)Y_{t-1} + U_t$, $U_t \sim$ IID $N(0, \Sigma(s_t))$ with $Y_t = (r_t\, z_t)'$, the short rate and spread, and $R(s_t) = \text{chol}(\Sigma(s_t))$, $s_t = 1, 2$. For the constant probability model $p(s_t = 1|s_{t-1} = 1) = P$ and $p(s_t = 2|s_{t-1} = 2) = Q$. For the time-varying probability model, $p(s_t = i|s_{t-1} = i; \mathcal{I}_{t-1}) = \frac{\exp(a_i + b_i r_{t-1} + c_i z_{t-1})}{1+\exp(1+a_i+b_i r_{t-1}+c_i z_{t-1})}$ $i = 1, 2$. The first subscript denotes the regime s_t, and the second subscript (or numbers in square brackets) denote the matrix element. The models are estimated cross-sectionally across the US, Germany and the UK. A likelihood ratio test for $b_i = c_i = 0$ yields a p-value of 0.0000.

Table 6: Bivariate Spread Model: US Estimation

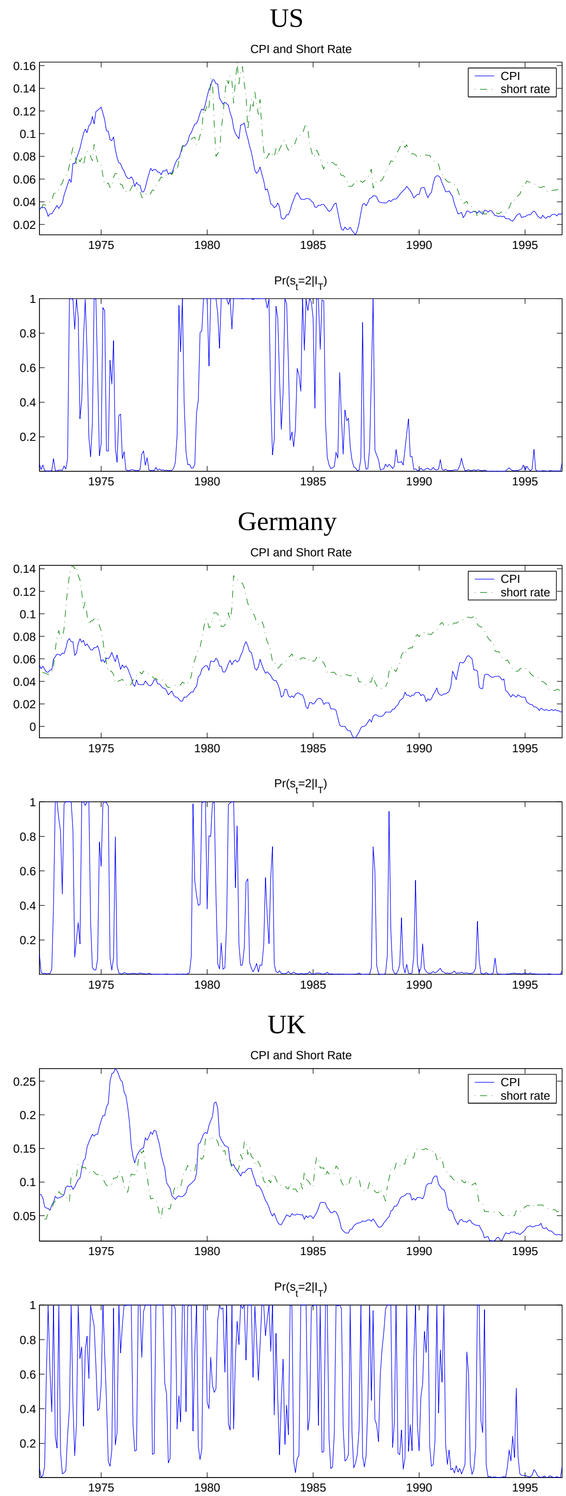
Constant Probs				Time-Varying Probs			
Parameter	Estimate	Std Error	p value	Parameter	Estimate	Std Error	p value
μ_{11}	-0.1140	0.1307	0.3831	μ_{11}	-0.1049	0.1236	0.3963
μ_{12}	0.5683	0.1142	0.0000	μ_{12}	0.6217	0.2746	0.0236
μ_{21}	1.3415	0.7799	0.0854	μ_{21}	1.0415	0.7398	0.1592
μ_{22}	-1.0236	0.8885	0.2493	μ_{22}	-0.8008	0.6406	0.2112
$A_1[1,1]$	1.0123	0.0156	0.0000	$A_1[1,1]$	1.0140	0.0154	0.0000
$A_1[1,2]$	0.0456	0.0266	0.0865	$A_1[1,2]$	0.0383	0.0367	0.2959
$A_1[2,1]$	-0.0621	0.0137	0.0000	$A_1[2,1]$	-0.0707	0.0369	0.0553
$A_1[2,2]$	0.8430	0.0271	0.0000	$A_1[2,2]$	0.8351	0.0531	0.0000
$A_2[1,1]$	0.8723	0.0647	0.0000	$A_2[1,1]$	0.8958	0.0695	0.0000
$A_2[1,2]$	-0.1371	0.0886	0.1218	$A_2[1,2]$	-0.1201	0.0823	0.1447
$A_2[2,1]$	0.1129	0.0735	0.1246	$A_2[2,1]$	0.0958	0.0616	0.1197
$A_2[2,2]$	0.9119	0.0890	0.0000	$A_2[2,2]$	0.9055	0.0820	0.0000
$R_1[1,1]$	0.2859	0.0135	0.0000	$R_1[1,1]$	0.2819	0.0202	0.0000
$R_1[1,2]$	-0.1620	0.0239	0.0000	$R_1[1,2]$	-0.1598	0.0252	0.0000
$R_1[2,2]$	0.2870	0.0139	0.0000	$R_1[2,2]$	0.2867	0.0194	0.0000
$R_2[1,1]$	1.1847	0.1065	0.0000	$R_2[1,1]$	1.1688	0.1381	0.0000
$R_2[1,2]$	-1.0106	0.1197	0.0000	$R_2[1,2]$	-0.9901	0.1477	0.0000
$R_2[2,2]$	0.6516	0.0577	0.0000	$R_2[2,2]$	0.6422	0.0719	0.0000
P	0.9691	0.0139	0.0000	a_1	2.7599	0.8448	0.0011
				b_1			
				c_1	0.7599	0.4368	0.0819
Q	0.9004	0.0423	0.0000	a_2	3.5121	3.6956	0.3419
				b_2			
				c_2	-0.6516	1.1189	0.5603

