Working Paper 01-21
Business Economics Series 06
March 2001

Departamento de Economía de la Empresa
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 6249608

A NONPARAMETRIC DIMENSION TEST OF THE TERM STRUCTURE[1]

Javier Gil-Bazo* and Gonzalo Rubio

Abstract

This paper addresses the problem of conducting a nonparametric test of the dimension of the state variable vector in a continuous-time term structure model. The paper shows that a bivariate diffusion function of the short rate process is a sufficient condition for the term structure to be driven by two stochastic factors. Using an easy-to-implement kernel smoothing method the number of state variables can be tested under very unrestrictive assumptions. The results suggest that continuous-time models for the US interest rates should contain at least two stochastic factors.

JEL Classification: E43; C14

[1]We would like to thank Manuel Moreno, Rafael Repullo, Ignacio Peña, David Musto and José Marín for helpful comments and discussions. Javier Gil-Bazo received funding from Programa de Formación de Profesorado Universitario y Personal Investigador. Gonzalo Rubio acknowledges the financial support provided by Dirección Interministerial Científica y Técnica (DGICYT) grant PB97-0621

*Universidad Carlos III de Madrid. Departamento de Economía de la Empresa. C/ Madrid, 126, 28903 Getafe (Madrid), Spain. E-mail:jjgil@emp.uc3m.es. Tel.: +34-91-624.58.44. Fax: +34-91-624.9608

1 Introduction

The tractability that Ito calculus permits, as well as the possibility to allow for continuous portfolio rebalancing, explain the popularity of the continuous-time setting in …nancial economics[1]. Interest rate modeling in particular has bene…ted profoundly from the use of Ito processes to describe the behavior of state variables. As a result, an enormous amount of models have followed the seminal work of Vasicek (1977) and Cox, Ingersoll, and Ross (henceforth CIR) (1985b). These one-factor models have a single state variable that can be identi…ed with the instantaneous riskless interest rate. Equilibrium or no-arbitrage conditions then make it possible to price assets whose payo¤s are contingent on the short term interest rate such as default-free bonds or options on those bonds. The whole term structure of interest rates can thus be derived from the prices of zero-coupon bonds with di¤erent periods to maturity. However, the inadequacy of the Vasicek and CIR models to reproduce the observed shape of the yield curve, along with the observation that yields do not have the same autocorrelation as the single state variable, have motivated the inclusion of two or more state variables, such as the two-factor model proposed by Longsta¤ and Schwartz (1992).

The purpose of this paper is to conduct a test of the number of state variables in a general continuous-time term structure model. The standard parametric approach …rst speci…es a multifactor continuous time interest rate model that nests a lower dimension model. The model parameters are then estimated, and …nally a test is performed on whether the relevant parameters are statistically di¤erent from the values that make the model one-dimensional. This approach entails three main problems.

First of all, a model must be speci…ed ex ante from the wide array that exists in the literature. Di¤erent speci…cations of the dynamic behavior of the state variables, together with di¤erent approaches to solve the model, namely equilibrium or arbitrage-free, have resulted in an extraordinarily large family of interest rate models. For a review of term structure models see, for instance, Du¢e (1996, Ch. 7), or the discrete-time versions by Backus, Foresi and Telmer (1998).

Second, stochastic processes expressed in continuous-time have to be estimated using data observed only at discrete times. The problem arises because exact maximum likelihood estimation is generally not possible since the transition density between two discrete times is not known explicitly for most of the di¤usions employed in the literature. As an alternative to exact maximum likelihood, some authors have used a discretized version of the di¤usion process using the Euler scheme and then they have applied a (quasi) maximum likelihood or a generalized method of moments approach[2]. Using Monte Carlo simulation, Honoré (1997) shows that the …nite sample properties of both approaches are inappropriate due to the approximation error. In order to avoid this discretization bias, Du¢e and Singleton (1993) and Gallant and Tauchen (1997)

[1] For an extensive survey of continuous-time methods in …nance, see Sundaresan (2000).

[2] See for example, Chan, Karolyi, Longsta¤, and Sanders (1992)

use simulations to obtain unconditional moments, whereas, Pedersen (1995) and Santa-Clara (1995) develop a simulated maximum likelihood method. Despite the success of simulation-based estimation methods, they have a clear disadvantage in terms of computing time.

Finally, many of the multi-factor interest rate models are formulated in terms of an unobservable state variable: the stochastic volatility of the instantaneous interest rate. Although the dynamics of the volatility factor has been modeled as a process of the ARCH family[3], this approach is not consistent with stochastic volatility. Again, we must rely on simulation-based techniques from the stochastic volatility literature, such as Gouriéroux, Monfort and Renault's (1993) indirect inference or the simulated maximum likelihood estimator by Danielsson (1998).

The paper addresses these problems. First, by studying the relationship between the number of factors that determine interest rates and the short term continuous-time process, we seek a hypothesis that can be tested without incurring model misspecification error. It is shown that if the diffusion coefficient is a function of the model's two state variables then the whole term structure is driven by the two stochastic factors. Although in general failure to reject the univariate diffusion does not imply that the state variable vector is one-dimensional, it is shown that this is the case for the two-factor CIR model. Thus, we may use a method designed by Aït-Sahalia, Bickel and Stoker (1998) to test the dimension of the diffusion function of the short rate dynamics. The justification relies on the work of Stanton (1997) and Boudoukh, Richardson, Stanton and Whitelaw (1998) who use the infinitesimal generator of the diffusion to prove that the first and second conditional moments of discrete changes in the short rate converge to the actual drift and diffusion functions as the observation frequency grows to infinity. The approach we follow avoids the need to parameterize the model since the method is nonparametric, and at the same time provides us with a test that is consistent with continuous-time modeling.

Second, the kernel method that we use has a substantial advantage over simulation techniques in terms of computing time. Whereas a standard Monte Carlo-based method may require hundreds of thousands of simulated paths, the kernel method can be easily implemented on a spreadsheet. Despite this advantage, Chapman and Pearson (2000) find that Stanton's (1997) estimator –on which the test relies– may show spurious nonlinearities –especially in the drift term function– when only small samples are available. However, as we will discuss in the conclusions, this bias has a presumably minimal impact on the paper's main results.

Finally, if we assume that the process is globally invertible so that it can be expressed in terms of two observable variables, e.g. the short rate and the yield spread as an approximation of the slope of the yield curve, we avoid the problem of the unobservability of the instantaneous volatility.

In our empirical application, we use a series of weekly observations of annualized discount rates on the US Treasury Bill with three months to maturity

[3] See for instance Andersen and Lund (1997).

and on the 10-year US Government Bond, covering the period from January 1962 to May 1999. The results show that the drift of the short rate process is a function of just the current short rate, whereas the diffusion is a function of both the short rate and the yield spread.

The rest of the paper is organized as follows: section 3.2 shows the approach taken to test the dimension of the instantaneous riskless interest rate process; section 3.3 explains the econometric method employed for conducting the test; section 3.4 presents the results; and finally section 3.5 concludes.

