INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

A LINEAR MODEL FOR THE TERM STRUCTURE OF INTEREST RATES

Monia Mazigh

Faculty of Management

McGill University, Montreal

October, 2000

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Doctor of Philosophy.

©

Monia Mazigh 2000

National Library of Canada

Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services

Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-70095-X

Canada

Abstract

The term structure of interest rates shows the relationship between yields of zero-coupon bonds and their maturities. The empirical performance of the single-factor model of the affine term structure models, such as Vasicek (1977) and Cox, Ingersoll, and Ross (1985), has not been entirely satisfactory. The curve fitting methods, and particularly the spline method, used in practice to estimate the term structure are *ad hoc* and thus subject to arbitrage opportunities. Guo (1998) used the fundamental Partial Differential Equation (PDE) for bond pricing to derive a linear discount function, which is consistent with no-arbitrage. He showed that this is the unique linear solution to the PDE. This solution, the exponential-polynomial model or EP model for short, has n unobserved state factors that drive a stochastic discount process for pricing bonds so as to rule out arbitrage opportunities. In this thesis, we conduct an extensive cross-sectional analysis of the EP model on two different data sets: prices for daily Treasury bills, notes and bonds from the New York Federal Reserve Bank quotation sheets from July 1989 to October 1996, and daily Canadian bills, notes and bonds prices for the time period from June 1992 to May 1995. We estimate the model by applying a minimization criterion. The cross-sectional analysis shows that the EP model is able to describe adequately the term structure of interest rates. For the US data, we find that every term structure from the sampling period can be fully represented by either nine or ten state factors. Eigenvalue analysis indicates that the first three principal components are underlying the term structure movements. We conduct a time series analysis on the three principal components. They are found to be best described by ARMA/GARCH processes. We form two types of GARCH forecasts of the three principal components and test their out-of-sample performance. We conclude that the three principal components are predictable in a statistical sense. The use of an arbitrage strategy that attempts to take advantage of such predictive power generates some economic profits. However given the relative small size of the arbitrage profits, they will tend to vanish after transaction costs are considered.

Résumé

La structure à terme des taux d'intérêt étudie la relation entre les rendements des obligations à zéro-coupons et leurs maturités. Les test empiriques effectués sur les modèles à un seul facteur de la structure à terme des taux d'intérêt, tel quel le modèle de Vasicek (1977) ou le modèle de Cox, Ingersoll et Ross (1985) ont démontré que ces derniers ne décrivent pas adéquatement la courbe des taux d'intérêt. Par ailleurs, les méthodes empiriques, telle que la méthode spline, sont généralement dérivées d'une manière arbitraire et donc permettent la réalisation d'opportunités d'arbitrage. Guo (1998) a utilisé l'Équation aux Différentielles Partielles pour l'évaluation des obligations, afin de dériver une fonction d'actualisation linéaire qui soit consistante avec les conditions de non-arbitrage. Il a démontré qu'une telle solution est unique. Cette solution, désormais dénotée modèle exponentiel-polynomial (EP), décrit la fonction d'actualisation par n facteurs non observables. Dans notre thèse, nous effectuons une étude empirique exhaustive sur le modèle EP. Pour cela, nous utilisons deux bases de données distinctes. La première base de données est constituée de prix journaliers américains sur les bons du trésor, notes et obligations, répertoriés par la Banque Fédérale de New York. Elles couvrent la période de juillet 1989 jusqu'à octobre 1996. La seconde base de données est constituée de prix journaliers de bons, notes et obligations du gouvernement canadien. Ces données sont cueillies par la Banque du Canada. Elle s'étend de juin 1992 jusqu'à mai 1995. Nous avons estimé le modèle en appliquant un critère de minimisation des erreurs au carré. L'analyse en-section nous a permis de conclure que le modèle EP décrit adéquatement la courbe des taux. Chaque structure à terme de notre échantillon américain est décrite par neuf ou dix facteurs d'état. L'analyse des valeurs propres de ces facteurs nous indique que trois composantes principales sont essentielles pour expliquer la variation de la structure à terme. Par ailleurs, nous avons démontré que les séries temporelles de ces trois composantes principales du modèle EP, sont décrites par des processus ARMA/GARCH. Ce résultat est intéressant dans le sens où il est en accord avec des études précédentes sur le processus de certaines variables d'état des modèles de structure à terme. Nous avons utilisé les processus en question pour construire des prévisions hors-échantillon et étudier leur performance. Nous avons trouvé que les prévisions des principales composantes à partir des modèles ARMA/GARCH contiennent des informations substantielles. De plus, l'inclusion de ces prévisions dans des stratégies d'arbitrage a permis de générer des profits. Toutefois ces profits ne sont pas élevés et nous pensons qu'ils auront tendance à disparaître après que les coûts de transaction soient pris en compte.

Acknowledgement

I am most grateful to God for having helped me to complete this work.

I am indebted to Professor G. A. Whitmore, chairman of my dissertation committee, for his guidance and helpful comments.

I would like to thank Professor Chen Guo, for his comments and his enlightening discussions.

I would like to thank my dissertation committee members, Professor Jean-Louis Goffin, Professor Peter Ryan and Professor Jean-Guy Simonato for their valuable comments.

I would like to thank Professor Richard Loulou, Professor Allan Lee and Professor Jan Jorgensen for their help and support as Associate Deans of the Ph.D. Program.

Of course, I would like to thank my parents and my brother for their love and encouragement. Last but not least, I would like to thank my husband Maher for his help, support and patience while I completed this thesis.

For Barâa,

Contents

List of Figures

List of Tables

Introduction

1.1 Term Structure of Interest Rates

The term structure of interest rates is an important subject in finance. In very simple words, it shows the relationship between yields of zero-coupon bonds and their maturities. In this chapter, we will use the following notation:

- $P(x, t, \tau)$ denotes the price at date t of a discount bond with time τ to maturity. The price is assumed to depend on a state vector x. The bond pays one dollar (in all states x) at date $T = t + \tau$. By definition

$$P(x, t, 0) = 1.$$

- $D(x, t, \tau)$ denotes the discount function at time t. It corresponds to a discount bond price. It depends on a state vector x. It pays one dollar at date $T = t+\tau$.

- $y(x, t, \tau)$ denotes the yield at time t on a bond of maturity τ in state x. By definition

$$y(x, t, \tau) = -\frac{\log P(x, t, \tau)}{\tau} \quad \text{for} \quad \tau > 0. \tag{1.1}$$

- $r(x, t)$ denotes the short interest rate at time t, in state x. By definition,

$$r(x, t) = \lim_{\tau \to 0} y(x, t, \tau).$$

- $f(x, t, \tau + \delta t)$ denotes the forward rate with term δt. It is derived from the prices of two bonds maturing a δt period apart, as follows:

$$f(x, t, \tau + \delta t) = \log\left(\frac{P(x, t, \tau)}{P(x, t, \tau + \delta t)}\right). \tag{1.2}$$

  f is the rate of return that one can earn from $(t + \tau)$ to $(t + \tau + \delta t)$, with a long position in the $(\tau + \delta t)$-period bond and a short position in the τ-period bond.

- $F(t, \tau)$ denotes the instantaneous forward rate as seen at time t for a contract maturing at time τ. By definition,

$$F(t, \tau) = \lim_{\delta t \to 0} f(x, t, \tau + \delta t)$$

From definitions 1.1 and 1.2, one can deduce that yields are averages of forward rates:

$$y(x, t, \tau) = \frac{1}{\tau}\sum_{j=0}^{\tau-1} f(x, t, j). \tag{1.3}$$

Thus, the maturity structure of discount bonds can be expressed in three equivalent ways: prices, yields or forward rates. The use of yield curves is standard in monetary policy analysis in central banks and elsewhere. However, the use of forward rates, among other indicators, has started to be used by some financial institutions (see Svensson (1994)).

1.2 Motivation for Research

The subject of term structure is directly related to the bond market and real economic activity. Furthermore, it is used by central banks as an economic indicator for setting monetary policy. Traditionally, it is believed that central banks mainly affect short-term interest rates, such as yields on Treasury bills, whereas real economic activity is more linked to yields on bonds with the same maturity as physical capital, in the range of 10 to 20 years. Thus, it is important to understand the factors which affect the yields on these securities with different maturities and hence have a better understanding of the central bank role for affecting the state of the economy and the stance of monetary policy.

More recently, a big push for term structure research has come from the world of practice, which has experienced an explosion in interest rate derivative products. In fact, a thorough understanding of the empirical properties of term structure becomes more and more desirable because term structure conveys information about market expectations for the behavior and future course of interest rates. This information is essential for the pricing of interest rate-contingent claims.

Models of the term structure of interest rates range from simple curve fitting techniques to sophisticated theoretical models. In 1977, Vasicek launched the study of term structure models. His model and, later, one by Cox, Ingersoll, and Ross (1985), focused on describing the dynamics of the short rate. However some restrictions have been placed on the form of the stochastic process of the short rate in order to derive closed-form solutions for bond prices and the prices of contingent claims. Unfortunately, tractable models sometimes have undesirable economic properties. For instance, the assumption that interest rates follow an Ornstein-Uhlenbeck process (e.g. in the Vasicek model) leads to a closed-form solution for bond prices and interest rate derivatives but allows negative interest rates. Other models such as that of Heath, Jarrow, and Morton (1992) take into account the shape and the dynamics of the entire term structure. For this, the latter authors specify the current term structure, which is considered as the underlying asset in this model, in terms of forward interest rates. The derivation of results on the equivalent martingale measure is easier with the forward rate specification than with the spot rate specification. However, the forward rate specification turns out to be more difficult to implement. Even so, these bond price models have not been able to explain observed term structures; see for instance Brown and Dybvig (1986), and Gibbons and Ramaswamy (1993), and Kaushik and Morton (1994).