Model is $Y_t = A(s_t)Y_{t-1} + U_t$, $U_t \sim$ IID $N(0, \Sigma(s_t))$ with $Y_t = (r_t\, z_t)'$, the short rate and spread, and $R(s_t) = \text{chol}(\Sigma(s_t))$, $s_t = 1, 2$. For the constant probability model $p(s_t = 1|s_{t-1} = 1) = P$ and $p(s_t = 2|s_{t-1} = 2) = Q$. For the time-varying probability model, $p(s_t = i|s_{t-1} = i; \mathcal{I}_{t-1}) = \frac{\exp(a_i + b_i r_{t-1} + c_i z_{t-1})}{1+\exp(1+a_i+b_i r_{t-1}+c_i z_{t-1})}$ $i = 1, 2$. The first subscript denotes the regime s_t, and the second subscript (or numbers in square brackets) denote the matrix element. The models are estimated using only US data. A full estimation failed to converge so we report a restricted estimation dependent only on the spread in the time-varying probabilities. A likelihood ratio test for $c_i = 0$ yields a p-value of 0.0167.



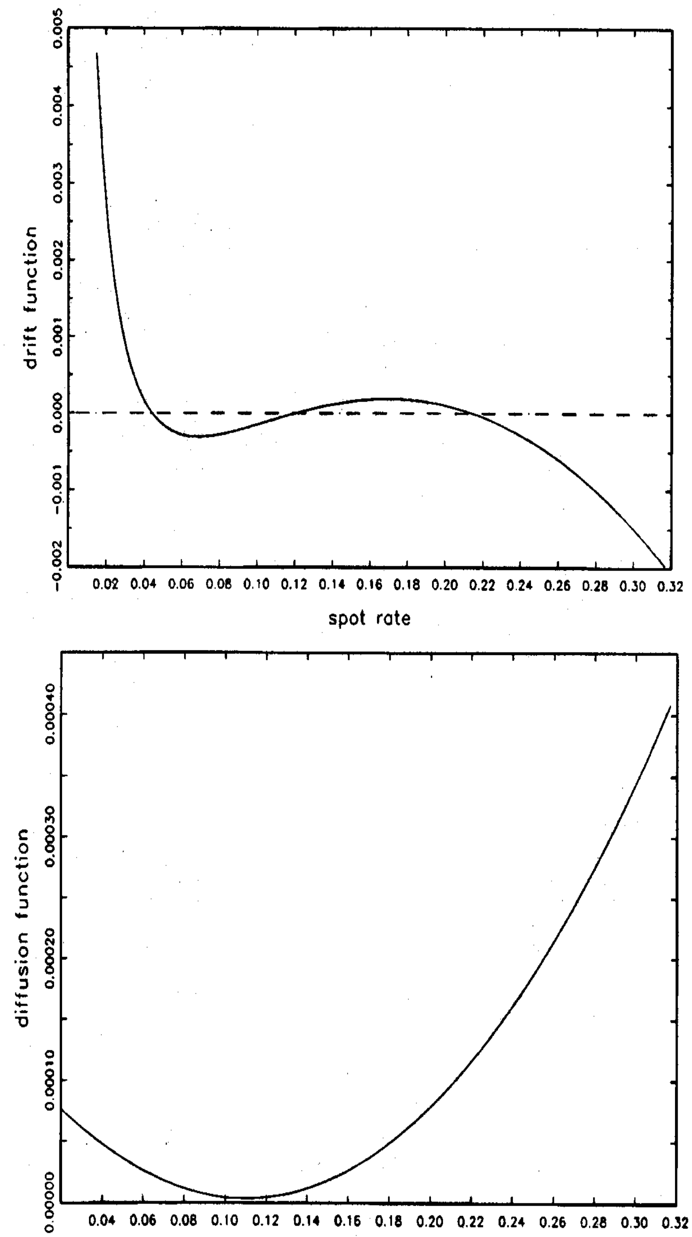
Time-varying transition probabilities from the univariate short rate model estimated jointly across the US, UK and Germany. The transition probabilities are logistic functions given by $p(s_t = i|s_{t-1} = i; \mathcal{I}_{t-1}) = \frac{e^{a_i+b_i r_{t-1}}}{1+e^{a_i+b_i r_{t-1}}}$, $i = 1, 2$, where $a_1 = 4.0530$, $b_1 = -0.2700$, $a_2 = -0.0124$, $b_2 = 0.1182$, as estimated in Table (2). The short rate r_{t-1} is on the x-axis.

