N° 9415

ESTIMATION OF THE TERM STRUCTURE FROM BOND DATA

C. GOURIEROUX*
O. SCAILLET**

May 1994

* CREST and CEPREMAP

** CREST

GOURIEROUX, C., et O., SCAILLET [1]

Estimation d'une structure par terme à partir de prix d'obligations

Résumé

Nous présentons diverses approches statistiques de reconstitution de structure par terme des taux d'intérêt à partir de prix d'obligations. Ces approches permettent toutes de corriger des effets de coupons et de maturité des obligations. Elles peuvent être paramétriques : modèles de régression, ou non paramétriques : splines et régressions locales. Une application aux données françaises est ensuite effectuée.

Estimation of the term structure from bond data

Abstract

We present and compare different statistical approaches to determine the term structure of interest rates from bond data. All these approaches allow for correction of cash flows and maturity effects. They may be parametric : regression models, or non parametric : splines and local regressions. An application to French data is described.

mots clés : structure par terme, obligation, correction de flux, spline, régression locale.

key words : term structure, bond, cash flow correction, spline, local regression.

JEL : G 12

[1]The second author gratefully acknowledges financial support from ICM, Belgium.

1 Introduction

At each date t, a coupon bond (j,t) is characterized by the deterministic cash flows $F_{j,t,t+h}, h=1,\dots$ which are delivered at the various dates $t+h, h=1,\dots$ Such a bond appears as a portfolio made of elementary bonds : the discount bonds. At date t, the discount bond of term h is the asset (eventually fictitious in the sense that it is not necessarily traded on the market), which delivers one money unit at date $t+h$, zero at the other dates.
Classical financial literature, which either relies on arbitrage arguments or on equilibrium models, assumes usually that the discount bond markets are all open. Let us denote $C(t,t+h)$ the price at date t of the discount bond of maturity h, $P_{j,t}$ the coupon bond price (j,t) assumed to be tradable on the market. It then results from arbitrage free conditions that :

$$P_{j,t} = \sum_{h=1}^{H} F_{j,t,t+h} C(t,t+h), \tag{1}$$

where H is the maximal maturity for all considered bonds. This arbitrage free condition in a frictionless market is introduced by Ross (1976) from the dual theory of linear programming. When the true bond maturity is strictly less than H, we can still use equation (1) thanks to the introduction of zero cash flows for the relatively large h indexes. So, if discount bonds were traded and by neglecting the differences which could possibly appear from transaction costs or tax differentials, the price of a portfolio of discount bonds should be the same as the price of the corresponding bond, and the pricing of bonds is equivalent to the pricing of discount bonds. We can then introduce models explaining how the discount bond prices $C(t,t+h)$ (or equally the discount rates $r(t,t+h) = -\frac{1}{h}\log C(t,t+h)$) are modified by the maturity h and/or date t, and directly estimate these models from observed discount bond prices.
Such an approach is however very delicate to use in practice for two very different reasons.
(i) Very often, term structure models depend on a number of independent factors much inferior than the maximum H of maturities. Generally, one specifies one factor or two factors models (see appendix 1). This induces strong restrictions between discount bond prices, deterministic relations for instance (a perfect correlation between all discount bond prices in one factor models), which are clearly not satisfied by the data. It appears from this feature that it is necessary to introduce additional noises, i.e. to enlarge a little the class of usual models in order to solve this difficulty.
(ii) On another hand, even if some coupon bonds are stripped (for instance at the moment the French OAT ("obligations assimilables du Trésor") of maturities 2003, 2008, 2019, 2023), it cannot be reasonably considered that the liquidity is sufficient for all various term discount bonds, which means that the quoted prices cannot be viewed as prices in an economic sense. In practice, markets where the liquidity is sufficient, where the number of participants is large enough, where they have suf-

ficiently weak correlated behaviors, exist only for a very thin number of bonds (in France between 10 and 15 approximatively at each date), bonds which moreover change with the chosen date, partly because of new bond issues. It is still possible to write arbitrage free conditions for these bonds (see DUFFIE (1992)). These conditions imply equation (1) which means that :
it exist real numbers $C(t, t+h) \geq 0$ such that

$$P_{j,t} = \sum_{h=1}^{H} F_{j,t,t+h} C(t, t+h).$$

Obviously, these real numbers are not uniquely defined, since the market completeness hypothesis is not fulfilled. We call such a set $(C(t, t+h),\ t,\ h$ varying) the discount bond price set, in order to make the presentation easier. It should be stressed that such an interpretation is usually false. Indeed, if all discount bond markets were suddenly open and with enough liquidity, a change of the price of bond (j, t) would probably take place due to this market change. The set of all $(C(t, t+h),\ t,\ h$ varying) satisfying (1) for tradable bonds does not include without doubt the set of these discount bonds with a larger set of open markets. In fact, these underlying variables $C(t, t+h)$ should be viewed as intermediate variables, which allow to correct observed bond price comparisons from cash flow effects, i.e. from the structure of the $(F_{j,t+h})$ and the bond maturity. Such a bond price analysis method should indeed be the least contingent to the cash flow structure of the bonds on which the estimation is based. Such latent variable approach is analogous to methods used to construct hedonic price indexes in order to correct the quality type effect or the effects created by the introduction of new products (GRILICHES (1961), GRILICHES and ADELMAN (1961), ROSEN (1974)). Here, the cash flows play the role of the "characteristics" (see BECKER (1965), LANCASTER (1971), ANDERSON, DE PALMA and THYSSE (1989)) for the use and definition of these notions.
In order to solve the two previous difficulties : lack of degree of freedom in these models, non observability of discount bond prices, the practice is usually the following one. One introduces a term structure model, a deterministic one or allowing some randomness (see appendix 1 for the description of usual models). Let us for instance consider the model :

$$C(t, t+h) = g(\theta, t, t+h), \tag{2}$$

where g is a given function and θ a vector of parameters to be determined. The parameter is afterwards estimated by a least squares calibration technique, based on the observed bond prices. This technique is often applied date by date. In this case, a solution $\widehat{\theta}_t$ is obtained from the minimization problem :

$$\min_{\theta} \sum_{j=1}^{J_t} \left[P_{j,t} - \sum_{h=1}^{H} F_{j,t,t+h}\, g(\theta, t, t+h) \right]^2, \tag{3}$$

where J_t is the number of liquid bonds at date t. Every estimation technique is linked with an implicit underlying model. The model associated with the calibration technique (3) is :

$$\left\{ \begin{array}{l} P_{j,t} = \sum\limits_{h+1}^{H} F_{j,t,t+h} C(t,t+h) + \epsilon_{j,t} \ , j = 1, \ldots, J_t, \\ C(t,t+h) = g(\theta_t, t, t+h), \\ E\epsilon_{j,t} = 0, V\epsilon_{j,t} = \sigma_t^2, Cov(\epsilon_{j,t}, \epsilon_{k,t}) = 0 \ , j \neq k. \end{array} \right. \tag{4}$$

This approach implicitly based on model (4) leads generally to a too strong instability of the interest rate forecasts (BROWN and DYBVIG (1986), DEMUNNIK and SCHOTMAN (1993)). How could it be explained? Three reasons may be called upon.
(i) The fact that the parameter θ, often independent of time in the underlying theoretical term structure models, is here taken as evolutionary. In fact, the possibility of a variable parameter is introduced to allow better adjustments. It essentially translates the fact that the constant parameter model (2) is misspecified, and that the calibrated values $\widehat{\theta}_t$ will then take all the effects due to the misspecification of the initial model.
(ii) The fact that the error magnitude is taken independent of the bond (σ_t independent of j), and that non correlation is assumed. But, clearly, these effects should depend on the cash flow structure. For instance if bonds $(1,t)$ et $(2,t)$ were mixed to get a portfolio (along proportions α_1 and α_2), we should naturally have :

$$\begin{aligned} P_{(1,2),t} &= \alpha_1 P_{1,t} + \alpha_2 P_{2,t} \\ &= \sum_{h=1}^{H} (\alpha_1 F_{1,t,t+h} + \alpha_2 F_{2,t,t+h}) g(\theta_t, t, t+h) + \epsilon_{(1,2),t}, \end{aligned} \tag{5}$$

with $\epsilon_{(1,2),t} = \alpha_1 \epsilon_{1,t} + \alpha_2 \epsilon_{2,t}$ and clearly $V\epsilon_{(1,2),t} \neq V\epsilon_{1,t}$. The model is not coherent with the constitution of bond portfolios.
(iii) Finally, we have to note that the selected implicit model is not compatible with the constrained (1) coming from the arbitrage free conditions.

In this paper, we present simple ways of interest rate determination based on latent variable models (CLEMENT, GOURIEROUX and MONFORT (1993)a, (1993)b). The first studied approach relies on a modified version of model (4) :

$$\left\{ \begin{array}{l} P_{j,t} = \sum\limits_{h=1}^{H} F_{j,t,t+h} C(t,t+h), \\ C(t,t+h) = g(\theta_t, t, t+h) + \eta_{h,t}, \\ E\eta_{h,t} = 0, \ V\eta_{h,t} = \sigma_h^2, \\ Cov(\eta_{h,t}, \eta_{k,t}) = 0, \ h \neq k, \ Cov(\eta_{h,t}, \eta_{k,\tau}) = 0, \ t \neq \tau. \end{array} \right. \tag{6}$$

The introduction of a noise at the level of the discount bonds and not at the level of the coupon bonds allows at the same time to have a model compatible with the arbitrage free conditions, to correct the major part of the cash flow effects and to have a more usual interpretation of the noise as a modelization error. It allows among others to avoid differences between the observed bond prices and the theoretical prices rebuild from the discount bonds. Moreover, the noise introduction at the discount bond level is at the basis of the development of several stochastic duration concepts where random shocks on the term structure are introduced either in an additive or multiplicative way or by combination of both (see BIERWAG (1987)). In section 2, we study such parametric models, we examine in this setting the properties of the usual approaches in order to compare them with the proposed approach and we discuss several extensions.

In section 3, we are interested in discount bond models which have a smoother structure. The idea is to look for discount bond prices which have a smooth enough variation both as function of h and t and are compatible with observed coupon bond prices. We discuss especially methods extending spline approaches and kernel estimators in this partial observability framework.

Section 4 is devoted to discount bond models and insists more on the parametric links between price evolutions than on the real price dependency on t and h. We examine the links between models based on estimating relations and factor models. The last section is devoted to empirical applications and to the practical comparison of the various approaches.

2 Regression model estimation

2.1 The model and the application of least squares

As previously considered, we deal with a model, where the additional noises are directly introduced at the level of discount bond prices :

$$C(t,t+h) = g(\theta,t,t+h) + \eta_{h,t}, \tag{7}$$

where the errors have zero expectation, are not correlated and have variance $V\eta_{h,t} = \omega^2\lambda_h$, where λ_h is given and where ω^2 is an unknown scaling factor. We suppose that we have at our disposal bond prices $P_{j,t}$, $t = 1,\dots,T$, $j = 1,\dots,J_t$, which satisfy equation (1) :

$$P_{j,t} = \sum_{h=1}^{H} F_{j,t,t+h}C(t,t+h). \tag{8}$$

It is useful for our purpose to introduce matrix notations. We denote by F_t the $J_t \times H$ cash flow matrix of the J_t bonds considered at date t, P_t the vector of size J_t of the bond prices, C_t the vector of size H of the discount bond prices at that date, $g(\theta,t)$ the vector of same size of means of C_t, η_t the corresponding error vector and Λ the diagonal matrix of size H with diagonal elements λ_h. The model is then written as :

$$\begin{cases} P_t = F_tC_t, \\ C_t = g(\theta,t) + \eta_t, \\ E\eta_t = 0,\ V\eta_t = \omega^2\Lambda,\ Cov(\eta_t,\eta_{t-k}) = 0\ ,k \neq 0. \end{cases} \tag{9}$$

The prices P_t only being observable, it is interesting to eliminate the intermediate variables C_t. We obtain :

$$P_t = F_t g(\theta, t) + F_t \eta_t, \tag{10}$$

so that :

$$\begin{cases} EP_t = F_t g(\theta, t), \\ VP_t = \omega^2 F_t \Lambda F_t', \\ Cov(P_t, P_{t-k}) = 0 \ , k \neq 0. \end{cases} \tag{11}$$

We can immediately remark the double effect of the bond cash flow structures : to create heteroscedasticity, because $VP_{j,t} = \omega^2 \sum_{h=1}^{H} (F_{j,t,t+h})^2 \lambda_h$ depends on the considered bond; to create correlations between bond prices of the same date, because $Cov(P_{j,t}, P_{k,t}) = \omega^2 \sum_{h=1}^{H} F_{j,t,t+h} F_{k,t,t+h} \lambda_h$ is generally not equal to zero.

The parameters θ, ω^2 can then be estimated date by date or globally. They will be obtained by a generalized least squares method in order to take into account the cash flow effects.

2.1.1 Date by date estimation

This estimation can be performed for date t as soon as the number of bonds J_t exceeds $dim\, \theta + 1$. It is important to note that this identifiability condition may be satisfied in a framework of incomplete markets. This situation arises when $dim\, \theta + 1 \leq J_t < H$. Such estimation is also compatible with a model of type (9), where the parameters θ and ω^2 would be taken independent of time. The application of least squares requires that F_t' is of full column rank, i.e. that the introduced bond set does not include redondant bonds in the sense that it might be interpreted as a portfolio built from other bonds. At date t, the G.L.S. estimator of θ is the solution of the minimization problem :

$$\widehat{\theta}_t = \underset{\theta}{\text{argmin}} (P_t - F_t g(\theta, t))' (F_t \Lambda F_t')^{-1} (P_t - F_t g(\theta, t)), \tag{12}$$

and the variance estimator is obtained from :

$$\widehat{\omega}_t^2 = \frac{1}{J_t} (P_t - F_t g(\widehat{\theta}_t, t))' (F_t \Lambda F_t')^{-1} (P_t - F_t g(\widehat{\theta}_t, t)). \tag{13}$$

The estimator $\widehat{\theta}_t$ has usually to be numerically determined thanks to Gauss-Newton optimization algorithm.

