A Semiparametric Single-Factor Model of the Term Structure*

Dennis Kristensen†, London School of Economics

June 4, 2004

Abstract

We propose a semiparametric single-factor diffusion model for the term structure of interest rate. The This model is highly flexible and encompasses most parametric single-factor models proposed in the literature. We fit the semiparametric model to a proxy of the Eurodollar short term interst rate and compare it with the most flexible parametric model found in the literature: First directly, by testing the fully parametric model against the semiparametric one. Secondly, we look at how much the bond prices predicted by the competing models differ; this yields an alternative measure of the performance of the models. The fitted semiparametric model picks up nonlinearities which the fully parameteric model cannot capture. This leads to a rejection of the parametric model in favour of the semiparametric model in the direct comparison of the two fitted models. Moreover, the calculated bond prices implied by the two competing models are shown to be significantly different.

*This paper is part of my PhD-thesis at the LSE. I would like to thank my supervisor, Oliver Linton, for valuable suggestions and comments.

†E-mail: d.kristensen@lse.ac.uk Address: R4z19C, FMG, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom

1 Introduction

The short-term interest rate is an important variable in many different areas of economic and financial theory. It has strong implications for the pricing of fixed income securities and interest rate derivatives, e.g. bonds, options, futures, swaps. But it is also used in general asset pricing, and as an input in macroeconomic models, e.g. in the analysis of the business cycle. It is therefore of great interest to obtain a suitable model describing its dynamics. Diffusion processes are widely used for this purpose, which owes to the fact that continuous-time models greatly facilitate the theoretical analysis of financial markets. They prove particularly useful in derivative pricing since continuous-time arbitrage arguments then can be applied, allowing for a relative simple, and yet elegant, solution to the problem, see e.g. Duffie (1996). The theoretical option prices turn out to functionals of the underlying short-term interest rate, so in order to apply these, one has to (i) set up an appropriate model for the interest rate and (ii) calibrate this to the market of interest. There is a huge literature dealing with (i), ranging from relatively simple Markov models of the short-term rate (so-called single-factor models), over multi-factor models where the short-term rate is assumed to depend on several (potentially unobserved) factors, to the class of Heath-Jarrow-Morton (1992) [HJM] type models where a continuum of factors drives the yield curve.

It is however still an open question which of the many proposed models is the most adequate when calibrating it to interest rate data, see Rogers (1995) for a discussion of these issues. There is a large number of studies where different diffusion models of the term structure are implemented using historical interest rate data. These studies have mainly focused on parametric specifications of the diffusion model. But there appears to be no universal model which fits all interest rate data equally well. It is therefore still an open question which model one should choose given a specific interest rate data set. Given this problem, one may benefit from using non- or semiparametric methods since these allow for a degree of flexibility compared to the parametric case.

Here, we focus on the class of single-factor models of the short-term interest rate. This class of term structure models assumes that the short-term interest rate is a Markov process solving a SDE, in which case this state variable drives the whole yield curve. Economic theory puts no restrictions on the drift and diffusion term, except that the resulting short-term interest rate should be positive. Thus, a model for the short-term interest rate can only be judged by how well it fits the available data. In such situations, instead of restricting oneself to a parametric model, it would be more appropriate to "let data speak for itself", which is exactly what non- and semiparametric models do. A number of different parametric models for the SDE has been proposed and tried out on historical data with varying degrees of success. Aït-Sahalia (1996b) fitted some of these parametric models to an interest rate data set, and then tested each model against a nonparametric alternative. The striking conclusion was that none of these could be accepted as the true model; more flexible models were needed. A number of other empirical studies have found similar evidence of nonlinearities both in the drift and diffusion term for this type of data, which the models of Aït-Sahalia (1996a) and Conley et al. (1997) cannot capture, see Ahn and Gao (1999), Bandi (2002), Jiang and Knight (1997, Stanton (1997), Tauchen (1995). The class of single-factor diffusion models are characterised by its drift and diffusion function which can be interpreted as the instantaneous mean and variance respectively. We propose a semiparametric diffusion model where we choose a very flexible parametric form for the diffusion term while leaving the drift term unspecified. The chosen parameterisation of the diffusion term is highly flexible, and the model encompasses most of the parametric models found in the literature. This model is very general; in particular, it includes most of the parametric models suggested in the literature as special cases.

The model can be estimated using the general estimation procedure proposed in Kristensen (2004a). The parametric part is estimated by a profiled version of the log-likelihood, while the drift term is estimated using kernel methods. Since the semiparametric model nests most of the parametric specifications as special cases, we are able to perform a specification test of each of these models against the semiparametric alternative.

We fit our model to a proxy of the short-term Eurodollar interest rate. Various parametric single-factor models have been fitted to this interest rate in a number of empirical studies, see for example Aït-Sahalia (1996b), Elerian et al (2001), Durham (2002). The conclusions drawn in the different studies are not conclusive, and it is not clear which model should be preferred. We reexamine the data set, fitting the proposed semiparametric diffusion model to it. We find nonlinearities in the drift function that even the most flexible parametric model cannot capture. The performance of the semiparametric model is then compared with the parametric model proposed in Aït-Sahalia (1996b); this is the most flexible parametric single-factor model found in the literature. The comparison is made along two lines: First we test the parametric diffusion models against the semiparametric alternative, using the test statistic proposed in Kristensen (2004a). Second, we calculate a range of bond prices predicted by the competing models and see whether they are statistically significant; this is done using the results of Kristensen (2004b). The second comparison is the most useful for practitioners if the end goal with the model is to price bonds and derivatives.

In Section 2, we give an overview over the various parametric single factor models proposed in the literature. In order to apply a single-factor model to price bonds and options the market price for risk has to be determined; in Section 3, we present a general estimation procedure to do this. The Eurodollar interest rate data set is presented in Section 4 along with a discussion of the various studies who have previously examined this. In Section 5, the estimation results for the semiparametric model and the implied bond and derivative prices are presented. We conclude in Section 6.

2 A Semiparametric Single-Factor Model

Single-factor models constitute a relatively simple class of models where the whole term structure is driven by one single state variable, the short-term interest rate. More advanced models such as multi-factor models, and HJM-models might be a more plausible way of describing the dynamics of the term structure, but this comes at the cost of a more difficult and computationally intensive implementation. Most of the applied studies of multi-factor and models only consider linear specifications in order to overcome the difficulties of estimating the model. Prominent examples are Brennan and Schwartz (1979), Chen and Scott (1992, 1993), Dai and Singleton (2000), Longstaff and Schwartz (1992). As an alternative, Ahn, Dittmar & A.R. Gallant (2002) consider a quadratic specification. In these type of models, while a large number of empirical studies have argued that at least two- or three-factor models are needed to fit the term structure properly, it is also found in that the short rate accounts for up to 90 percent of the variation in the data, see e.g. Litterman and Scheinkman (1991). Thus, it is of interest even within a multi-factor framework to find a suitable model for the short rate.

In the class of single-factor models we consider here, the short-term rate solves a time-homogenous stochastic differential equation (SDE) of the form

$$dr_t = \mu(r_t)\,dt + \sigma(r_t)\,dW_t, \tag{1}$$

where $\{W_t\}$ is a standard Brownian motion. We then wish to model the drift term, $\mu : \mathbb{R}_+ \mapsto \mathbb{R}$, and the diffusion term, $\sigma^2 : \mathbb{R}_+ \mapsto \mathbb{R}_+$. A more flexible class of single-factor models can be constructed by using time-inhomogenous SDE's where μ and σ^2 are allowed to depend on time t. These are widely used in the financial industry, since these can be calibrated on a daily basis to deliver a perfect fit of the current yield curve, see e.g. Ho and Lee (1986), Hull and White (1990). But it is not evident that this leads to better out of sample performance and more correct pricing. In particular, these models do not specify the dynamics of the time varying coefficients. This and other arguments against time-inhomogenous models can be found in Dybvig (1997) and Backus et al. (1998).

Within the framework of single-factor models, the price of any bond or interest rate derivative can be shown to be a functional of μ and σ^2, see for example Björk (1998, Chapter 16). So in order to be able to price such claims correctly, one needs to specify μ and σ^2 correctly. The traditional models normally assume a linear drift and diffusion term, but in the past decade a large body of empirical work has indicated that such specifications do not fit observed interest rates very well. Using a misspecified term structure model can have serious implications. For example, as observed in Chapman and Pearson (2001), "the existence and strength of nonlinear mean reversion have important implications for the likelihood of extreme interest rate changes and for the distribution of interest rate changes over long time horizon. As a result, they have significant implications for value-at risk calculations over long horizon and asset-liability management. Moreover, nonlinear mean reversion may also have implications for pricing long-term bonds and interest rate options". We may add, that the presence of nonlinearities in the diffusion term of course also will have important implications in the aforementioned applications.

A large part of the finance literature has focused on σ^2 as the crucial parameter of interest in derivative pricing, while to a certain extent neglecting the role of μ. In a Black-Scholes setting where the underlying variable is a traded asset this focus is correct since only σ^2 enters the derivative pricing formula. But in interest rate derivative pricing, the drift will also enter the formula and can have important effects on the prices. Moreover, in the calibration of the model, both the drift and diffusion term has to be specified correctly in order to avoid biased estimates. Given discrete observations, one can in most cases not separate the estimation of σ^2 from μ, these are invariably linked together in the estimation. So even if one has correctly specified σ^2, misspecification of μ will lead to a biased estimator of σ^2, which in turn will have implications for the pricing of derivatives.