On another level, curve fitting techniques have been used to give a more practical way of describing the term structure. In 1975, McCulloch suggested the use of polynomial spline models to estimate the term structure from observed treasury security prices. Later, Vasicek and Fong (1982) proposed the exponential polynomial spline function to estimate the term structure of interest rates. These techniques, although useful in practice, have never been proved to be consistent with the absence of riskless arbitrage in the bond market. The latter feature is indeed the key distinguishing point of term structure models.

Guo (1998) addresses this particular issue. He suggests taking the unspecified Partial Differential Equation (PDE), previously derived by many authors, as a no-arbitrage condition and determining if any linear discount function of the term structure is consistent with no-arbitrage. His concern is whether the PDE has a linear solution. He shows that a linear model does exists, derives it, and shows its uniqueness. His solution will be referred to as the exponential-polynomial (EP) model of term structure.

The purpose of this thesis is to test the empirical performance of the EP model and to study the implications of the empirical findings. We conduct an extensive empirical analysis of this particular model. We test the ability of this model to describe the observed term structure of interest rates by fitting it, first to 1805 daily cross-sections of nearly 388,082 U.S. bonds over the period from 1989 to 1996, and second to 800 daily cross-sections of nearly 60,667 Canadian bond prices over the period from 1992 to 1995. Hence, this study provides a comprehensive empirical test of the EP term structure model using different data sets of traded bonds across a broad maturity spectrum. Since our analysis is extensive, it will allow us to draw important conclusions about the empirical tractability of the EP model and its potential use as a model for term structure.

1.3 Survey of the Literature

Various term structure models for interest rates have been suggested by academicians and practitioners since the early part of this century. There has been many attempts to estimate it with various degrees of success. Durand (1942) was among the first authors to study this subject. He measured the term structure of interest rates by fitting a "smooth" curve to the average yields to maturity of the observed securities. The technique he used for that purpose is based on hand fitting and one can easily understand that any small error which occurs while fitting the curve will be magnified, especially for long maturities, if the yield curve is used to infer forward rates.

This literature survey is structured around two main classes in the term structure literature: theoretical models and curve fitting models. The theoretical literature is generally formulated in terms of general equilibrium or partial equilibrium approach. It is concerned with the determination of stochastic processes that are suitable for the state variables. Moreover, it is interested in the economic identification of the state variables. On the other hand, the curve fitting literature is interested in fitting model parameters to the data in order to determine the shape of the observed term structure. So far, there is a wide gap between the methodologies and predictions of the theoretical and curve fitting models. We will present a review of the best known models in each class as well as a summary of their empirical performance in describing the term structure of interest rates.

1.3.1 The theoretical models

This class of models constitutes the largest part of the development of the subject of term structure. Most of the theoretical contributions in this class can be broadly included in two categories. The first one was initiated by Vasicek (1977) and then thoroughly developed by Cox, Ingersoll, and Ross (1985) (hereafter CIR). It emphasizes the description of the dynamics of the short rate. These two models and many variations of them, are referred to as the single-factor models. Indeed, they only use information on the short-term rate and ignore information from other rates drawn from the yield curve. This category of models has been extended to include many of the term structure models that have been proposed so far. Brown and Schafer (1993) call this category the affine yield class of term structure models. They provide a complete description of this vast class. Later, Duffie and Kan (1996) extended this category to include the multifactor models.

1.3.1.1 The affine class of term-structure models

The class of time homogeneous single factor models "have (these models) the property that the yield curve at any point in time depends only on the state variable, e.g. the short rate, and not on calendar time[1]". Many processes have been suggested for the short term interest rate $r(t)$. A general formulation frequently presented in the literature is the following

$$dr(t) = \kappa(l - r(t))dt + \sigma_r r^{\psi} dz(t), \tag{1.4}$$

where κ is the "rate" of the mean reversion term, l the long term mean towards which the short rate is pulled, σ_r the volatility of the short rate, and $z(t)$ a standard Wiener process. In this framework, Vasicek and CIR models are considered as special cases of this class of short rate model. In fact, both models agree on having the short-term interest rate as the only state variable. However, each model assumes a different process for the short-rate. While Vasicek model is recovered from equation 1.4 by setting ψ equal to 0, CIR assumed that $\psi = 1/2$, which implies a square-root

[1] See Brown and Schafer (1993).

process. CIR showed that the price $P(r,t,\tau)$ of an interest rate contingent claim must satisfy

$$\alpha P_r(r,t,\tau) + P_t(r,t,\tau) + \frac{1}{2}\sigma_r^2 P_{rr}(r,t,\tau) - rP(r,t,\tau) = 0, \qquad (1.5)$$

where α is the "risk-adjusted" drift of the short rate process (*i.e.*, $\kappa(l - r(t))$ plus a market risk premium). $P_r(r,t,\tau)$ is the first partial derivative of $P(r,t,\tau)$ with respect to the variable r, $P_t(r,t,\tau)$ is the first derivative of $P(r,t,\tau)$ with respect to t, and $P_{rr}(r,t,\tau)$ is the second derivative of $P(r,t,\tau)$ with respect to r. Since the zero-coupon bond is considered as a contingent claim which pays 1 at maturity and 0 elsewhere, then its price $P(r,t,\tau)$ is obtained by solving equation 1.5 such that

$$P(r,t,0) = 1. \qquad (1.6)$$

Many authors (including Merton (1973), Vasicek and CIR) showed that the solution for a zero-coupon bond has this particular form

$$P(r,t,\tau) = A(\tau)\exp^{-B(\tau)r(t)}, \qquad (1.7)$$

where $A(\tau)$ and $B(\tau)$ are functions of time-to-maturity, τ [2]. In equation 1.7, the zero-coupon bond price is expressed as an exponential function of the short rate r. From 1.7, the zero-coupon yield, $y(r,t,\tau)$, is derived as

$$y(r,t,\tau) = -1/\tau(\log[P(r,t,\tau)]) = -\log(A(\tau))/\tau + B(\tau)r(t)/\tau. \qquad (1.8)$$

As can be seen from equation 1.8, the zero-coupon yield from this class is affine in the short rate, r; hence, the name affine yield models.

Other models such as Chen and Scott (1992) or Longstaff and Schwartz (1992) are considered as multifactor models from the affine class of term structure models.

Most of the empirical tests conducted on the exponential affine class of models have used various approximations for the short rate process in order to estimate

[2]See Appendix A page 197 for an explicit expression of $A(.)$ and $B(.)$ for the Vasicek and CIR models.

the models. These approximations have led to biased and inconsistent parameter estimates. Thus, in general, the empirical performance of these models has not been entirely satisfactory. Indeed, Brown and Dybvig (1986), and Gibbons and Ramaswamy (1993), two famous empirical studies on the CIR model, concluded that the CIR model has poor parameter stability and produces unreasonable parameter estimates (i.e., negative variances).

1.3.1.2 The no-arbitrage models of term structure

The second category of models contains no-arbitrage models. It was initiated by Ho and Lee (1986) and then extended by Heath, Jarrow, and Morton (1992) (hereafter HJM). The general idea is to take bond prices as inputs and then try to price derivatives based on bond prices. Their mechanism for pricing interest claims is similar to the Black-Scholes stock option pricing model with an arbitrage-free argument as developed by Harrison and Kreps (1979) and Harrison and Pliska (1981). Instead of reasoning from the short-rate, they choose the forward rate process and a measure of its volatility. The process for the instantaneous forward rate is written as

$$dF(t,T) = \alpha(t,T)dt + \beta(t,T)dW_t, \tag{1.9}$$

where α and β are, respectively, the drift and the standard deviation of the forward process, t is the current date, T is the maturity date and W_t is a Wiener process at time t. At a first glance, the HJM model resembles the Vasicek model with the important difference that we are dealing with forward rates rather than spot rates. However, it turns out that using the forward rate is a more general approach to price bonds than using the short interest rate as the only state variable of the term structure. Indeed, the short rate is a particular forward rate

$$F(t,t) = r(t). \tag{1.10}$$

From equation 1.9, the forward rate can also be written as

$$F(t,T) = F(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \beta(s,T)dW_s. \qquad (1.11)$$

HJM used equation 1.9 to derive bond-price processes as Ito processes whose drifts and diffusions are in terms of α and β. They also derived sufficient conditions on α and β for the absence of arbitrage. The instantaneous forward rate F is said to be Gaussian if α and β are deterministic for each t and T. This condition implies that the forward rates are normally distributed as well as the short rate. This result relative to the Gaussian forward rate has been studied by many authors, and specially when dealing with the option pricing of interest sensitive contingent claims (see Jamashidian (1988)). The first comprehensive test conducted on the HJM model is by Kaushik and Morton (1994). They assumed six alternative special cases of the HJM formulation. These special cases are classified into two categories of models: one-parameter models and two-parameter models. They used bond option prices to compute implied volatilities. Their results indicate that the implied volatilities from all six alternative special cases are unstable. Their empirical study suggests that the one-parameter models fit the term structure slightly less well than the two-parameter models. However, the implied parameter values of the one-parameter models are more stable over time than the ones implied from the two-parameter models.