2.1.2 Global estimation

The estimator is determined by simultaneously taking into account all dates. It is defined by :

$$\widehat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{t=1}^{T} (P_t - F_t g(\theta,t))'(F_t \Lambda F_t')^{-1}(P_t - F_t g(\theta,t)), \tag{14}$$

and the corresponding variance factor is estimated by :

$$\widehat{\omega}^2 = \frac{\sum_{t=1}^{T} (P_t - F_t g(\widehat{\theta},t))'(F_t \Lambda F_t')^{-1}(P_t - F_t g(\widehat{\theta},t))}{\sum_{t=1}^{T} J_t}. \tag{15}$$

2.2 Particular case of a linear formulation of discount bond prices

The preceding results take a simplified form, when the function $g(\theta,t)$ is linear in the parameters, i.e. can be written as :

$$g(\theta,t) = X_t \theta, \tag{16}$$

where X_t is a known $H \times p$ matrix with $p = dim\,\theta$. The estimators of θ, date by date or globally, admit an explicit form :

$$\widehat{\theta}_t = [X_t' F_t' (F_t \Lambda F_t')^{-1} F_t X_t]^{-1} X_t' F_t' (F_t \Lambda F_t')^{-1} P_t, \tag{17}$$

$$\widehat{\theta} = \left[\sum_{t=1}^{T} X_t' F_t' (F_t \Lambda F_t')^{-1} F_t X_t\right]^{-1} \sum_{t=1}^{T} X_t' F_t' (F_t \Lambda F_t')^{-1} P_t. \tag{18}$$

$\widehat{\theta}$ appears as a weighted average of estimators computed date by date, an average with matrix weights :

$$\widehat{\theta} = \left[\sum_{t=1}^{T} X_t' F_t' (F_t \Lambda F_t')^{-1} F_t X_t\right]^{-1} \sum_{t=1}^{T} X_t' F_t' (F_t \Lambda F_t')^{-1} F_t X_t \widehat{\theta}_t. \tag{19}$$

These estimators of θ are unbiased and admit as covariance matrices :

$$V\widehat{\theta}_t = \omega^2 [X_t' F_t' (F_t \Lambda F_t')^{-1} F_t X_t]^{-1}, \tag{20}$$

$$V\widehat{\theta} = \omega^2 \left[\sum_{t=1}^{T} X_t' F_t' (F_t \Lambda F_t')^{-1} F_t X_t\right]^{-1}. \tag{21}$$

2.3 Reconstitution of discount bond prices

In the preceding approach, discount bond prices usually appear as non observable random variables. The reconstitution of their corresponding values is made in two steps. We begin by determining the best (linear) expectation of the discount bond prices knowing the observed coupon bond prices. These expectations depend on the parameters θ and ω^2. In a second step, we replace the parameters by their estimations.

2.3.1 Optimal linear expectation of discount bond prices

It is given by :

$$\begin{aligned}
E(C_t \mid P_\tau, \tau = 1, \ldots, T) \\
&= E(C_t \mid P_t) \\
&\quad \text{(from the temporal non correlation hypothesis)} \\
&= E(C_t \mid F_t C_t) \\
&= EC_t + Cov(C_t, F_t C_t)[V(F_t C_t)]^{-1}[F_t C_t - E(F_t C_t)] \\
&= EC_t + V(C_t)F_t'[F_t V(C_t) F_t']^{-1} F_t (C_t - EC_t).
\end{aligned}$$

Replacing expectation and variance by their expressions, we find :

$$\widehat{C}_t = E(C_t \mid P_t) = g(\theta, t) + \Lambda F_t' (F_t \Lambda F_t')^{-1} [P_t - F_t g(\theta, t)]. \tag{22}$$

We can see that the expectations are such that :

$$P_t = F_t \widehat{C}_t, \tag{23}$$

i.e. are compatible with values satisfying arbitrage free conditions.
In a second step, we replace the parameter θ by its estimator. For example, reasoning date by date, we have :

$$\widehat{\widehat{C}}_t = g(\widehat{\theta}_t, t) + \Lambda F_t' (F_t \Lambda F_t')^{-1} [P_t - F_t g(\widehat{\theta}_t, t)]. \tag{24}$$

What can be said about the statistical properties of these expectations? For sake of simplicity, we consider a linear formulation. We have then :

$$\begin{aligned}
\widehat{\widehat{C}}_t &= X_t \widehat{\theta}_t + \Lambda F_t' (F_t \Lambda F_t')^{-1} [P_t - F_t X_t \widehat{\theta}_t] \qquad (25) \\
&= M_t P_t,
\end{aligned}$$

where :

$$\begin{aligned}
M_t = & \left[Id - \Lambda F_t' (F_t \Lambda F_t')^{-1} F_t \right] X_t \left[X_t' F_t' (F_t \Lambda F_t')^{-1} F_t X_t \right]^{-1} X_t' F_t' (F_t \Lambda F_t')^{-1} \\
& + \Lambda F_t' (F_t \Lambda F_t')^{-1}.
\end{aligned}$$

We have :

$$\begin{aligned}
E\widehat{\widehat{C}}_t &= M_t E P_t \\
&= M_t F_t X_t \theta \\
&= X_t \theta \\
&= EC_t.
\end{aligned}$$

Even after replacement of the parameter θ by its estimator, $\widehat{\widehat{C_t}}$ remains an unbiased predictor of C_t. Moreover we have :

$$\begin{aligned}\widehat{\widehat{C_t}} - C_t &= M_t P_t - X_t\theta - \eta_t \\ &= M_t F_t X_t \theta - X_t\theta + M_t\eta_t - \eta_t \\ &= (M_t - Id)\eta_t.\end{aligned}$$

We deduce the prediction error :

$$V(\widehat{\widehat{C_t}} - C_t) = \omega^2 (M_t - Id)\Lambda(M_t - Id)'.$$

2.4 Comparaison of the cash flow correction approach and the usual approach

Let us consider the case of a linear formulation in order to be able to compare analytically and not only numerically the two approaches. We reason globally and take $\Lambda = Id$. The cash flow corrected estimator (cf. equation (18)) is :

$$\widehat{\theta} = \left[\sum_{t=1}^{T} X_t' F_t' (F_t F_t')^{-1} F_t X_t\right]^{-1} \sum_{t=1}^{T} X_t' F_t' (F_t F_t')^{-1} P_t.$$

The non corrected estimator is obtained by minimizing $\sum_{t=1}^{T} (P_t - F_t X_t \theta)^2$ and is given by :

$$\tilde{\theta} = \left[\sum_{t=1}^{T} X_t' F_t' F_t X_t\right]^{-1} \sum_{t=1}^{T} X_t' F_t' P_t. \quad (26)$$

These two estimators are unbiased and their covariance matrices are respectively :

$$V\widehat{\theta} = \omega^2 \left[\sum_{t=1}^{T} X_t' F_t' (F_t F_t')^{-1} F_t X_t\right]^{-1}, \quad (27)$$

$$V\tilde{\theta} = \omega^2 \left[\sum_{t=1}^{T} X_t' F_t' F_t X_t\right]^{-1} \sum_{t=1}^{T} X_t' F_t' F_t F_t' F_t X_t \left[\sum_{t=1}^{T} X_t' F_t' F_t X_t\right]^{-1}. \quad (28)$$

In order to illustrate the differences between the two estimation techniques, let us consider the case where the observed bond prices correspond to a same bond, with a coupon C and a face value 1. The corresponding cash flows are :

$$F_t = (C, C, \ldots, C, 1, 0, \ldots, 0),$$

where 1 is placed at index $H + 1 - t$. Let us suppose moreover $g(\theta, t, t+h) = a\rho^h$, $h \geq 1$, where ρ is given and a is to be determined. It is a model where a breakpoint may appear in the pattern of the discount bond prices as soon as a is not equal to 1. In this case, $X_t = (\rho, \ldots, \rho^H)'$ is independent of t. We have :

$$\begin{aligned}
F_t X_t &= C(\rho + \ldots + \rho^{H-t}) + \rho^{H+1-t} \\
&= C\rho \frac{1-\rho^{H-t}}{1-\rho} + \rho^{H+1-t}, \\
F_t F_t' &= 1 + C^2(H-t).
\end{aligned}$$

We can compute the asymptotic relative efficiency of the two estimators, i.e. the ratio $\alpha = \frac{V\hat{\theta}}{V\tilde{\theta}}$, independent of ω^2 and always less than one. It is given by :

$$\begin{aligned}
\alpha(C,\rho,H,T) &= \frac{\left[\sum_{t=1}^{T} X_t' F_t' F_t X_t\right]^2}{\left[\sum_{t=1}^{T} X_t' F_t' F_t F_t' F_t X_t\right]\left[\sum_{t=1}^{T} X_t' F_t' (F_t F_t')^{-1} F_t X_t\right]} \\
&= \frac{\left[\sum_{t=1}^{T}\left(C\rho\frac{1-\rho^{H-t}}{1-\rho} + \rho^{H+1-t}\right)^2\right]^2}{\left[\sum_{t=1}^{T}\left(C\rho\frac{1-\rho^{H-t}}{1-\rho} + \rho^{H+1-t}\right)^2\left(1 + C^2(H-t)\right)\right]} \\
&\quad \times \frac{1}{\left[\sum_{t=1}^{T}\frac{\left(C\rho\frac{1-\rho^{H-t}}{1-\rho} + \rho^{H+1-t}\right)^2}{1+C^2(H-t)}\right]}.
\end{aligned}$$

We give in Figures 1 and 2 for $H = 10$, $T = 10$ and $H = 10$, $T = 5$, the values of this ratio for various C and ρ. We can see that the precision loss due to the omission of cash flow effects essentially depends on the coupon and can, even in this simple case, be important.

3 Non parametric methods of term structure reconstitution

The regression approach exposed in the previous section lies greatly on the retained specification $g(\theta, t)$ and can from this fact be sensitive to this choice even chosen with care (see section 4). It is so useful to have non parametric type methods lying on the arbitrage free conditions (1), and providing smooth enough evolutions for the discount bonds together with evolutions close to given parametric ones. In order to take simultaneously into account these three aspects we may use spline approaches or local regressions. We successively detail these two approaches.

3.1 Spline reconstitution

We first present this method reasoning date by date; only the argument h varies.

3.1.1 Discount bond observation

If some discount bonds corresponding to terms $h_j, j = 1, \dots, J_t$ were observable, we could directly apply a spline approach to rebuild the curve $h \longrightarrow C(t, t+h)$. The classical idea is the following one. Let us denote by $g(t,h)$ the approximation looked for $C(t, t+h)$. We would then impose to this approximation :
(i) to be compatible with the observations,
(ii) to be as smooth as possible.
By analogy with the continuous case, the smoothing degree may be measured by (see GOOD and GASKINS (1971), EUBANK (1988)) :

$$\int_0^H \left[\frac{\partial^2 g}{\partial h^2}(t,h)\right]^2 dh. \tag{29}$$

The proximity to available data may be measured by :

$$\sum_{j=1}^{J_t}[C(t,t+h_j) - g(t,h_j)]^2. \tag{30}$$

Then we can look for a function g solution of :

$$\min_g \sum_{j=1}^{J_t}[C(t,t+h_j) - g(t,h_j)]^2 + \lambda_1 \int_0^H \left[\frac{\partial^2 g}{\partial h^2}(t,h)\right]^2 dh, \tag{31}$$

where $\lambda_1 > 0$ is a smoothing constant.
In this case, we know (SCHOENBERG (1964)) that problem (31) admits a unique solution $\widehat{g}(t,h)$ characterized by the conditions :
(i) $\widehat{g}(t,h)$ is a third degree polynomial on the interval $[h_j, h_{j+1}), j = 1, \dots, J_t - 1$.
(ii) $\widehat{g}(t,h)$ and its derivatives of order one and two are continuous at points $h_j, j = 1, \dots, J_t$.
(iii) $\widehat{g}(t,h)$ is a first degree polynomial on each side of the interval $[h_1, h_{J_t}]$.
This last condition does evidently not take into account the constraint $g(t,0) = 1$ specific to our problem (see appendix 2).
The problem (31) can be approximated by another problem where the integral is replaced by a finite sum and the second derivative of g by the second order increment. The problem becomes :

$$\begin{aligned} \min_g \mathcal{L}_1 \quad = \quad & \sum_{j=1}^{J_t}[C(t,t+h_j) - g(t,h_j)]^2 \\ & + \mu_1 \sum_{h=1}^{H-1}[g(t,h+1) - 2g(t,h) + g(t,h-1)]^2, \end{aligned} \tag{32}$$

with $g(t,0) = 1$.
It is a quadratic form in parameters $g(t,h)$, $h = 1, \dots, H$, for which the solution is easily obtained.

3.1.2 Coupon bond observation

Having at our disposal only coupon bond prices, we can follow an approach of the same kind by modifying the first part of the criterion. As in (32) where we implicitly suppose equivalent approximation degrees for all terms, i.e. λ_h independent of h in section 2 model, it is natural to make cash flow structure corrections corresponding to this homoscedasticity hypothesis. If we first consider the discrete case, the objective function becomes :

$$
\begin{aligned}
\mathcal{L}_2 &= (P_t - F_t g_t)'(F_t F_t')^{-1}(P_t - F_t g_t) \\
&\quad + \mu_1 \begin{pmatrix} 1 \\ g_t \end{pmatrix}' A \begin{pmatrix} 1 \\ g_t \end{pmatrix}, \qquad (33)\\
\text{where } g_t &= (g(t,1), \dots, g(t,H))', \\
\text{and } A &= \begin{pmatrix} 0 & 0 & 0 & & & \\ 1 & -2 & 1 & & 0 & \\ & & & \ddots & & \\ & 0 & & 1 & -2 & 1 \\ & & & 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
$$

With obvious notations, the last term may be decomposed in :

$$
\begin{pmatrix} 1 \\ g_t \end{pmatrix}' A \begin{pmatrix} 1 \\ g_t \end{pmatrix} = c + 2b'g_t + g_t' B g_t.
$$

The solution of problem (33) satisfies the first order condition :

$$
F_t'(F_t F_t')^{-1}(P_t - F_t g_t) + \mu_1 b + \mu_1 B g_t = 0. \qquad (34)
$$

It is given by :

$$
\widehat{g}_t = \left(F_t'(F_t F_t')^{-1} F_t - \mu_1 B\right)^{-1} \left[F_t'(F_t F_t')^{-1} P_t + \mu_1 b\right]. \qquad (35)
$$

This estimator depends on the magnitude of the retained smoothing parameter μ_1. When μ_1 increases, we take less and less into account the observed prices and the reconstitution is smoother and smoother. In the limit case where $\mu_1 = 0$, the matrix $F_t'(F_t F_t')^{-1} F_t - \mu_1 B = F_t'(F_t F_t')^{-1} F_t$ is singular; it is a consequence of the incompleteness of the market.
How to choose μ_1 in practice? Let us first remark that components $\widehat{g}(t,h)$ are not necessarily positive for all values μ_1. It may seem natural to keep only a value for which $\widehat{g}(t,h) \geq 0$, $\forall h$, which will require a minimal smoothing.