In the past decade, a number of empirical studies have been directed towards finding an appropriate specification of μ and σ^2. Economic theory puts no restrictions on the drift and diffusion term, except that the resulting short-term interest rate should stay positive. Thus, a model for the short-term interest rate can only be judged by how well it fits the available data. In such situations, instead of restricting oneself to a parametric model, it would be more appropriate to "let data speak for itself", which is exactly what non- and semiparametric models do. A number of different parametric models for the SDE has been proposed and tried out on historical data with varying degrees of success. Aït-Sahalia (1996b) fitted some of these parametric models to an interest rate data set, and then tested each model against a nonparametric alternative. The striking conclusion was that none of these could be accepted as the true model; more flexible models were needed. A number of other empirical studies have found similar evidence of nonlinearities both in the drift and diffusion term for this type of data, which the models of Aït-Sahalia (1996a) and Conley et al. (1997) cannot capture, see Ahn and Gao (1999), Bandi (2002), Jiang and Knight (1997, Stanton (1997), Tauchen (1995). The importance of nonlinearities in the two terms should not be downplayed.

We here propose the following semiparametric model where the drift term is unspecified, while the diffusion term follows the flexible parameterisation,

$$dr_t = \mu(r_t)\,dt + \sqrt{\sigma_0 + \sigma_1 r_t + \sigma_2 r_t^{\gamma}}\,dW_t. \tag{2}$$

Some of the most popular (stationary) models are quoted in Table 1. As can be seen, the semiparametric model encompasses most of these models. The most flexible of these is the model proposed by Aït-Sahalia (1996b),

$$dr_t = \left\{\beta_0 + \beta_1 r_t + \beta_2 r_t^2 + \beta_3 r_t^{-1}\right\} dt + \sqrt{\sigma_0 + \sigma_1 r_t + \sigma_2 r_t^{\gamma}}\,dW_t. \tag{3}$$

Observe that the parameterisation of the diffusion term,

$$\sigma^2(x;\theta) = \sqrt{\sigma_0 + \sigma_1 r_t + \sigma_2 r_t^{\gamma}},$$

is the same for the two models, so (3) is a submodel of (2). We then wish to calibrate the semiparametric model to the Eurodollar-rate to see how this perform in comparison to standard parametric models. As a benchmark, we shall use (3).

Table 1: Parametric specifications of the spot rate.

$\mu(r)$	$\sigma^2(r)$	Reference
$\beta(\alpha - r)$	σ^2	Vasicek (1977)
$\beta r(\alpha - \log(r))$	$\sigma^2 r^2$	Brennan and Schwartz (1979)
$\beta(\alpha - r)$	$\sigma^2 r^2$	Courtadon (1982)
$\beta_1 r + \beta_2 r^{-(1+\gamma)}$	$\sigma^2 r^{\gamma}$	Marsh and Rosenfeld (1983)
$\beta(\alpha - r)$	$\sigma^2 r$	Cox et al. (1985)
$\beta(\alpha - r)$	$\sigma^2 r^{\gamma}$	Chan et al. (1992)
$\beta_0 + \beta_1 r + \beta_2 r^2$	$(\sigma_0 + \sigma_1 r)^2$	Constantinides (1992)
$\beta(\alpha - r)$	$\sigma_0^2 + \sigma_2^2 r$	Duffie and Kan (1996)
$\beta_0 + \beta_1 r + \beta_2 r^2 + \beta_3 r^{-1}$	$\sigma_0 + \sigma_1 r + \sigma_2 r^{\gamma}$	Aït-Sahalia (1996b)
$\beta_0 + \beta_1 r + \beta_2 r^2 + \beta_3 r^{-1}$	$\sigma^2 r^{\gamma}$	Tauchen (1995), Conley et al. (1997)
$\beta_0 + \beta_1 r + \beta_2 r^2$	$\sigma_0 + \sigma_1 r + \sigma_2 r^3$	Ahn and Gao (1999)
$\beta_0 + \beta_1 r + \beta_2 r^2 + \beta_3 r^{-1}$	$\sigma_0 + \sigma_1 r + \sigma_2 r^2 + \sigma_2 r^3$	Elerian et al (2001)

We now give a brief description of our estimator in the semiparametric model; for a detailed treatment, we refer to Kristensen (2004a). The main problem is that μ is unspecified so standard parametric methods are obviously not directly applicable. Under the assumption of stationarity, one can however identify the drift in terms of the parametric version of the diffusion term together with the marginal density. This will be the basis of our estimation strategy. The semiparametric model in (2), is characterised by a parametric diffusion term,where $\theta = (\sigma_0, \sigma_1, \sigma_2, \gamma)'$, and an unspecified drift term, $\mu(x)$. Assuming stationarity of the short-term interest rate process $\{r_t\}$, one can express μ as a functional of σ^2 and the stationary, marginal density, π:

$$\mu(x) = \mu(x; \theta, \pi) = \frac{1}{2\pi(x)} \frac{\partial}{\partial x} \left[\sigma^2(x; \theta)\, \pi(x)\right]. \tag{4}$$

We then propose to estimate the marginal density π by nonparametric kernel methods; see Silverman (1986) for an introduction. For a kernel function $K(\cdot)$ and a bandwidth h, we define

$$\hat{\pi}(x) = \frac{1}{nh} \sum_{i=0}^{n-1} K\left(\frac{x - X_i}{h}\right). \tag{5}$$

Under suitable regularity conditions, $\hat{\pi}(x) \to^P \pi(x)$. We then plug this estimator into (5) to obtain

$$\hat{\mu}(x; \theta) = \frac{1}{2\hat{\pi}(x)} \frac{\partial}{\partial x} \left[\sigma^2(x; \theta)\, \hat{\pi}(x)\right]. \tag{6}$$

We now have an estimator of the drift for any given value of θ. So in a sense, we are now back in the fully parametric framework, and can estimate θ using standard methods. Here, we suggest to use likelihood-methods: Let $p_s(y|x)$ denote the conditional density of r_{t+s} conditional on $r_t = x$. The conditional density proves to be a functional of μ and σ^2, $p_s(y|x; \mu, \sigma^2)$. We then substitute in the parametric version of σ^2 and the semiparametric estimator of μ,

$$\hat{p}_s(y|x; \theta) = p_s\left(y|x; \hat{\mu}(\cdot; \theta), \sigma^2(\cdot; \theta)\right).$$

Given discrete observations of the short-term interest rate, $r_0, r_\Delta, r_{2\Delta}, ..., r_{n\Delta}$, where $\Delta > 0$ is the time distance between, we may then estimate θ by

$$\hat{\theta} = \arg\max \frac{1}{n} \sum_{i=1}^{n} \log \hat{p}_\Delta\left(r_{i\Delta}|r_{(i-1)\Delta}; \theta\right).$$

Kristensen (2004a, Theorem 5) gives regularity conditions under which

$$\sqrt{n}(\hat{\theta} - \theta) \to^d N\left(0, H_{\theta\theta}^{-1}\left(H_{\theta\theta} + V\right) H_{\theta\theta}^{-1}\right),$$

where $H_{\theta\theta} = E\left[\partial^2 \log p_\Delta\left(r_{i\Delta}|r_{(i-1)\Delta}; \theta\right)/\partial\theta\partial\theta'\right]$ is the information matrix for known π, and V is an adjustment term due to the fact that we are using an estimator of π instead of the actual density itself.

Given the asymptotic properties of $\hat{\theta}$, one can easily show that

$$\sqrt{nh^3}(\hat{\mu}(x; \hat{\theta}) - \mu(x)) \to^d N(0, V_\mu(x)),$$

where $V_\mu(x) = \frac{1}{4} \int |K'(z)|^2 dz \sigma_0^4(x)/\pi_0(x)$, under smoothness conditions on π and regularity conditions on the bandwidth and the kernel. A consistent estimator of the pointwise variance can be obtained by

$$\hat{V}_{\mu}(x) = \frac{[\int K'(y)^2 \, dy]\sigma^4(x;\hat{\theta})}{4\hat{\pi}(x)} \tag{7}$$

The fully parametric model (3) can also be estimated by likelihood-methods. In that model, $\gamma = (\beta^\top, \theta^\top)^\top \in \mathcal{G} = \mathcal{B} \times \Theta \subseteq \mathbb{R}^4 \times \mathbb{R}^4$ is the vector of unknown parameters, where θ is given as before and $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^\top$, and the conditional density is given by

$$p_s(y|x;\gamma) = p_s\left(y|x;\mu(\cdot;\beta), \sigma^2(\cdot;\theta)\right).$$

We may then estimate the unknown parameter by

$$\hat{\gamma} = \arg\max \frac{1}{n}\sum_{i=1}^{n} \log p_\Delta\left(r_{i\Delta}|r_{(i-1)\Delta};\gamma\right).$$

Under suitable regularity conditions, see e.g. Aït-Sahalia (2002, Proposition 3),

$$\sqrt{n}(\hat{\gamma} - \gamma) \to^d N\left(0, H^{-1}\right),$$

where

$$H = \begin{bmatrix} H_{\beta\beta} & H_{\beta\theta} \\ H_{\theta\beta} & H_{\theta\theta} \end{bmatrix}$$

with $H_{\beta\beta} = E\left[\partial^2 \log p_\Delta\left(r_{i\Delta}|r_{(i-1)\Delta};\theta\right)/\partial\beta\partial\beta'\right]$, $H_{\theta\beta} = E\left[\partial^2 \log p_\Delta\left(r_{i\Delta}|r_{(i-1)\Delta};\theta\right)/\partial\theta\partial\beta'\right]$, $H_{\beta\theta} = H_{\theta\beta}^\top$, and $H_{\theta\theta}$ is given as before. In particular, we have that $\sqrt{n}(\hat{\theta} - \theta) \to^d N(0, H_{\theta\theta}^{-1})$ with a smaller asymptotic variance than for the semiparametric model.