2 Testing for the dimension of the term structure

The presence of a second stochastic factor driving interest rates was suggested among others by Dybvig (1989) and Litterman and Scheinkman (1991), and has been found not only to easily account for the strong conditional heteroscedasticity observed in the short rate series but also to allow in theory for more realistic yield curves (Backus, Foresi and Telmer (1998)). Moreover, term structure models that include a stochastic volatility factor such as Longstaff and Schwartz (1992) and Andersen and Lund (1997) have been found to empirically accommodate the shortcomings of the single-factor CIR model. These findings however rely on the specification of a parametric model.

In a nonparametric continuous-time setting, Boudoukh, Richardson, Stanton and Whitelaw (1998) estimate a two-factor model assuming that the short rate and the slope of the yield curve, proxied by the yield spread, span the same space as the true state variables. Although their estimates suggest that the slope of the term structure plays a role in determining the diffusion coefficient, they do not provide a formal test for this relationship or explain how this relates to the dimension of the term structure model.

The purpose of this section is to show how a dimension test can be conducted in a nonparametric continuous-time framework.

First, it is shown that if the term structure is determined by two factors, then either the drift of the short rate process under the risk-neutral probability measure, the diffusion term, or both, must be a function of the short rate process and the second stochastic factor, provided the short rate and the second factor can be taken as state variables. Therefore, a bivariate diffusion function is a sufficient condition for the term structure to be bidimensional.

When two factors affect interest rates and the corresponding vector process is globally invertible, then any two interest-rate contingent factors can be taken as state variables. We will assume that the instantaneous riskless interest rate, r, and an observable variable X can be taken as state variables, and that their evolution through time is governed by the following joint Ito process:

$$dr = \mu_1(r; X; t)dt + \sigma_1(r; X; t)dw_1 \qquad (1)$$

$$dX = \mu_2(r; X; t)dt + \sigma_2(r; X; t)dw_2. \qquad (2)$$

where dw_1 and dw_2 are two independent Wiener processes. Next, we assume that there exists an asset whose price at time t is a function of the value of the two state variables at t. Denoting the asset's price by P,

$$P = P(r; X; t) \tag{3}$$

Assuming that $P(\cdot)$ is twice continuously differentiable in r and X, we may apply Ito's lemma to the previous expression in order to find the stochastic process governing the behavior of P. In particular, its drift is given by:

$$m = \frac{1}{2}\sigma_1^2(r; X; t)P_{rr} + \mu_1(r; X; t)P_r + \frac{1}{2}\sigma_2^2(r; X; t)P_{XX} + \mu_2(r; X; t)P_X + P_t \tag{4}$$

If the asset pays no dividends, the absence of arbitrage implies that the drift is a linear combination of the riskless interest rate and the market value of the bond's exposure to the sources of risk associated with r, and X:

$$m = rP + \lambda_r(r; X; t)P_r + \lambda_X(r; X; t)P_X \tag{5}$$

Equating both expressions we have the following valuation equation:

$$\frac{1}{2}\sigma_1^2(\cdot)P_{rr} + (\mu_1(\cdot) - \lambda_r(\cdot))P_r + \frac{1}{2}\sigma_2^2(\cdot)P_{XX} + (\mu_2(\cdot) - \lambda_X(\cdot))P_X + P_t - rP = 0 \tag{6}$$

If the asset is a zero-coupon bond that pays one unit at time T, then its price at time t is the solution to equation (6) with boundary condition:

$$P(r; X; T) = 1. \tag{7}$$

Subject only to technical conditions, the solution to (6) with (7) can be expressed in the form of the following expectation:

$$P_{t;T} = E_t\left[\exp\left(-\int_t^T \hat{r}_s ds\right)\right] \tag{8}$$

where $\hat{r}$ denotes the risk-adjusted process, given by:

$$d\hat{r} = (\mu_1(\hat{r}; \hat{X}; t) - \lambda_r(\hat{r}; \hat{X}; t))dt + \sigma_1(\hat{r}; \hat{X}; t)dw_1^* \tag{9}$$

$$d\hat{X} = (\mu_2(\hat{r}; \hat{X}; t) - \lambda_X(\hat{r}; \hat{X}; t))dt + \sigma_2(\hat{r}; \hat{X}; t)dw_2^*. \tag{10}$$

with:

$$\hat{r} = r, \text{ and } \hat{X} = X \tag{11}$$

From (8), (9), and (10) it is clear that the second stochastic factor X, affects bond prices indirectly through the drift and/or the diffusion term of the risk-adjusted short rate process. The problem is that we do not observe the risk adjusted process but the actual process that follows the instantaneous riskless rate. However, since the diffusion term in both processes is the same, a diffusion function that depends on r, and X is a sufficient condition for the term structure to be driven by two factors.

The appendix shows that a bivariate diffusion function is a necessary and sufficient condition for the presence of a second stochastic factor in the term structure model in the case of a two-factor CIR model.

3 Econometric approach

3.1 Testing for the dimension of the diffusion function

From the result of the previous section, we know that a bivariate diffusion function of the short rate process is a sufficient condition for the term structure to be driven by a second stochastic factor. Nevertheless, we will also test whether the drift function depends on the yield spread as well as on the short rate process.

Assume that the behavior of the short rate and the yield spread is described by the following time-homogeneous vector Ito process:

$$dr = \mu_r(r; S)dt + \sigma_{r1}(r; S)dW_1 + \sigma_{r2}(r; S)dW_2 \tag{12}$$

$$dS = \mu_S(r; S)dt + \sigma_{S1}(r; S)dW_1 + \sigma_{S2}(r; S)dW_2 \tag{13}$$

where S represents the yield spread and dW_1 and dW_2 are two independent Wiener processes.

We will test the following two hypotheses:

- Hypothesis 1 (henceforth H1):

$$\mu_r(r; S) = \mu_r(r) \tag{14}$$

- Hypothesis 2 (henceforth H2):

$$V(r; S) = \sigma^2_{r1}(r; S) + \sigma^2_{r2}(r; S) = V(r) \tag{15}$$

Stanton (1997) and Boudoukh, Richardson, Stanton, and Whitelaw (1998), show that first-order approximations to μ_r and V are given by the conditional first and second moments of the Euler discretization of the short rate process:

$$\mu_r = \frac{1}{\Delta} E_t \left[r_{t+\Delta} - r_t \right] + O(\Delta); \tag{16}$$

$$V = \frac{1}{\Delta} E_t \left[(r_{t+\Delta} - r_t)^2 \right] + O(\Delta): \tag{17}$$

Where Δ is the interval between the times when r_t and $r_{t+\Delta}$ are observed. As the observation frequency increases to infinity, $\Delta \to 0$, the approximations converge to the actual values of the drift and diffusion functions. This suggests that the hypotheses can be restated as:

- H1:

$$E\left[r_{t+\Delta} - r_t \mid r_t; S_t \right] = E\left[r_{t+\Delta} - r_t \mid r_t \right]; \tag{18}$$

- H2:

$$E\left[(r_{t+\Delta} - r_t)^2 \mid r_t; S_t \right] = E\left[(r_{t+\Delta} - r_t)^2 \mid r_t \right]: \tag{19}$$

Recall that if H1 is rejected, we cannot assure that interest rates are driven by two stochastic factors, unless the risk premiums are zero, in which case the drift of the actual and risk-adjusted short rate processes coincide. Also, failure to reject H1 does not imply that only one factor affects the term structure, as is the case with the two-factor CIR model. However, the test of H1 may have interesting implications regarding the way the actual short rate process should be modeled. In particular, if H1 is not rejected, efficiency gains could be attained when estimating continuous-time models of the term structure if the short-rate drift is modeled as a univariate function.