A second idea consists in examining properties of this estimator assuming that the model :

$$\begin{cases} P_t = F_t C_t, \\ C_t = g_t + \eta_t, \\ E\eta_t = 0, \ V\eta_t = \omega^2 Id, \end{cases}$$

is satisfied. We have :

$$\begin{aligned} E\widehat{g}_t &= \left(F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right)^{-1}\left[F_t'(F_tF_t')^{-1}F_t g_t + \mu_1 b\right], \\ E\widehat{g}_t - g_t &= \mu_1\left(F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right)^{-1}\left[Bg_t + b\right]. \end{aligned}$$

It is a biased estimator, the bias depending among others on the smoothing parameter μ_1 and on the curve $Bg_t + b$ of the true underlying structure.
The variance of this estimator is :

$$\begin{aligned} V\widehat{g}_t &= \omega^2\left(F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right)^{-1} F_t'(F_tF_t')^{-1}F_t \\ &\quad \times\left(F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right)^{-1}, \end{aligned}$$

and the quadratic risk is deduced by :

$$\begin{aligned} R(\widehat{g}_t) &= \omega^2\left(F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right)^{-1} F_t'(F_tF_t')^{-1}F_t\left(F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right)^{-1} \\ &\quad + \mu_1^2\left(F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right)^{-1}(Bg_t + b) \\ &\quad \times(Bg_t + b)'\left(F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right)^{-1}. \end{aligned}$$

Then we could choose μ_1 in order to minimize this risk. However this approach presents two difficulties.
i) The risk being under a matrix form, there exists no value μ_1 which minimizes it simultaneously in all directions. We can then either minimize a global risk measure as $trace(R(\widehat{g}_t))$, or be interested in the discount bond price corresponding to a specific term and choose μ_1 in relation with this term.
ii) This difficulty solved, it stays that the solution would depend on unknown parameters g_t, ω^2 and should then be approached by replacing g_t, ω^2 with first step estimators (with a non optimized choice of μ_1).
Let us finally note that the spline estimator of g_t can be compared with the parametric one introduced in the previous section. Let us assume $g_t = X_t\theta$. The parametric estimator of g_t, let us say $\widehat{\widehat{g}}_t$, is :

$$\widehat{\widehat{g}}_t = X_t\widehat{\theta}_t = X_t\left(X_t'F_t'(F_tF_t')^{-1}F_tX_t\right)^{-1}X_t'F_t'(F_tF_t')^{-1}P_t.$$

Its quadratic risk is equal to its variance and is given by :

$$R(\widehat{\widehat{g}}_t) = V\widehat{\widehat{g}}_t = \omega^2 X_t\left(X_t'F_t'(F_tF_t')^{-1}F_tX_t\right)^{-1}X_t'.$$

We can easily verify that this estimator, the best linear unbiased estimator (BLUE) from the Gauss-Markov theorem, is not systematically to be prefered to the spline estimator $\widehat{g}_t$ which is biased, but may have a smaller variance.

3.1.3 Introduction of a benchmark model

Finally, it is possible to modify the previous approach to be close to a benchmark model $g(\theta, t)$. The idea is analogous to the one in Poo (1992). It consists in looking for the solution $\widehat{g}_t, \widehat{\theta}$ of problem :

$$\begin{aligned} \min_{g_t,\theta} \mathcal{L}_3 &= (P_t - F_t g_t)'(F_t F_t')^{-1}(P_t - F_t g_t) \\ &+\mu_1 \begin{pmatrix} 1 \\ g_t \end{pmatrix}' A \begin{pmatrix} 1 \\ g_t \end{pmatrix} \\ &+\mu_2 \left(g_t - g(\theta,t)\right)' \left(g_t - g(\theta,t)\right). \end{aligned} \tag{36}$$

The solution depends on two auxiliary parameters : μ_1 the smoothing parameter, μ_2 the parameter measuring the a priori belief in the parametric model.
If $\mu_2 = 0$, the solution is the previously derived solution with splines.
If $\mu_1 = 0$ and μ_2 goes to infinity, we find the regression approximation of section 2.
The solution can explicitly be given in the linear case $g(\theta, t) = X_t\theta$. Indeed, we have :

$$\begin{aligned} \mathcal{L}_3 &= (P_t - F_t g_t)'(F_t F_t')^{-1}(P_t - F_t g_t) \\ &+\mu_1 \begin{pmatrix} 1 \\ g_t \end{pmatrix}' A \begin{pmatrix} 1 \\ g_t \end{pmatrix} \\ &+\mu_2 \left(g_t - X_t\theta\right)' \left(g_t - X_t\theta\right). \end{aligned} \tag{37}$$

It is a quadratic expression in $\begin{pmatrix} g_t \\ \theta \end{pmatrix}$. A first optimization with respect to θ leads to an OLS approximation :

$$\theta_t^* = (X_t'X_t)^{-1}X_t'g_t.$$

Replacing in the criterion function θ by its expression, we obtain an objective function concentrated in θ and which depends only on g_t :

$$\begin{aligned} \mathcal{L}_3^c &= (P_t - F_t g_t)'(F_t F_t')^{-1}(P_t - F_t g_t) \\ &+\mu_1 \begin{pmatrix} 1 \\ g_t \end{pmatrix}' A \begin{pmatrix} 1 \\ g_t \end{pmatrix} \\ &+\mu_2 g_t'(Id - \Pi_t) g_t, \end{aligned}$$

with $\Pi_t = X_t(X_t'X_t)^{-1}X_t'$ the orthogonal projector on the space spanned by the columns of X_t. The optimization solution of $\mathcal{L}_3^c$ is :

$$g_t^* = \left[F_t'(F_tF_t')^{-1}F_t - \mu_1 B + \mu_2(Id - \Pi_t)\right]^{-1} \left[F_t'(F_tF_t')^{-1}P_t + \mu_1 b\right]. \tag{38}$$

It is an estimator directly linked to the spline estimator which does not take into account the model, and which is deduced from it by a shrinkage operator :

$$g_t^* = N_t \hat{g}_t, \tag{39}$$

with :

$$\begin{aligned} N_t &= \left[F_t'(F_tF_t')^{-1}F_t - \mu_1 B + \mu_2(Id - \Pi_t)\right]^{-1} \left[F_t'(F_tF_t')^{-1}F_t - \mu_1 B\right] \\ &= Id - \mu_2 \left[F_t'(F_tF_t')^{-1}F_t - \mu_1 B + \mu_2(Id - \Pi_t)\right]^{-1} (Id - \Pi_t). \end{aligned}$$

The model influence comes from the shrinkage coefficient through matrix Π_t. Let us insist on the fact that the terminology "shrinkage", even though classical, is here misleading in the sense that matrix $F_t'(F_tF_t')^{-1}F_t - \mu_1 B + \mu_2(Id - \Pi_t)$ has generally simultaneously positive and negative eigenvalues. We do not give detailed expressions of the first two moments of g_t^*, which are easily computed from the moments of $\hat{g}_t$.

3.1.4 Continuous term splines

Finally, we could also study the implementation of the spline method which takes into account the continuous curve measure. The criterion would be :

$$\begin{aligned} \mathcal{L}_4 &= (P_t - F_t g_t)'(F_tF_t')^{-1}(P_t - F_t g_t) \\ &\quad + \mu_1 \int_0^H \left[\frac{\partial^2 g}{\partial h^2}(t,h)\right]^2 dh + \mu_2 \int_0^H \left[g(t,h) - g_0(\theta,t,h)\right]^2 dh, \end{aligned} \tag{40}$$

where g_0 is the approximated model. The resolution of such a model only appears easy in the case where $\mu_2 = 0$ (see appendix 2).

3.2 Local regressions

3.2.1 Discount bond prices observation

Another way to introduce more flexibility in model $g(\theta, t)$ consists in using different approximations of θ according to the term of interest. These approximations can be obtained by weighting more in the regression the terms close to the maturity we are interested in.

These approaches by local regression have for instance been used to determine series corrected for seasonal variations (KENDALL and STUART (1968)) and are the basis for non parametric functional estimation by kernel regression (CLEVELAND (1979), HAERDLE (1990), ROBINSON (1990), GOURIEROUX, MONFORT and TENREIRO (1994)). As in the previous section, we first expose the approach in the fictitious case, where all discount bonds might be observed. The reasoning is presented date by date.

We introduce a kernel, i.e. a real positive function K. We impose moreover that function K is maximal at zero, increasing before zero, decreasing after, going to zero at $\pm\infty$. If we are interested in a given term k, the discount bond price $C(t, t+k)$ will be approximated by :

$$\widehat{C}(t, t+k) = g(\widehat{\theta}_{k,t}, t, t+k), \tag{41}$$

where the approximation $\widehat{\theta}_{k,t}$ is solution to the problem :

$$\widehat{\theta}_{k,t} = \underset{\theta}{\operatorname{argmin}} \sum_{h=1}^{H} K(\frac{k-h}{\mu}) \left[C(t, t+h) - g(\theta, t, t+h)\right]^2. \tag{42}$$

The idea is then to give more weight in the regression to terms h close to the term of interest k. The auxiliary parameter μ is a smoothing parameter. If μ goes to zero, only the information in $C(t, t+k)$ will be taken into account and from this fact the reconstitution $g(\widehat{\theta}_{k,t}, t, t+k)$ will be very erratic. If μ goes to infinity, weights become equal; we find as solution the usual OLS estimator; this estimator is independent of k and the reconstitution presents the required smoothing degree introduced in model $g(\theta, t)$. In order to extend the method to the case where discount bond prices are not directly observable, we have to point out that the previous weighted regression may be viewed as corresponding to a model indexed by k :

$$C(t, t+h) = g(\theta_k, t, t+h) + \eta_{h,t},$$

$$\text{with } E\eta_{h,t} = 0,\ V\eta_{h,t} = \omega^2 \lambda_{k,h},\ Cov(\eta_{h,t}, \eta_{h',t}) = 0,\ h \neq h',$$
$$\text{and where } \lambda_{k,h} = \left[K(\frac{k-h}{\mu})\right]^{-1}.$$

3.2.2 Coupon bond prices observation

The previous remark and section 2 results lead naturally to extend the approach in the following way. We define :

$$\widehat{C}(t, t+k) = g(\widehat{\theta}_{k,t}, t, t+k), \tag{43}$$

$$\text{with } \widehat{\theta}_{k,t} = \underset{\theta}{\operatorname{argmin}} (P_t - F_t g(\theta, t))' (F_t \Lambda_k F_t')^{-1} (P_t - F_t g(\theta, t)),$$
$$\text{and } \Lambda_k = diag\left(K(\frac{k-h}{\mu})\right)^{-1}.$$

In the particular case of a linear formulation $g(\theta, t) = X_t \theta$, whose k^{th} component is $X_{k,t}\theta$, we obtain the approximation :

$$\widehat{C}(t, t+k) = X_{k,t} \left(X_t' F_t' (F_t \Lambda_k F_t')^{-1} F_t X_t\right)^{-1} X_t' F_t' (F_t \Lambda_k F_t')^{-1} P_t. \tag{44}$$

The reconstitution formula (44) is not compatible with the constraint $P_t = F_t \widehat{C}_t$. We could adequately modify it. We have at this level to remember that these local approaches are means to find the function $g(t, t+h) = EC(t, t+h)$ without judging of its parametric form and that the weighting does not mean that the true modelization errors have different variances. All reconstitutions being done from the same kernel without introducing a smoothing parameter μ function of k, it is natural to modify (44) by proposing (see (25)) :

$$\hat{\hat{C}}_t = \hat{C}_t + F_t'(F_tF_t')^{-1}[P_t - F_t\hat{C}_t]. \tag{45}$$

3.3 Empirical results evaluation

We deal with a descriptive reconstitution approach, for which statistical properties cannot be easily evaluated. This could only be done by assuming that the portfolio number J_t goes to infinity and that these portfolios have sufficiently independent cash flow structures. Such hypotheses are not realistic for our type of application and we will not develop them.

However, we can try to evaluate from a practical point of view the finite sample precision of estimators obtained by smoothing methods (splines or local regressions) by using simulations. This solution is similar to the one proposed by GREGORY and SMITH (1991) for checking the quality of calibration in macroeconomic models.

We may begin by considering a model of type :

$$\begin{cases} P_t = F_tC_t, \\ C_t = \hat{g}_t + \eta_t, \\ E\eta_t = 0, V\eta_t = \omega^2\Lambda, \eta_t \ i.i.d \ N(0,\omega^2\Lambda), \end{cases} \tag{46}$$

where Λ is a given diagonal matrix and where $\hat{g}_t$ indicates the approximation obtained by smoothing. As :

$$VP_t = \omega^2 F_t\Lambda F_t',$$

we may approach the variance factor by :

$$\hat{\omega}^2 = \frac{1}{\sum_{t=1}^{T} J_t} \sum_{t=1}^{T} (P_t - F_t\hat{g}_t)'(F_t\Lambda F_t')^{-1}(P_t - F_t\hat{g}_t).$$

From these estimators, we may deduce the approximated formulation :

$$\begin{cases} P_t = F_tC_t, \\ C_t = \hat{g}_t + \eta_t, \\ \eta_t \sim N(0,\hat{\omega}^2\Lambda), \end{cases}$$

and use this formulation to simulate errors values η_t^s, $t = 1,\dots,T$ and values P_t^s of bonds, where $s = 1,\dots,S$ denotes the independent simulations and S the number of replications.

The smoothing method (spline ou local regression) can then be reapplied to these fictitious data, which gives estimations :

$$\begin{aligned} &\hat{g}_t^s, \quad t = 1,\dots,T, \ s = 1,\dots,S, \\ &\hat{\omega}_s^2, \qquad\qquad s = 1,\dots,S. \end{aligned}$$

This Bootstrap type approach (EFRON and TIBSHIRANI(1986)) gives an idea of the initial estimator precision. The law of $\widehat{g}_t$ is indeed well approximated by the empirical distribution of the simulated values $\widehat{g}_t^s$, $s = 1, \dots, S$, and it is also true for the variance estimator $\widehat{\omega}^2$.

4 Factor models and estimating constraints

Several models introduced to describe discount bond prices rely on the notion of factors or the notion of estimating constraints (see appendix 1). In this section, we discuss an example. Let us consider discount bond price evolutions such as :

$$C(t, t+h) \;=\; \sum_{k=1}^{K} a_k^h Z_{k,t} + \epsilon_{t,k},$$

which gives in vectorial notation :

$$C_t \;=\; AZ_t + \epsilon_t,$$

where errors terms are time independent, centered and independent of the introduced "factors" $Z_{k,t}$, $k = 1, \dots, K$. These factors can themselves be time correlated, so that all discount bond price dynamics (time correlation) is driven by them. Let us introduce the lag polynomial defined by :

$$\begin{aligned} Q(L) &= \prod_{k=1}^{K} (L - a_k) \\ &= \sum_{k=0}^{K} \alpha_k L^k (\text{by definition}), \text{ with } \alpha_0 = 1. \end{aligned}$$

We have :

$$Q(L)C(t, t+h) = \sum_{k=0}^{K} \alpha_k C(t, t+h-k) = \sum_{k=0}^{K} \alpha_k \epsilon_{t,h-k}, \; h > k.$$

We deduce from the assumption on the error terms, that if I_{t-1} denotes the available information at time $t-1$, including the observed prices until $t-1$, we have :

$$E\Big(\sum_{k=0}^{K} \alpha_k C(t, t+h-k) \mid I_{t-1}\Big) \;=\; 0, \tag{47}$$

or

$$E\Big(V_{t-1} \sum_{k=0}^{K} \alpha_k C(t, t+h-k)\Big) \;=\; 0, \tag{48}$$

where V_{t-1} is an instrumental variable, function of I_{t-1}.

4.1 Use of estimating constraints

When discount bonds are observable, estimating constraints (48) can directly be used in order to obtain estimators of parameters α_k, $k = 1, \dots, K$, i.e. of parameters a_k, $k = 1, \dots, K$, without looking for the distribution of the factors themselves. It is enough for that to use instruments as for example lagged bond prices, and to apply an instrumental variable type or generalized moment type estimation method (see THEIL (1961) and HANSEN (1982)).