A problem with the approach outlined above is that p_Δ cannot be written on an analytical form. So in order to obtain $\hat{\theta}$ and $\hat{\gamma}$ in practice, one will normally substitue p_Δ for some approximation of it. This can either be done by numerical/approximative methods, see for example Aït-Sahalia (1999, 2000) and Lo (1988), or simulation-based methods, see for example Durham and Gallant (2002), Elerian et al (2001), Nicolau (2001), Pedersen (1995).

Observe that since the parametric model is nested within our semiparametric one, our semiparametric estimates will remain consistent under the parametric model. This allows us to test the parametric model against the semiparametric alternative. The crucial difference between the two models is the drift, μ, which is left unspecified in the semiparametric model while a specific parameterisation is assumed in the parametric one. We propose to test the parametric model against the semiparametric one by directly comparing the estimates of the drift obtaine from the two models against each other. We define the test statistic

$$T_n \equiv T_n \equiv nh^3 \sum_{j=1}^{d} \frac{\left[\mu(x_j;\hat{\beta}) - \hat{\mu}(x_j)\right]^2}{\hat{V}_\mu(x_j)} \tag{8}$$

where $\mu(x;\hat{\beta})$ and $\hat{\mu}(x;\hat{\theta})$ are the drift estimators in the parameteric and semiparametric model respectively, $\{x_i\}_{i=1}^N$ is N distinct points in $(0,\infty)$. Under the hypothesis $H_0 : \mu(x) = \mu(x;\beta)$ for some $\beta \in \mathcal{B}$, we obtain

$$\sqrt{nh^3} T_n \xrightarrow{d} \chi^2(N).$$

3 Estimation of the Risk Premium

Within the framework of the single-factor interest rate model, formulae for bond and option prices can be derived. An important ingredient in these is the so-called "market price of risk" process which we denote $\{\lambda_t\}$. In order to apply our calibrated models to the pricing of such securities we therefore have to obtain an estimate of this process. In the following, we go through some of the methods suggested in the literature. For a more detailed treatment, we refer to Garcia, Ghysels and Renault (2004).

We start out with a general model for an option price Π. Assume that the price satisfies

$$\Pi = \Gamma(X, Z) + \varepsilon,$$

where X is a collection of random variables (including, for example, the current value of the underlying asset), Z a collection of deterministic characteristics associated with the option (time to maturity, type of pay-off function etc.), and ε is an error term (present due to e.g. pricing errors, failure of the theoretical model to perfectly match the observed data).

A branch of the empirical option pricing literature proposes to estimate the function Γ non- or semiparametrically, thereby not having to specify the dynamics of the underlying variable. Prominent examples of this approach is Aït-Sahalia and Lo (1998), Aït-Sahalia and Duarte (2003), Bondarenko (2003). This approach has the advantage of not imposing any restrictions on the dynamics of the underlying variable, and not making the estimation of any risk premium necessary. On the other hand, the precision of predicted option prices will suffer from the slower convergence of nonparametric estimators.

Restricting the dynamics of the term structure to be of diffusion type gives us additional information about the function. In the arbitrage-free framework of the single-factor model, it holds that

$$\Gamma(x, z) = \Gamma(r, g, t, T) = E^Q\left[g(r_T)\exp\left[-\int_t^T r_u du\right] |r_t = r\right],$$

c.f. Chapter 3. Assuming that $\{r_t\}$ is time-homogenous under Q, the formula simplifies further to

$$\Gamma(r, g, \tau) = E^Q\left[g(r_\tau)\exp\left[-\int_0^\tau r_u du\right] |r_0 = r\right], \tag{9}$$

where $\tau = T - t$ is time to maturity. The dynamics of $\{r_t\}$ under Q depends on the risk premium λ,

$$dr_t = \{\mu(r_t) + \lambda(r_t)\sigma(r_t)\} dt + \sigma(r_t) dW_t. \tag{10}$$

Let $\{(\Pi_{ij}, r_i, g_j, \tau_j) | 1 \leq i \leq n,\ 1 \leq j \leq J\}$ be a collection of observed option prices together with the associated observed short rate and the characteristics. Taking μ and σ^2 for given (in our case, we will have preliminary estimates of these), the only unknown is λ. This yields the following regression model,

$$\Pi_{ij} = \Gamma(r_i, g_j, \tau_j; \lambda) + \varepsilon_{ij},$$

where the function Γ takes the form (9), and we assume that $E[\varepsilon_{ij}|r_i] = 0$. We may then estimate the unknown function λ by for example least squares. In practice this means that we choose λ such that the option prices implied by the single-factor model mimic the observed ones as closely as possible. It is still an open question whether λ can be identified in the above regression model. Observe that in fact the estimation problem here in some sense is the inverse of the one considered in Kristensen (2004b). While there we had derived the asymptotic properties of the solution to the PDE given estimators of μ and σ^2, we here have observed solutions to the PDE (bond and derivative prices) from which we wish to extract an estimator of the one of the coefficients driving the PDE. The main problem here is then the inversion of the functional $u = \Gamma\left(\mu - \lambda\sigma, \sigma^2\right)$, with Γ given in Kristensen (2004b), w.r.t. its first argument such that $\mu - \lambda\sigma = \Gamma^{-1}\left(u, \sigma^2\right)$. If the inverse of Γ is well-defined, identification of λ is ensured. Assuming an affine specification of μ, σ^2 and λ^2, a closed form expression of Γ is available, and one can in this setting show that the parameters entering λ are identifiable, c.f. Duffie and Kan (1996). In the general case however, the function Γ is a complicated functional of λ which cannot be written on analytical form, and the identification problem is not easily resolved. To the author's knowledge, no general results concerning identification of λ exist. Here, we shall therefore simply assume that λ is identified. A nonparametric estimator of λ can be obtained by the method of sieves. Normally however, one assumes a parametric version, $\lambda(r) = \lambda(r;\theta)$, where θ is an unknown finite-dimensional parameter. For example, $\lambda(r) = \lambda$ in Vasicek (1977) and Aït-Sahalia (1996a), and $\lambda(r) = \lambda/\sigma\sqrt{r}$ in Cox-Ingersoll-Ross (1985). The general estimation procedure presented above is also applicable to multifactor models.

Some alternative estimation procedure can be found in the literature. Consider the yield of a zero-coupon bond with maturity at time T, $Y_t(T) = \log(B_t(T))/(T-t)$ with $B_t(T)$ being the price of zero-coupon bond with maturity at time $T \geq t$. Vasicek (1977) observed that

$$\frac{\partial Y_t(T)}{\partial T}|_{T=t} = \frac{1}{2}\left[\mu(r_t) - \sigma(r_t)\lambda(r_t)\right],$$

and proposed to use observed yields to approximate the right hand side of the equation. Jiang (1998) derived an expression of the risk premium in terms of two different yields and the dynamics of those. By Itô's Lemma, $Y_t(T)$ solves an SDE,

$$dY_t(T) = \alpha(r_t, T)\,dt + \kappa(r_t, T)\,dW_t, \tag{11}$$

where the no-arbitrage assumption imposes restrictions on the drift and diffusion term. In fact, the following relation have to hold for any $T_1, T_2 > 0$,

$$\lambda(r_t) = \frac{\Delta Y_t(T_1, T_2) + \frac{1}{2}\left[\tau_1^2\kappa^2(r_t, T_1) - \tau_2^2\kappa^2(r_t, T_2) + \tau_2\alpha(r_t, T_2) - \tau_1\alpha(r_t, T_1)\right]}{\tau_2\kappa(r_t, T_2) - \tau_1\kappa(r_t, T_1)}, \tag{12}$$

where $\Delta Y_t(T_1, T_2) = Y_t(T_1) - \Delta Y_t(T_2)$ is the yield spread and $\tau_i = T_i - t$, $i = 1, 2$, is the time to maturity. Jiang (1998) then proposes to estimate λ by choosing two representative bonds, fit a diffusion model of the type (11) to their yields using nonparametric methods, and then plug these estimators into (12).

4 The Data

The data set consists of 5505 daily observations from June 1, 1973 to February 25, 1995 of the 7-day Eurodollar rate. Eurodollars are any dollar denominated deposit in commercial banks outside of the U.S. Eurodollar accounts are not transferable but banks can lend on the basis of the Eurodollar accounts they hold. The interest rate charged for Eurodollar loans is often based upon the London Interbank Offer Rate (LIBOR). The Eurodollar rate is considered the benchmark interest rate for corporate funding, and Eurodollar futures are by far the most actively traded interest-rate product. A more detailed account of the Eurodollar market is found in Burghardt (2003).

The 7-day rate should be a reasonable good proxy for the short-term interest rate.[1] We do not use a lower maturity since this might lead to various market micro structure problems. We shall not attempt to remove any seasonal effects, such as weekend effects, from the data, and simply treat Monday as the first day after Friday such that we have 252 observations per year. We measure time in days such that the time between observations, $\Delta = 1/252$.