1.3.1.3 Recent developments in term structure models

Besides these two well-known approaches, there has been a growing number of sophisticated approaches which have been developed in more recent years. Among others, the "potential" approach which attempts to model the state-price density of the short rate. In simple words, this approach models the intertemporal marginal rate of substitution (IMRS) or pricing kernel, derived in a representative consumer economy in which the consumer has specific preferences. The potential approach was first proposed by Constantinides (1992) and then by Flesaker and Hughston (1995), Rogers (1997) and Jin and Glasserman (1998). This approach starts by considering a consumer economy in which the representative consumer maximizes expected discounted utility with constant discount factor ρ.

$$E_t(\int_t^{\infty} e^{-\rho s} U(c_s) ds).$$

where $U(.)$ is the consumers's Von Neumann-Morgenstern preference function, and c_t is the consumption rate at time t. In equilibrium, the time-t price of a contingent claim that pays D units at future date s is

$$p(t) = E_t(\frac{Z_s}{Z_t} D) \text{ for } 0 \leq t \leq s \leq \infty.$$

Z_t is the marginal utility of optimal consumption at time t. It is also called the pricing kernel or state-price density. It is expressed as

$$Z_t = e^{-\rho t} \frac{\partial U}{\partial c_t^*}.$$

where c_t^* is the optimal consumption process. The stochastic process followed by the pricing kernel is described by the following equation (see Duffie (1992))

$$dZ_t = Z_t(-r(t)dt - \sum_{j=1}^{m} \phi_j dW_j(t)), \tag{1.12}$$

where $r(t)$ is the short rate and $\phi_j(t)$ is the market price of risk associated with the j-th random factor $W_j(t)$, a standard Brownian motion. Therefore, if the pricing kernel is modeled as

$$dZ_t = \mu_Z dt + \sum_{j=1}^{m} Y_j(t) dW_j(t), \tag{1.13}$$

then, the short rate is $r(t) = -\mu_Z(t)/Z_t$ and the market price of risk $\phi_j(t) = -Y_j(t)/Z_t$.

For instance, Constantinides has modeled positive interest rates through an explicit model of the pricing kernel. He defined the process of Z_t in terms of

$(N+1)$ independent processes, $x_i(t)$, $i = 0, 1, \ldots, N$, and constants g, σ_0^2 and α_i for $i = 0, 1, \ldots, N$, as

$$Z_t = \exp\left\{-(g + \frac{\sigma_0^2}{2}) + x_0(t) + \sum_{i=1}^{N}(x_i(t) - \alpha_i)^2\right\}. \tag{1.14}$$

The process $x_i(t)$, $i = 1 \ldots, N$, is assumed to be a continuous-time AR(1) process defined as the stochastic integral

$$x_i(t) = x_i(0)e^{-\lambda_i t} + \int_0^t e^{\lambda_i(t-s)} dW_i(s), \quad i = 0, \ldots, N, \tag{1.15}$$

where $\lambda_i > 0$. The $(N+1)$ Wiener processes are assumed to be mutually independent. Under some conditions that ensure the positivity of the nominal interest rates, Constantinides derives the expression of a zero-coupon bond price as well as the expression of the short rate. The author claims that his model can (theoretically) accommodate different shapes of the term structures. To our knowledge, however, the empirical performance of the model has not been investigated.

1.3.2 The curve fitting models

Besides these theoretical models, many empirical term structure or curve-fitting models have appeared. McCulloch (1975), Vasicek and Fong (1982) and Coleman, Fisher, and Ibbotson (1992) are a few of them. The empirical models usually start from a pricing function relating bond prices to a discount function and other factors. Then the discount function is approximated by an *ad hoc* functional form. Finally, the variables of the term structure function are estimated through an econometric method. These models offer the advantage of being flexible. However their choice is somewhat arbitrary and thus allow for possibilities of arbitrage. In chapter 3, we will be interested in comparing the empirical performance of some curve-fitting models. Here, we will review some of these curve-fitting models and will point out their main drawbacks.

1.3.2.1 Nelson and Siegel model

This model was first developed to fit only discount bonds. However, Bliss (1997) uses the extended Nelson-Siegel method to fit the discount rate function directly to bond prices (including coupon-bonds). In this model, the forward function can be mathematically described by[3]

$$f(t,\tau) = a_0 + (a_1 + a_2\tau)e^{-a_3\tau}, \tag{1.16}$$

where t is the current time, τ is the time to maturity of a zero-coupon bond, and (a_0, a_1, a_2, a_3) are parameters to be estimated. These parameters represent the current state of the economy. This model has been widely used in practice. However, it is not clear how it can be included in an arbitrage-free framework.

1.3.2.2 The recursive method

This method is known on Wall Street as the "bootstrap". It infers the consecutive forward rates $f(t,\tau)$ iteratively from observed bond prices. It has been clearly formalized by Fama and Bliss (1987). Let us consider one simple example. For the sake of exposition, we assume that there exists three distinct bonds $P_i = 1, 2, 3$. P_1 and P_2 are two zero-coupon bonds with respective time to maturity τ_1 and τ_2 $(\tau_1 < \tau_2)$. P_3 is a coupon bond with a face value of F. It pays an annual coupon, C, at times τ_1, τ_2 and τ_3 respectively $(\tau_1 < \tau_2 < \tau_3)$. The three forward rates $f(t,\tau_i)$ for $i = 1,\dots,3$, are the solution to the following system of equations

$$\begin{cases} P_1 &= e^{-f(t,\tau_1)\tau_1}, \\ P_2 &= e^{-f(t,\tau_2)\tau_2}, \\ P_3 &= Ce^{-f(t,\tau_1)\tau_1} + Ce^{-f(t,\tau_2)\tau_2} + (F+C)e^{-f(t,\tau_3)\tau_3}. \end{cases}$$

First, $f(t,\tau_1)$ is extracted, then $f(t,\tau_2)$ and finally $f(t,\tau_3)$. In this simple example, it is assumed that there is only one unknown in the third equation. However, if

[3] Here x is dropped from the functional notation for the forward rate.

for a given bond price there are several unknown forward rates (this tends to be encountered at the long end of the yield curve where observations are scarce) then they will be expressed in terms of adjacent forward rates using linear interpolation. Then, we solve for the unknown forward rate with an optimization procedure. In general, the Newton-Raphson method is used for that purpose. This method has the disadvantage of depending on a huge number of parameters (i.e., the extracted forward rates).

1.3.2.3 McCulloch model

The objective of McCulloch (1975) was to estimate the discount function $D(\tau)$ from market prices[4]. He suggested a method for fitting a smooth curve, the discount function, by a cubic spline. First, he assumed that the discount function can be written as follows

$$D(\tau) = 1 + \sum_{k=1}^{K} \beta_k f_k(\tau), \tag{1.17}$$

where $f_k(\tau)$ are functions specified such that

$$f_k(0) = 0.$$

The β_k are unknown parameters to be estimated by linear regression. Equation 1.17 means that when the maturity of each bond is evenly divided into K intervals, the discount function $D(\tau)$ is approximated by a distinct cubic polynomial function $f_k(.)$ over each interval. Usually, K is set to be equal to the nearest integer nearest to $\sqrt{N}$, where N is the number of bonds in the sample. The intervals are joined at knots (or break points) in such a manner that the spline's first and second derivatives are set equal at these points.

[4]For simplicity, t is dropped from the notation of $D(t, \tau)$.

According to this method, the price of a bond j with maturity τ and which pays a discrete constant coupon C_j and has a face value of F, can be expressed as

$$P_j(t,\tau) = \sum_{i=1}^{M} C_j(\mu_i)D(\mu_i) + FD(\tau) + \epsilon_j, \tag{1.18}$$

where ϵ_j is an error term. The symbol μ_i is the ith coupon payment date expressed in terms of fraction of a year. With this notation, we have $\mu_M = \tau$ at maturity. If we replace the discount function D in equation 1.18 by its representation in equation(1.17), we obtain

$$P_j(t,\tau) = X_0 + \beta_1(X_1 + L_1) + \ldots + \beta_K(X_K + L_K) + \epsilon_j, \tag{1.19}$$

where:

- $X_0 = F + \sum_{i=1}^{M} C_j(\mu_i)$.
- $X_k = \sum_{i=1}^{M} f_k(\mu_i)C_j(\mu_i),\ k = 1\ldots K$.
- $L_k = f_k(\tau)F,\ k = 1\ldots K$.

In his paper, McCulloch replaced $P_j(t,\tau)$ by the average of the bid and ask prices. Then, he estimated the β's by the following minimization procedure:

$$\min_{\beta_k} \sum_{j=1}^{N} (w_j\epsilon_j^2) \qquad k = 1\ldots, K.$$

The w_j are weights computed as follows

$$w_j = \frac{2}{(P_j^a - P_j^b)},$$

where P_j^a and P_j^b are ask and bid prices of bond j, respectively. McCulloch chose to use the weighted least square method in order to prevent the estimates from being affected by large errors that are solely caused by transaction costs.

Unfortunately, the forward curves produced by the models of McCulloch and Vasicek and Fong tend to oscillate and reach negative values. To solve this problem, Shea (1985) suggested to use constraints on the splines and varying the number of break points. Fisher, Nychka, and Zervos (1995) implemented some of these suggestions.

1.3.2.4 Fisher et al. cubic spline model

Fisher, Nychka, and Zervos (1995) use spline functions to estimate the term structure. However, instead of using regression splines as McCulloch did, they introduced smoothing splines. The advantage of smoothing splines is that the number and location of knots is chosen optimally rather than predetermined by the user. Moreover, Fisher et. al choose to place the spline on the forward function instead of the discount function. Their methodology minimizes what they call a criterion function specified as follows

$$\sum_{i=1}^{N} \epsilon_i^2 + \lambda \int_0^T D(\tau)'' d\tau, \tag{1.20}$$

where N is the number of bonds in the sample, $D(\tau)$ the discount function, $D(\tau)''$ the second derivative of $D(.)$, and λ a weight parameter. The second term in equation 1.20 is added in order to penalize the roughness of the approximating discount function $D(\tau)$.

All these previously mentioned methods are based on curve fitting procedures. They are designed to match bond price observations. It is not clear how they can be placed into a no-arbitrage framework.

1.3.3 Empirical studies of term structure

Empirical studies of term structure can be classified in three distinct categories:

- cross-sectional studies,
- time series studies,

- density function estimation studies.