On the other hand, if H2 is rejected, then as explained above we know that the price of interest-rate contingent assets is driven by (at least) two factors, which by assumption can be taken as the short rate and the yield spread. Furthermore, since the diffusion term is a key variable in option pricing, the result would imply that the price of options on fixed-income assets is affected by the current level of the short-term interest rate as well as the slope of the yield curve.

3.2 The method

The method described in this subsection enables us to test nonparametrically the following hypothesis:

$$E[Y \mid W, V] = E[Y \mid W], \qquad (20)$$

so it serves our purposes in a straightforward fashion:

- H1 can be tested by setting:

$$\begin{aligned} Y &= r_{t+\Delta} - r_t, \\ W &= r_t, \\ V &= S_t, \end{aligned}$$

- H2 can be tested by setting:

$$\begin{aligned} Y &= (r_{t+\Delta} - r_t)^2, \\ W &= r_t, \\ V &= S_t. \end{aligned}$$

The test in question was proposed by Aït-Sahalia, Bickel and Stoker (1998) and uses kernel methods to estimate the regression under the restricted specification and under the alternative. Then, the difference between the restricted and the unrestricted kernel regression is measured via the residual sum of squares.

It should be noted that although the test in principle applies to a data sample of independent and identically distributed observations, the distribution of the test statistic is unchanged by serial dependence in the data provided that this is strictly stationary ergodic and the amount of serial dependence in the data decays sufficiently fast.[4]. We will assume that this condition holds true for the data set.

If the sample data consists of $Z_i = (Y_i, V_i, W_i)$, $i = 1, ..., N$ the test answers the question of whether the predictor variables V can be omitted from the regression of Y on (W, V). The regression function of Y on (W, V) is defined by

$$m(w, v) \equiv E(Y \mid W = w, V = v) = \frac{\int y f(y, w, v) dy}{f(w, v)} \qquad (21)$$

and the regression function of Y on W by

[4] More technically, if Z_i is the vector of observations at time i, then it must be the case that:

1. The data $\{Z_i, i = 1, ..., N\}$ are strictly stationary and β-mixing with $\beta_N = O(N^{-k})$, $k > 19/2$.
2. The joint density $f_{1,j}(\cdot,\cdot)$ of (Z_1, Z_{1+j}) exists for all j and is continuous on $(R \times S)^2$.

$$M(w) \equiv E(Y \mid W = w) = \frac{\int y f(y;w)dy}{f(w)} \tag{22}$$

These conditional moments may be consistently estimated using the Nadaraya-Watson kernel regression method:

$$\hat{m}_h(w;v) \equiv \frac{\sum_{i=1}^{N} K_h(w - W_i; v - V_i) Y_i}{\sum_{i=1}^{N} K_h(w - W_i; v - V_i)} \tag{23}$$

$$\hat{M}_H(w) \equiv \frac{\sum_{i=1}^{N} K_H(w - W_i) Y_i}{\sum_{i=1}^{N} K_H(w - W_i)} \tag{24}$$

where $K_h(u) = h^{-d} K(u/h)$ and $K_H(u) = H^{-d} K(u/H)$, d being the dimension of the vector u that measures the distance of the observed regressor data to the design point. The shape of the kernel weights is determined by K, whereas the size of the weights is parameterized by the bandwidth, denoted by h[5].

The test statistic is based on the distance, measured in a mean squared error way, between both regression functions or more precisely their estimates. If we define the following statistic:

$$\tilde{\Gamma} \equiv \frac{1}{N} \sum_{i=1}^{N} \left\{ \hat{m}_h(W_i;V_i) - \hat{M}_H(W_i) \right\}^2 A_i \tag{25}$$

where A_i is the value that a weighting function[6] takes for $W_i; V_i$, then $\tilde{\Gamma}$ is a consistent estimator of the weighted expected squared difference between $m(W;V)$ and $M(W)$. Under the null hypotheses $\tilde{\Gamma}$ is asymptotically zero.

The distribution of the test statistic is derived under standard assumptions about the density functions and the kernel. Specially relevant are those concerning the kernel function and the bandwidth choice:

1. The kernel K is a bounded function on R, symmetric about 0, with $\int |K(z)| dz < \infty$; $\int K(z)dz = 1$; $\int z^j K(z)dz = 0$ for $1 \le j \le r$. Further,

$$r > 3(p+q)/4 \tag{26}$$

where p and q are the dimensions of W and V respectively.

2. As $N \to \infty$, the unrestricted bandwidth sequence $h = O(N^{-1/\delta})$ is such that

[5] More details on kernel smoothing techniques can be found in Härdle (1990).

[6] This weighting function allows us to test goodness-of-fit for particular value ranges and/or avoid technical problems such as the estimation of conditional expectation in areas of low density. In this application A_i is an indicator function that equals 1 if both W_i and V_i are between the 5th and 95th percentiles of their respective samples, and zero otherwise.

$$2(p + q) < \delta < 2r + (p + q)/2 \tag{27}$$

while the restricted bandwidth $H = O(N^{-1/\delta})$ satisfies

$$p < \Delta \quad 2r + p \tag{28}$$

as well as

$$\delta p/(p + q) \quad \Delta < \delta \tag{29}$$

The authors show that under the null hypothesis that V can be omitted from the regression:

$$\hat{\tau} \equiv \sigma_{11}^{-1}(N h^{(p+q)/2} \cdot \tilde{I} - h^{-(p+q)/2}\hat{}_{12} - h^{(q-p)/2}\hat{}_{22} - h^{(p+q)/2}H^{-p}\hat{}_{32}) \to N(0;1) \tag{30}$$

where the critical values are calculated in the following way:

$$\sigma_{11}^2 = \frac{2C_{11}}{N}\sum_{i=1}^{N}\frac{\sigma_h^4(W_i;V_i)A_i^2}{\hat{f}_h(W_i;V_i)}; \qquad \hat{}_{12} = \frac{C_{12}}{N}\sum_{i=1}^{N}\frac{\sigma_h^2(W_i;V_i)A_i}{\hat{f}_h(W_i;V_i)}$$

$$\hat{}_{22} = -\frac{2C_{22}}{N}\sum_{i=1}^{N}\frac{\sigma_h^2(W_i;V_i)A_i}{\hat{f}_H(W_i;V_i)}, \qquad \hat{}_{32} = \frac{C_{32}}{N}\sum_{i=1}^{N}\frac{\sigma_H^2(W_i)A_i}{\hat{f}_H(W_i)}$$

with $\sigma_h^2(W_i;V_i)$ and $\sigma_H^2(W_i)$ being the conditional variances of Y estimated nonparametrically. The constants C_{ij} are determined by the choice of kernel. In our application:

$$C_{12} = 1/(2\sqrt{\pi})^2; C_{22} = 1/\sqrt{2\pi}; C_{32} = 1/(2\sqrt{\pi}); C_{11} = 1/(2\sqrt{2\pi})^2$$