A number of single-factor models have been fitted to this particular data set. Aït-Sahalia (1996a) fitted a semiparametric model with linear drift and unspecified diffusion term to the 7-days Eurodollar rate and found strong nonlinearities in the diffusion term. He compared his semiparametric model to the CIR and Vasicek-model (which both are nested within his semiparametric model), both in terms of the actual model fit but also the resulting bond and option prices. He concluded that the two parametric model were significantly different from his semiparametric specifications, both in terms of model estimates and implied prices.

Aït-Sahalia (1996b) reinvestigated the data set, setting up a nonparametric specification test which allowed him to test any parametric (stationary) diffusion model against a nonparametric alternative. He tested a number of parametric single-factor models; all of these were rejected by his test. As an answer to the failure of the standard models, Aït-Sahalia (1996b) then proposed The drift term is parameterised such that the drift can be nearly zero in a large part of the domain, but still ensure mean reversion in the tails. The specification test accepted this new model as giving an adequate description of the data. The conclusions in Aït-Sahalia (1996b) have been questioned in a number of later studies however due to poor finite sample performance of his proposed test. Chapman and Pearson (2000) and Pritsker (1998) presented evidence of that the asymptotic distribution of the test statistic delivers a very poor approximation of the finite sample distribution in the presence of strong serial correlation. And this is exactly the case with the interest rate data set used in Aït-Sahalia (1996b). In particular, Pritsker (1998) demonstrated in a simulation study that the test statistic is prone to reject correctly specified model when using the asymptotic critical values.

Hong and Li (2002) and Thompson (2000) have suggested alternative specification tests which should exhibit improved finite sample properties compared to the one proposed by Aït-Sahalia (1996b). The main idea in both studies is to apply a transformation of the data which should decrease the serial correlation. Hong and Li (2002) applied their test to the same data set as used in Aït-Sahalia (1996b), and still rejected all the standard parametric models examined in Aït-Sahalia (1996b), but also the model in (3). They argued that the data exhibits strong non-Markovian behaviour, and that the class of single-factor (Markov) models is too restrictive in this sense. The application of the specification test in Thompson (2000) to the Eurodollar 7-days rate data set also lead to the rejection of all the parametric models, including the one in (3). His explanation for the failure of the parametric models

[1] Chapman et al (1999) find that using the 7-day rate as a proxy when fitting standard parametric short-term interest rate model does apparently not lead to significantly different implied bond prices.

differs from the one of Hong and Li (2002) however. He argued that the problem is that the driving noise process is misspecified. By using either jump- or gamma-processes instead of the Brownian motion, he was able to accept a relatively simple parametric model.

In Bandi (2002), the kernel estimators of Bandi and Phillips (2003) were employed to fit the model (1) nonparametrically to the Eurodollar data set. Nonlinearities were present in both the estimated drift and diffusion term. In particular, the drift estimate was nearly zero in a major part of the data domain, but exhibited mean-reversion in its right tail while the behaviour in the left tail was inconclusive. The kernel estimator is robust to departures from the stationarity assumption normally imposed in single-factor models. There is no clear-cut evidence of non-stationarity in the data however.

In a fully parametric framework, a number of studies have reexamined the model (3). Elerian et al (2001) fitted the model to the same Eurodollar data set in a Bayesian framework using simulated maximum-likelihood techniques. They found that the parameters σ_2 and γ in (3) were difficult to identify in the data, and instead proposed the following slightly different parameterisation,

$$dr_t = \left\{\beta_0 + \beta_1 r_t + \beta_2 r_t^2 + \beta_3 r_t^{-1}\right\} dt + \sqrt{\sigma_0 + \sigma_1 r_t + \sigma_2 r_t^2 + \sigma_3 r_t^3} dW_t. \tag{13}$$

While the parameter estimates were easier to pin down, this new model did not appear to fit the data very well, and they found that any significant mean reversion only appeared in the estimated drift when this was assumed in the prior.[2] Elerian et al (2001) concluded that a single-factor model was not a very appropriate description of the data, and conjectured that a stochastic volatility model of the type in Andersen and Lund (1997) would be needed. This conjecture was confirmed in Durham (2002) where the model (3) was compared with a stochastic volatility model using simulated maximum-likelihood methods. He found that the single-factor model fit the data poorly, while the stochastic volatility (2-factor) model on the other hand did a good job. Similarly, Hurn and Lindsay (2002) found that the parameters in both the drift and diffusion term of (3) were difficult to identify and suggested the use orthogonal polynomials to amend this problem. They estimated the model using discrete time approximations however which means that their estimates are very likely to suffer from discretisation bias.

The general conclusion to be drawn form the empirical studies seems to be that the Eurodollar short-term rate is difficult to model properly within a single-factor framework and that the use of multi-factor models improve on the fit. If one restricts attention to single-factor models, inconclusive results have been obtained about the degree of (non-linear) mean reversion. Some studies have found evidence of nonlinear mean-reversion while others have rejected this hypothesis. Moreover, the more advanced parametric models seems to suffer from poor identification of the parameters. In the following, we shall re-examine the single-factor models using the above semiparametric diffusion model as a starting point.

The raw data is plotted in levels and differences in Figure 1 and 2 respectively. Figure 1 shows that the data exhibits a very strong correlation over time as is usually found in interest rate data. Taking differences, we see in Figure 2 that the short-term interest rate behaves as heteroskedastic white noise, which indicates that it is close to being a random walk. One should also notice the significant different behaviour of the rate in the period 1979-1981. The significant break in the data set in this period is due to the so-called Fed-Experiment where the U.S. Federal Reserve targeted monetary aggregates instead

[2] Similar findings are reported in Jones (2003)

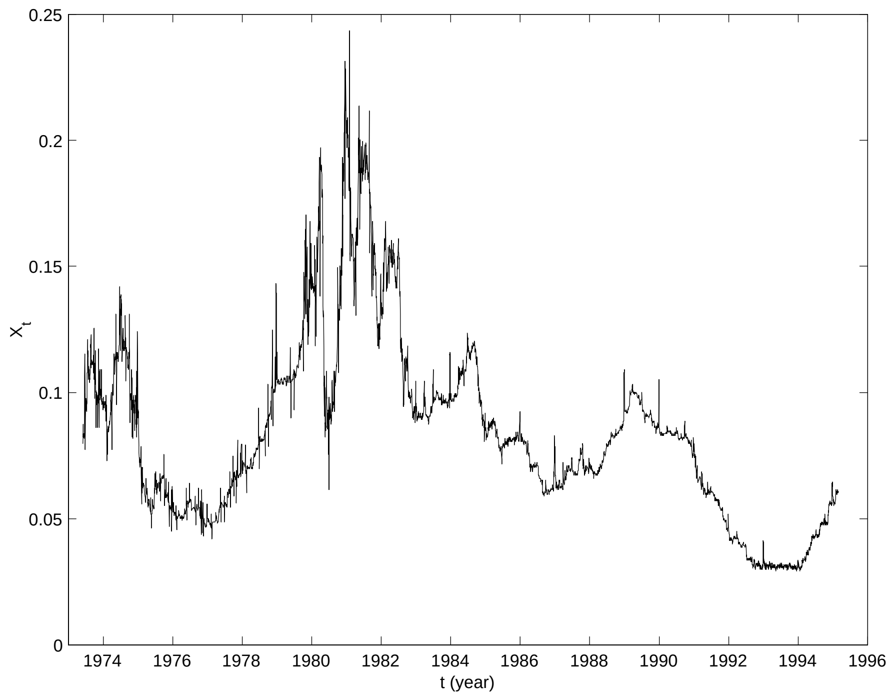
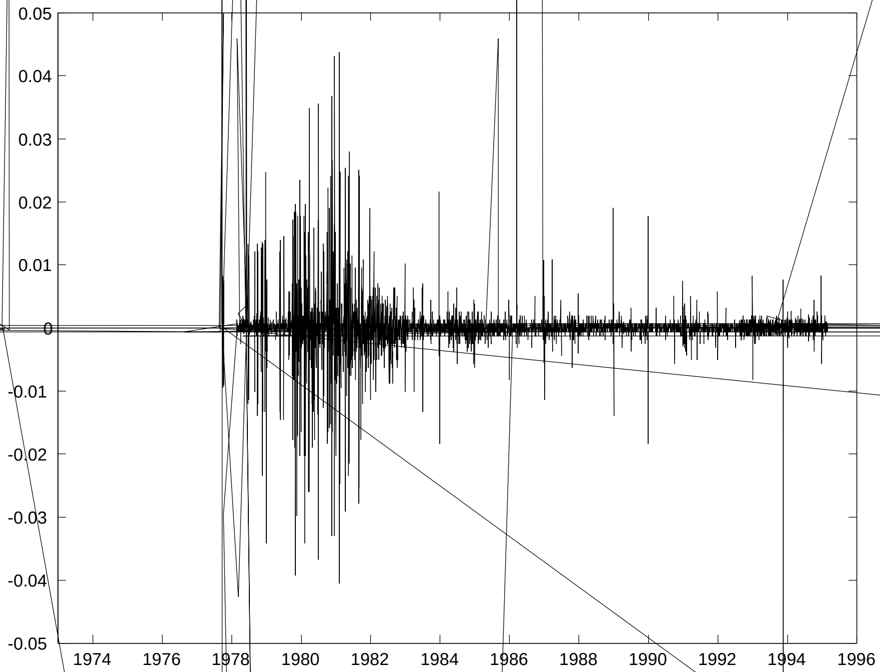


Figure 1: The Eurodollar spot rate in levels, 1973-1995.

of, as done before and after, interest rate levels. One may argue that the data from this period should be left out or a Markov-switching model should be used; the latter alternative is pursued in Ang and Bekaert (2002). We shall in the following try to fit a model both to the full sample and the subsample 1982-1995, the latter excluding the period of the Fed-Experiment.