1.3.3.1 Cross-sectional Studies

Brown and Dybvig (1986) are considered to be the pioneers of this first category. They tested the one factor CIR model with U.S coupon-bearing bonds. They fitted the yield curve by pooling data and kept the structural parameters constant for several days. Under the one factor CIR model, the price of a zero-coupon bond can be written as

$$P(r,t,\tau) = A(\tau)\exp^{-B(\tau)r(t)}, \tag{1.21}$$

where $A(\tau) = f_1(\kappa, \lambda, \sigma)$ and $B(\tau) = f_2(\kappa, \lambda, \sigma, \theta)$, with f_1 and f_2 being two specific functions (see the details of CIR model in Appendix A page 197). Brown and Dybvig assumed that the market price of a zero coupon bond, P_i, is the sum of a theoretical price (here the CIR model) plus an error term. Thus,

$$P_i = P_i^{CIR}(r, \kappa, \lambda, \theta, \sigma) + \epsilon_i.$$

To conduct their cross-sectional study, they chose to minimize the sum of squared errors defined as

$$S_t^2 = \sum_{i=1}^{n}(P_i - P_i^{CIR})^2 \text{ for } i = 1, \ldots, n, \tag{1.22}$$

over the structural parameters κ, λ, σ, θ and the risk free rate r for n different traded bonds at time t.

The Nonlinear Least Square (NLS) method that they used allows the estimation of $(\kappa + \lambda)$, σ and the long rate l (defined in Appendix A), but the mean reversion parameter κ, the risk premium parameter λ and the unconditional mean of the spot rate θ are *not separately* identified. Since the short rate is also one of the estimated parameters, Brown and Dybvig compared its successive values to the time series of yields on two-week US bills. A systematic deviation was noted between the two series. Moreover, they found that the CIR model fits generally the T-bills better than longer Treasury issues.

Brown and Schaeffer, Barone, Cuoco, and Zautzik (1991), de Munnik and Schotman (1994) used a similar econometric approach but applied respectively to British, Italian and Dutch data. Like others, they found in their empirical study of the CIR model that it exhibits parameter instability. The results point to the need to include more than one state variable in these models to be able to fully describe the term structure.

More recently, Jordan and Kuipers (1997) presented an empirical study of the CIR model as well as the Vasicek and Merton models. Their study is in the spirit of Brown and Dybvig. However, it presents the advantage of using daily data over a horizon of five years. The data consists of prices STRIPS for bonds (a sort of synthetically created zero coupon bonds of longer maturities than a year) and thus avoid complications arising from estimating coupon bonds. Once again, the CIR model was found to exhibit high parameter instability.

Chacko (1998) suggests the use of a technique based on Fourier transforms to estimate affine term structure models (e.g., CIR model and Vasicek models). This technique relies on maximum likelihood estimation. The author claims that it can *separately* estimate the risk premium parameter (that we referred to as λ in the CIR model) while simultaneously estimating other parameters. Until now, we did not encounter an empirical study that implemented this technique.

In general, the common criticism addressed at this category is that the models are unable to separate out the interest rate risk premium from the individual parameters of the interest rate process. Moreover, this approach does not constrain the interest rate parameters to be stable over time. Most of the empirical evidence indicates that the parameters estimated with this technique are quite unstable.

1.3.3.2 Density function estimation

This approach covers studies that usually infer information about the term structure from a specific likelihood function.

Gibbons and Ramaswamy (1993) conducted an empirical test on the CIR model using the Generalized Method of Moments (GMM). This method has the advantage of avoiding assumptions relative to the stochastic process of the aggregate price. Their result concluded that parameter estimates from CIR model preclude a humped shape for the term structure. Moreover, the estimated autocorrelation coefficient of the short rate implied from CIR model was too small compared to the corresponding sample autocorrelation coefficient computed from US T-bills prices.

Chen and Scott (1992) extended one-factor CIR model by including additional factors which follow square-root processes. For the estimation, they used the time series of four distinct bond maturities. In order to derive the likelihood function of the unobservable state variables, they expressed the underlying state variables in terms of the observed bond prices and some errors terms. Their findings were in favor of two or three state variables to fully characterize the movements of the term structure.

Pearson and Sun (1994) used the conditional density of the state variables to estimate and test a two-factor extension of CIR model. Their idea is to infer the conditional density of the unobservable state variables from the bond pricing formula. For this, they use N observations of two different bonds available at different points in time. Their results showed that estimates, based on bills only, imply unreasonable large price errors for longer maturities.

The density function estimation methodology is theoretically attractive since it includes all of the relevant information about the stochastic process and allows separate identification of all model parameters. Unfortunately, from an empirical point of view, it suffers from some drawbacks. For example, the difficulty with implementing the test procedure requires limited sample sizes or simple treasury securities such as short-term Treasury bills.

1.3.3.3 Time-series approach

Time-series studies are not as computationally burdensome as cross-sectional studies. They start by assuming a stochastic process for the short rate. Then, using some optimization technique, the parameters of such a process are estimated. Usually, this approach is used to validate whether interest rates follow the hypothesized process. However, it provides no way to estimate the risk premium parameter, λ. Moreover, it ignores information available from bond prices and the estimated parameters may imply theoretical bond prices totally different from their observable counterparts. One famous study in this category is that of Chan, Karolyi, Longstaff, and Sanders (1992). They compared eight competing models of short-term interest rate dynamics. All the models were nested within a framework that allows comparison among the models. Their study was interesting in highlighting the most important features of the short rate process. However, their study is subject to many criticisms especially with respect to its use of the short rate. In fact, the short rate is an unobservable variable and any empirical research on the time-series properties of the short rate typically requires a proxy. Thus, choosing the one-month Treasury bill yield or the weekly Eurodollar rate as proxies for the short-rate is always a subjective choice that has an impact on the term structure models, see for instance Chapman, Long, and Pearson (1999).

Table 1.1 presents a summary of some of the empirical studies of term structure that are mentioned above.

In general, it seems that the theoretical development of term structure models has followed a rapid path whereas their empirical testing and practical implementation have remained far behind. The curve-fitting models, although simple and empirically satisfactory, lack an underlying theory. In the next chapter, we will present a term structure model proposed by Guo (1998) that is conceived in the curve fitting spirit but that has the significant advantage of being a linear solution to the Partial Differential Equation of bond pricing and, hence, has a strong theoretical basis.

1.4 Organization of the thesis

The rest of this thesis is organized as follows. Chapter 2 presents a derivation of the EP model developed by Guo. Chapter 3 describes the cross-sectional estimation of the EP model using U.S bonds data. Chapter 4 reports the main results obtained for the cross-sectional study using Canadian daily data set. Chapter 5 examines the time series properties of the principal components of the state factors of the EP model. Chapter 6 includes the out-of-sample results from the estimated principal components. It also examines some arbitrage strategies that employ forecasts of the three principal components. Finally, Chapter 7 summarizes the results and contains concluding remarks.

Authors	**Data**	**Model(s)**	**Method**	**Conclusion**
Brown and Dybvig (1986)	U.S. Treasuries monthly data	CIR	NLS	model rejected parameter instability
Brown and Schaefer (1994)	British monthly data	CIR	NLS	model rejected parameter instability
Cuoco et. al. (1991)	Italian monthly data	CIR	NLS	model rejected parameter instability
deMunnick and Shotman (1994)	Dutch monthly data	CIR	NLS	model rejected
Gibbons and Ramaswamy (1993)	U.S. T-bills monthly data	CIR	GMM	model partially rejected unrealistic parameter estimates
Chen and Scott (1992)	U.S. T-bills monthly data	multifactor CIR	MLE	more than two factors needed for term structure
Pearson and Sun (1994)	U.S. T-bills monthly data	two-factor CIR	MLE	model rejected unrealistic price errors
Jordan and Kuipers (1997)	U.S. STRIPS daily data	CIR, Vasicek and Merton	NLS and iterative GMM	poor out-of- sample performance

Table 1.1: Summary of previous empirical studies of term structure models.

Chapter 2

A Linear Model for Term Structure

2.1 Introduction

The concept of term structure is usually expressed in terms of three functions that are interrelated: the discount function, the discount rate function (zero-coupon yield curve) and the forward function (forward-rate curve). The discount function, which relates the zero-coupon bond prices to different maturities, has been an important measure of the term structure of interest rates. Through the years, discount functions have been estimated mostly with *ad hoc* smoothing techniques. The usual practice among authors is to select an approximating function for the discount function and then estimate the parameters of this function. Spline models were originally brought to term structure estimation by McCulloch. Vasicek and Fong introduced the exponential spline model. However, the emphasis was directed to the fitting performance of the empirical models of term structure with little attention granted to their consistency with the absence of riskless arbitrage in the bond market. In his paper, Guo (1998) has raised this particular issue. He derived an empirical linear model of the term structure consistent with the Arbitrage Pricing Theory (hereafter the APT) of Ross (1976). First, I will present Guo's derivation of this model as a solution of the fundamental Partial Differential Equation (PDE). Second, following Guo's arguments, I will show the uniqueness of this solution. Finally, I will highlight the relationship between this solution and a special version of the HJM model.

2.2 The Exponential Polynomial (EP) Model

2.2.1 Notation

Here is the notation that is used in this chapter.

- $D(x, T, t)$ is a zero-coupon bond price at time t. It pays one unit at the maturity calendar date T.
- $x(t) = [x_1(t) \dots x_n(t)]$ is a state vector.
- $x_i(t)$ is the ith state factor.

Following McCulloch, the discount function has been modeled as a time-homogeneous state factor model. It has the following general linear form:

$$\begin{aligned} D(x,T,t) &= h_0(T) + \sum_{i=1}^{n} x_i(t) g_i(T), \\ D(x,T,T) &= 1, \end{aligned} \tag{2.1}$$

where:

- $h_0(T)$ is an arbitrary function of T, with $h_0(0) = 1$.
- $g_i(T)$ is a basis function designed to satisfy $g_i(0) = 0$ for $i = 1 \dots n$.

This system of equations expresses the discount function as a linear combination of future cash flows discounted by some functions $g_i(t)$, with the condition that the discount function reaches the value 1 at maturity. Before presenting the exponential polynomial model, I will briefly review the derivation of the fundamental PDE of bond pricing.