Following Aït-Sahalia, Bickel and Stoker (1998), the following bandwidth functions are used for the unrestricted and restricted regressions:

$$h = h_0 N^{-1/\delta} \quad \text{with } \delta = 4.75 \tag{31}$$

$$H = H_0 N^{-1/\Delta} \quad \text{with } \Delta = 4.25 \tag{32}$$

Although there is not one single commonly agreed upon method, a cross-validation procedure was employed to determine the values of h_0 and H_0 that minimize the mean squared error of the estimates. The method employed here works as follows: for every observation, Z_i use all observations except for those in its neighborhood to estimate $m(W;V)$ or $M(W)$ and then compute the sum of squared differences between Y and $m(W;V)$ or $M(W)$. Finally h_0 and H_0 are chosen to minimize the sum of squared residuals[7]. Results strongly suggest that h_0 in the first test should be set equal to 0.19 and H_0 ought to be 0.15. For the second test, h_0 was chosen to equal 0.11 and H_0 was chosen to equal 0.09.

Finally, the test was performed using an independent Gaussian kernel

$$\begin{aligned} K_H(w - W_i) &= H^{-1}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{w-W_i}{H}\right)^2\right) \\ K_h(w - W_i; v - V_i) &= h^{-2}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{w-W_i}{h}\right)^2\right) \times \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{v-V_i}{h}\right)^2\right) \end{aligned} \tag{33}$$

4 Data and results

A series of daily observations of annualized discount rates on US Treasury Bills with three months to maturity and on 10-year US Government Bonds was obtained from the Internet Site of the Federal Reserve. The rates correspond to secondary market closing bid rates. The missing observations problem was overcome by constructing a weekly series in which each observation corresponds to the rate quoted on Wednesday (the day of the week with the least number of missing observations). If the observation on a given Wednesday was not available, the quote from the day before was used[8]. Finally, discount rates were converted to annualized continuously compounded yields. The resulting series covers the period from January 1962 to May 1999 – a total of 1,948 observations– and is displayed on Figure 1. Table 1 shows descriptive statistics of the dataset. Figure 2 contains the plot of weekly changes in the short-term interest rate. As Gray (1996) observes, periods of high volatility coincide with periods of high interest rates, whereas periods of low volatility happen when the short rate is lower. He consequently identifies two different regimes. In particular, the high variance/high interest rate regime comprises the 1973-1975 period (the time of the OPEC oil crisis), the 1979-1982 period (when the Fed targeted nonborrowed reserves instead of interest rates), and the months after the 1987 stock market crash.

Next, we test whether the drift function depends only on the current level of the short-term interest rate.

The test statistic τ associated with H1 and its corresponding p-value is shown in Table 2 for the values obtained from the cross validation approach ($H_0 = 0.15$; $h_0 = 0.19$) as well as for neighboring values of the parameters that determine the kernel bandwidths through (32) and (31). The results imply that we cannot reject the null hypothesis H1 at the 5% or 10% significance levels. Even when considering bandwidth departures, the test suggests that the slope

[7] In particular, the cross-validation methods employed left between 2 and 6 observations out on each side of the observation point, that is roughly between half a month and one and a half months. Optimal bandwidth estimates appeared to be robust to the number of observation left out.

[8] In constructing the series we have followed Andersen and Lund (1997).

of the yield curve contains no predictive power about future changes of interest rates. This result is consistent with most model specifications and empirical findings.

Next, we test whether the diffusion function depends only on the current level of the instantaneous riskless interest rate.

The test statistic $\hat{\tau}$ associated with H2 and its corresponding p-value is shown in Table 3 for the values obtained from the cross validation approach ($H_0 = 0.09$; $h_0 = 0.105$) as well as for neighboring values of the parameters that determine the kernel bandwidths through (32) and (31). The test shows that the null hypothesis that the conditional variance of instantaneous changes in interest rates is a function of only the short rate is rejected at the 10% level in all cases and at the 5% level in all cases considered but one. Although the spread's additional effect on the conditional second moment of the short rate changes appears to be significant, the analysis performed suggests that this result is sensitive to bandwidth choices. In fact, p-values range from practically zero to almost 8% for small changes in the bandwidths.

In order to explore the reason for this lack of robustness we need to analyze the estimates of the diffusion coefficient under the restricted and unrestricted specifications. Figure 3 shows the shape of the estimated diffusion function when only the short rate is used as the conditioning variable and figure 4 shows the shape of the diffusion function when dependence on the yield spread is allowed. Comparing the two, we see that a univariate diffusion function underestimates the instantaneous volatility of the short rate process when the short rate level is high and the yield curve has a steep negative slope. Therefore, given that periods of high short-term interest rates usually happen when the slope of the yield curve is negative, an increase in interest rates will most likely result in an increase in the conditional variance that is higher than that predicted by single-factor models. Allowing for a second stochastic factor therefore accounts for the larger conditional volatility observed in the high variance regime, as well as for the larger degree of dependence of future volatility upon the short rate for high levels of the short-term interest rate (see Gray (1996)).

Furthermore, figure 4 shows that the dependence of the short rate diffusion upon the yield spread is higher the higher the short rate. Since the short-term interest rate spends most of the time in the low variance/low short rate regime, it appears that we would need a larger sample to provide stronger evidence for a bivariate diffusion function, especially considering that the tails have been trimmed out of the sample in order to perform the test. Even so, the test rejects the null hypothesis H2 for the optimal bandwidth values, so we may conclude that an adequate US term structure continuous-time model should contain at least two state variables.

5 Conclusions

The main contribution of this paper is two-fold: it explains how to conduct a nonparametric test of the number of factors underlying the short rate process and explores the implications of such a test regarding the dimension of the state variable vector in a continuous-time term structure model. Furthermore, the analysis provides a rationale for using the yield spread as a second stochastic factor in testing for the dimension of the term structure.

A major finding is that the dependence of the diffusion function on the short rate and the yield spread is generally a sufficient condition for the presence of a second stochastic factor. Moreover, it is a necessary and sufficient condition only in special cases, such as the two-factor CIR model.

The econometric approach is appropriate for three reasons. First, because it is nonparametric it is free of model misspecification problems. This is of particular importance given the broad range of modeling possibilities that one finds in the interest rate literature. Second, the estimation method attached to the test yields consistent estimations of the conditional mean and variance of changes in the short rate so results apply directly to continuous-time models provided the discrete approximation is accurate enough. In addition, the test does not rely on computationally intensive simulations and can be easily implemented on a spreadsheet. Finally, under the assumption that the short rate and the yield spread completely describe the state of the system, we have avoided the need to adjust an ARCH-type model to the short rate or to perform a simulation-based estimation of the conditional volatility.

The results suggest that the drift of the continuous-time stochastic process governing the short rate dynamics should be modeled as a function of just one factor, whereas the diffusion term does depend on the spread as well as on the level of the short rate. There is however a number of reasons why this result should be taken with caution.