In Table 2, descriptive statistics of the data set is presented for the full period and the subperiod 1982-1995. Notably is the very strong autocorrelation in levels, while the correlation in the differenced data is decreasing fairly quickly. This is the case for both the full sample and the subsample, but it is less pronounced in the full sample due to the Fed-Experiment. In a linear modelling framework, the persistent autocorrelation in the levels would lead one to conclude that the process is non-stationary; the resulting estimate of the drift term is insignificantly different from zero such that the interest rate would be deemed to be driven by a random walk and thereby being non-stationary. Moreover, all the descriptive statistics of the full sample are significantly different from the ones of the subsample. This confirms that the period of the Fed-Experiment lead to significantly different behaviour of the interest rate.

We now investigate further the seemingly nonstationary behaviour of the data in a linear framework. This is done by unit root tests: We set up a standard AR model, $\Delta r_i = \alpha_0 + (\phi - 1)\, r_{i-1} + \sum_{k=1}^{20} \alpha_k \Delta r_{i-k} + \varepsilon_i$, estimate the parameters by least squares, and then perform the Augmented Dickey-Fuller (ADF) test as outlined in Said and Dickey (1984). We also implement the $Z(t)$-test as proposed in Phillips (1987) using the model $r_i = \alpha_0 + \phi r_{i-1} + \varepsilon_i$; this should have less distortions in the presence of MA(1)-errors (see Phillips and Perron, 1988). The results are reported in Table 3 where we reject the hypothesis of a unit root for large negative values of the test statistic. The ADF test leads to


0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996

4

non-conclusive results in the full sample with rejecting a unit root at a 10% level while accepting the hypothesis at a 5% level. The $Z(t)$ for the full sample on the other hand clearly rejects the hypothesis on a 5% level. In the subsample, both tests clearly accept the unit root hypothesis. So in a linear framework there is mixed evidence of non-stationarity in the full sample, while the subsample appear to be driven by a random walk. But if the drift is non-linear the above regression model is misspecified, and the estimation and test results invalid. A non-linear drift and diffusion term may lead to different results. As we shall see in the next section, the estimated drift and diffusion term generate processes with seemingly non-stationary behaviour in a major part of its domain with the process as a whole being stationary.

Table 3: Unit root test results.

		$\hat{\phi}$	Test Statistic	5% critical value	10% critical value
1973-1995	ADF test	0.9970	-2.60	-2.87	-2.59
	$Z(t)$ test	0.9935	-3.27	-2.87	-2.59
1982-1995	ADF test	0.9991	-1.47	-2.87	-2.59
	$Z(t)$ test	0.9963	-2.98	-2.87	-2.59

5 Empirical Results

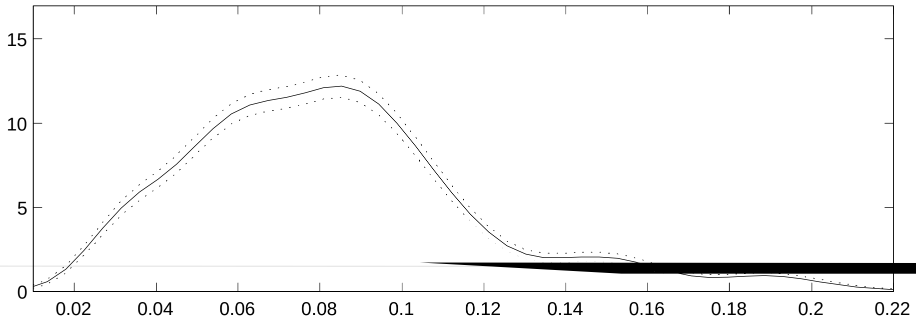
5.1 Estimation of the Single-Factor Model

In this section, we estimate the semiparametric model in (2) together with the fully parametric submodel (3).

In Figure 3, we report the nonparametric kernel estimate of the marginal density as given in (5) for the data, using both the full sample and the subsample. A Gaussian kernel was used while the bandwidth was chosen by cross validation. The density estimate has a highly non-Gaussian shape for both samples. In the full sample, the density has a very long right tail due to the Fed-Experiment during which very high levels of interest rates were observed. Excluding these, one is left with the density in the second plot which is bimodal; a slightly smaller choice of bandwidth will give a trimodal shape of the density with the last mode being around 0.12. If $\{r_t\}$ is stationary, this indicates that contrary to what most of the models in Table 1 would suggest, the interest rate here evolves around not just one but two-three steady states. This is a strong indication of nonlinearities in the drift and the diffusion term.

We now report the estimates for our semiparametric model. As a benchmark, we also fit the model in (3) to the data set. Both models are estimated using the approximate log-likelihood suggested by Aït-Sahalia (2002) with order of approximation $M = 6$. For the semiparametric model, the bandwidths are chosen as described below, and we trim the data at the 1st and 99th empirical percentile. The reported standard errors

In the fully parametric model we had problems obtaining a precise estimate since the likelihood curve is relatively flat in the vicinity of the optimum. This was particularly a problem along the


15
10
5
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
0.22

not influence the estimates.

The semiparametric estimates proved to be fairly robust over a range of bandwidth choices. An initial set of bandwidths was chosen by using standard cross-validation methods. We then generated a grid of bandwidth centered around this initial one. For each set of bandwidths, we obtained an estimate of θ. In Table 4, the results of this sensitivity check for the subsample are reported for seven different bandwidth sets. The bandwidths decrease as one moves from left to right in the table with the 7th bandwidth being the cross-validated one. Relatively large bandwidths choices, slightly bigger than the initial cross-validated ones, gave the most reliable estimates for our sample. This probably stems from the fact that the cross-validation procedure does not take into account the dependence in the data which is very strong in our case.

Table 4: Sensitivity check of semiparametric estimates.

Bandwidth	1	2	3	4	5	6	7
σ_0	0.0112	0.0110	0.0113	0.0110	0.0104	0.0103	0.0109
σ_1	0.0129	0.0124	0.0127	0.0131	0.0120	0.0121	0.0116
σ_2	21.5643	24.1624	23.1004	22.3427	24.0175	23.3548	21.9449
γ	4.2699	4.2598	4.2978	4.2296	3.8650	3.7434	3.9873

Notes: Bandwidth 1 ($10^{-3}\times$): h_1= 10.09, h_2= 12.23, h_3= 13.89, h_4= 15.10, h_5= 17.22, h_6= 18.74, h_7= 19.97, h_8= 21.11.
Bandwidth 7 ($10^{-3}\times$): h_1= 5.17, h_2= 6.33, h_3= 6.97, h_4= 7.49, h_5= 7.85, h_6= 8.84, h_7= 9.60, h_8= 10.36.

A plot of the semiparametric estimates of σ^2 for the first four set of bandwidths can be found in Figure 4. The estimates do not vary a great deal across the different bandwidths which indicates that our estimation procedure is fairly robust. Similar results were obtained for the full sample.

Next, we report our nonparametric estimates of the drift function. Here, we use slightly higher bandwidths compared to the ones we used in the estimation of θ. In Figure 5, the nonparametric estimator of μ is plotted with pointwise 95%-confidence bands for both the full sample and the subsample. The confidence bands are calculated using (7). The range over which we plot the estimates was chosen as the one of the data from the subsample. We see that for both periods, the nonparametric estimate of μ is insigni

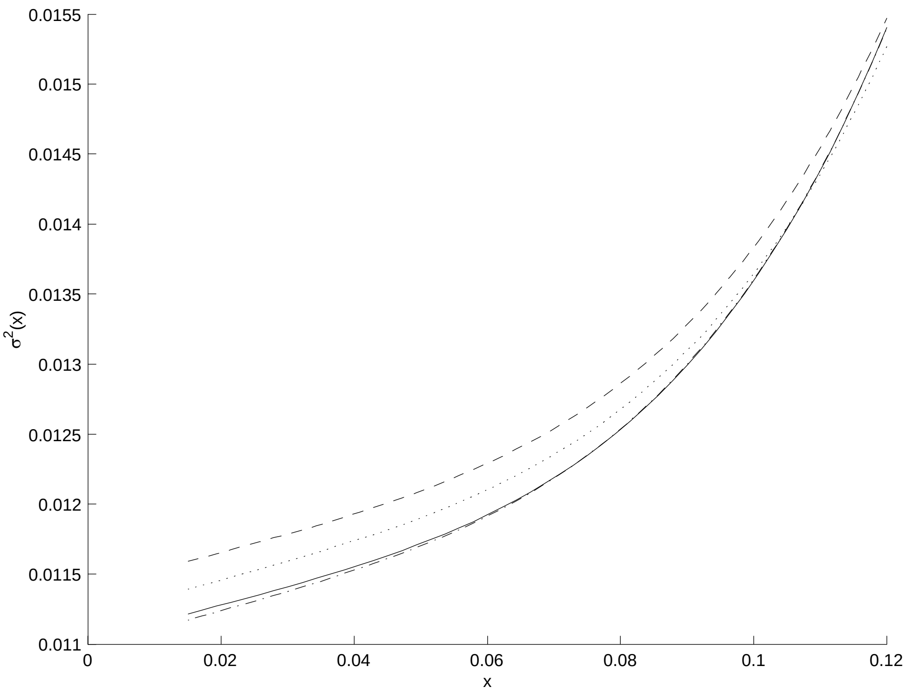
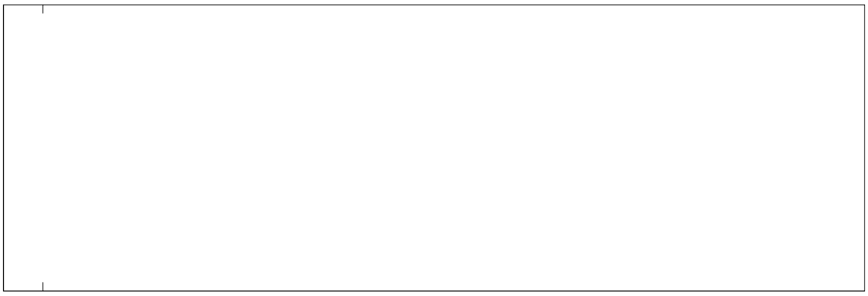