What can be said about such a method when only coupon bond prices are observable? We have :

$$\begin{aligned} P_{j,t} &= \sum_{h=1}^{H} F_{j,t,h} C(t, t+h) \\ &= \sum_{k=1}^{K} \left(\sum_{h=1}^{H} F_{j,t,h} a_k^h \right) Z_{k,t} + \sum_{h=1}^{H} F_{j,t,h} \epsilon_{t,h}. \end{aligned}$$

It immediatly appears that, because of cash flow effects, it is no more possible to find a fixed coefficient combination $\sum_{j=1}^{J} \beta_j P_{j,t}$, for which factor effects could be eliminated, i.e. such that :

$$\sum_{j=1}^{J} \sum_{h=1}^{H} F_{j,t,h} a_k^h \beta_j = 0, \; k = 1, \dots, K,$$

or equivalently to exhibit estimating constraints :

$$E \left(\sum_{j=1}^{J} \beta_j P_{j,t} \mid I_{t-1} \right) = 0,$$

in the observable variables alone. Therefore it is impossible to extend moment type approaches.

4.2 Use of factor models

On the other hand, if the factor form is completed with a dynamic specification for the factors :

$$Z_t = \begin{pmatrix} Z_{1,t} \\ \vdots \\ Z_{K,t} \end{pmatrix} = \Phi_0 + \Phi_1 Z_{t-1} + \nu_t,$$

where ν_t is a centered error term, time independent of Z_{t-1} and of ϵ_t, we may rewrite the associated model with observable variables under the form :

$$\begin{cases} P_t = F_t A Z_t + F_t \epsilon_t, \\ Z_t = \Phi_0 + \Phi_1 Z_{t-1} + \nu_t. \end{cases}$$

We obtain a state space form (see HARVEY (1989)) which can be estimated by pseudo maximum likelihood under a normal distribution, by using the Kalman filter. The method allows to determine jointly parameters a_k, $k = 1, \dots, K$, parameters Φ_0, Φ_1 which give the first order dynamics of Z_t, and parameters of variance (variances of ϵ_t and η_t). The approach does not require a complete specification of the distribution of the error terms.

5 Empirical results

In order to illustrate the various proposed methods, we present an application to French data. We first briefly describe the data.

5.1 Data description

Since 1986, the French Treasury issues standardized bonds. Among these assets, we use two types of bonds : the OAT, "obligations assimilables du Trésor" and the BTAN, "bons du Trésor à taux annuel". The OAT constitute the long term debt instrument of the French government (their maturity is up to 30 years); they are issued in French francs or in ecus, redeemable in fine and with a fixed or variable annual interest rate. The BTAN are issued for 3 or 5 years, in french francs, with a fixed interest rate and are redeemable in fine; they are used for medium term financing.

The data concern a total of 37 bonds (OAT in francs with a fixed rate and BTAN). The covered period is from 14/12/1992 until 27/12/1993 (259 dates). The quotes are daily and taken around 4 p. m.; they are made of prices given by the French central bank after consultation of some specialized operators : the SVT's ("spécialistes en valeurs du Trésor") for the BTAN and prices of the "Caisse des dépôts et consignations" for the OAT. We only retain for estimation purpose the liquid bonds (about 15 for each date) and we use data which have preliminary been corrected for quote differences between OAT and BTAN. We have not taken into account tax effects since most of the banks and institutional investors (which are the usual buyers of OAT and BTAN) are exonerated.

5.2 Comparison of methods

5.2.1 Cubic splines

We begin with a modelization of discount bond prices by a cubic polynomial spline function (see MCCULLOCH (1975a,b)). Such functions are easily estimated by least squares (SUITS, MASON and CHAN (1978)) because discount bond prices can be expressed as :

$$C(t, t+h, \theta_{it}\, i = 1, \dots, k+4) = 1 + \theta_{1t} h + \theta_{2t} h^2 + \theta_{3t} h^3 + \sum_{i=4}^{k+4} \theta_{it} (h - h_i)^3 \mathbb{1}_i(h),$$

where k is the number of knots placed at terms h_i, $i = 4, \ldots, k+4$ and $\mathbb{1}_i(h)$ is the indicator function corresponding to the chosen knots : $\mathbb{1}_i(h) = 0$ if $h < h_i$; $\mathbb{1}_i(h) = 1$ if $h \geq h_i$.

The predictions of the discount bonds made with cash flow correction and without correction are very close (see Figures 3 et 4). The number of knots we first use is 5; knots are placed at maturities of 1, 5, 10, 15 and 20 years (in days : 365, 1825, 3650, 5475 and 7300 days). The corresponding figures for the yield curve are given in Figures 5 and 6.

The oscillations which appear at long maturities are due to the cubic polynomial specification used to fit the data and might be suppressed by diminishing the number of knots and by changing their location as can be seen on Figures 7 and 8 where two knots located at 1 and 10 years are used.

The mean and standard deviation of parameters θ_{it} computed on dates with 5 and 2 knots are given respectively in Tables 1 and 2. It is seen from time standard deviation that the yield curve without correction has been artificially smoothed with respect to time. These time standard deviations cannot be considered as a measure of the precision of the date by date predictors, since we know that this precision is better with correction than without correction.

We can easily see from the figures that the curve effect is better taken into account when evaluated with correction. We have computed durations corresponding to the set of all bonds for each date and for various rates. These durations are given by :

$$dur_t = \frac{\displaystyle\sum_{j=1}^{J_t} \sum_{h=1}^{H} \frac{hF_{j,t,t+h}}{(1+r)^h}}{\displaystyle\sum_{j=1}^{J_t} \sum_{h=1}^{H} \frac{F_{j,t,t+h}}{(1+r)^h}}$$

and are shown on Figure 9 for $r = 0$, 0.10 and 0.15. When the duration for a given date is low, it means that the short term cash flows are not as large as the long term cash flows. For these dates, the estimation without correction gives too much importance to the long term cash flows and the curve effect appears less clearly, while the estimation with correction takes into account this structure effect.

parameters	with correction		without correction	
	mean	stand. dev.	mean	stand. dev.
θ_{1t}	-3.7313 E-04	8.2211 E-05	-2.4857 E-04	7.4084 E-05
θ_{2t}	4.3487 E-07	3.0584 E-07	2.6939 E-07	1.8768 E-07
θ_{3t}	-3.2119 E-10	3.0683 E-10	-2.3919 E-10	1.7773 E-10
θ_{4t}	2.9590 E-10	3.1554 E-10	2.3835 E-10	1.8097 E-10
θ_{5t}	3.4423 E-11	1.5648 E-11	1.8142 E-12	8.3685 E-11
θ_{6t}	-1.3343 E-11	1.2186 E-11	-1.1156 E-12	1.1635 E-11
θ_{7t}	6.1964 E-12	1.5651 E-11	-3.0976 E-13	1.5492 E-11
θ_{8t}	-3.7476 E-12	1.4141 E-11	-7.6972 E-13	1.4074 E-11

Table 1: Parameter means and standard deviations : cubic polynomial function (5 knots)

parameters	with correction		without correction	
	mean	stand. dev.	mean	stand. dev.
θ_{1t}	-5.5320 E-04	9.1908 E-05	-2.5240 E-04	7.1433 E-05
θ_{2t}	1.1761 E-06	2.7159 E-07	2.8784 E-07	1.6217 E-07
θ_{3t}	-1.07285 E-09	2.5203 E-10	-2.5910 E-10	1.4612 E-10
θ_{4t}	1.0732 E-09	2.5253 E-10	2.5952 E-10	1.4599 E-10
θ_{5t}	-1.9471 E-13	1.8814 E-12	-5.9137 E-13	8.6534 E-13

Table 2: Parameter means and standard deviations : cubic polynomial function (2 knots)

5.2.2 Vasicek model

The second modelization which we consider concerns the model of VASICEK (1977). In this model, short term interest rate is assumed to follow an Ornstein Uhlenbeck process ($dr_t = (\alpha + \beta r_t)dt + \sigma d\mathcal{W}t$) and the discount bond price of maturity h is :

$$\begin{aligned} C(t, t+h, \lambda-\alpha, \beta, \sigma) &= \exp\left[\frac{\left(\left(\frac{1-e^{\beta h}}{-\beta}\right) - h\right)\left(\beta(\lambda-\alpha) - \sigma^2/2\right)}{\beta^2} + \frac{\sigma^2\left(\frac{1-e^{\beta h}}{-\beta}\right)^2}{4\beta}\right] \\ &\times \exp\left[-\left(\frac{1-e^{\beta h}}{-\beta}\right) r_t\right]. \end{aligned}$$

where λ represents the risk premium. The identifiable functions of the parameters are β, $\lambda - \alpha$, σ^2, r_t and are estimated date by date. As shown by Figures 10 and 11 where the date by date predicted yield curves are given, the results greatly differ between the approaches with and without correction. It seems that there exist some regime shifts between the different dates (3 regimes in this case). We give below (Figures 12-15) the values of the estimated parameters $\widehat{\beta}_t$, $\widehat{\lambda_t - \alpha_t}$, $\widehat{\sigma}_t$ and $\widehat{r}_t$ as functions of time and also the difference between these values and the values obtained by estimation without correction. The values for coefficient β_t are often outside the stationarity region ($-2 < \beta_t < 0$) and the short term interest rate are quite high. These elements lead to think that the VASICEK model is probably misspecified. Indeed, we can see that the observed regimes are linked with the cash flow structure summarized by the durations. The predicted yield curve of the cubic spline methods show that we are in presence of an inverse yield curve. Since this type of yield curve shape is ill fitted by the VASICEK model, the method with correction has implicitly estimated several VASICEK models depending on the weight of the short term cash flows relative to the long term cash flows.

5.2.3 Local regression based on the Vasicek model

In order to examine in greater details this model and to see whether the weighting of a specific maturity has a large influence, we look at its behavior when using a local regression approach. Table 3 gives the various possible kernels $K(u)$ which may be used. We examine the case of the Epanechnikov kernel (since the other

Epanechnikov	$(3/4)(-u^2+1)\mathbb{1}(\mid u \mid \leq 1)$
Quartic	$(15/16)(1-u^2)^2\mathbb{1}(\mid u \mid \leq 1)$
Triangular	$(1-\mid u \mid)\mathbb{1}(\mid u \mid \leq 1)$
Gaussian	$(1/\sqrt{2\pi})\exp(-\frac{u^2}{2})$
Uniform	$(1/2)\mathbb{1}(\mid u \mid \leq 1)$

Table 3: Kernels $K(u)$

kernels have given similar results) for a given date corresponding to parameter values satisfying stationarity conditions ($\widehat{r}_t$ = 0.11779, $\widehat{\beta}_t$ = -1.4548, $\widehat{\lambda_t - \alpha_t}$ = -0.35258, $\widehat{\sigma}_t$ = 0.85772). The values of the benchmark term k (see equation (42)) we use correspond to 1, 3, 5, 10, 15 and 20 years and we try several values of the bandwidth μ (500, 1000, 2500 and 5000). The parameter values are given in Table 4.
We find that the choice of μ has little influence for great values of k but that the predicted discount bonds are very sensitive for the short maturities (k = 365)(see Figures 16 and 17). It is due to the fact that the VASICEK model has some difficulties to fit the very short end of the term structure. However these results have to be taken with great care. Indeed, we use the Bootstrap type method proposed in section 3.3 to check the finite sample accuracy of our estimations. We take $k = 1095$, $\mu = 2500$ and we perform 1000 simulations. We use the quantile method i.e. we take $(1/2)(\widehat{C}(t,t+h)(25) + \widehat{C}(t,t+h)(26))$ as the lower limit of the confidence interval and $(1/2)(\widehat{C}(t,t+h)(975) + \widehat{C}(t,t+h)(976))$ as the upper limit where $\widehat{C}(t,t+h)(x)$ is the quantile at $x/10\%$ computed from the simulated values. Even for this "well

behaved" case, the confidence intervals are relatively large as shown in Figure 18. It has to be noted that the use of kernels may also be seen as a way to examine whether models based on a single state variable (e. g. the short term interest rate) are able to fit the term structure. Indeed in the case of one state variable model, all discount bonds of various maturities are perfectly correlated and the weighting of one or another maturity should not change the estimated parameters since the information contained in each bond should be the same if the model were well specified.

	parameters			
	r_t	β_t	$\lambda_t - \alpha_t$	σ_t
term k	bandwidth $\mu = 500$			
365	0.19889	-2.6618	-0.79320	1.7700
1095	0.11550	-1.4551	-0.35098	0.85823
1825	0.11625	-1.4550	-0.35053	0.85860
3650	0.11838	-1.4547	-0.35205	0.85813
5475	0.11960	-1.4546	-0.25413	0.85716
7300	0.11934	-1.4546	-0.35337	0.85755
term k	bandwidth $\mu = 1000$			
365	0.10668	-1.4564	-0.33374	0.86621
1095	0.11714	-1.4549	-0.35384	0.85805
1825	0.11626	-1.4550	-0.35090	0.85841
3650	0.11758	-1.4548	-0.35143	0.85835
5475	0.12026	-1.4545	-0.35442	0.85712
7300	0.11931	-1.4546	-0.35322	0.85763
term k	bandwidth $\mu = 2500$			
365	0.12308	-1.4541	-0.35504	0.85726
1095	0.12325	-1.4541	-0.35545	0.85706
1825	0.12343	-1.4541	-0.35604	0.85676
3650	0.11624	-1.4550	-0.35012	0.85883
5475	0.11955	-1.4546	-0.35284	0.85789
7300	0.11919	-1.4546	-0.35265	0.85795
term k	bandwidth $\mu = 5000$			
365	0.12277	-1.4542	-0.35621	0.85654
1095	0.12260	-1.4542	-0.35616	0.85654
1825	0.12251	-1.4542	-0.35605	0.85658
3650	0.12219	-1.4542	-0.35564	0.85676
5475	0.11626	-1.4550	-0.35141	0.85814
7300	0.11878	-1.4547	-0.35218	0.85813

Table 4: Parameter values estimated with an EPANECHNIKOV kernel

6 Conclusions

We describe several estimation procedures based on cash flow correction. Both parametric and non parametric models are discussed in the case of discount bond or coupon bond observability. Some empirical results are reported for the French bond market. Three techniques of term structure estimation are used. The first one fits a cubic polynomial spline to the discount function, the second is based on the parametric bond model of VASICEK and the third one considers local regressions based on the VASICEK model. For the non parametric techniques, we obtain rather stable parameters. But in the second case, misspecification elements are observed and in some estimations with and without cash flow correction are very different. This possibility to detect the misspecification of the VASICEK model by introducing different weighting is a consequence of the time dependent structure of the set of the bonds used for the estimations as it is can be seen by the evolution of the duration. Needless to say that in these cases of misspecification, the VASICEK model has to be used with caution when pricing bonds or bond derivative securities such as options or futures contracts.