2.2.2 The fundamental PDE

The fundamental PDE for bond pricing models has been derived by many authors, e.g. Langetieg (1981) and Cox, Ingersoll, and Ross (1985). The PDE has been a starting point for many equilibrium term structure models, including the multi-factor model of Langetieg and the single factor models of Vasicek and CIR. The differences between all these models stem from further assumptions made about the identity of the state factors $x_i(t)$, their stochastic processes and the market price of risk. Guo has insisted on the use of the fundamental PDE as a way "...to check if a given model is consistent with no-arbitrage, regardless of how the model is derived." Under the assumption of market perfection, Langetieg assumed that the state vector follows a joint Ito process

$$dx_i = \mu_i(x(t),t)dt + \sigma_i(x(t),t)dz_i, \qquad i = 1 \ldots n, \tag{2.2}$$

where:

- $\mu_i(x(t),t)$ is the drift of the process.
- $\sigma_i(x(t),t)$ is the diffusion coefficient of the process.
- z_i is a standard Wiener process.

By applying Ito's formula, the instantaneous change of the bond price is written

$$dD(x,T,t) = (\frac{\partial D}{\partial t} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\sigma_i\sigma_j\frac{\partial^2 D}{\partial x_i \partial x_j} + \sum_{i=1}^{n}\mu_i\frac{\partial D}{\partial x_i})dt + \sum_{i=1}^{n}\sigma_i\frac{\partial D}{\partial x_i}dz_i, \tag{2.3}$$

where ρ_{ij} is the correlation coefficient between x_i and x_j. Equation 2.3 can be simplified to

$$dD(x,T,t) = \bar{D}_t dt + \sum_{i=1}^{n} \sigma_i \frac{\partial D}{\partial x_i} dz_i, \tag{2.4}$$

with the drift of the instantaneous change of the bond price expressed as

$$\bar{D}_t = \frac{\partial D}{\partial t} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}\sigma_i\sigma_j \frac{\partial^2 D}{\partial x_i \partial x_j} + \sum_{i=1}^{n} \mu_i \frac{\partial D}{\partial x_i}.$$

Consider $(n+2)$ zero-coupon bonds, equation 2.4 can be expressed in vector form as

$$d\mathbf{D} = \bar{\mathbf{D}}_{\mathbf{t}} dt + \sum_{i=1}^{n} \sigma_i \mathbf{D}_{x_i} dz_i, \tag{2.5}$$

where:

- $\mathbf{D}$ is a price vector of $(n+2)$ zero-coupon bonds of different maturities.
- $\bar{\mathbf{D}}$ is the vector drift.
- $\mathbf{D}_{\mathbf{x}_i}$ is the partial derivative vector of $\mathbf{D}$ with respect to x_i.

At this stage, the fundamental PDE for bond pricing can be derived invoking the APT theory. Consider a portfolio of $(n+2)$ zero-coupon bonds with a weight vector w_t. Assuming unrestricted short sales, the vector of weight can be chosen such that the portfolio requires zero investment

$$w_t'\mathbf{D} = \mathbf{0},$$

and bears zero risk

$$w_t'\mathbf{D}_{\mathbf{x}_i} = \mathbf{0}.$$

In perfect markets (i.e., no commissions, taxes,...) and in equilibrium, Ross argues that such a portfolio must earn a zero rate of return. Thus,

$$w_t' d\mathbf{D} = w_t'\bar{\mathbf{D}}_{\mathbf{t}} = 0.$$

The above equality implies that the $(n+2)$ vectors $\mathbf{D}$, $\bar{\mathbf{D}}_\mathbf{t}$ and $\mathbf{D}_{\mathbf{x_i}}$, must be linearly dependent (see Langetieg and CIR). In turn, this linear dependence implies that there exists at most $(n+1)$ independent scalars such that

$$\bar{\mathbf{D}}_\mathbf{t} = \phi_0(x(t),t)\mathbf{D} + \sum_{i=1}^{n} \phi_i(x(t),t)\sigma_i(x(t),t)\mathbf{D}_{\mathbf{x_i}}, \tag{2.6}$$

where $\phi_0(x(t),t)$ can be interpreted as the instantaneous interest rate (short rate) and $\phi_i(x(t),t), i = 1,\ldots,n$, a market price of risk related to the state factor x_i. Thus, equation 2.6 implies that

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 D}{\partial x_i \partial x_j}\sigma_{ij}(t,x(t)) \ + \\ &\sum_{i=1}^{n}\frac{\partial D}{\partial x_i}(\mu_i(x(t),t) - \phi_i(x(t),t)\sigma_i(x(t),t)) + \frac{\partial D}{\partial t} - \phi_0(x(t),t)D \ = \ 0, \end{aligned} \tag{2.7}$$

where σ_{ij} is the covariance between x_i and x_j. Equation 2.7 is the fundamental PDE for bond pricing. It has the following boundary condition at maturity

$$D(x,T,T) = 1.$$

To further explain equation 2.7, let's denote

$$\eta_i(x(t),t) = \mu_i(x(t),t) - \phi_i(x(t),t)\sigma_i(x(t),t). \tag{2.8}$$

In equation 2.8, the market price of risk for state factor ϕ_i, is replaced by an equivalent expression. If we substitute the expression for η_i in equation 2.7, the PDE becomes

$$\phi_0(x(t),t)D = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 D}{\partial x_i \partial x_j}\sigma_{ij}(x(t),t) + \sum_{i=1}^{n}\frac{\partial D}{\partial x_i}\eta_i(x(t),t) + \frac{\partial D}{\partial t}. \tag{2.9}$$

Having reviewed the derivation of the PDE, we now will present the exponential polynomial model of Guo as a unique solution to the PDE 2.9.

2.2.3 The exponential polynomial solution and its uniqueness

Here, the purpose is to derive expressions for both $h_0(T)$ and $g_i(T)$ of equation 2.1, that are consistent with the absence of riskless arbitrage. For this, the PDE will be the main equation. Taking the first order partial derivative of equation 2.1 with respect to the state factors x_i, we get

$$\frac{\partial D}{\partial x_i} = g_i(T).$$

From the linearity of 2.1 in x_i, the second order partial derivative of D with respect to the state factors vanishes

$$\frac{\partial D^2}{\partial x_i \partial x_j} = 0.$$

The partial derivative of D with respect to t is

$$\frac{\partial D}{\partial t} = -\frac{\partial D}{\partial T} = -h_0'(T) - \sum_{i=1}^{n} x_i g_i'(T),$$

where the apostrophe denotes the derivative with respect to T. Substituting these partial derivatives in equation 2.9 and re-arranging terms, we obtain

$$\sum_{i=1}^{n}\{[\eta_i(x(t),t) - \phi_0(x(t),t)x_i]g_i(T) - x_i g_i'(T)\} - \{\phi_0(x(t),t)h_0(T) + h_0'(T)\} = 0. \tag{2.10}$$

At the beginning of this chapter, we claimed that the discount function has been modeled as a time-homogeneous state factor function. Thus, equation 2.10 can be simplified to

$$\sum_{i=1}^{n}\{[\eta_i(x) - \phi_0(x)x_i]g_i(T) - x_i g_i'(T)\} - \{\phi_0(x)h_0(T) + h_0'(T)\} = 0. \tag{2.11}$$

Assume that the last term in braces in equation 2.11 is equal to zero, i.e.,

$$\{\phi_0(x)h_0(T) + h_0'(T)\} = 0,$$

the solution $h_0(T) = e^{-\phi_0(x)T}$ would be in contradiction with the linearity of the discount function 2.1. In fact the exponent term ϕ_0 which represents the short rate is not assumed constant. Thus in order to be independent of the state factors x_i, the basis function must be written as

$$g_i(T) = h_i(T) - h_0(T). \qquad i = 1, \dots, n, \tag{2.12}$$

where $h_i(T)$ is an arbitrary function with $h_i(0) = 0$. If we substitute expression 2.12 into equation 2.11, we get

$$\begin{aligned}
&\sum_{i=1}^{n}\{[\eta_i(x) - \phi_0(x)x_i]h_i(T) - x_i h_i'(T)\} \; - \\
&\{h_0'(T)[1 - \sum_{i=1}^{n} x_i] + h_0(T)[\phi_0(x) + \sum_{i=1}^{n}[\eta_i(x) - \phi_0(x)x_i]]\} \;=\; 0.
\end{aligned} \tag{2.13}$$

This equation is very important. It will allow us to derive the expressions for $h_0(T)$ and for $h_i(T)$. The necessary and sufficient condition that equation 2.13 holds for any x is that both terms in braces equal zero. First, let the first sum be equal to zero, then

$$[\eta_i(x) - \phi_0(x)x_i]h_i(T) - x_i h_i'(T) = 0. \tag{2.14}$$

Introducing the natural logarithm of $h_i(T)$, we obtain

$$d \ln h_i(T)/dT = [\eta_i(x) - \phi_0(x)x_i]/x_i. \tag{2.15}$$

Equation 2.15 is a differential equation. It can be solved subject to ($h_i(0) = 1$). The unique solution to this equation is

$$h_i(T) = e^{-\lambda_i(x)T}, \tag{2.16}$$

with the expression of λ_i written as

$$\lambda_i(x) = -[\eta_i(x) - \phi_0(x)x_i]/x_i. \tag{2.17}$$

The discount function expressed in equation 2.1 is linear in x. Thus, $\lambda_i(x)$ is required to be a positive constant that is $\lambda_i(x) = \lambda_i$ for all x. From equation 2.17 the first-order coefficient $\eta_i(x)$ is the unique solution to the following equation

$$\eta_i(x) = [\phi_0(x) - \lambda_i]x_i. \tag{2.18}$$

Following similar reasoning as before, the expression for $h_0(T)$ can be derived from the last braced term of equation 2.13 as

$$h_0(T) = e^{-\lambda_0 T}, \tag{2.19}$$

with the expression of λ_0 written as

$$\lambda_0 = \frac{\phi_0 + \sum_{i=1}^{n}[\eta_i(x) - \phi_0(x)x_i]}{1 - \sum_{i=1}^{n} x_i}. \tag{2.20}$$