Figure 4: Sensitivity check: Semiparametric estimates of $\sigma^2(r;\theta)$ for different bandwidths.

interest rate in the full sample and the subsample are not the same, and we are only able to compare the estimates within the domain of the subsample. Mean reversion for the full sample first appears further out in its right tail, c.f. figure 6. In total however the two drift estimates are not significantly different from each other in the domain of the data in the subsample.

The results found here are compatible with other empirical studies of the short-term interest rate. Jiang and Knight (1997) and Bandi (2002) obtain nonparametric kernel estimates of μ that exhibit a similar behaviour. In a discrete time framework, parametric Markov switching AR-models have proved to be able to generate the same type of dynamics as the ones we have found here. For a recent application of this type of models to the short term interest rate see Ang and Bekaert (2002).

One should however be careful with the tailbehaviour of the nonparametric estimates, which may be an artifact of the use of kernel estimators. These are known for not being precise in the tails and outside of the support of the data as a combination of their local nature and the sparsity of data there. For the subsample, the data support is within $[0.03, 0.12]$ so one may question the quality of the local estimates of $\mu(x)$ outside of this interval, where it may be prone to a certain degree of erratic behaviour. Chapman and Pearson (2000) report in a simulation study that this is the case when using the nonparametric kernel estimator of Jiang and Knight (1997). Another issue related to the behaviour at the left tail is that the domain in our case is $I = (0, \infty)$. It is by now recognised that kernel densities estimates based on symmetric kernels may perform poorly near the boundary of the support, which may be another source of bias in our case near 0^+. Recall from the simulation study that $\hat{\mu}(x)$ did not perform well near the lower boundary for the CIR-model; one could suspect the same to be the case here. Observe however that in order for the short-term interest rate to remain positive, the drift has to be positive near zero. So at least our estimates have the right sign. Bandi (2002), when applying the nonparametric kernel estimator of Bandi and Phillips (2003), reports similar estimates of the drift function in the right tail of the data support. He however only reports estimates within the data support and is not able to conclude what the drift looks like close to 0^+. The tailbehaviour may also be a result of us imposing the assumption of stationarity on the process as argued by Jones (2003). In conclusion, the estimates for $x \notin [0.03, 0.12]$ should be interpreted with care. We plan to carry out further monte Carlo studies of these issues, and also apply bootstrapping when calculating the confidence bands.



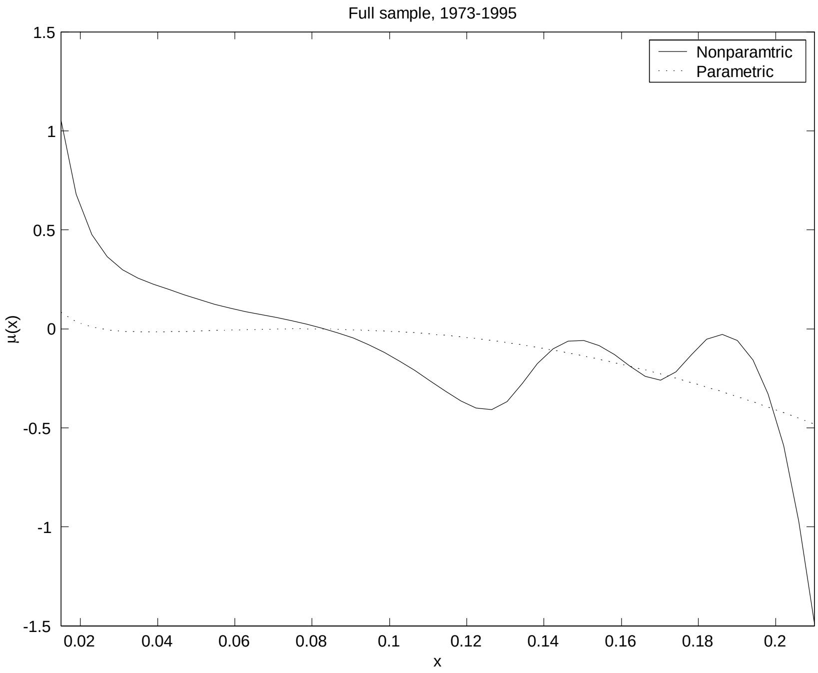


Figure 6: Comparison of nonparametric and parametric estimate of μ for the full sample, 1973-1995.

continuous-time equivalent of a such.

We test the parametric specification against the semiparametric alternative to see whether the latter supplies us with an adequate description of data. We do this using the test statistic in (8). The realised value of the test statistic is $T_n = 280.65$ and 552.96 for the full sample and the subsample respectively, while the critical value at a 1%-level is 63.6907. So we clearly reject the hypothesis that the parameterisation of the drift considered here is appropriate.[3] It could now be of interest to set up a parsimonious parametric model which was able to generate the same behaviour of the drift as our semiparametric estimate does. This is left for future work.

For both samples, the estimates for the parametric model are both qualitatively and quantitatively very different from the ones reported in Aït-Sahalia (1996b). In particular, while the estimated drift in Aït-Sahalia (1996b) predicts two turning points (steady-states), we here only have one. The second steady-state is partially reinforced in the semiparametric model fitted to the full sample. The estimated diffusion term in Aït-Sahalia (1996b) has a smile while ours is monotonously increasing. Hurn and Lindsay (2002) report qualitatively similar estimates to ours with their drift and diffusion term having the same shape as ours, but on a much smaller scale. Our estimates are fairly close to the ones obtained in Durham (2002) though.

[3] The distribution of test statistics for parametric vs. nonparameteric alternatives are known to be poorly approximated by their asymptotic distribution. So it would probably be more appropriate to perform bootstrap here, see e.g. Fan (1994, 1995). This will be done in a future version of the paper.

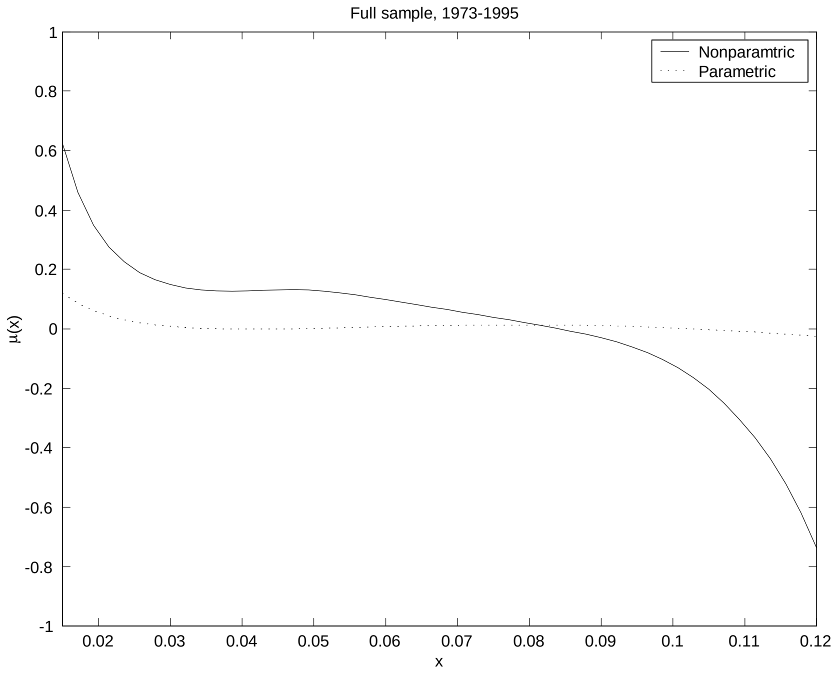


Figure 7: Comparison of nonparametric and parametric estimate of μ for the subsample, 1982-1995.

5.2 Implied Bond and Derivative Prices

Given the calibrated models obtained in the previous section, we now wish to see what implications the competing models will have on bond and interest rate derivative prices. We compute bond prices implied by the competing models to see if they generate significantly different prices. This is only done for the models fitted using data in the period 1982-1995.