Appendix 1 : The models in the literature

We can mainly distinguish two approaches in the modelling of discount bond prices. The first approach consists in estimating "directly" the term structure by various calibration methods. This approach is essentially a descriptive one. The second approach consists in introducing models based on some assumptions on the evolution of state variables and on asset pricing methods using either equilibrium or arbitrage arguments. The estimation will afterwards be based on various econometric techniques : nonlinear least squares (possibly generalized), generalized method of moments, maximum likelihood, ...

The first approach is followed by McCulloch (1971, 1975a,b), Vasicek and Fong (1982), Chambers, Carleton and Waldman (1984). This approach tries to adjust a discount function giving a smooth enough approximation of the data. The calibration is made thanks to interpolating functions (spline functions) linking the data. McCulloch (1971) uses second degree spline function to approximate the discount function and extends the method to third degree spline function when looking at tax effect adjusted functions (1975a,b). Vasicek and Fong (1982) use cubic polynomial function after linearization of the discount function (which is assumed to be a negative exponential function : $C(t, t+h) = e^{-\gamma_t h}$). The linearization is made after the variable change : $h = -\frac{1}{\alpha_t}\log(1-x)$ $0 \leq x < 1$, and the function which has to be adjusted is : $G_t(x) = (1-x)^{\frac{\gamma_t}{\alpha_t}}$. Chambers, Carleton and Waldman (1984) prefer to use the exponential polynomial function : $g_t(h) = \exp(-\sum_{j=1}^{J} a_{tj}h^j)$.

As we mention, the second approach is based on term structure models derived from a specification of the dynamic followed by one or several state variables and from asset pricing equilibrium or arbitrage arguments. These models are mainly developed in a continuous time framework.

We propose herafter a classification of the main models following various criteria. The criteria we use are : discrete time vs continuous time, equilibrium pricing vs arbitrage pricing, one factor model vs multifactor model (the underlying factors are given between brackets), analytical form vs numerical evaluation. The models classified in table (5) are the ones developped by Vasicek (V) (1977), Richard (R) (1978), Brennan and Schwartz (BS) (1979), Cox, Ingersoll and Ross (CIR1, CIR2) (1985), Ho and Lee (HL) (1986), Longstaff and Schwartz (LS) (1992), Heath, Jarrow and Morton (HJM) (1992), Clément, Gouriéroux and Monfort (CGM) (1993b), Duffie and Kan (DK) (1993), El Karoui and Lacoste (EK) (1993).

continuous time	**discrete time**
V, R, BS, CIR1, CIR2, LS, HJM, DK, EL	HL, CGM
pricing by equilibrium	**pricing by arbitrage**
CIR1, CIR2, LS	V, R, BS, HL, HJM, CGM, DK, EL
one factor model	**multifactor model**
V (instantaneous rate) CIR1 (instantaneous rate) HL (discount bond price)	R (instantaneous rate and inflation) BS (instantaneous rate and long term rate CIR2 (expected inflation and price level) LS (instantaneous rate and variance of instantaneous changes of rate) HJM (unspecified factors) CGM (unspecified factors) DK (yields) EL (unspecified factors)
explicit form	**numerical method**
V, R, CIR1, CIR2, HL, LS, HJM, CGM, EL	BS, DK

Table 5: Term structure models

Appendix 2 : splines with continuous term

1) Curvature optimization on interval $[h, h+1]$

This part of the proof is classical. Let us consider given values $g_0(t,h)$, $g_0(t,h+1)$, $\frac{\partial g_0}{\partial h}(t,h)$, $\frac{\partial g_0}{\partial h}(t,h+1)$ and denote : $\tilde{g}(h) = ah^3 + bh^2 + ch + d$, the third degree polynomial such that : $\tilde{g}(h) = g_0(t,h)$, $\tilde{g}(h+1) = g_0(t,h+1)$, $\frac{\partial \tilde{g}}{\partial h}(h) = \frac{\partial g_0}{\partial h}(t,h)$, $\frac{\partial \tilde{g}}{\partial h}(h+1) = \frac{\partial g_0}{\partial h}(t,h+1)$. We will establish that the polynomial function $\tilde{g}$ is solution of problem :

$$\min_g \int_h^{h+1} \left[\frac{\partial^2 g}{\partial u^2}(t,u)\right]^2 du,$$

submitted to :

$$\begin{aligned} g(t,h) &= g_0(t,h), \\ g(t,h+1) &= g_0(t,h+1), \\ \frac{\partial g}{\partial h}(t,h) &= \frac{\partial g_0}{\partial h}(t,h), \\ \frac{\partial g}{\partial h}(t,h+1) &= \frac{\partial g_0}{\partial h}(t,h+1). \end{aligned}$$

Equivalently, we have to establish that the problem :

$$\begin{aligned} &\min_f \int_h^{h+1} \frac{\partial^2}{\partial u^2}\left[\tilde{g}(u) + f(u)\right] du, \\ \text{s. t. } & f(h) = f(h+1) = \frac{\partial f}{\partial u}(h) = \frac{\partial f}{\partial u}(h+1) = 0, \end{aligned}$$

admits as solution the zero function.
Let us consider the function to optimize. It is written :

$$\int_h^{h+1} \left[6au + b + \frac{\partial^2 f}{\partial u^2}(u)\right]^2 du,$$

which gives :

$$\int_h^{h+1} [6au+b]^2\, du + \int_h^{h+1} \left[\frac{\partial^2 f}{\partial u^2}(u)\right]^2 du,$$

because of the constraints on function f. The solution is then obtained by imposing : $\frac{\partial^2 f}{\partial u^2}(u) = 0$, $\forall u \in [h, h+1]$, or equivalently $f(u) = 0$, $\forall u \in [h, h+1]$, taking into account the limit constraints.

2) Curvature expression on interval $[h, h+1]$

Let us define the third degree polynomial with respect to point h :

$$\begin{aligned}
\tilde{g}(u) &= a(u-h)^3 + b(u-h)^2 + c(u-h) + d, \\
\frac{\partial^2 \tilde{g}}{\partial u^2}(u) &= 6a(u-h) + 2b, \\
\int_h^{h+1} \left[\frac{\partial^2 \tilde{g}}{\partial u^2}(u)\right]^2 du &= \int_h^{h+1} [6a(u-h)+2b]^2\, du \\
&= 12a^2 + 4b^2 + 12ab.
\end{aligned}$$

We have still to explicit the expressions of coefficients a and b with help of the limit constraints at the interval ends. The conditions $\tilde{g}(h) = g(t,h), \frac{\partial \tilde{g}}{\partial u}(h) = \frac{\partial g}{\partial h}(t,h)$, provide :

$$d = g(t,h), \quad c = \frac{\partial g}{\partial h}(t,h).$$

The two other constraints give :

$$\left\{\begin{array}{l}
\tilde{g}(h+1) = g(t,h+1) = a + b + \frac{\partial g}{\partial h}(t,h) + g(t,h), \\
\frac{\partial \tilde{g}}{\partial h}(h+1) = \frac{\partial g}{\partial h}(t,h+1) = 3a + 2b + \frac{\partial g}{\partial h}(t,h),
\end{array}\right.$$

and so :

$$\left\{\begin{array}{l}
a = \frac{\partial g}{\partial h}(t,h+1) + \frac{\partial g}{\partial h}(t,h) - 2[g(t,h+1) - g(t,h)], \\
b = -\frac{\partial g}{\partial h}(t,h+1) - 2\frac{\partial g}{\partial h}(t,h) + 3[g(t,h+1) - g(t,h)].
\end{array}\right.$$

3) Criterion expression

The criterion $\mathcal{L}_4$ with $\mu_2 = 0$ is of the form :

$$\begin{aligned}
\mathcal{L}_4 = \; & (P_t - F_t g_t)'(F_t F_t')^{-1}(P_t - F_t g_t) \\
& +4\mu_1 \sum_{h=0}^{H-1} \left\{ 3\left[\frac{\partial g}{\partial h}(t,h+1) + \frac{\partial g}{\partial h}(t,h) - 2[g(t,h+1) - g(t,h)]\right]^2 \right. \\
& + \left[-\frac{\partial g}{\partial h}(t,h+1) - 2\frac{\partial g}{\partial h}(t,h) + 3[g(t,h+1) - g(t,h)]\right]^2 \\
& +3\left[\frac{\partial g}{\partial h}(t,h+1) + \frac{\partial g}{\partial h}(t,h) - 2[g(t,h+1) - g(t,h)]\right] \\
& \left. \times \left[-\frac{\partial g}{\partial h}(t,h+1) - 2\frac{\partial g}{\partial h}(t,h) + 3[g(t,h+1) - g(t,h)]\right] \right\}.
\end{aligned}$$

It is a quadratic form in unknown values $g(t,h),\ h=1,\ldots,H$, $\dfrac{\partial g}{\partial h}(t,h),\ h=0,\ldots,H$, and it has to be optimized under constraints $g(t,0)=1$. This allows to find the term structure for integer values. It is then deduced for intermediate values of $[h,h+1]$ by using the third degree polynomial compatible with the ending values, which have previously been found.

References

ANDERSON, S., DE PALMA, A., and F. THYSSE (1989) : "Demand for Differentiated Products, Discrete Choice Models and the Characteristic Approach", *Review of Economic Studies*, 56,.

BECKER, G. (1965) : "A Theory of the Allocation of Time", *Economic Journal*, 75, 493-517.

BIERWAG, G. (1987) : "Duration Analysis", Cambridge, Ballinger.

BRENNAN, M. E. SCHWARTZ (1979) : "A Continuous Time Approach to the Pricing of Bonds", *Journal of Banking and Finance*, 3, 135-155.

BROWN, S., and P. DYBVIG (1986) : "The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates", *Journal of Finance*, 41, 617-630.

COX, J., INGERSOLL, J., and S. ROSS (1985) : "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, 385-407.

CHAMBERS, D., CARLETON, W., and D. WALDMAN (1984) : "A New Approach to the Estimation of the Term Structure of Interest Rates", *Journal of Financial and Quantitative Analysis*, 19, 233-252.

CLEMENT, E., GOURIEROUX, C., and A., MONFORT (1993a) : "Modèles Statistiques de Valorisation par Arbitrage", Document de Travail INSEE, no 9315.

CLEMENT, E., GOURIEROUX, C., and A., MONFORT (1993b) : "Linear Factor Models and the Term Structure of Interest Rates", Document de Travail INSEE, no 9331.

CLEVELAND, W. (1979) : "Robust Locally Weighted Regression and Smoothing Scatterplots", *Journal of the American Statistical Association*, 74, 823-836.

DE MUNNIK, J., and P. SCHOTMAN (1993) : "Cross Sectional Versus Time Series Estimation of Term Structure Models : Empirical Results for the Dutch Bond Market", WP, University of Limburg.

DUFFIE, D. (1992) : "Dynamic Asset Pricing Theory", Princeton, Princeton University Press.

DUFFIE, D., and R. KAN (1993) : "A Yield-Factor Model of Interest Rates", Working Paper, Graduate School of Business, Stanford University.

EFRON, B., and R. TIBSHIRANI (1986) : "Bootstrap Method for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy", *Statistical Science*, 1, 54-77.

EL KAROUI, N., and V. LACOSTE (1992) : "Multifactor Models of the Term Structure of Interest Rates", Working Paper, University Paris IV.

EUBANK, R. (1988) : "Spline Smoothing and Nonparametric Regression", New York, Marcel Dekker.

GOOD, I., and R. GASKINS (1971) : "Non Parametric Roughness Penalties for Probability Densities", *Biometrika*, 5, 255-277.

GOURIEROUX, C., MONFORT, A., and C. TENREIRO (1994) : "Kernel M-Estimators : Nonparametric Diagnostics for Structural Models", DP INSEE.

GREGORY, A., and G. SMITH (1991) : "Calibration as Testing : Inference in Simulated Macroeconomic Models", *Journal of Business and Economic Statistics*, 3, 297-303.

GRILICHES, Z. (1961) : "Hedonic Price Indexes for Automobiles : An Econometric Analysis of Quality Change", in "The Price Statistics of the Federal Government", General Series, 73, New York, Columbia University Press for the NBER.

GRILICHES, Z., and I. ADELMAN (1961) : "On an Index of Quality Change", *Journal of the American Statistical Association*, 56, 535-548.

HAERDLE, W. (1990) : "Applied Nonparametric Regression", Cambridge, Cambridge University Press.

HANSEN, L. (1982) : "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50, 1029-1054.

HARVEY, A. (1989) : "Forecasting, Structural Time Series Models and the Kalman Filter", Cambridge, Cambridge University Press.

HEATH, D., JARROW, R., and A. MORTON (1992) : "Bond Pricing and the Term Structure of Interest Rates : A New Methodology for Contingent Claims Valuation", *Econometrica*, 60, 77-106.

HO, T., and S. LEE (1986) : "Term Structure Movements and the Pricing of Interest Rate Contingent Claims", *Journal of Finance*, 41, 1011-1030.

KENDALL, G., and A., STUART (1968) : "The Advanced Theory of Statistics", vol. 3, London, Charles Griffin.

LANCASTER, K. (1971) : "Consumer Demand : A New Approach", New York, Columbia University Press.

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

MCCULLOCH, J. (1971) : "Measuring the Term Structure of Interest Rates", *Journal of Business*, 44, 19-31.

MCCULLOCH, J. (1975a) : "The Tax Adjusted Yield Curve", *Journal of Finance*, 30, 811-829.

MCCULLOCH, J. (1975b) : "An Estimate of the Liquidity Premium", *Journal of Political Economy*, 83, 95-119.

POO, J.-R. (1992) : "Constrained Nonparametric Regression", PhD Thesis, Louvain-la-Neuve.

RICHARD, S. (1978) : "An Arbitrage Model of the Term Structure of Interest Rates", *Journal of Financial Economics*, 6, 33-57.

ROBINSON, P. (1990) : "Time Varying Non Linear Regression", in "Economic Structural Change, Analysis and Forecasting", P. Hackl and A. Westland eds, Springer Verlag.

ROSEN, S. (1974) : "Hedonic Prices and Implicit Markets Product Differentiation in Pure Competition", *Journal of Political Economy*, 82, 34-55.

ROSS, S. (1976) : " Return Risk and Arbitrage" in "Risk and Return in Finance", Vol. 1, I. Friend and J. Bicksler eds, Cambridge, Ballinger.

SCHOENBERG, I. (1964) : "Spline Functions and the Problem of Graduation", *Mathematics*, 52, 947-950.

SUITS, D., MASON, A., and L. CHAN (1978) : "Spline Functions Fitted by Standard Regression Methods", *Review of Economics and Statistics*, 60, 132-139.
THEIL, H. (1961) : "Economic Forecasts and Policy", Amsterdam, North Holland.
VASICEK, O. (1977) : "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, 5, 177-188.
VASICEK, O., and H. FONG (1982) : "Term Structure Modeling with Exponential Splines", *Journal of Finance*, 37, 339-356.