Figure 1: Transition Probabilities



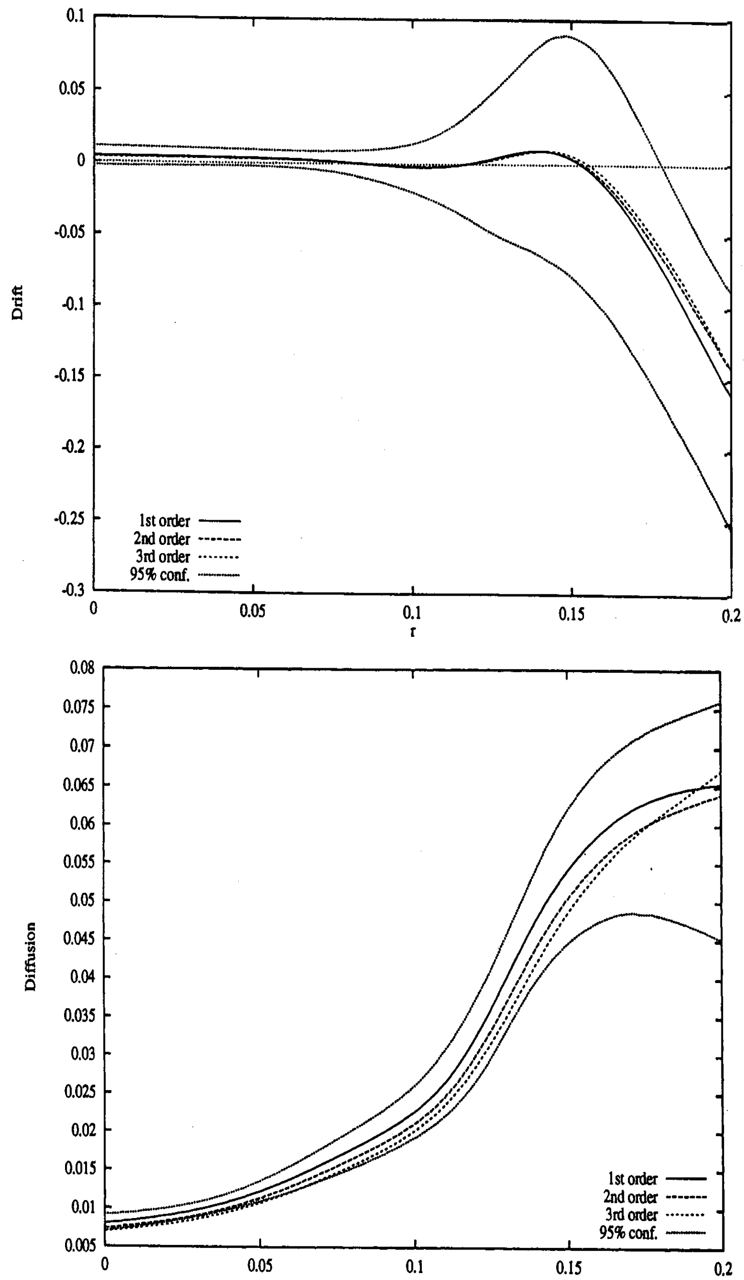
In each plot we show CPI inflation (percentage changes over the last 12 months) and short rates (top subplot) and smoothed probabilities $p(s_t = 2|\mathcal{I}_T)$ (bottom subplot). The correlation between inflation and the smoothed probabilities is 0.5420, 0.4994 and 0.3610 for the US, Germany and UK respectively. The smoothed probabilities are calculated using the bivariate spread model with time-varying probabilities estimated cross-sectionally across the US, Germany and UK.