First of all, the method may be biased in small samples due to the strong persistence in the short rate and spread series. In fact, Chapman and Pearson (2000) have shown that non-linearities displayed by Stanton's kernel drift estimates may be spurious. This problem is however minimized on this application. First, according to Chapman and Pearson (2000), the larger bias in Stanton's estimates occurs for values of the conditioning variable close to the support of the distribution of the sample. Our choice of the weighting function A_i; however, ensures that only observations in the center of the distribution are taken into account when constructing this test. Second, Chapman and Pearson (2000) also show that more accurate estimates can be attained by using bandwidths obtained through cross-validation[9], which is the method chosen in this application. Finally, this problem can neither be attributed to a discretization bias nor to the use of a kernel method, but to a truncation of the sample that affects the estimation of the drift exclusively. Our main result however rests on the test concerning the diffusion function.

Another source of concern is the great subjectivity underlying the choice of H_0 and h_0; aggravated by the sensitivity of the test statistic to the bandwidth values. This problem, associated with kernel regressions, becomes especially

[9] It should be noted however that the method they use is based on a function that penalizes the sum of squared residuals for small bandwidths, rather than the "leave-out" approach employed here.

important when the goal is hypothesis testing. The sensitivity analysis performed shows that conclusions for the conditional mean are quite consistent across different bandwidth values, whereas results for the conditional variance are somewhat less clear. This problem however could be overcome if the tails of the distribution were appropriately accommodated rather than trimmed away. Also, using two distinct bandwidths rather than a single smoothing parameter for the bivariate regressor would presumably result in the test detecting a higher dependence on the second stochastic factor. The reason is that the term spread's influence is underestimated when a single smoothing parameter is employed given that the spread series exhibits less dispersion.

Finally, controlled experiments should be carried out in order to assess the finite-sample properties of the method for persistent series, such as the ones usually found in Finance. It could be sensible to use the empirical rather than the asymptotic distribution given the high serial autocorrelation in the regressors. Therefore, a bootstrap approach seems like a plausible alternative to the asymptotic test.

A natural extension of this paper would be a test for a third factor driving interest rates. A problem likely to arise is the well known curse of dimensionality associated with fully nonparametric estimators: as the dimension of the independent variables grows, the speed of convergence decreases exponentially.

Appendix: the two-factor CIR model

Let us assume the following joint diffusion:

$$dx_1 = \theta_1(\bar{x}_1 - x_1)dt + C_1\sqrt{x_1}dW_1 \quad (34)$$

$$dx_2 = \theta_2(\bar{x}_2 - x_2)dt + C_2\sqrt{x_2}dW_2; \quad (35)$$

where dW_1 and dW_2 are independent pure Brownian motions, x_1 and x_2 are the unobservable state variables, and θ_1, θ_2, $\bar{x}_1$, $\bar{x}_2$, C_1, C_2 are constants. Equations (34) and (35) with θ_1, $\theta_2 > 0$, imply that the state variables revert to their means: $\bar{x}_1$ and $\bar{x}_2$.

The model is completed with the following implications for the instantaneous riskless interest rate and the risk-premium specifications:

$$r = x_1 + x_2; \quad (36)$$

$$\lambda_1(x_1; x_2) = \lambda_1 x_1; \quad (37)$$

$$\lambda_2(x_1; x_2) = \lambda_2 x_2; \quad (38)$$

where λ_1 and λ_2 are constants.

The model given by (34), (35), (36), (37), and (38) is known as the two-factor CIR model, a special case of which is the Longstaff and Schwartz (1992) model: the state variables in their model represent economic factors affecting expected returns on physical investment and λ_1 equals zero.

Let us assume that the model can be specified in terms of the short rate, r, and the instantaneous variance of changes in the interest rate, V, as in Longstaff and Schwartz (1992). Let us consider as state variables, however, the short rate and the yield spread. First, it is necessary to write the state variable vector process under the risk neutral measure Q:

$$dx = \mu(x)dt + \sigma(x)dW^* \quad (39)$$

where

$$\mu(x) = \begin{bmatrix} \theta_1(\bar{x}_1 - x_1) - \lambda_1 x_1 \\ \theta_2(\bar{x}_2 - x_2) - \lambda_2 x_2 \end{bmatrix} \quad (40)$$

$$\sigma(x) = \begin{bmatrix} C_1\sqrt{x_1} & 0 \\ 0 & C_2\sqrt{x_2} \end{bmatrix} \quad (41)$$

$$dW^* = [dW_1^*; dW_2^*]^T \quad (42)$$

with dW_1^*, and dW_2^*, being the Brownian motions under Q.

Duffie and Kan (1996) show that when μ, $\sigma\sigma^T$, and r are affine, i.e., linear in the vector of state variables x, then the solution to the valuation equation for zero-coupon bonds is exponential affine:

$$P(x; T - t) = \exp[A(T - t) + B(T - t) \cdot x]. \tag{43}$$

The terms $A(\cdot)$ and $B(\cdot)$ are constant for a given maturity, and can be obtained solving an ordinary differential equation.

The solution (43) implies that the yield on the zero-coupon bond is affine:

$$L = \frac{-\log P(x; T - t)}{T - t} = \frac{-A(T - t) - B(T - t) \cdot x}{T - t} \tag{44}$$

As Duffie and Kan observe, we may specify a d-factor term structure model in terms of the yields on any d bonds, ruling out singularities. For the model considered here, Duffie and Kan's condition holds[10] as can be verified from (40) and (41) so we have that

$$L = -a - b_1^* x_1 - b_1^* x_2; \tag{45}$$

where $a \equiv \frac{A(T-t)}{T-t}$ and $b^* = (b_1^*; b_2^*) \equiv \frac{B(T-t)}{T-t}$.

We can thus write the following equation system:

$$Z^* = K^* x \tag{46}$$

with

$$Z^* = (r; L + a)^T \tag{47}$$

$$K^* = \begin{pmatrix} 1 & 1 \\ -b_1^* & -b_2^* \end{pmatrix} \tag{48}$$

Provided that K^* is invertible, i.e. $b_1^* \neq b_2^*$, we can make a change of variables by noting that

$$x = (K^*)^{-1} Z^* \tag{49}$$

Similarly, we can also take the short rate and the yield spread, S, as state variables, since:

$$\begin{aligned} S &= L - r \\ &= -a - (1 + b_1^*) x_1 - (1 + b_2^*) x_2 \\ &= -a - b_1 x_1 - b_2 x_2. \end{aligned} \tag{50}$$

[10] The two-factor CIR model is in fact a special case of the affine-yield models.