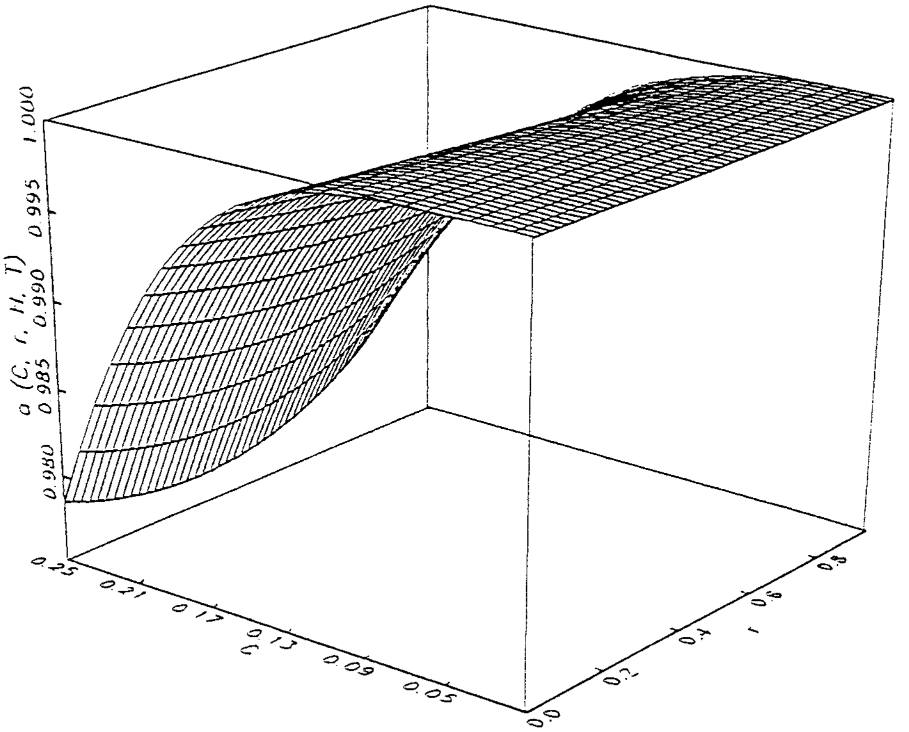

Figure 1: Ratio $\alpha(C, \rho, H, T)$: $H = 10$, $T = 10$.

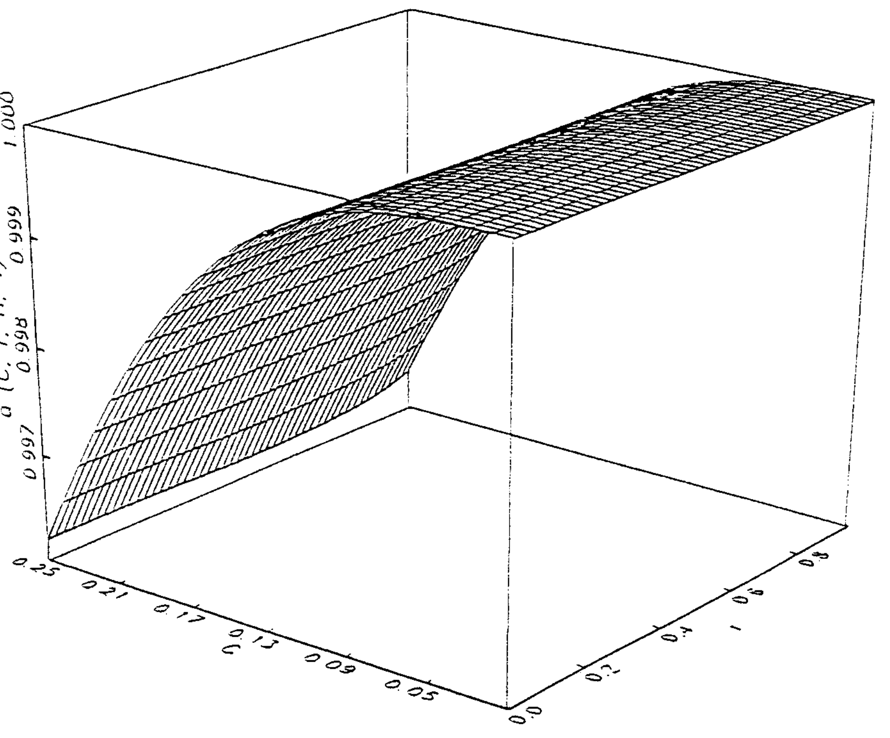

Figure 2: Ratio $\alpha(C, \rho, H, T)$: $H = 10$, $T = 5$.

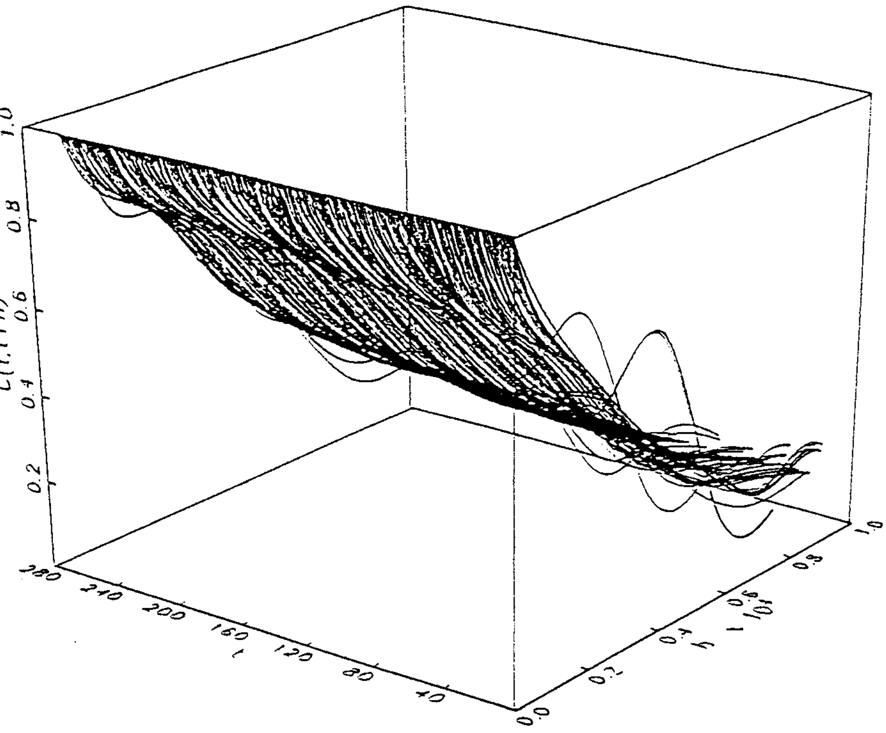

Figure 3: Discount bond predictions : spline with cash flow correction (5 knots)

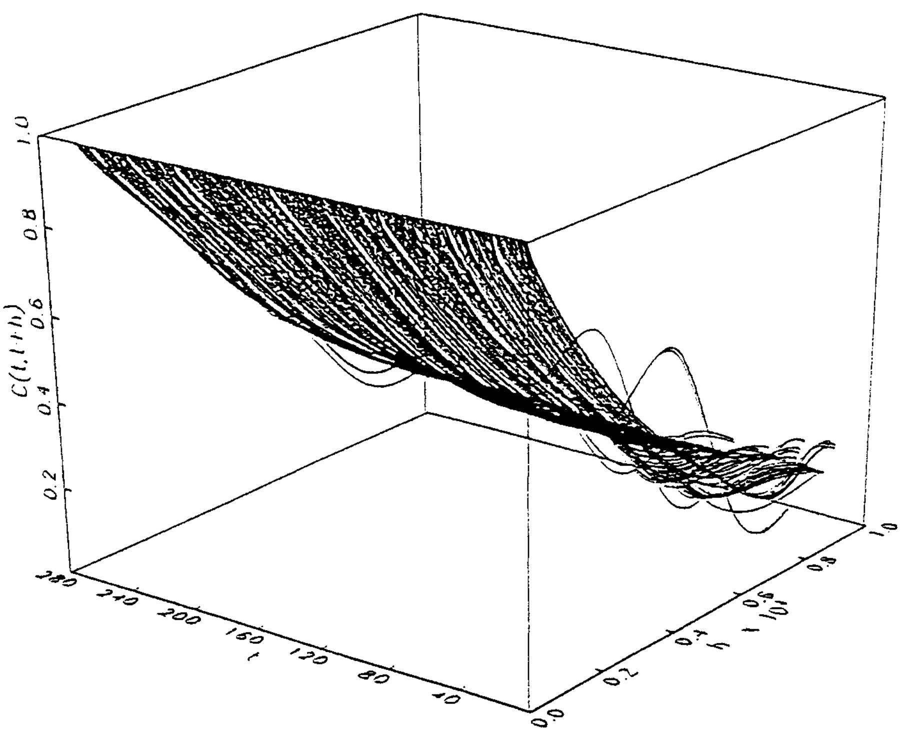

Figure 4: Discount bond predictions : spline without cash flow correction (5 knots)

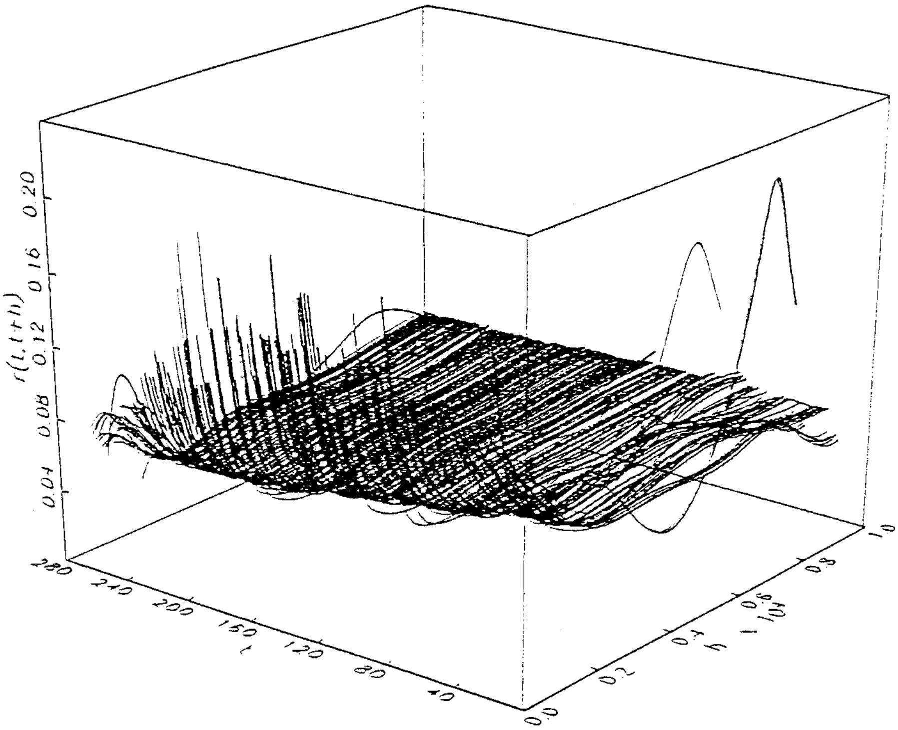

Figure 5: Predicted yield curve : spline with cash flow correction (5 knots)

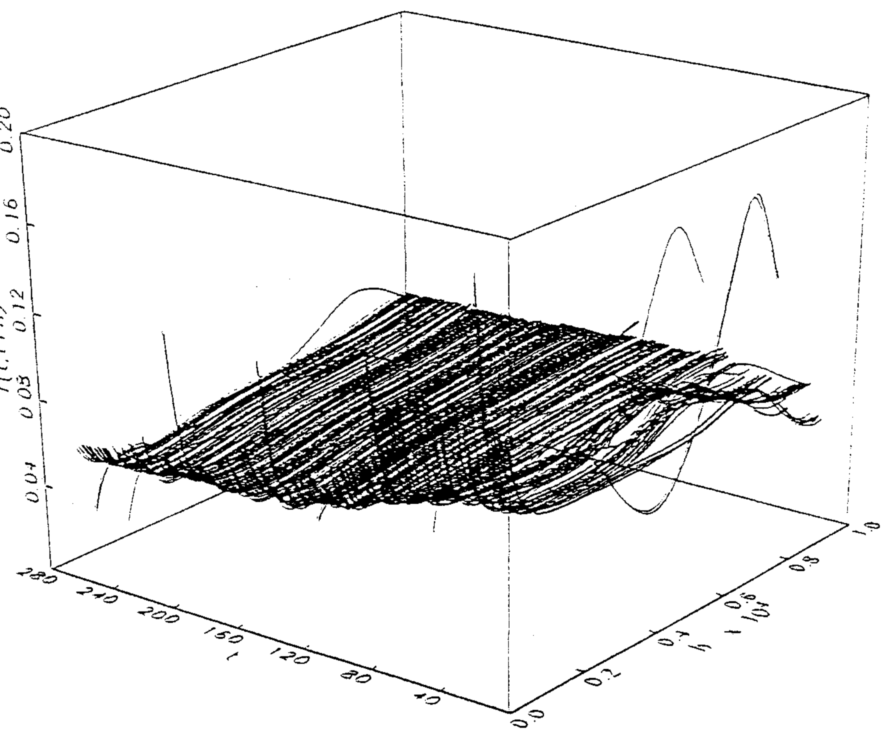

Figure 6: Predicted yield curve : spline without cash flow correction (5 knots)

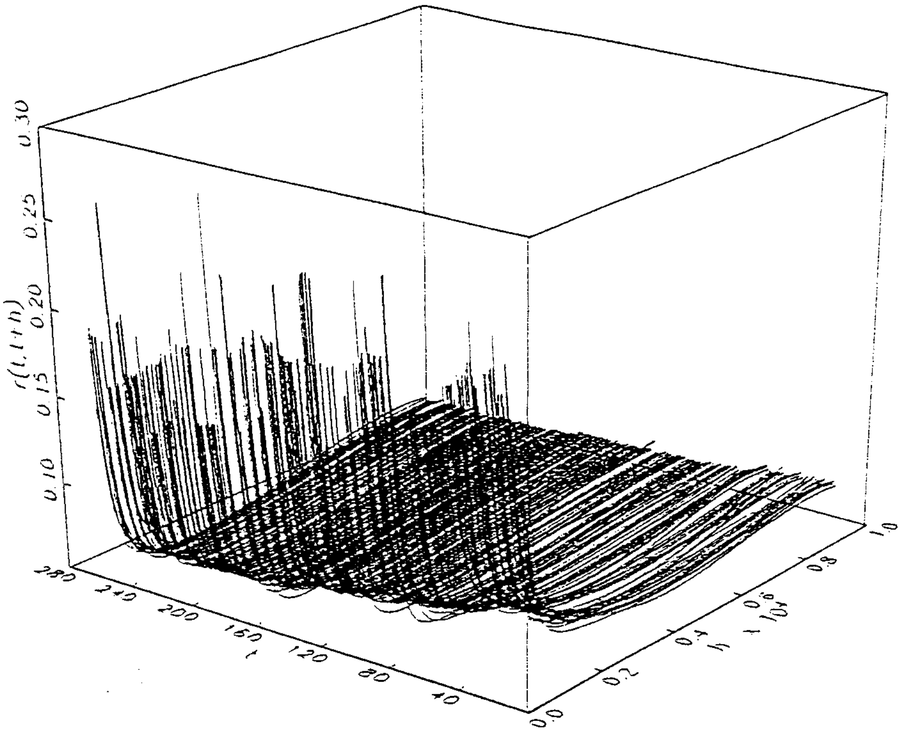

Figure 7: Predicted yield curve : spline with cash flow correction (2 knots)