Figure 2: Inflation and Smoothed Probabilities



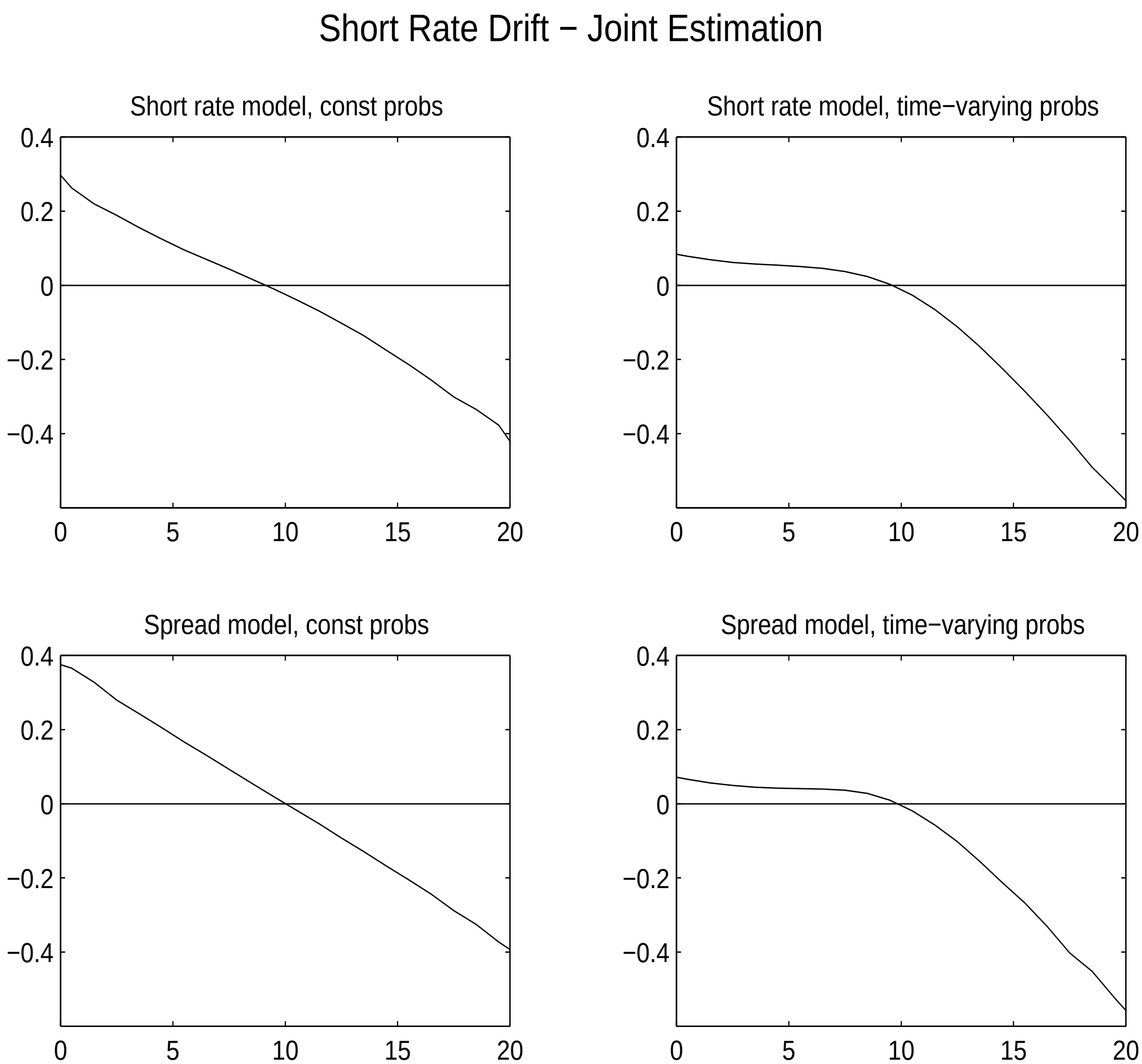
We reproduce Figures 4b and 4c of Aït-Sahalia (1996a). The top plot shows Aït-Sahalia's estimated non-linear drift of the short rate and the bottom plot shows Aït-Sahalia's nonlinear diffusion of the short rate.

Figure 3: Aït-Sahalia (1996a)'s Short Rate Drift and Diffusion.



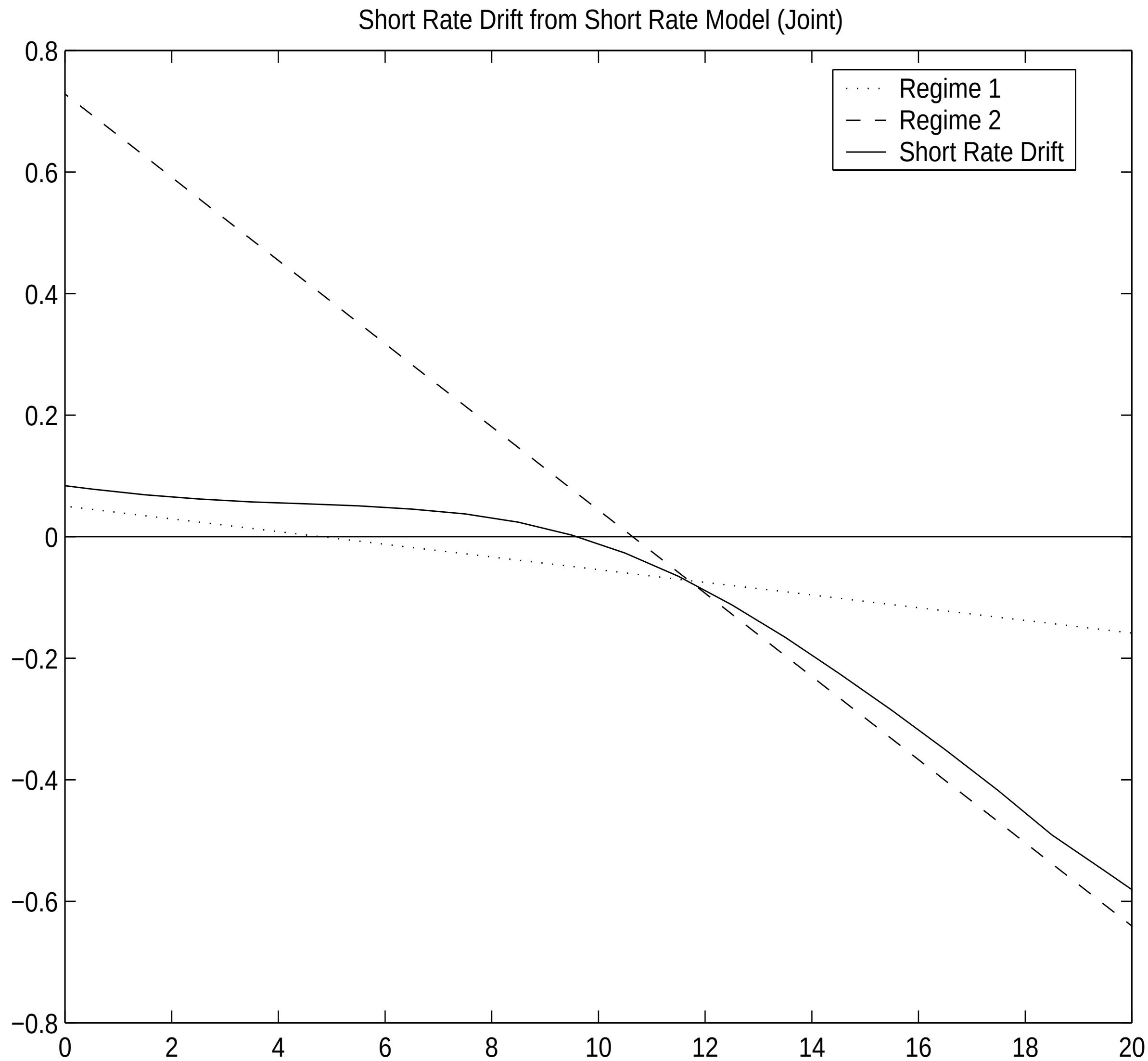
We reproduce Figures 4 and 5 of Stanton (1997). The top plot shows Stanton's estimated non-linear drift of the short rate with estimates of first, second and third order approximations to the drift. The bottom plot shows Stanton's nonlinear diffusion of the short rate with estimates of first, second and third order approximations to the diffusion.