λ_0 is required to be a positive constant. Simply substituting 2.18 into 2.20 the expression for the coefficient ϕ_0 can be uniquely determined as

$$\phi_0 = \lambda_0 + \sum_{i=1}^{n}(\lambda_i - \lambda_0)x_i. \tag{2.21}$$

Having the exact expressions of $h_0(T)$ and $h_i(T)$ respectively, the linear discount function can be expressed as follows

$$D(x, s, t) = (1 - \sum_{i=1}^{n} x_i)e^{-\lambda_0(s-t)} + \sum_{i=1}^{n} x_i e^{-\lambda_i(s-t)}, \tag{2.22}$$

where s is a variable time to maturity. Equation 2.22 is the unique linear solution to the fundamental PDE derived by Guo. Since, the instantaneous (short) interest rate is, by definition, nothing more than a limiting value of the zero-coupon discount function, then from equation 2.22, the expression for the short rate $r(x(t), t)$ can be derived as

$$r(x(t), t) \equiv \frac{\partial D(x, s, t)}{\partial t}|_{s=t} = \lambda_0 + \sum_{i=1}^{n}(\lambda_i - \lambda_0)x_i. \tag{2.23}$$

It is clear from equations 2.21 and 2.23 that ϕ_0 is indeed the instantaneous interest rate. It appears that the short rate is a linear combination of n unobservable state factors. It is well known that the inflation rate or the real interest rate are some factors influencing the short rate. However, at this stage of work, we are unable to attribute any economic meaning to the state factors. Moreover, the exact number of these factors is not a theoretical matter, it must be empirically specified. In fact, the model offers a wide latitude in defining the term structure problem. Vasicek defined the term structure in terms of a single factor, the short rate. Brennan and Schwartz (1979) suggested a two-factor specification, the short rate and the long rate. In general, a multifactor specification is attractive in the sense it can match various shapes of term structure over time. Unfortunately, a multifactor specification is usually difficult to implement. In the empirical part of this thesis, we will discuss these issues further.

Equation 2.22 represents a linear discount function. Guo referred to it as the exponential polynomial (EP) model of term structure. Indeed, it describes the discount function as a linear combination of state factors discounted at different exponential rates. The EP model is different from spline functions. It is an unconditionally arbitrage-free model, derived as the unique linear solution to the fundamental PDE of bond pricing. The EP model is defined on the entire maturity range of term structure ($s \in [0, \infty)$). Unlike spline functions, it does not need to be defined on some subintervals. Moreover, as we will see in the following chapters, the EP model can easily be used for the empirical estimation of the term structure.

2.2.4 The exponential basis

In its expression, the EP model is related to the exponential spline model of Vasicek and Fong. However, the Vasicek and Fong model is defined on two consecutive knots. It has the following expression

$$D(s,t) = a_0 + a_1 e^{-R(s-t)} + a_2 e^{-2R(s-t)} + a_3 e^{-3R(s-t)}, \quad s \in [s_i, s_{i+1}], \qquad (2.24)$$

where:

- R denotes the asymptotic forward rate (long rate).
- $i = 1, \ldots n$. n is the number of knots.
- a_0, a_1, a_2, a_3 are linear coefficients.

If the subinterval is extended to the entire maturity and the linear coefficients are modified to satisfy the boundary condition ($D(T,t) = 1$), equation 2.24 would be equivalent to the EP model. In the EP model, Guo specifies the exponential parameters in a more flexible way than in equation 2.24:

$$\lambda_0 = R, \quad \lambda_i = l_i + R, \qquad i = 1, \ldots, n, \qquad (2.25)$$

where $l_i, i = 1, \ldots, n$ are positive constants. The choice of l_i is motivated by some empirical issues that we will discuss in the next two chapters. In general, we choose a decreasing, or increasing, suite for l_i, and this is to avoid multicollinearity between successive state factors. As the term to maturity goes to infinity, the discount function tends to

$$D(x,s,t) \to (1 - \sum_{i=1}^{n} x_i) e^{-R(s-t)} \text{ as } s \to \infty. \qquad (2.26)$$

Thus, the state factor R has an economic interpretation as the asymptotic forward rate or the long rate. In some empirical studies R was fixed as the yield of a long-maturity discount bond, see Brennan and Schwartz (1979). From expression 2.26, it is clear that the state factors must satisfy the following constraint

$$\sum_{i=1}^{n} x_i < 1. \qquad (2.27)$$

This constraint ensures the discount function to be strictly positive. Before concluding this chapter, we would like to show the relationship between the EP model and an HJM specification.

2.3 Relation between the EP model and an HJM specification

HJM defined a family of forward rate processes as follows

$$f(t,T)-f(0,T)=\int_0^t \alpha(v,T)dv+\sum_{i=1}^{n}\int_0^t \sigma_i(v,T)dW_i(v) \quad \text{for all } 0\le t\le T, \quad (2.28)$$

where $f(t,T)$ is the forward rate at time t, with maturity time T, α and σ are respectively the drift and the volatility of the forward process, and the W_i are independent standard Brownian motions.

HJM treat the variable T as the calendar maturity date. Instead of using this notation, we will follow the term structure parametrization proposed by Musiela (1994), Brace and Musiela (1994) and Guo (1997). We denote by $f_m(t,\tau)$ the forward rate at time t with a relative maturity τ defined as $\tau = T - t$. There is an obvious relationship between the HJM forward rate $f(t,T)$ and notation $f_m(t,\tau)$. Indeed:

$$f(t,T)=f(t,\tau+t)=f_m(t,\tau).$$

Under this new notation, the HJM model 2.28 can be expressed as

$$f_m(t,\tau)-f_m(0,\tau)=\int_0^t \alpha(v,\tau)dv+\sum_{i=1}^{n}\int_0^t \sigma_i(v,\tau)dW_i(v) \quad \text{for all } \tau\ge 0, \quad (2.29)$$

which simply involves replacing T by τ. In what follows, we will refer to this model as the HJM model re-indexed to relative maturity. As specified in equation 2.29, the HJM model is path-dependent. In other words, it permits the future term structure to depend on the entire path of prices since the term structure is initialized. For instance. the interest rate at date t_1 may not only depend on its level at date t_0 (with $t_1 > t_0$) but on previous levels along its path going all the way back to the initial date. Moreover, the interest rate at date t may not be sufficient for the determination of all the forward rates at date t. Thus, the knowledge of many points on the term structure at date t are not always sufficient for the identification of other forward rates at that time. This is called the path-dependency property. This property together with the fact that there may not be a finite number of state variables, creates pricing difficulties. In particular lattice procedures may not recombine and can become exploding.

In order to avoid the problem posed by path-dependency and exploding lattices, HJM proposed a two-factor version of their general model. This special case has a constant volatility long run factor $W_0(t)$ and a spread factor $W_i(t)$ with an exponentially decaying volatility function. For convenience of notation, the subscripts of the Brownian motions in equation 2.29 are replaced by

$$W_0(v), W_1(v), \ldots, W_k(v).$$

The volatility functions are defined as

$$\begin{cases} \sigma_0(v,\tau) &= \sigma_0, \\ \sigma_i(v,\tau) &= \sigma_i e^{-\gamma_i \tau} \quad \text{for } i = 1, \ldots, k, \end{cases}$$

where γ_i is a parameter. The stochastic integrals in 2.29 can be solved explicitly, because they are no longer expressed with respect to the relative maturity τ. Thus, equation 2.29 can be simplified to[1]

$$f_m(t,\tau) - f_m(0,\tau) = \int_0^t \alpha(v,\tau)dv + \sigma_0 W_0(t) + \sum_{i=1}^{k} \sigma_i e^{-\gamma_i \tau} W_i(t). \qquad (2.30)$$

The possibility that the drift function of the re-indexed HJM model vanishes is in sharp contrast to the original drift function in HJM. In the empirical section, we

[1] The Brownian motions $W_i(t)$ are initialized at zero by HJM.

will show that the re-indexed HJM model can be tested by estimating a proxy of the EP model and thus there is no need to stipulate the drift function.

In what remains we will show the close relation between the EP model and the re-indexed HJM model. Under the EP framework, the yield function can be found from the following transformation

$$y(x,s,t) = -\frac{\ln D(x,s,t)}{s-t} \quad s > t, \tag{2.31}$$

where $y(.)$ is the yield function. The time-homogeneity of the discount function allows the corresponding forward function to be obtained from one of the following transformations:

$$f(x,s,t) = -\frac{\partial \ln D}{\partial s} = -\frac{\partial D/\partial \tau}{D}, \text{ for } \tau = s - t. \tag{2.32}$$

From equation 2.22, the EP model can be re-written as:

$$D(x,s,t) = e^{-R(s-t)}[1 + \sum_{i=1}^{n} x_i(e^{-l_i(s-t)} - 1)]. \tag{2.33}$$

Or equivalently:

$$D(x,\tau) = e^{-R\tau}[1 + \sum_{i=1}^{n} x_i(e^{-l_i\tau} - 1)]. \tag{2.34}$$

The EP forward function can be obtained using the transformation in equation 2.32

$$f(x,\tau) = R - \frac{\partial \ln(1 + \sum_{i=1}^{n} x_i(e^{-l_i\tau} - 1))}{\partial \tau}. \tag{2.35}$$

An interesting question can be asked at this stage. Does there exist variables ξ_i with $i = 1, \ldots, n$, such that

$$1 + \sum_{i=1}^{n} x_i(e^{-l_i\tau} - 1) \simeq \exp[\sum_{i=1}^{n} \xi_i(e^{-l_i\tau} - 1)]?$$

Or, equivalently, such that

$$\ln[1 + \sum_{i=1}^{n} x_i(e^{-l_i\tau} - 1)] \simeq \sum_{i=1}^{n} \xi_i(e^{-l_i\tau} - 1)? \tag{2.36}$$

If yes, then the EP forward function can be approximated by a proxy forward function denoted by

$$\hat{f}(\varphi, \tau) = R + \sum_{i=1}^{n} \varphi_i e^{-l_i \tau}, \tag{2.37}$$

where $\varphi_i = \xi_i l_i$, for $i = 1, \ldots, n$. Hence the empirical existence of φ_i would imply the existence of ξ_i. Now, if we assume that[2]

$$\begin{cases} R &= \sigma_0 W_0(t), \\ \varphi_i &= \sigma_i W_i(t), \text{for } i = 1, \ldots, n \end{cases}$$

then the forward function as defined from the EP model can be approximated–subject to the existence of φ_i–by a proxy forward rate expressed as

$$\hat{f}(\tau) = \sigma_0 W_0(t) + \sum_{i=1}^{n} \sigma_i e^{-l_i \tau} W_i(t), \tag{2.38}$$

which is no more than a simple form of the re-indexed HJM model in 2.30 with a slightly different notation (i.e., $\gamma_i = l_i$ and $k = n$) and without a drift, (i.e., with $\alpha(v, \tau) = 0$).