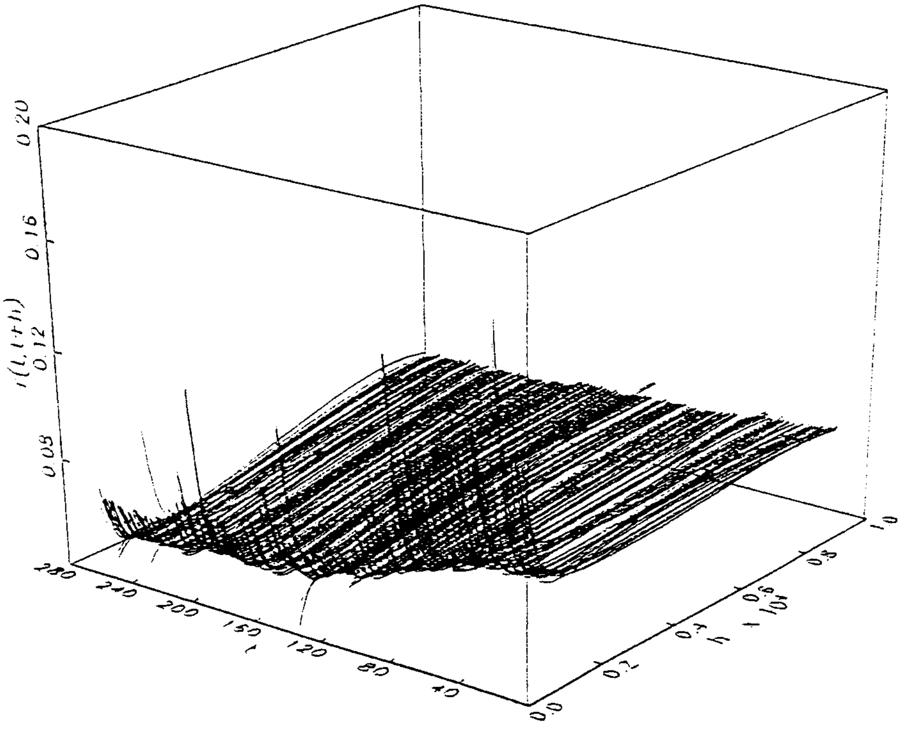

Figure 8: Predicted yield curve : spline without cash flow correction (2 knots)