Figure 4: Stanton (1997)'s Short Rate Drift and Diffusion



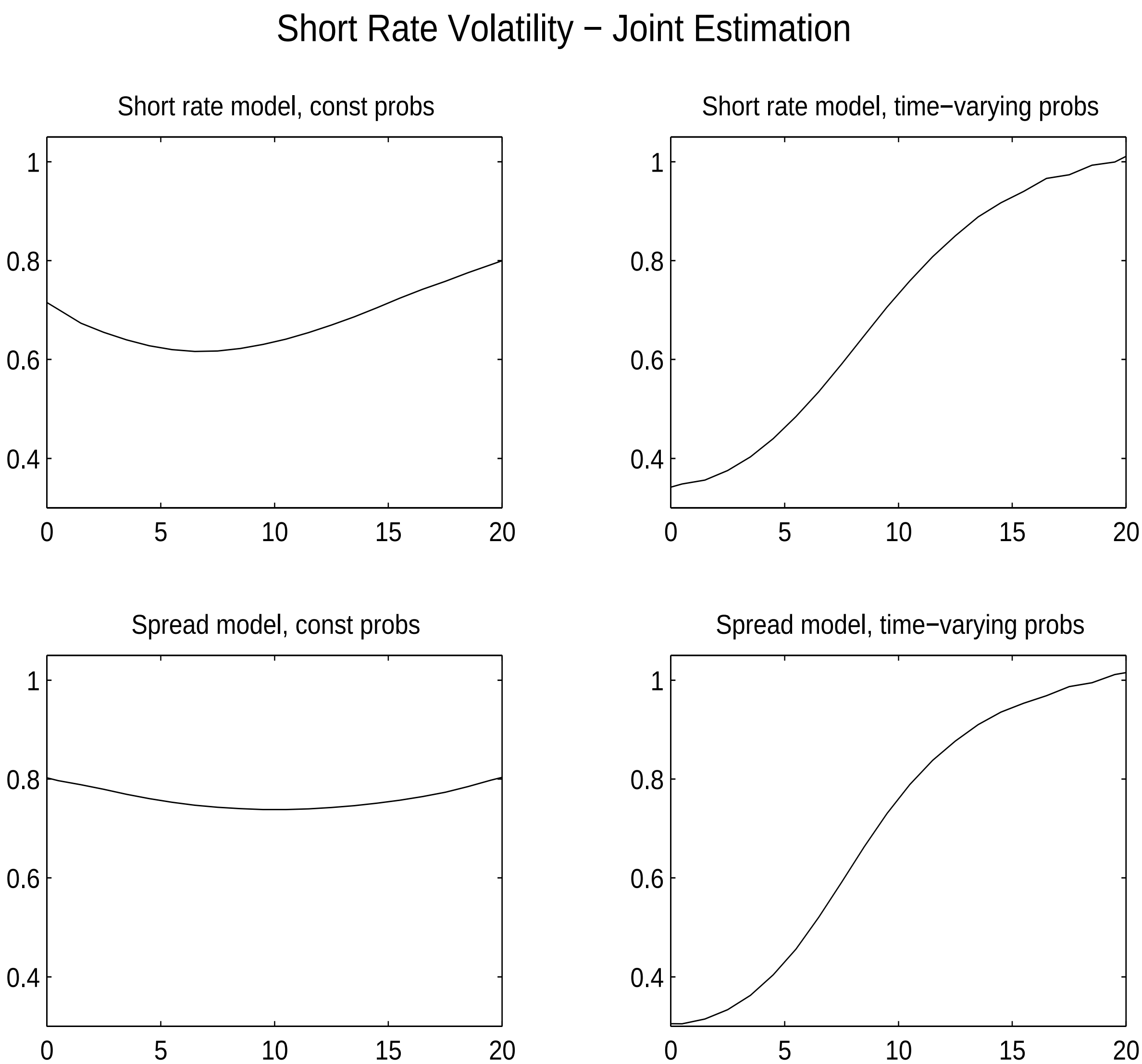
Conditional drift of the short rate for the joint estimation over US, UK and Germany. From top left, clockwise: the univariate model with constant transition probabilities; the univariate model with time-varying probabilities; the bivariate spread model with time-varying probabilities; the bivariate model with constant probabilities. The short rate is on the x-axis.