[2]With this notation R and φ_i are seen as changing with respect to time t.

Chapter 3

Empirical Performance of the EP Model: Results for U.S. Government Bonds

This chapter presents the use of the EP model in an extensive cross-sectional investigation involving data for U.S. government bonds.

This investigation answers the following general question:

- Can the proposed EP model presented in Chapter 2, fit the observed term structures?

By answering this question, we will be able to know whether the state factors of the EP model can be measured with constant exponential basis. Indeed, the EP model describes the term structure space as a linear combination of some state factors on a basis of exponential functions. These state factors are measured relative to the basis. Thus, the validity of the EP solution will be related to the question of constancy of the parameters of the basis over time.

3.1 Notation

Here, we present the notation that we will use in this chapter.

- $P_j(s,t)$ is the price of a semiannual coupon-bond j at time t, with s as the calendar time to maturity. This bond has a face value of 100.
- C_j is the jth bond future cash flow (the annual coupon) paid at the calendar date m_q.
- $q = 1, \ldots, s$. Note that $m_s = s$ at maturity .
- $D(s,t)$ is the discount function. It relates the spot rates at different maturities. It is also the price of a discount bond, which pays 1 unit at s and 0 at other times.

In the absence of arbitrage opportunities, the value of the jth coupon bond at time t is supposed to equal the sum of the present values of all its future cash flows

$$P_j(s,t) = C_j \sum_{q=1}^{s} D(m_q, t) + 100D(s,t). \tag{3.1}$$

However, when it comes to fit a set of market price data to a pricing relation, it becomes necessary to include an error term to account for the difference between actual and theoretical prices. Thus, the bond pricing relation can be expressed as

$$P_j(s,t) = C_j \sum_{q=1}^{s} D(m_q, t) + 100D(s,t) + \epsilon_j(s,t), \tag{3.2}$$

where ϵ_j(s,t) is an error term. Next, we discuss equation 3.2 and point out its mains assumptions.

3.1.1 Assumption about the pricing relation and the errors

The error term represents the statistical error due to model approximation but also accounts for other factors such as tax effects, transaction costs and measurement and mispricing errors. Predictable $\epsilon_j(s,t)$'s suggest that there is additional information that could be included in equation 3.2. Equation 3.2 makes two important assumptions.

1. **The pricing relation** assumes that the price of a treasury security is solely the sum of all the promised individual cash flows. Unfortunately, the frictionless markets assumption is not supported by the facts. Attempts to fit discount functions to sets of government bond prices find that no discount function, D, exists to exactly price all bonds, even when bid-ask spreads are taken into account. However, many studies have showed that the existence of disparities are real and not quotation errors. The term structure estimation literature has sought to explain these disparities in terms of friction such as liquidity premia, taxes or short sale constraints. Indeed, these phenomena are studied in Daves and Ehrhardt (1992). Other examples involving disparities between short-term notes and bills have been studied in the U.S. treasury market by Amihud and Mendelson (1991) and Kamara (1994). Beim (1992) and Bliss (1997) find differential liquidity value at longer horizons. Constantinides and Ingersoll (1984) and Jordan and Jordan (1991) investigate the role of tax-timing options and Ronn (1987) finds tax-clientele effects.

2. **The error term** assumes that measurement errors are additive in the price. Some authors such as Jordan and Kuipers (1997) suggest that a proper specification of the errors should be in terms of the quoted yield-to-maturity, or a function of the yield and price. Other authors suggest a proportional rather than an additive error term. Brown and Dybvig assumed that the error terms ϵ_j are independent and identically distributed normal random variables. This assumption is relatively strong. In fact, one can predict that since treasury securities with different maturities are exchanged with different frequencies

the variance from quotation errors may differ with maturity. Moreover, assuming that ϵ_j accounts for mispricing errors, then errors will be related to the treasury security price, implying that errors increase with maturity. In the following, we will assume that the errors have zero expectation and are independently distributed. We find the normality assumption very strong and do not invoke it. Later in our empirical analysis we will check whether the estimated errors are independently distributed or not.

In equation 2.22 from Chapter 2 , we showed that the solution to the fundamental PDE is an exponential polynomial (EP) which we also showed to be the unique linear solution. Thus, the discount function can be expressed as

$$D(\beta, s, t) = (1 - \sum_{i=1}^{n} \beta_i) e^{-R(s-t)} + \sum_{i=1}^{n} \beta_i e^{-\lambda_i (s-t)}. \tag{3.3}$$

Here, we change the notation slightly: The ith state factor of the EP model is now denoted by β_i instead of x_i.

Equation 3.3 can be considered as a discrete version of the functional

$$D(\beta, s, t) = \int_0^\infty \beta(t, \lambda) e^{-\lambda T} d\lambda, \tag{3.4}$$

where $\beta(t, \lambda)$ is normalized at each time point t as follows

$$\int_0^\infty \beta(t, \lambda) d\lambda = 1 \quad \text{for all} \quad t.$$

Now, assuming that $\beta(t, \lambda)$ is non-negative for all λ at any time point t, then $D(\beta, t, T)$ can be interpreted as the Laplace transform (moment generating function) of a random variable λ that is defined by the probability measure $\beta(t, \lambda)$. This continuous representation is helpful for understanding our future empirical work.

From another perspective, equation 3.3 is similar to equation (7) of Vasicek and Fong (1982). However, Vasicek and Fong divide the maturity range into several intervals. Within each maturity interval and between two consecutive knots or breaks, a spline function, as a linear combination of three exponential functions, is fitted to the discount price. The EP model is much simpler in the sense that it does not use spline functions. It is a linear combination of n exponential functions defined on the entire maturity range of the discount price. Ferguson and Raymar (1998) used the Vasicek and Fong model without splines, which is no more than the EP model. Using simulated data, they conclude that a six-factor EP specification is a good description of the term structure (i.e., $n = 6$).

A linear discount function is convenient for pricing coupon bonds. In fact, it has been shown that a coupon bond is a linear combination of zero-coupon bonds. Substituting the discount function $D(\beta, s, t)$ in 3.3 into equation 3.2 and re-arranging yields

$$\begin{aligned} P_j(s,t) &= C_j \sum_{q=1}^{s} e^{-R(m_q-t)} + 100e^{-R(s-t)} + C_j \sum_{q=1}^{s}\sum_{i=1}^{n} \beta_i e^{-\lambda_i(m_q-t)} \\ &+100 \sum_{i=1}^{n} \beta_i e^{-\lambda_i(s-t)} - C_j \sum_{q=1}^{s}\sum_{i=1}^{n} \beta_i e^{-R(m_q-t)} \\ &-100 \sum_{i=1}^{n} \beta_i e^{-R(s-t)} + \epsilon_j. \end{aligned} \tag{3.5}$$

Let's denote the following respective expressions by P_j^0 and P_j^i.

$$P_j^0(s,t) = C_j \sum_{q=1}^{s} e^{-R(m_q-t)} + 100e^{-R(s-t)} \tag{3.6}$$

$$P_j^i(s,t) = C_j \sum_{q=1}^{s}\sum_{i=1}^{n} e^{-\lambda_i(m_q-t)} + 100 \sum_{i=1}^{n} e^{-\lambda_i(s-t)} \tag{3.7}$$

Substituting these two expressions in equation 3.5 gives

$$P_j(s,t) - P_j^0(s,t) = \sum_{i=1}^{n} \beta_i [P_j^i(t,s) - P_j^0(t,s)] + \epsilon_j. \tag{3.8}$$

This last equation is very important because it expresses the price of a coupon-bond as a linear combination of the "hypothetical" prices: P_j^0 and P_j^i. These prices are the present values of the cash flows of the bond discounted by the long rate R and by the rate λ_i respectively, where in essence $R = \lambda_0$. See Appendix B page 200 for details of the present value computation, including the handling of accrued interest.

3.2 Estimation procedure

Our estimation procedure fits equation 3.8 to the market data of bond prices. The bond price $P_j(s,t)$ will be approximated by the simple average of the bid and ask prices, plus accrued interest. Hence

$$P_j(s,t) = P_j^M(s,t) + AI_j,$$

where

- $P_j^M(s,t)$ is the average price of the bond, with $j = 1\ldots, N$.
- N is the number of treasury securities in each cross-sectional sample.
- AI_j is the accrued interest that corresponds to security j, computed as follows

$$AI_j = \frac{C_j}{2}(1 - t_c),$$

- t_c is the time to the next coupon payment (expressed as a fraction of a half-year).