We invert the system

$$Z = Kx \tag{51}$$

with

$$Z = (r, S + a)^T \tag{52}$$

$$K = \begin{pmatrix} 1 & 1 \\ -b_1 & -b_2 \end{pmatrix}; \tag{53}$$

to obtain:

$$x_1 = -\frac{b_2}{b_1 - b_2} r - \frac{1}{b_1 - b_2}(S + a) \tag{54}$$

$$x_2 = \frac{b_1}{b_1 - b_2} r + \frac{1}{b_1 - b_2}(S + a). \tag{55}$$

Applying Ito's lemma to the system given by (52) and (53), we get:

$$dr = dx_1 + dx_2 \tag{56}$$

$$dS = -b_1 dx_1 - b_2 dx_2; \tag{57}$$

substituting from (34) and (35),

$$dr = \theta_1(\bar{x}_1 - x_1)dt + C_1\sqrt{x_1}dW_1 + \theta_2(\bar{x}_2 - x_2)dt + C_2\sqrt{x_2}dW_2 \tag{58}$$

$$dS = -b_1[\theta_1(\bar{x}_1 - x_1)dt + C_1\sqrt{x_1}dW_1] - b_2[\theta_2(\bar{x}_2 - x_2)dt + C_2\sqrt{x_2}dW_2] \tag{59}$$

Finally, we replace the values of x_1 and x_2 obtained in (54) and (55) into (58) and (59):

$$dr = \mu_r(r, S)dt + \sigma_{r1}(r, S)dW_1 + \sigma_{r2}(r, S)dW_2 \tag{60}$$

$$dS = \mu_S(r, S)dt + \sigma_{S1}(r, S)dW_1 + \sigma_{S2}(r, S)dW_2 \tag{61}$$

with:

$$\mu_r(r, S) = \frac{\theta_2 b_1 - \theta_1 b_2}{b_1 - b_2}(\bar{r} - r) + \frac{\theta_2 - \theta_1}{b_1 - b_2}(\bar{S} - S) \tag{62}$$

$$\begin{aligned} var(dr) &= \sigma_{r1}^2(r, S) + \sigma_{r2}^2(r, S) \\ &= \frac{C_2^2 b_1 - C_1^2 b_2}{b_1 - b_2} r - \frac{C_2^2 - C_1^2}{b_1 - b_2}(S + a) \end{aligned} \tag{63}$$

$$\mu_S(r;S) = \frac{b_1 b_2(\theta_1 - \theta_2)}{b_1 - b_2}(\bar{r} - r) + \frac{b_1\theta_1 - b_2\theta_2}{b_1 - b_2}(\bar{S} - S) \tag{64}$$

$$\begin{aligned} var(dS) &= \sigma^2_{S1}(r;S) + \sigma^2_{S2}(r;S) \\ &= \frac{C_2^2 b_1 b_2^2 - C_1^2 b_1^2 b_2}{b_1 - b_2} r - \frac{C_2^2 b_2^2 - C_1^2 b_1^2}{b_1 - b_2}(S + a) \end{aligned} \tag{65}$$

It is interesting to consider the special case when the mean reversion parameters of the state variables have the same value:

$$\theta_1 = \theta_2 = \theta$$

Then we have that

$$\mu_r(r;S) = \mu_r(r) = \theta(\bar{r} - r) \tag{66}$$

$$\mu_S(r;S) = \mu_S(r;S) = \theta(\bar{S} - S) \tag{67}$$

which clearly shows that neither the dimension of the drift function of the actual short rate process nor the dimension of the drift function of the yield process necessarily coincide with the number of factors driving interest rates.

Similarly, when $C_1^2 = C_2^2$; from (63) the instantaneous variance is a function of only the short rate. This case is however ruled out by the assumption that the volatility can be taken as a state variable since $C_1^2 \neq C_2^2$ is a necessary and sufficient condition for the following system to be invertible:

$$\begin{pmatrix} r \\ V \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ C_1^2 & C_2^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Therefore, under the model assumptions, the diffusion coefficient of the short rate process is a bivariate function if and only if interest rates are driven by two stochastic factors.

References

Aït-Sahalia, Yacine, 1996, Testing continuous-time models of the spot interest rate, Review of Financial Studies 9(2):385-426.

Aït-Sahalia, Yacine, Peter J. Bickel and Thomas M. Stoker, 1998, Goodness-of-…t tests for regression using kernel methods, Working paper, Princeton University.

Andersen, T. and J. Lund, 1997, Estimating continuous-time models of the spot interest rate, Review of Financial Studies 9(2):385-426.

Backus, David, Silverio Foresi, and Chris Telmer, 1998, Discrete-time models of bond pricing, NBER Working paper.

Boudoukh, J., M. Richardson, R. Stanton and R. F. Whitelaw, 1998, The stochastic behavior of interest Rates: implications from a multifactor, nonlinear continuous-time model, Working paper, University of California at Berkeley.

Broze, Laurence, Olivier Scaillet, Jean-Michel Zakoïan, 1995, Testing for continuous-time models of the short-term interest rate, Journal of Empirical Finance 2, 199-223.

Chan, K. C., G. Andrew Karolyi, Francis A. Longsta¤, and Anthony B. Sanders, 1992, An empirical comparison of alternative models of the short-term interest rate, Journal of Finance 47, 1209-1228.

Chapman, D. A. and N. D. Pearson, 2000, Is the Short Rate Drift actually Nonlinear?, Journal of Finance 55, 355-388.

Cox, J. C., J. E. Ingersoll, and S. Ross, 1985a, An intertemporal general equilibrium model of asset prices, Econometrica 53, 363-384.

———–, J. E. Ingersoll, and S. Ross, 1985b, A theory of the term structure of interest rates, Econometrica 53, 385-407.

Danielsson, J., 1998, Multivariate stochastic volatility models: estimation and a comparison with VGARCH models, Journal of Empirical Finance 5, 155-173.

Du¢e, Darrel, 1996a, Dynamic Asset Pricing Theory, Second Edition, Princeton University.

Du¢e, Darrell and Rui Kan, 1996b, A yield-factor model of interest rates, Mathematical Finance 6, 379-406.

Du¢e, Darrell and Kenneth J. Singleton, 1993, Simulated moments estimation of markov models of asset prices, Econometrica 50, 987-1007.

Dybvig, Philip H., 1989, Bond and bond option pricing based on the term structure, unpublished paper.

Fama, Eugene F., 1989, Term-structure forecasts of interest rates, in‡ation, and real returns, Journal of Monetary Economics 25, 59-76.

Gallant, A. Ronald, and George Tauchen, 1997, Estimation of continuous-time models for stock returns and interest rates, Macroeconomic Dynamics 1, 135-168.

Gouriéroux, C., A. Monfort, and E. Renault, 1993, Indirect inference, Journal of Applied Econometrics 8, 85-118.

Gray, Stephen F., 1996, Modeling the conditional distribution of interest rates as a regime-switching process, Journal of Financial Economics 42, 27-62.

Härdle, W., 1990, Applied Nonparametric Regression, Cambridge University Press.

Honoré, Peter, 1997, Maximum likelihood estimation of non-linear continuous-time term-structure models, Working paper, Aarhus School of Business.

Ingersoll, Jonathan E. Jr., 1987, Theory of Financial Decision Making, Rowman & Little…eld.

Lo, Andrew W., 1998, Maximum likelihood estimation of generalizzed Itô processes with discretely sampled data, Econometric Theory 4, 231-247.

Litterman, R. and J. Scheinkman, 1991, Common factors a¤ecting bond returns, Journal of Fixed Income 1, 54-61.