Figure 5: Short Rate Drifts from the Joint Estimation



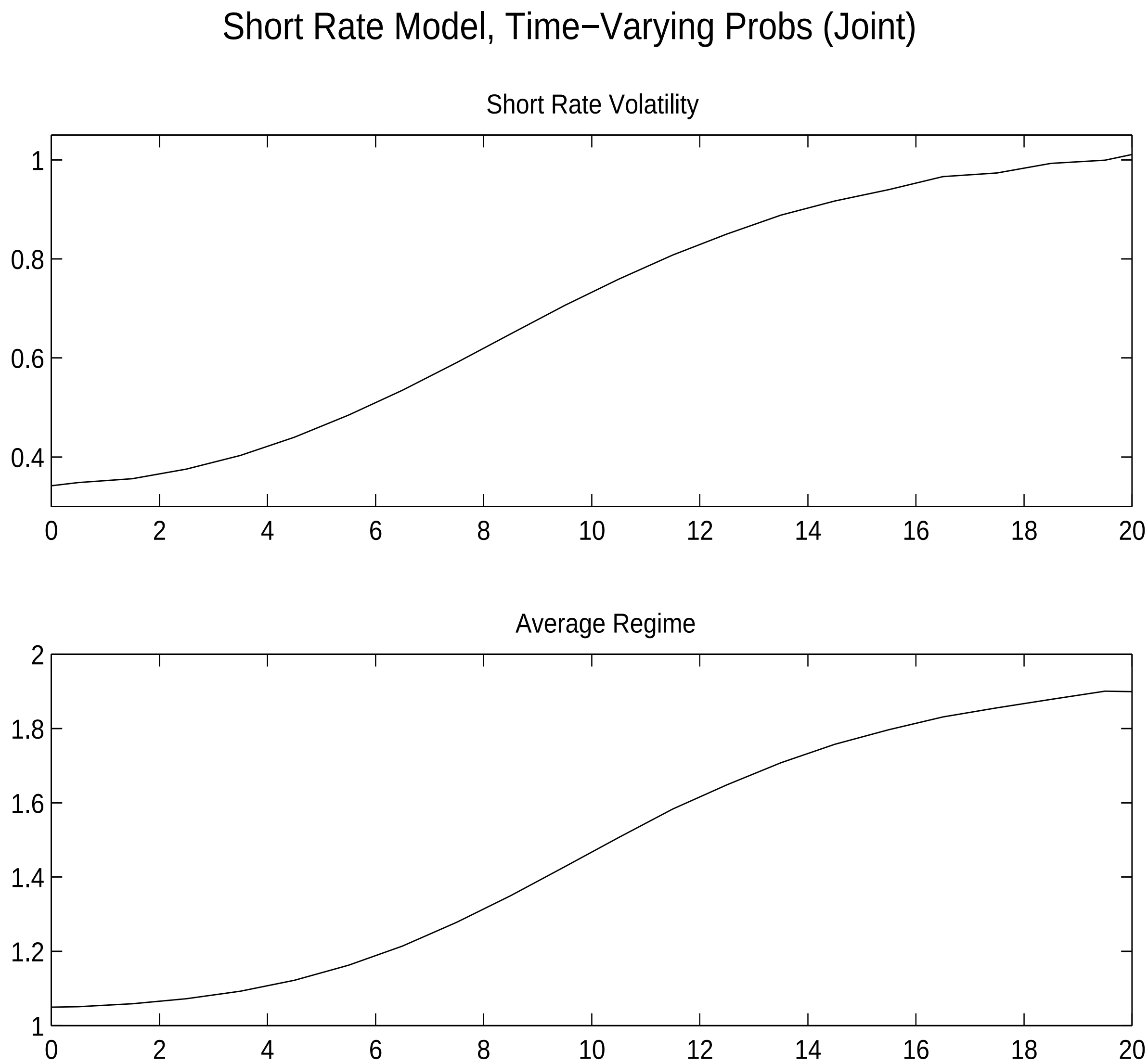
We show the regime-dependent drifts, and the overall short rate drift for the univariate short rate model with time-varying probabilities, jointly estimated over the US, UK and Germany. The short rate is on the x-axis.