In order to do this, we first need to estimate the risk premium as discussed in Section 3. We use the general estimation procedure where the risk premium is chosen to minimise the squared difference between implied and observed bond prices. Here, we follow Aït-Sahalia (1996) and Vasicek (1977) amongst others and assume that the risk premium is constant over time. This facilitates the estimation since we then only have one parameter to optimise with respect to. We estimate this parameter, λ, using the least squares method presented in Section 3: Given the estimates of θ and the drift function and any value of λ, we calculate a set of implied bond prices and compare them to the observed prices. The data set used in the calibration consists of daily observations of 1-, 3- and 6-month Eurodollar bond prices, $(B_{i\Delta}(\tau_j))$, and the daily observations of the short-term interest rate, $r_{i\Delta}$, $i = 1, ..., n$, $j = 1, 2, 3$, where $\tau_1 = 1/12$, $\tau_1 = 1/4$ and $\tau_3 = 1/2$ are the times to maturities. We assume the following model for the observed bond prices,

$$B_{i\Delta}(\tau_j) = \Pi_B(r_i, \tau_j; \lambda) + \varepsilon_{ij}, \quad E^P[\varepsilon_{ij}] = 0, E^P\left[\varepsilon_{ij}^2\right] = \sigma_\varepsilon^2$$

where

$$\Pi_B(r, \tau; \lambda) = E^Q\left[\exp\left[-\int_0^\tau r_s ds\right] | r_0 = r\right]$$

and $\{r_t\}$ has dynamics (10) with $\lambda(r) = \lambda$. The estimate of λ is the obtained by running non-linear least squares on the above regression model. The (approximated) implied bond prices, $\Pi_B(r,\tau)$, are obtained by Monte Carlo simulation of the short term interest rate process under the risk-neutral measure.

Here, we use zero-coupon prices from the period December 1, 1994 to February 25, 1995 in our calibration of λ. This gives us $n = 154$ observations; we have actually a much larger data set of bond prices available which could be used, but the Monte Carlo simulations are fairly time-consuming so we here choose to only use a small proportion of this. In Table 5, the estimated market price of risk premiums are reported when one uses the semiparametric and parametric fit respectively. In Kristensen (2004b), it was demonstrated that under weak regularity conditions,

$$\sqrt{n}(\hat{\lambda} - \lambda_0) \to^d N\left(0, \sigma_\varepsilon^2 H^{-1}(\lambda_0)\right),$$

where

$$H(\lambda) \equiv E_\theta\left[\dot{\Pi}_B(r_{i\Delta};\lambda)^\top \dot{\Pi}_B(r_{i\Delta};\lambda)\right], \quad \Pi_B(r;\lambda) = \{\Pi_B(r,\tau_j)\}_{j=1}^3.$$

The standard errors were calculated using consistent estimators derived in Kristensen (2004b). They were not adjusted for the fact that the drift and diffusion term came from a preliminary estimation step. That is, reported standard errors of $\hat{\lambda}$ are calculated conditional on the estimated drift and diffusion term.

Table 5: Estimate of the market price of risk, λ.

	Semiparametric	Parametric
Estimate of λ	-8.5494 10^{-2}	-1.0596 10^{-1}
Standard Error	(2.0603 10^{-2})	(3.0651 10^{-2})

Notes: Standard errors are reported in parentheses.

We are now able to calculate bond prices under the risk neutral measure. In Table 7 in the appendix, we report the implied bond prices given either the semiparametric or the parametric fit of the short-term interest rate. We report prices for four different times to maturities, 0.5, 1, 5 and 10 years, and 4 different levels of the current short rate. In general as the maturity increases, the prices difference between the two competing models increase. For most maturities, the prices implied by the parametric model fall outside one or more standard deviations of the semiparametric prices. So we get significantly different prices when applying the semiparametric model compared to the parametric one.

6 Conclusion

We have proposed a new semiparametric diffusion model for the short-term interest rate which contains most parametric models found in the term structure literature. We estimated the semiparametric model using a data set of daily observations of the 7-day Eurodollar rate in the period 1973-1995. For comparison, we also fitted the parametric model proposed by Aït-Sahalia (1996b) to the data, this being the most flexible parametric single-factor model found in the literature. Due to so-called Fed-Experiment in 1979-1981, there is a break in the data in this period compared to before and after.

So we estimated the models using both the full sample and the subsample 1982-1995. Estimating the models using either of the two samples lead to markedly different results. The two sets of estimates were however not statistically different from each other. If one wishes to model the full period, one may gain from either including dummies or using a Markov-switching model.

Both the drift and diffusion term in the fitted semiparametric model exhibited a nonlinear behaviour which most of the parametric models in the literature cannot problems generate. The nonparametric estimate of the drift were significantly different from the one of the most flexible parametric model in both the full sample and the subsample. In both samples, this was due to a much stronger degree of mean reversion in the nonparametric estimates. The parametric form was not able to allow for a mean reversion in the tails of this type, and could therefore not mimic the shape of the nonparametric drift properly there. We tested the parametric model against the semiparametric alternative and rejected it in favour of the semiparametric one. The implications on the pricing of bonds were also examined. We found that the implied bond prices predicted by the two models were significantly different from each other. So not only do the models differ on a basic level, but they will also lead to different prices. This is particularly important for market practitioners.

As a next step, it would be interesting to develop further statistical tests to evaluate the performance of the model when using it for pricing derivatives. Such would be a useful tool in model selection and evaluation, helping the practitioner to choose a parsimonious model without misspecifications, and which at the same time performs well in a pricing scenario. Also, we plan to set up a fully parametric model which is consistent with the semiparametric estimates.

How to extend our approach to cover semiparametric multifactor models is not obvious since our estimator is not easily extended to multivariate diffusion models. If one is ready to restrict one's attention to a smaller class of multifactor models however, it should be possible to adapt the approach used here to this more general case, c.f. Chen et al (2000b).

References

Ahn, D.-H., R.F. Dittmar & A.R. Gallant (2002) Quadratic Term Structure Models: Theory and Evidence. *Review of Financial Studies* 15, 243-288.

Ahn, D.-H. & B. Gao (1999) A Parametric Nonlinear Model of Term Structure Dynamics. *Review of Financial Studies* 12, 721-762.

Aït-Sahalia, Y. (1996a) Nonparametric Pricing of Interest Rate Derivative Securities. *Econometrica* 64, 527-560.

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

Aït-Sahalia, Y. (1999) Transition Densities for Interest Rates and Other Nonlinear Diffusions. *Journal of Finance* 54, 499-547.

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

Aït-Sahalia, Y. & J. Duarte (2003) Nonparametric Option Pricing under Shape Restrictions. *Journal of Econometrics* 116, 9-47.

Aït-Sahalia, Y. & A.W. Lo (1998) Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *Journal of Finance* 53, 499-547.

Andersen, T.G. & J. Lund (1997) Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate. *Journal of Econometrics* 77, 343-377.

Ang, A. and G. Bekaert (2002) Short Rate Nonlinearities and Regime Switches. *Journal of Economic Dynamics and Control* 26, 1243-1276.

Backus, D., Foresi, S. & Zin, S. (1998) Arbitrage Opportunities in Arbitrage-Free Models of Bond Pricing. *Journal of Business and Economic Statistics* 16.

Bandi, F.M. (2002) Short-Term Interest Rate Dynamics: A Spatial Approach. *Journal of Financial Economics* 65, 73-110.

Björk, T. (1998) *Arbitrage Theory in Continuous Time*. Oxford: Oxford University Press.

Black, F. & M. Scholes (1973) The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81, 637-654.

Bondarenko, O. (2003) Estimation of Risk-Neutral Densities Using Positiv

Broze, L., O. Scaillet & J.-M. Zakaïan (1998) Quasi-Indirect Inference for Diffusion Processes. *Econometric Theory* 14, 161-186.

Burghardt, G. (2003) *The Eurodollar Futures and Options Handbook*. New York: McGraw-Hill.

Chacko, G. & S. Das (2002) Pricing Interest Rate Derivatives: A General Approach. *Review of Financial Studies* 15,195-242.

Chan, K.C., Karolyi, G.A., Longstaff, F.A. & Sanders, A.B. (1992) An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *Journal of Finance* 47, 1209-1227.

Chapman, D., J.B. Long Jr & N. Pearson (1999) Using Proxies for the Short Rate: When Are Three Months Like an Instant? *Review of Financial Studie*s 12, 763-806.

Chapman, D. & N. Pearson (2000) Is the Short Rate Drift Actually Nonlinear? *Journal of Finance* 55, 355-388.

Chapman, D. & N. Pearson (2001) Recent Advances in Estimating Term-Structure Models. *Financial Analysts Journal* 57, 77-95.

Chen, R.R. & L. Scott (1992) Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross models of the Short-Term Interest Rates. *Review of Financial Studies* 5, 613-636

Chen, R.R. & L. Scott (1993) Maximum Likelihood Estimation for a Multifactor General Equilibrium Model of the Term Structure of Interest Rates *Journal of Fixed Income* 3, 14–31.

Conley, T., L.P. Hansen, E. Luttmer & J. Scheinkman (1997) Short-term Interest Rates as Subordinated Diffusions. *Review of Financial Studies* 10, 525-577.

Constantinides, G.M. (1992) A Theory of the Nominal Term Structure of Interest Rates. *Review of Financial Studies* 5, 531-552.

Courtadon, G. (1982) The Pricing of Options on Default Free Bonds. *Journal of Financial and Quantitative Analysis* 17, 531-552.

Cox, J.C., J.E. Ingersoll & S. Ross (1985) A Theory of the Term Structure of Interest Rates. *Econometrica* 53, 373-384.