Longsta¤, F. and E. Schwartz, 1992, Interest rate volatility and the term structure: a two-factor general equilibrium model, Journal of Finance 47(4):1259-1282.

Pearson, N. and Tong-Shen S., 1994, Exploiting the conditional density in estimating the term structure: an application to the Cox, Ingersoll, Ross model,Journal of Finance 49, 1279-1304.

Pedersen, A. R., 1995, A new approach to maximum likelihood estimation of stochastic di¤erential equations based on discrete observations, Scandinavian Journal of Statistics 22, 55-71.

Santa-Clara, P., 1995, Simulated likelihood estimation of di¤usion with an application to the short term interest rate, Working paper, University of California at Los Angeles.

Stanton, R., 1997, A nonparametric model of term structure dynamics and the market price of interest rate risk, Journal of Finance 52, 1973-2002.

Sundaresan, Suresh M., 2000, Continuous-time methods in …nance: A review and an assessment, Journal of Finance 55, 1569-1622.

Vasicek, Oldrich, 1977, An equilibrium characterization of the term structure, Journal of Financial Economics 5, 177-188.

Table 1. Summary statistics of the 3-month T-Bill yield, the yield on the 10-year US Bond, the yield spread, and weekly changes in the 3-month T-Bill. The series covers the period from January 1962 to May 1999.

Variable	Mean	Std. Dev.	Autocorr.
3-m. yield	0.06142	0.02567	0.99494
10-yr. yield	0.05537	0.01371	0.99815
Spread	-0.00604	0.01545	0.98931
3-m. weekly changes	-0.000013	0.00277	0.08328

Correlation Matrix	3-m. yield	10-yr. yield	Spread
3-m. yield	1		
10-yr. yield	0.8641	1	
Spread	-0.8945	-0.5479	1
3-m. changes	-0.0539	-0.0310	0.0620

Table 2. Test of H1. This table shows the test statistic values and the corresponding p-values for the null hypothesis that the yield spread does not contribute to explaining changes in the drift function of the short rate process. The test statistic was computed for di¤erent values of H_0 (the smoothing parameter in (32)) and $\mathbf{h_0}$ (the smoothing parameter in (31)). The cell in the center of the table corresponds to the optimal smoothing parameter values obtained through cross validation.

		h_0		
		0.18	0.19	0.2
H_0	0.14	0:3035 (0:38)	0:2268 (0:41)	0:1539 (0:43)
	0.15	0:4006 (0:34)	0:3252 (0:37)	0:2533 (0:40)
	0.16	0:4896 (0:31)	0:4151 (0:33)	0:3438 (0:36)

Table 3. Test of H1. This table shows the test statistic values and the corresponding p-values for the null hypothesis that the yield spread does not contribute to explaining changes in the conditional instantaneous variance of the short rate process. The test statistic was computed for di¤erent values of H_0 (the smoothing parameter (32)) and h_0 (the smoothing parameter in (31)). The cell in the center of the table corresponds to the optimal smoothing parameter values obtained through cross validation.

		h_0		
		0.095	0.105	0.115
H_0	0.08	2:3512 (0:0094)	3:3059 (0:0004)	4:5669 (0)
	0.09	1:7674 (0:0386)	2:4475 (0:0072)	3:4059 (0:0003)
	0.10	1:3866 (0:0828)	1:8194 (0:0344)	2:5024 (0:0062)

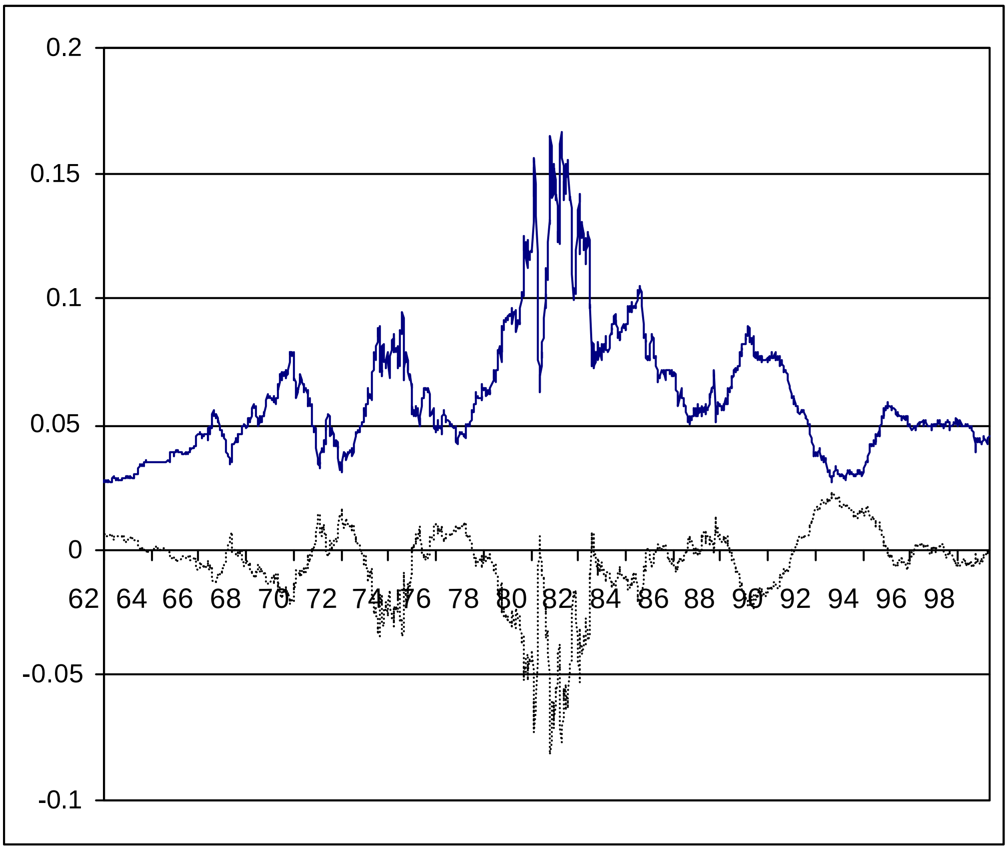

Figure 1. Time series plot of the short-term interest rate (solid line) and the spread between the three-month yield and the 10-year yield (dashed line) for the 1962-1999 period.