Figure 6: Regime-Dependent Drifts



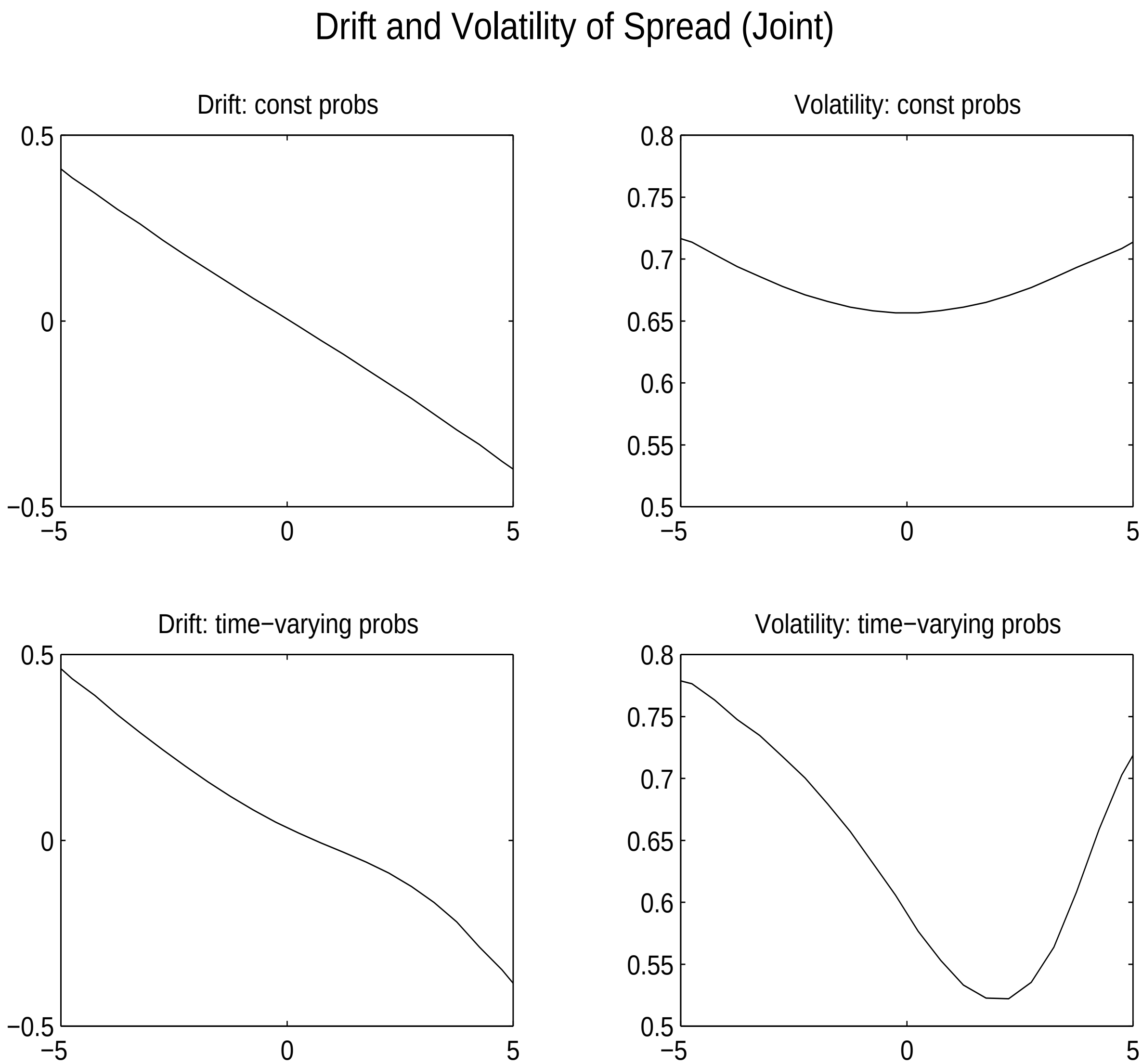
Conditional volatility of the short rate for the joint estimation over US, UK and Germany. From top left, clockwise: the univariate model with constant transition probabilities; the univariate model with constant probabilities; the bivariate spread model with time-varying probabilities; the bivariate model with constant probabilities. The short rate is on the x-axis.

Figure 7: Short Rate Volatility from the Joint Estimation



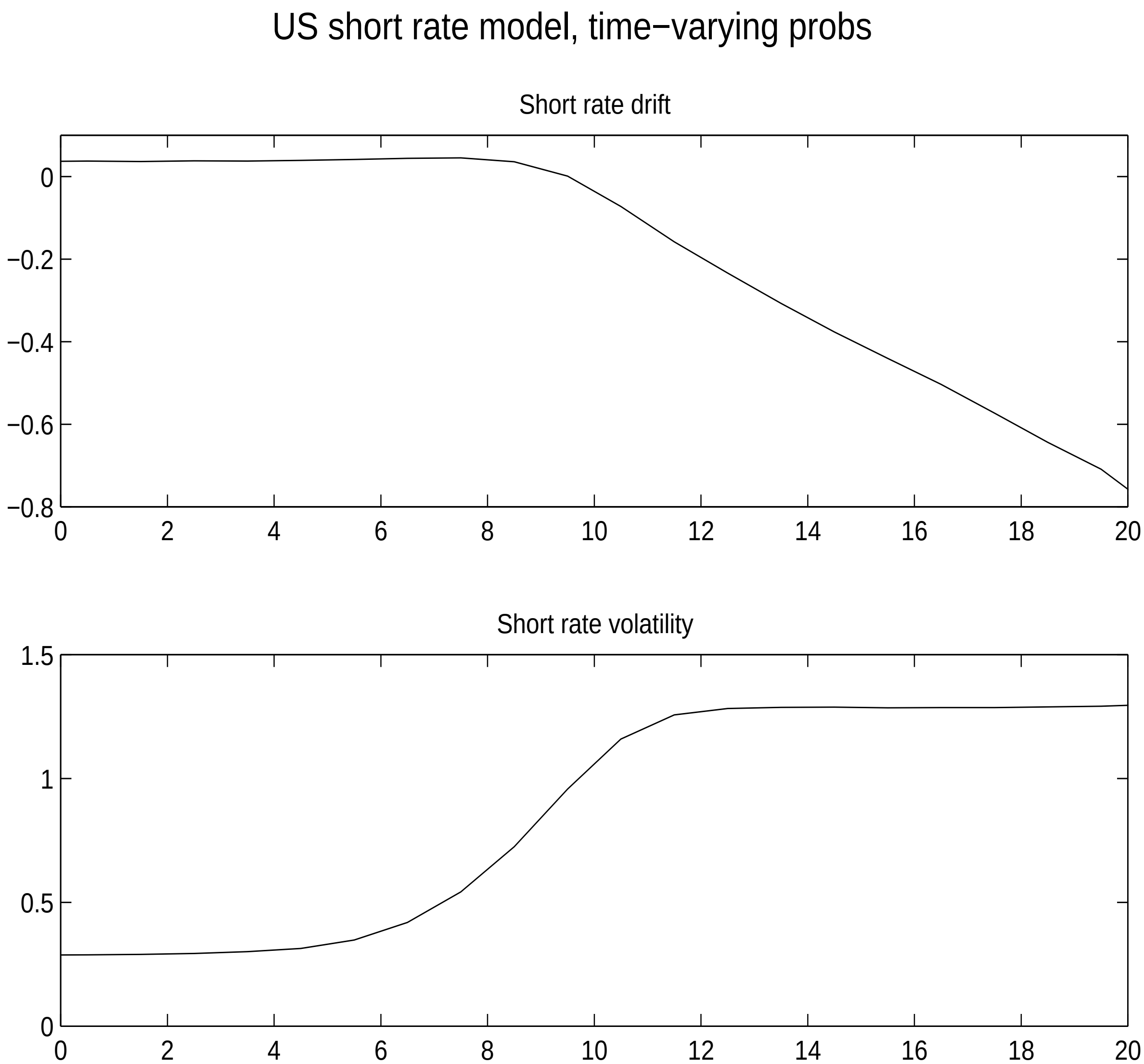
Plots are done for the univariate short rate model with time-varying probabilities estimated jointly over the US, UK and Germany. The top panel shows the volatility of the short rate. The bottom panel shows the average regime associated with each short rate (the regime can either be 1 or 2). The short rate is on the x-axis.