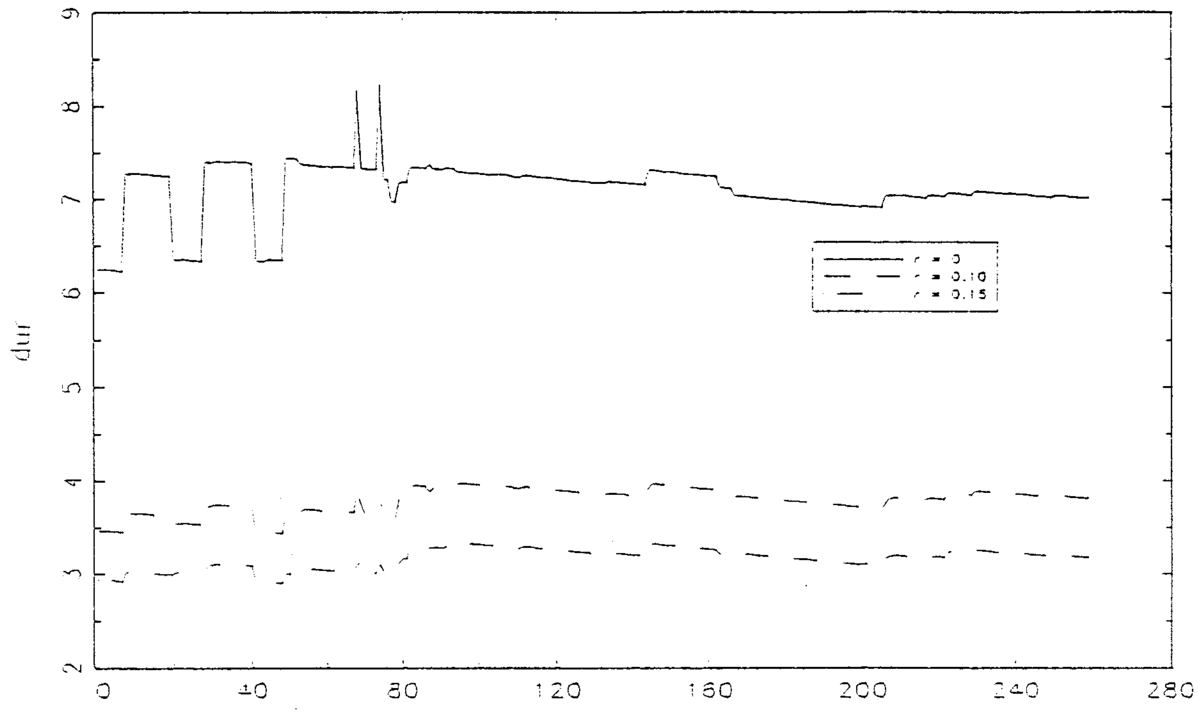

Figure 9: Durations

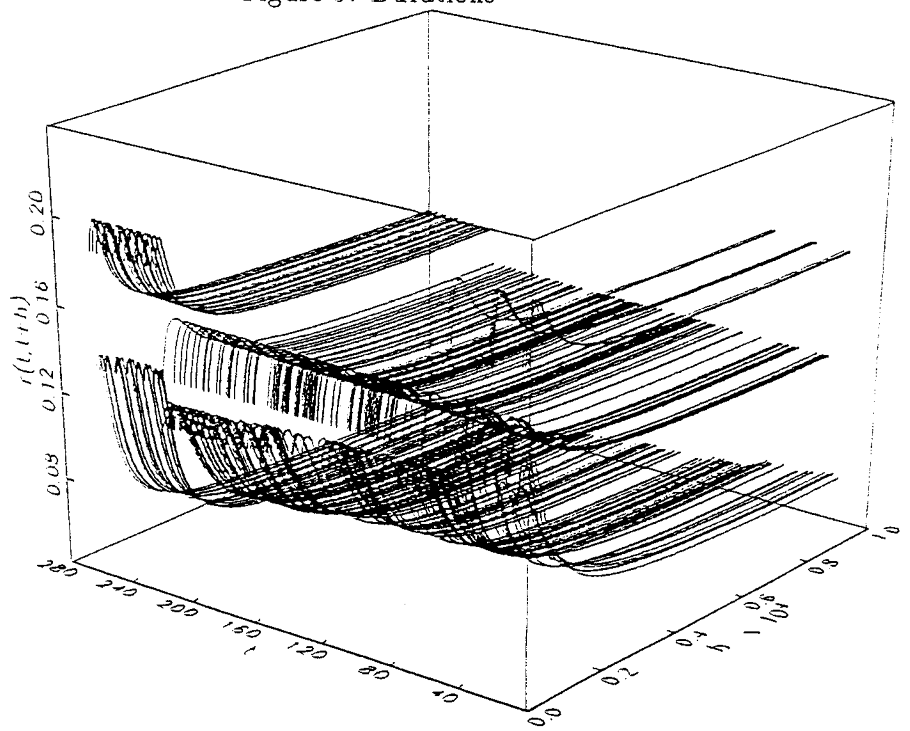

Figure 10: Predicted yield curve : VASICEK model with cash flow correction

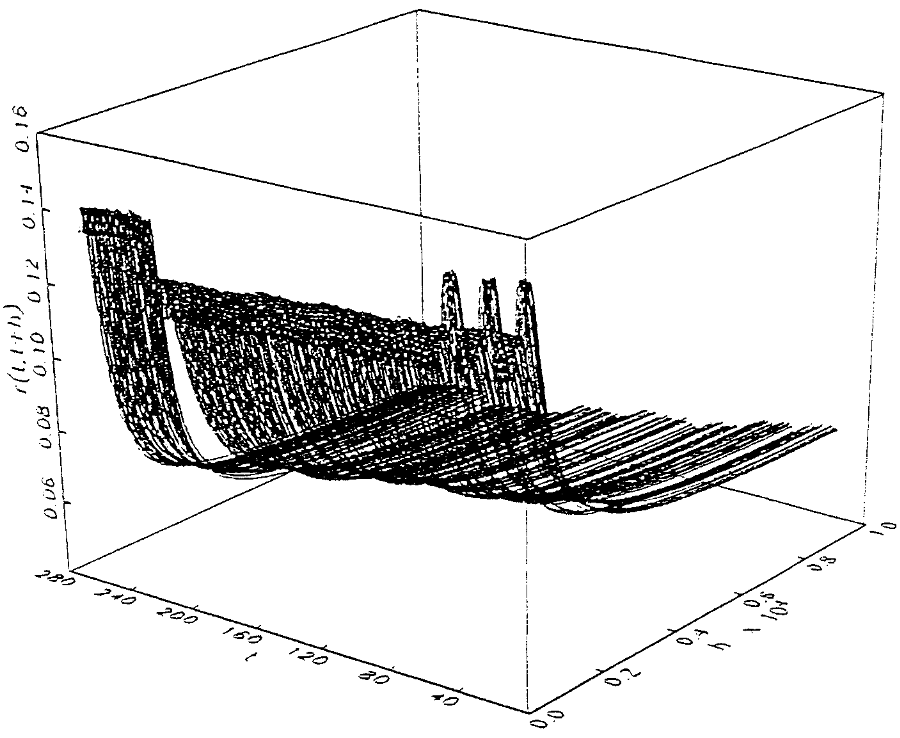

Figure 11: Predicted yield curve : VASICEK model without cash flow correction

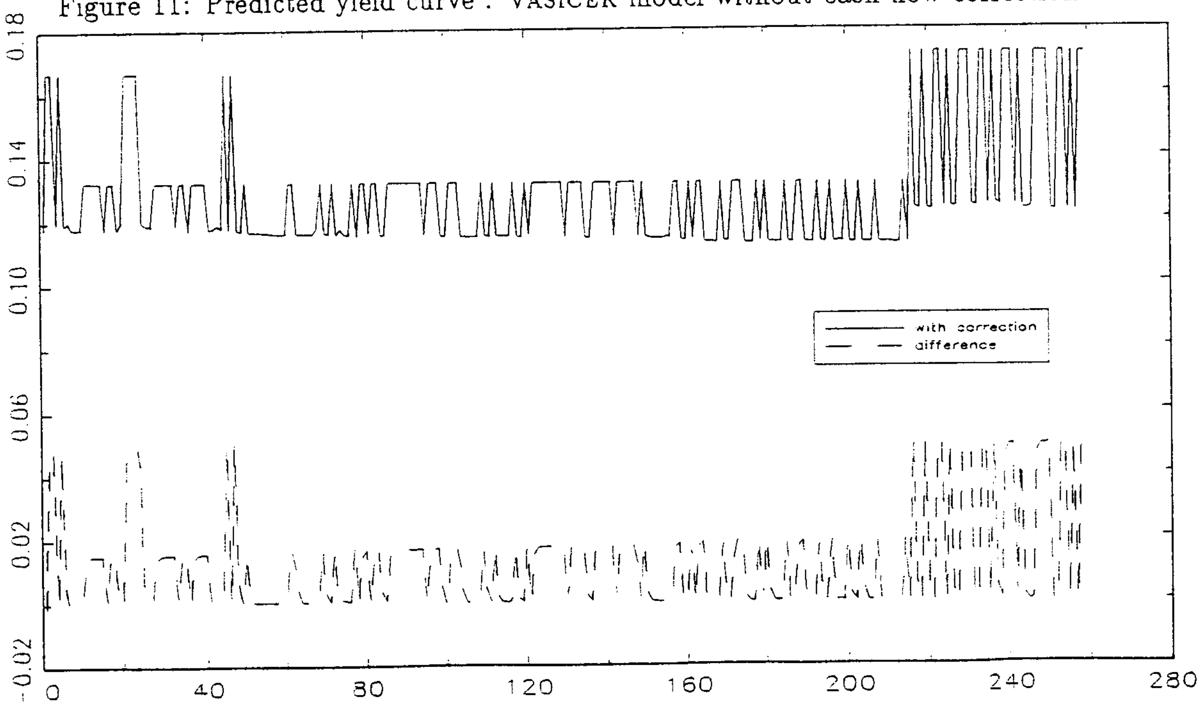

Figure 12: Predicted values of r_t

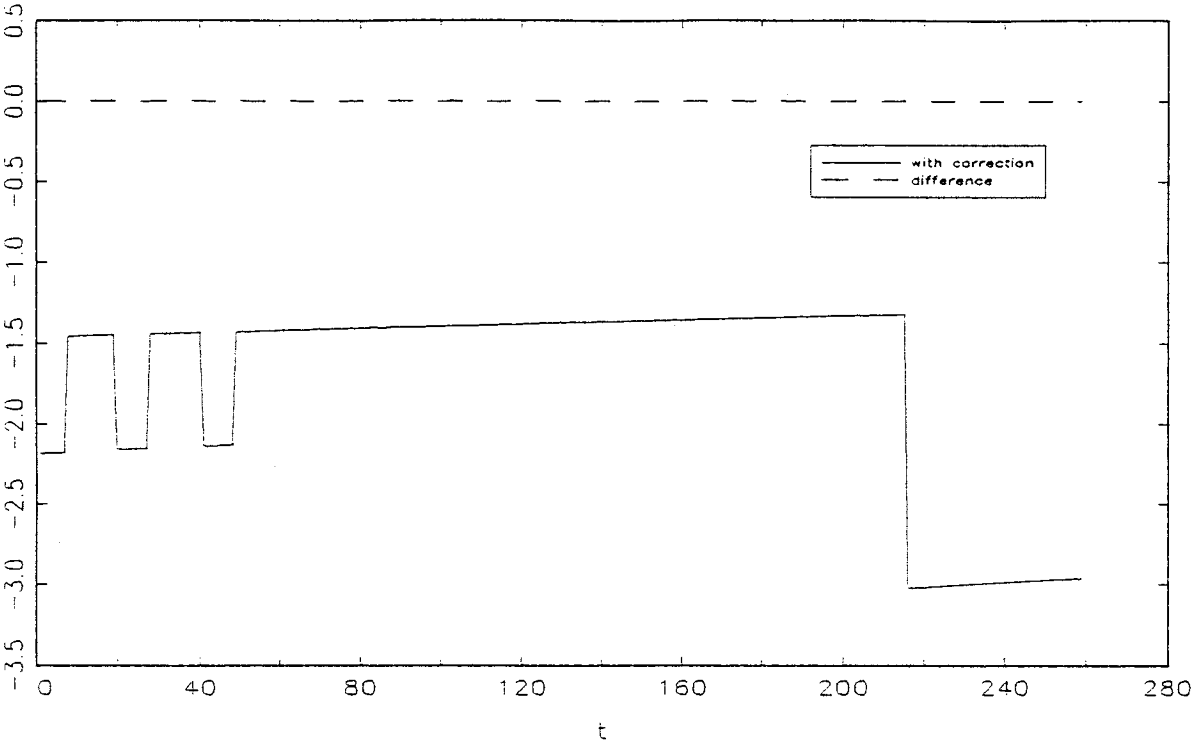

Figure 13: Predicted values of β_t

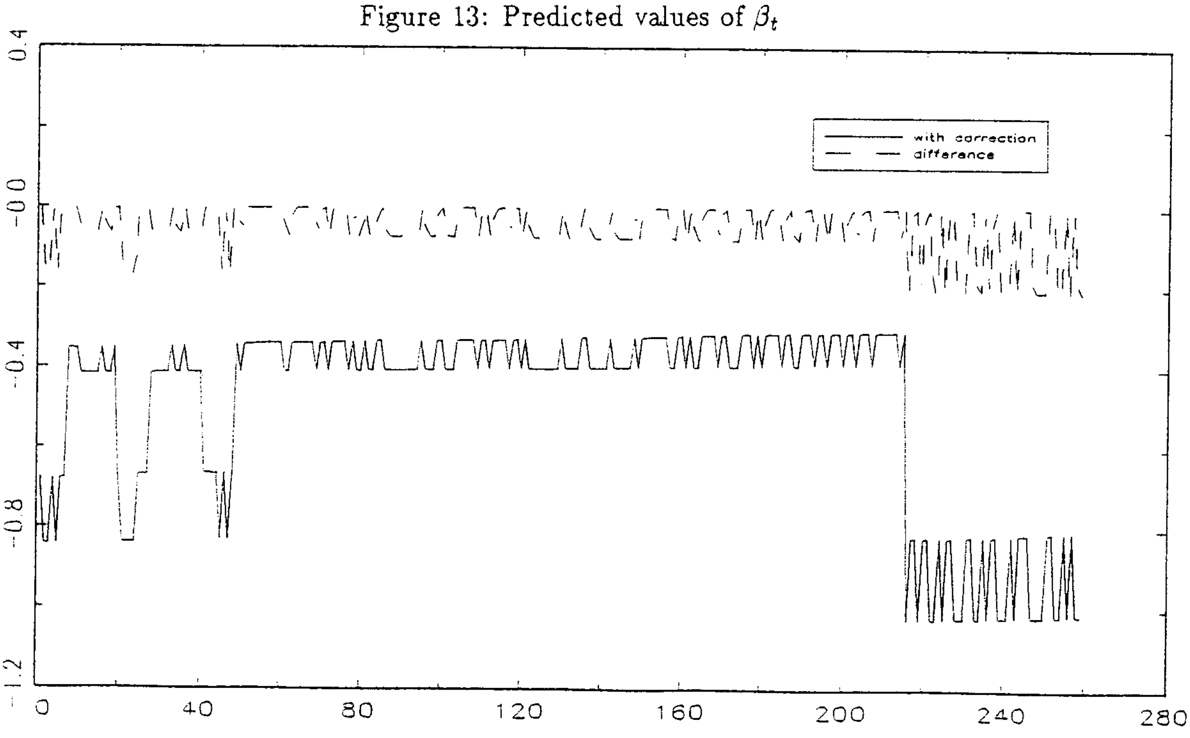

Figure 14: Predicted values of $\lambda_t - \alpha_t$

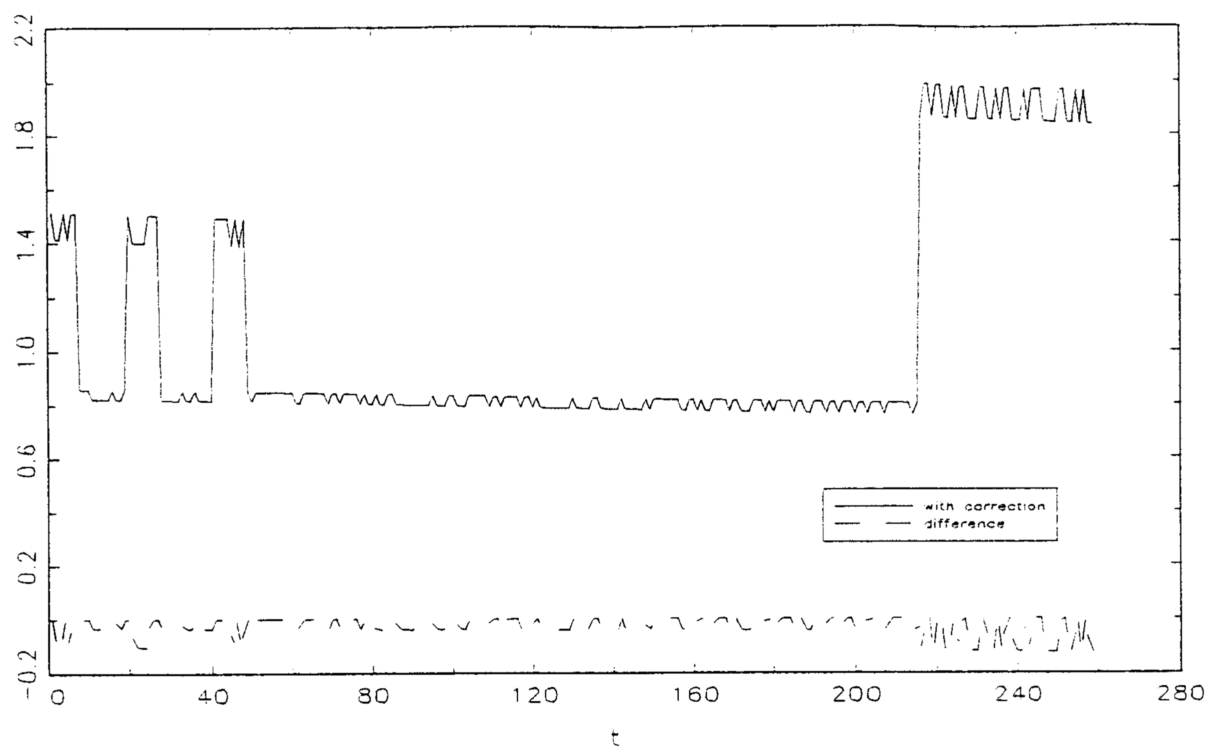

Figure 15: Predicted values of σ_t

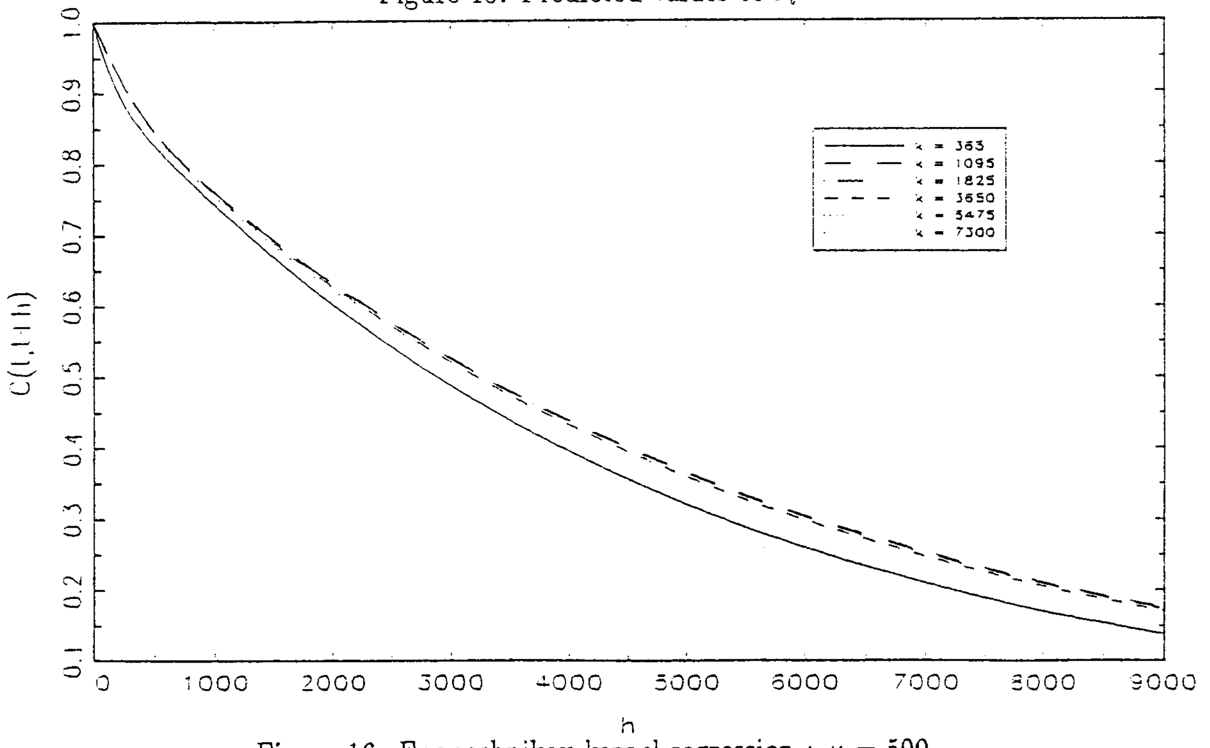

Figure 16: Epanechnikov kernel regression : $\mu = 500$

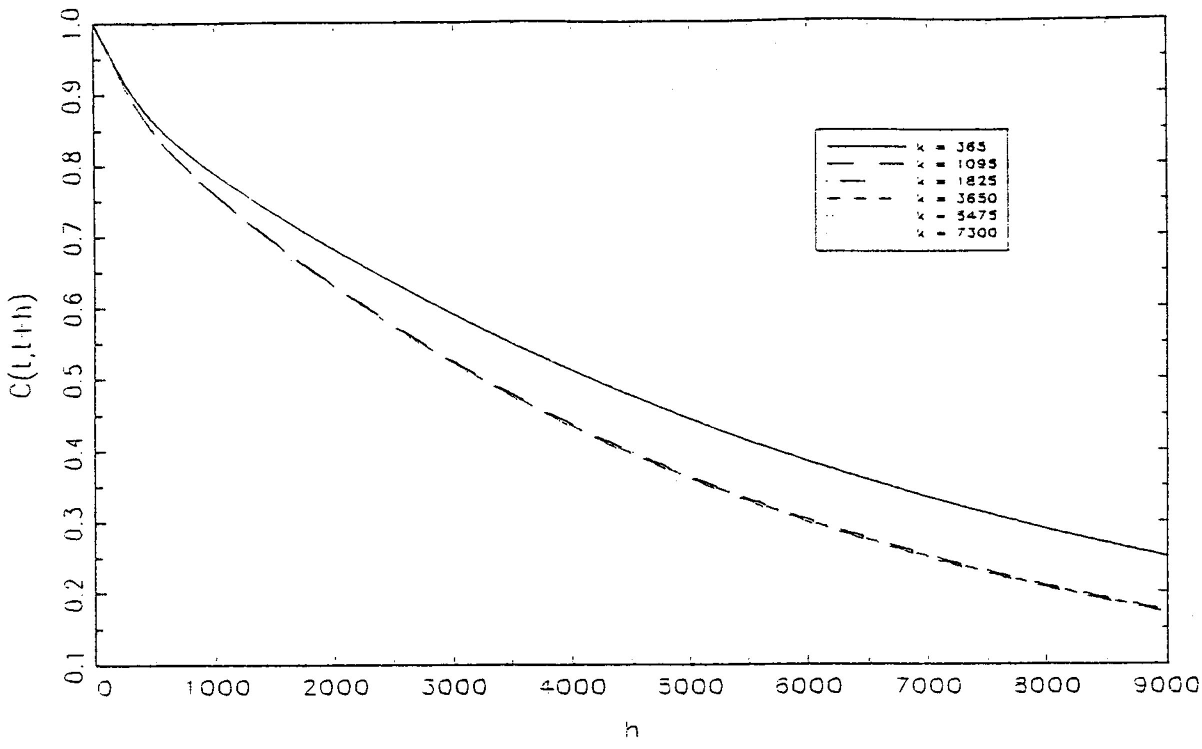

Figure 17: Epanechnikov kernel regression : μ = 1000

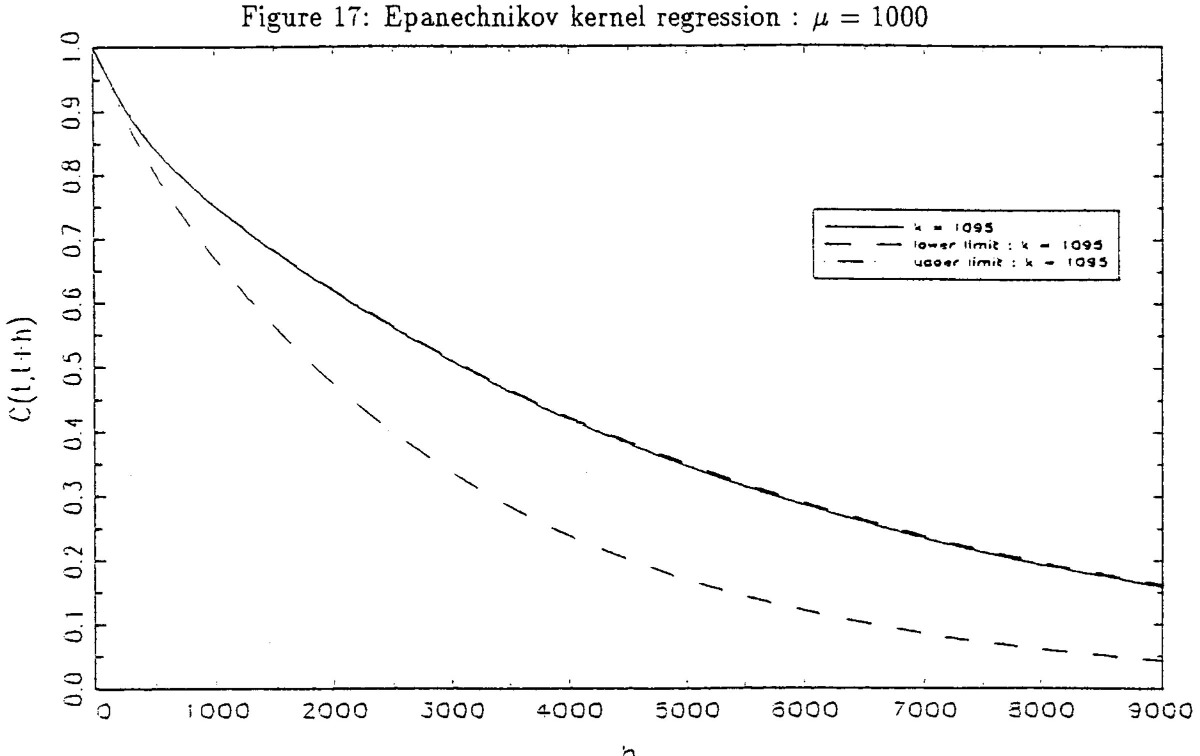

Figure 18: Epanechnikov kernel regression (μ = 2500) : confidence intervals