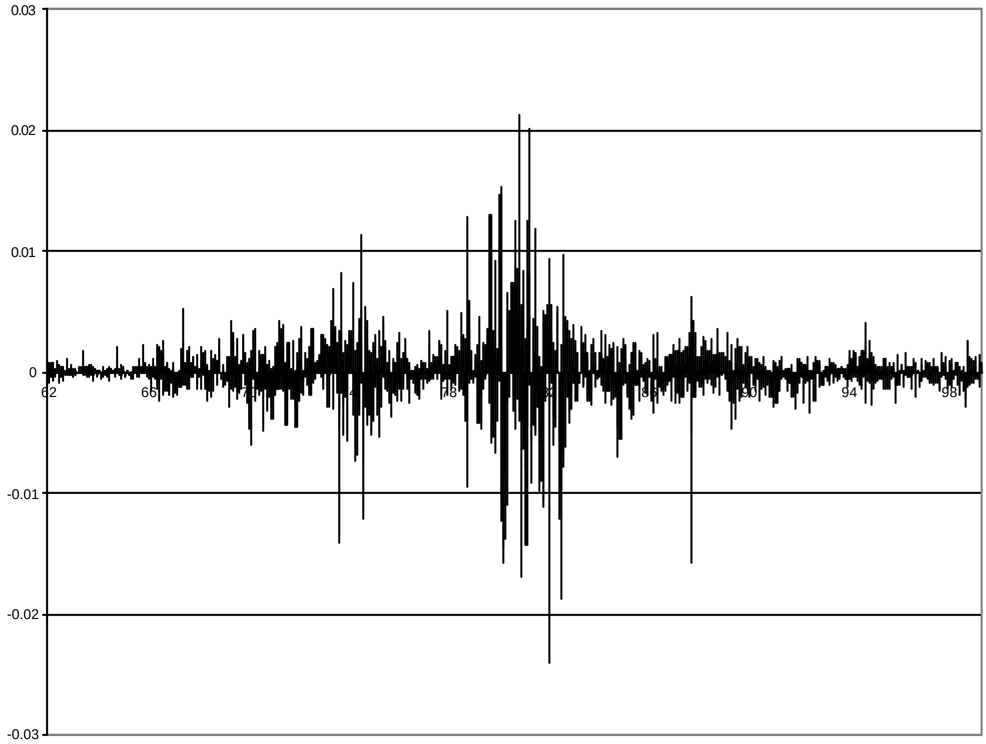

Figure 2. First differences of the time series of the short-term interest rate.

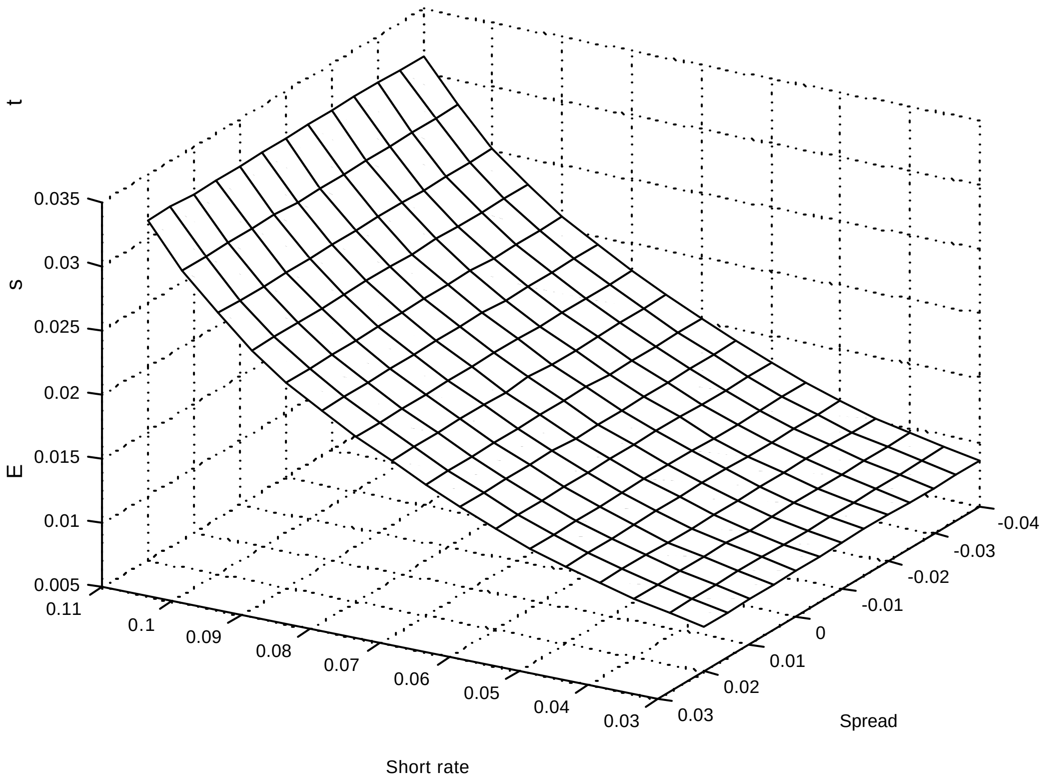

Figure 3. Estimated …rst-order approximation to the di¤usion as a function of the short rate using the Nadaraya-Watson estimator.

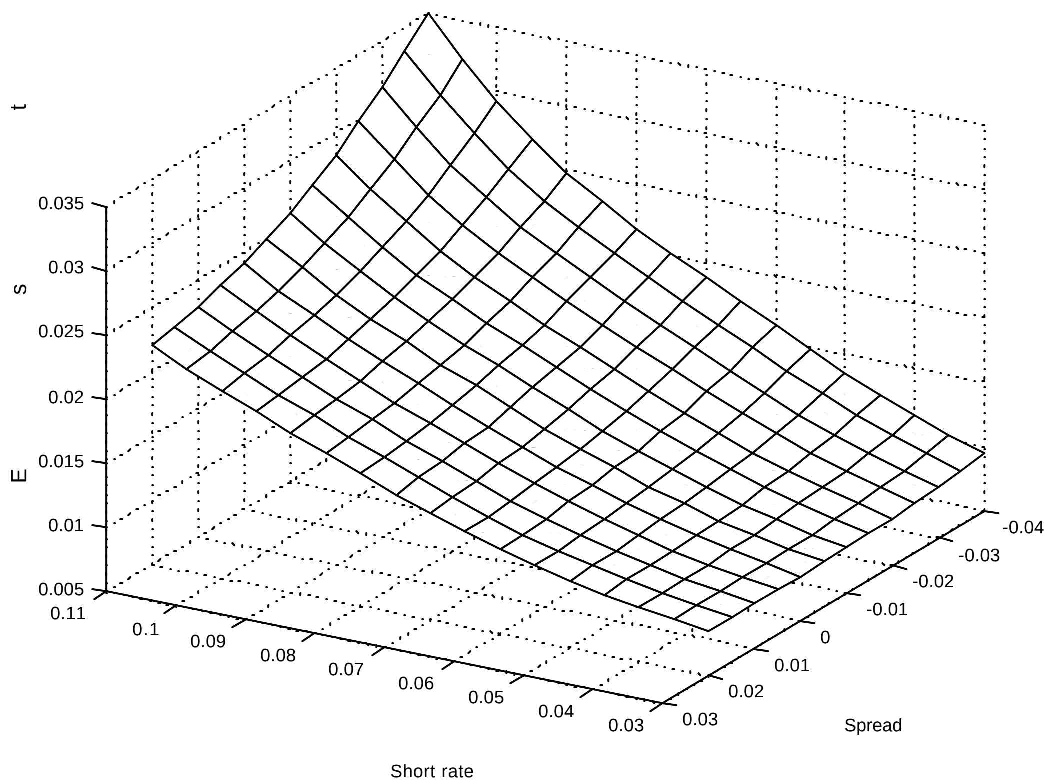

Figure 4. Estimated …rst-order approximation to the di¤usion as a function of the short rate and the yield spread using the Nadaraya-Watson estimator.