Figure 8: Comparison with the Average Regime



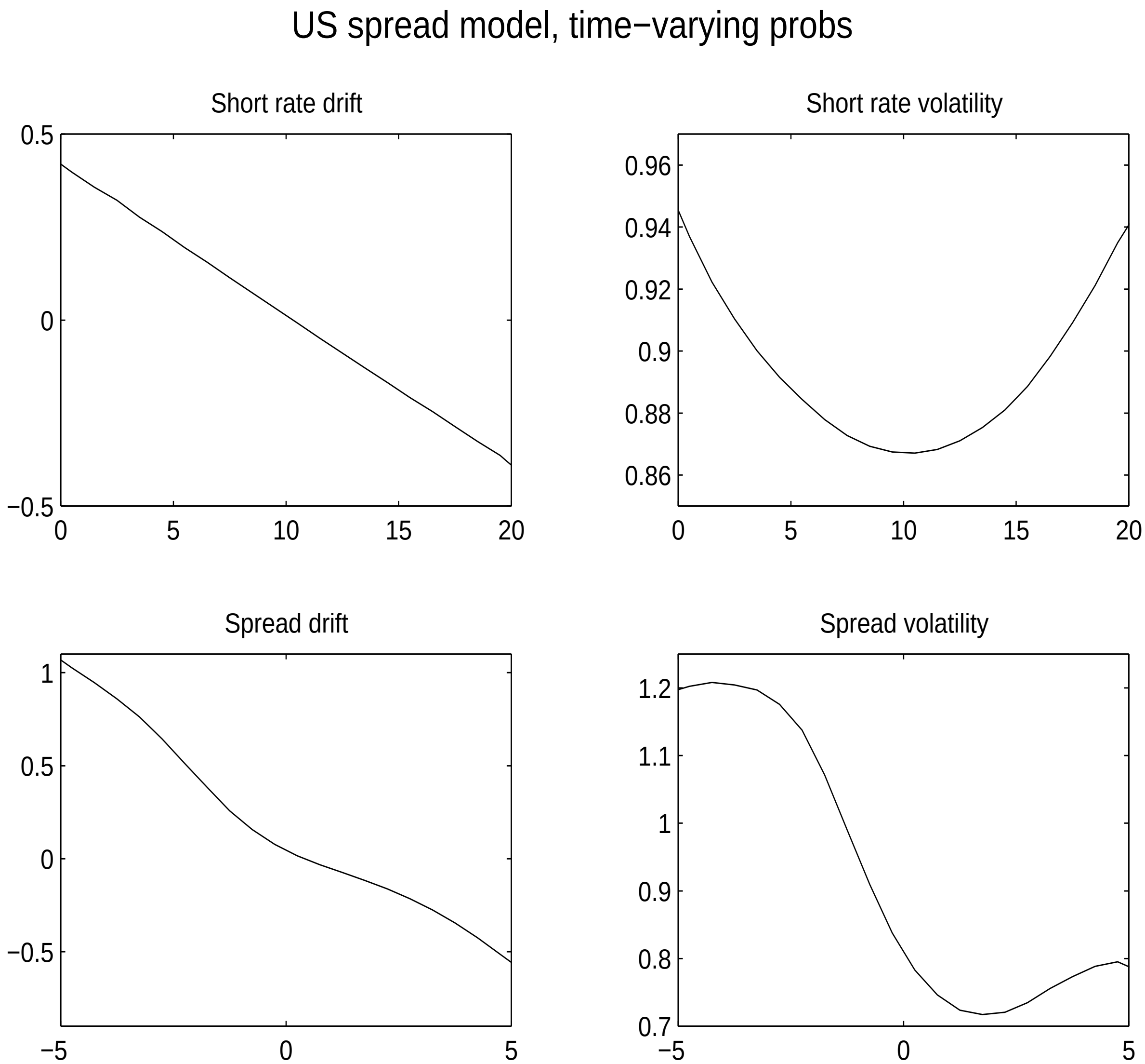
Conditional drift and volatility of the spread for the joint estimation over US, UK and Germany using the bivariate spread model with constant and time-varying probabilities. The drift and volatility from the bivariate model with constant (time-varying) probabilities is shown in the top (bottom) row. The spread is on the x-axis.

Figure 9: Spread Drifts and Volatilities from the Joint Estimation



Conditional drift (top panel) and volatility (bottom panel) of the short rate for the univariate short rate model with time-varying probabilities estimated using only US data. The US short rate is on the x-axis.

Figure 10: US Short Rate Model



Conditional drifts and volatilities of the short rate and spread for the bivariate spread model with time-varying probabilities estimated using only US data. From top left, clockwise: short rate drift, short rate volatility, spread volatility, and spread drift. The US short rate or spread is on the x-axis.

Figure 11: US Spread Model