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

Duffee, G.R. (2002) Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance* 57, 405-443.

Duffie, D. (1996) *Dynamic Asset Pricing Theory*, 3rd edition. New Jersey: Princeton University Press.

Duffie, D. & R. Kan (1996) A Yield Curve Model of Interest Rates. *Mathematical Finance*, 6, 379-406.

Duffie, D. & K.J. Singleton (1993) Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica* 61, 929-952.

Durham, G.B. (2002) Likelihood-Based Specification Analysis of Continuous-Time Models of the Short-Term Interest Rate. Working paper, University of Iowa. Forthcoming in *Journal of Financial Economics.*

Durham, G.B. & R.A. Gallant (2002) Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes. *The Journal of Business and Economic Statistics* 20, 297-316.

Dybvig, P. (1989) Bond and Bond Option Pricing Based on the Current Term Structure. In *Mathematics of Derivative Securities* (eds. M.A.H. Dempster & S. Pliska). Cambridge University Press.

Elerian, O., S. Chib, S. & N. Shephard (2001) Likelihood Inference for Discretely Observed Non-linear Diffusions. *Econometrica* 69, 959-993.

Fan, Y. (1994) Testing the Goodness-of-Fit of a Parametric Density Function by Kernel Method. *Econometric Theory* 10, 316-356.

Fan, Y. (1995) Bootstrapping a Consistent Nonparametric Goodness-of-Fit Test. *Econometric Reviews* 14, 367-382.

Florens-Zmirou, D. (1989) Approximate Discrete Time Schemes for Statistics of Diffusion Processes. *Statistics* 20, 547-557.

Gallant, A.R. & J.R. Long (1997) Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square. *Biometrika* 84, 125-141.

Garcia, R., E. Ghysels and E. Renault (2004) The Econometrics of Option Pricing. Working paper 2004s-04, CIRANO. Forthcoming in *Handbook of Financial Econometrics.*

Hall, P. (1992) *The Bootstrap and Edgeworth Expansion.* New York: Springer-Verlag.

Heath, D.C., R.A. Jarrow & A. Morton (1992) Bond Pricing and the Term Structure of Interest Rates: A New methodology for Contingent Claims Valuation. *Econometrica* 60, 77-105.

Ho, T.S.Y. & S.-B. Lee (1986) Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance* 41, 1011-1029.

Hong, Y. & Li, H. (2002) Nonparametric Specification Testing for Continuous-Time Models with Application to Spot Interest Rates. Working paper, Cornell University.

Horowitz, J. (2003) Bootstrap Methods for Markov Processes. *Econometrica* 71, 1049-1082.

Hull, J.C. & A.D. White (1993) One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities. *Journal of Financial and Quantitative Analysis* 28, 235-254.

Hurn, A. S. & K. A. Lindsay (2002) On the Specification of the Drift and Diffusion Functions for Continuous-time Models of the Spot Interest Rate. *Oxford Bulletin of Economics and Statistics* 64, 547-564.

Hurn, A.S., K.A. Lindsay & V.L. Martin (2003) On the efficacy of simulated maximum likelihood for estimating the parameters of stochastic differential Equation. *Journal of Time Series Analysis* 24, 45-63.

Jiang, G (1998) Nonparametric Modeling of U.S. Interest Rate Term Structure Dynamics and Implications on the Prices of Derivative Securities. *Journal of Financial and Quantitative Analysis* 33, 465-497.

Jiang, G. & J. Knight (1997) A Nonparametric Approach to the Estimation of Diffusion Processes - with an Application to a Short-term Interest Rate Model. *Econometric Theory* 13, 615-645.

Jones, C.S. (2003) Nonlinear Mean Reversion in the Short-Term Interest Rate. *Review of Financial Studies* 16, 793-843.

Karatzas, I. & S.E. Shreve (1991) *Brownian Motion and Stochastic Calculus,* 2nd Edition. New York: Springer Verlag.

Karlin, S. & H.M. Taylor (1981) *A Second Course in Stochastic Processes*. Academic Press.

Karlsen, H. and D. Tjøstheim (2001) Nonparametric Estimation in Null Recurrent Time Series. *Annals of Statistics* 29, 372-416.

Kloeden, P.E. & E. Platen (1999) *Numerical Solution of Stochastic Differential Equations* (2nd ed.). Heidelberg: Springer-Verlag.

Kristensen, D. (2004a) Estimation in Two Classes of Semiparametric Diffusion Models. Working paper, LSE.

Kristensen, D. (2004b) Estimation of Partial Differential Equations with Applications in Finance. Working paper, LSE.

Litterman, R. & J. Scheinkman (1991) Common Factors Affecting Bond Returns. *Journal of Fixed Income*, 3, 54-61.

Lo, A. (1986) Statistical Tests of Contingent-Claims Asset-Pricing Models. *Journal of Financial Economics* 17, 143-173.

Lo, A. (1988) Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data. *Econometric Theory* 4, 231-247.

Longstaff, F.A. & E.S. Schwartz (1992) Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model. *Journal of Finance* 47, 1259-1282.

Marsh, T.A. & E.R. Rosenfeld (1983) Stochastic Processes for Interest Rates and Equilibrium Bond Prices. *Journal of Finance* 38, 635-646.

Merton, R.C. (1973) Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science* 4, 141-183.

Newey, W.K. & K.D. West (1987) A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55, 703-708.

Pagan, A. and A. Ullah (1999) *Nonparametric Econometrics.* Cambridge: Cambridge University Press.

Parzen, E. (1962) On Estimation of a Probability Density Function and Mode. *Annals of Mathematical Statistics* 33, 1065-1076.

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

Phillips, Peter (1987) Time Series Regression with a Unit Root. *Econometrica* 55, 277-301.

Phillips, P.C.B. & J.Y. Park (1998) Nonstationary Density Estimation and Kernel Autoregression. Cowles Foundation Discussion Paper no. 1181, Yale University.

Phillips, P. & P. Perron (1988) Testing for a Unit Root in Time Series Regression. *Biometrika* 75, 335-346.

Politis, D.N, J.P. Romano and M. Wolf (1999): *Subsampling.* New York: Springer-Verlag.

Pritsker, M.G. (1998) Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models. *Review of Financial Studies* 11, 449–487.

Robinson, P.M. (1988) Root-N-Consistent Semiparametric Regression. *Econometrica* 56, 931-954.

Rogers, L.C.G. (1995) Which Model For Term-Structure of Interest Rates Should One Use?. In *Mathematical Finance* (eds. M.H.A. Davis, D. Duffie, W.H. Fleming and S.E. Shreve). New York: Springer-Verlag.

Said and Dickey (1984) Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order. *Biometrika* 71, 599-607.

Santa-Clara, P. & D. Sornette (2001) The Dynamics of the Forward Interest Rate Curve with Stochastic String Shocks. *Review of Financial Studies* 14, 149-185.

Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis.* Chapman and Hall.

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

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, S.M. (2000) Continuous-Time Methods in Finance: A Review and an Assessment. *Journal of Finance* 55, 1569-1622.

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

A Tables

Table 6: Estimates of θ and β.

	1973-1995		1982-1995	
	Parametric	Semiparametric	Parametric	Semiparametric
β_0	-0.3054	-	-0.3106	-
	(0.1590)	-	(0.1484)	-
β_1	5.5729	-	5.6579	-
	(1.2301)	-	(1.1872)	-
β_2	-30.9985	-	-30.4436	-
	(2.0307)	-	(3.6436)	-
β_3	0.0047	-	0.0053	-
	(0.0011)	-	(0.0019)	-
σ_0	-0.0006	0.0120	-0.0006	0.0110
	(0.0005)	(0.0009)	(0.0004)	(0.0011)
σ_1	0.0210	0.0213	0.0221	0.0123
	(0.0090)	(0.0108)	(0.0122)	(0.0134)
σ_2	30.8413	28.0874	32.5880	23.0874
	(7.8007)	(7.7745)	(8.7490)	(8.6923)
γ	4.2573	4.3019	4.2166	4.0019
	(1.5422)	(1.5361)	(2.7491)	(2.1923)

Notes: Standard errors are reported in parentheses. In the parametric model these were estimated by standard covariance estimation methods. In the semiparametric model the estimator proposed in Theorem **??** was used

Table 7: Implied bond prices of the semiparametric and parametric model.

	Short rate level			
Maturity (years)	0.04	0.06	0.08	0.10
0.5	97.2866	99.1478	96.1416	94.2130
	(0.1054)	(0.0868)	(0.0883)	(0.0697)
	97.1404	98.3368	96.5284	92.7327
1	94.9322	93.0204	91.0659	88.3761
	(0.1147)	(0.0906)	(0.0835)	(0.0657)
	92.2625	92.6393	91.2850	91.0595
5	65.6758	65.2582	62.6428	60.8854
	(0.1131)	(0.1078)	(0.0995)	(0.0656)
	65.3361	63.6363	62.7447	60.7810
10	41.7999	40.0005	39.1834	38.7795
	(0.0916)	(0.0803)	(0.0721)	(0.0681)
	42.9604	40.5908	40.0661	39.0452

Notes: (i) All prices correspond to a face value of the bond equal to $100. Each cell has three elements: The first and third are the implied prices of the semiparametric and parametric model respectively; the second the associated standard error of the semiparametric model.
(ii) The s.e.'s were calculated by using the estimator outlined in Section 5.3.2.