Simplifying the notation, now equation 3.8 can be written as

$$P_j^M + AI_j - P_j^0 = \sum_{i=1}^{n} \beta_i [P_j^i - P_j^0] + \epsilon_j. \tag{3.9}$$

Let

$$\begin{cases} y_j &= P_j^M + AI_j - P_j^0, \\ x_j^i &= P_j^i - P_j^0. \end{cases}$$

If we substitute the terms y_j and x_j^i in equation (3.9), we obtain

$$y_j = \sum_{i=1}^{n} \beta_i x_j^i + \epsilon_j \quad \text{for all} \quad j = 1, \ldots N. \tag{3.10}$$

Equation 3.10 can be written in matrix notation as follows

$$Y = X\beta + \epsilon. \tag{3.11}$$

As we mentioned previously, the error terms ϵ are assumed to be independently distributed. However, we do not assume they are normal. Instead, their distribution form will be the subject of an empirical investigation. For a cross-sectional study, Y is the $(N \times 1)$ vector of observations on the dependent variable. β is the vector of linear regression coefficients, X is the $(N \times n)$ matrix of observations on the independent variables and ϵ is the $(N \times 1)$ vector of errors. The state vector β is estimated by ordinary least square (OLS), yielding the estimate

$$\hat{\beta} = (X'X)^{-1}X'Y. \tag{3.12}$$

The application of OLS to the above linear model does not require any assumption about the probability distribution of the errors terms ϵ_j. The only assumptions made about the ϵ_j were that they have zero expectations and are uncorrelated. Indeed, since bonds of different maturities are traded with different frequencies, it is anticipated that the variance of the errors ϵ_j may vary with maturity. This fact will be confirmed empirically.

3.3 The cross-sectional analysis

3.3.1 Data

Our data consist of daily prices on US Treasury bills, notes and bonds, for the period from 27 July 1989 to 15 October 1996. The quotations are provided by the New

York Federal Reserve Bank (NYFRB). Every daily sample consists of the bid and ask quotes of the treasury securities at 3:30 P.M. Eastern Time. The cross-sectional samples have an average of 215 observations (bills, notes and bonds). We excluded two types of bonds from the data set: callable bonds and flower bonds (bonds with special tax status). Unlike other studies, e.g. Bliss (1997), we did not eliminate bills under one month to maturity from our samples. Indeed, as will be confirmed by our in-sample results, the EP model does not "suffer" from the presence of such securities. The time to maturity of each observation is computed as the difference between the maturity date indicated by the NYFRB and the settlement date. The basis year is 365 days. The yields to maturity (ytm) provided by the NYFRB are computed assuming one business day delivery and at the bid ask average[1]. These yields to maturity are compared to our estimated yields to maturity from the EP model.

3.3.2 Procedure

In order to compute the independent variables x_j^i, we must specify the row vector $\lambda = (\lambda_1, \ldots, \lambda_n)$, as well as the long-rate $\lambda_0 = R$. The parameters λ_i's are chosen to make the basis exponential functions as distinct as possible in order to avoid extreme multicollinearity in the OLS regression. The length of this vector depends on the number of state factors chosen for the regression. For our U.S. data sample we use from 8 to 9 state factors plus the long rate. This choice is motivated by the quality of the cross-sectional fitting, assessed by criteria discussed later, as well as the multicollinearity encountered among the distinct factors. We tested different specifications for the λ vector. For instance, when we use nine factors in regression

[1] In the Center for Research in Security Prices bond file, the ytms are computed assuming two-business day delivery. According to Duffie (1996), this assumption is incorrect. Instead, one business day must be assumed for delivery.

3.12, the cross-sectional fitting is always good, however in some samples, we notice a high level of multicollinearity among the factors which translates into very large values of the state factors. In those cases, using 8 factors seems to be a better choice for the cross-sectional estimation. After many trials, we conclude that using 8 factors, or for some samples 9 factors plus the long rate, brings a better result in terms of multicollinearity as well as cross-sectional fitting. Setting $\lambda_i = l_i + R$, then we let the vector $l = (l_1, \dots, l_9)$ take the following component values

i :	1	2	3	4	5	6	7	8	9
l :	2.75	0.17	0.115	0.065	0.04	0.025	0.015	0.01	0.005

Here we note that we have chosen the vector l as a decreasing series of numbers. It is clear that empirical consideration alone guide our choice of λ. However, as we pointed out on page 39, the λ vector is considered as a discrete approximation to some continuous function. The connection between the discrete and continuous representations of the model will be the subject of future research.

3.3.2.1 Optimization procedure

The cross-sectional estimation, using the EP model, of the term structure implies the estimation of the long rate R as well as the other state factors β_i for $i = 1, \dots, 8$, or 9. How are these factors estimated? We suggest to implement a two-stage optimization procedure. We determine the optimal value of R in a first-stage analysis. Then, in a second-stage analysis, the state factors β_i of the model are estimated as linear coefficients of regression 3.11. Here are the details of the two distinct stages:

1. Let us consider again the following equation:

$$P_j^M + AI_j - P_j^0 = \sum_{i=1}^{n} \beta_i [P_j^i - P_j^0] + \epsilon_j. \tag{3.13}$$

Note that P_j^M and AI_j are picked from data available from the NYFRB, the P_j^is are computed using expressions in Appendix B. However, the computation of P_j^0 requires a value of R. Thus, for each sample we initialize R in the interval [3.5% and 18 %]. This choice in particular is motivated by some empirical significance. The core OLS regression 3.11 is running repeatedly inside a loop for the long rate. Each iteration will find the estimates of $\beta_i, i = 1, \ldots, n$ and the pricing errors from the EP model. The root mean squared error (RMSE) of the sample is computed as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{P}_j - P_j^M)^2}{N}}, \quad j = 1 \ldots N,$$

where $\hat{P}_j$ is the estimated price from the EP model, and P_j^M is the corresponding bond market price.

There is a wide choice of measures to fit prices. Two measures are frequently used in term structure studies. The mean absolute fitted-price errors (MAE) and the RMSE. In our study, we will define the short SRMSE as the RMSE of all the bonds in the sample with maturity less than one year. Hence, we choose the long rate R to minimize the SRMSE. We use a Golden Section Search for this minimization.

At this stage, one can ask for the reason behind choosing R to minimize the SRMSE. Our argument is the following: Almost all observations with maturities less than one year are Treasury bills (zero-coupon bonds) or notes or bonds with just one or two coupon payments left. Thus, the yields to maturity of such securities, provided by the NYFRB, should coincide with the estimated yield function in the absence of pricing errors. In other words, the pricing errors of the short term securities (with maturities less than one year) contribute little to the overall sample RMSE. Thus, the accuracy of the estimation especially in the short range maturity range cannot be based solely on the sample RMSE. The SRMSE is used to guide the choice of R. The optimal choice of R must fit the short end of the yields and simultaneously minimizes the SRMSE.

2. Given the optimal R, found in the previous step, all the variables of equation 3.9 are now computed. The OLS regression defined in equation 3.11 is now ready to be estimated. The state factors β_i's are the estimated coefficients of this OLS regression. Thus, the β_i's are chosen to minimize the sum of squared errors. The RMSE as well as the SRMSE computed at this stage are retained as results of the cross-sectional fitting.

3.3.3 Results

Our analysis of the cross-sectional results will be conducted from three perspectives. First, we will examine the general performance of the EP model as measured by the average price errors and the RMSE. Then, we present some specific examples as a demonstration of term structure estimation using the EP model. Along with the term structure curves, we report the estimated state factors of each sample. These examples are chosen to illustrate the ability of the EP model to model various shapes encountered in reality. The CIR model, for instance, implies that the yield curve attains only three shapes: uniformly rising, humped, and uniformly declining. However, an examination of U.S. yield curves in the last 40 years reveals inverted-humped yield curves as well. In the examples presented at the end of this chapter, one case (Figure 3.3) shows an inverted-humped yield curve. Finally, we reserve a subsection to examine and discuss the residuals of the EP model.

3.3.3.1 EP pricing results

From the examination of Table 3.1, we can conclude that the EP model fits the Treasury coupon bonds well. The average pricing error is less than 1 cent on Treasury bonds with $100 of face value. The mean error is not the only criterion for measuring model performance. Thus, we also report the RMSE and the MAE. We can see that the average RMSE is less than 15 cents. The average MAE is less than 5 cents. Since typical market spreads are between 20 cents to 30 cents, it is clear from our results that the EP model fits the term structure, on average, within the bid-ask range.

Figure 3.1 represents the evolution of the $RMSE$ for all the daily cross-sectional results. It can be noted that the RMSE ranges between 7 cents per $100 par value to 27 cents per $100 par value, with an average RMSE near 14 cents per $100 par value which is less than a round-commission for ordinary investors. In other studies such that of McCulloch, an equivalent measure of the RMSE was found to be $2 per $100 par value. Figure 3.2 depicts the evolution of the SRMSE. As we noted above, the SRMSE is the root-mean squared error for all Treasury bills, notes and bonds with a maturity less than one-year. In general, these securities are accurately priced and thus provide a good check on the reliability of a model. As can be seen from Figure 3.2, the SRMSE values range between 2 cents and 18 cents with an average near 4 cents per $100 of value. This confirms our previous remark regarding the performance of the EP model. However, the presence of some sharp spikes in the evolution of the SRMSE can only be explained by data errors.

3.3.3.2 Examples of curve shapes

In order to examine more closely the cross-sectional performance of the EP model, we choose four cases representing different shapes of the term structure.

In Tables 3.2 to 3.5, we report, for each sample, the long rate R obtained from the adjustment procedure, the short rate r as computed from equation 2.16 in Chapter 2, the β_i, for $i = 1, \ldots n$, the $RMSE$, the $SRMSE$ and the MAE. We report the standard errors of the estimates. These standard errors are inflated by the presence of multicollinearity among the state factors. We note from these tables that standard errors have been computed for all the state factors except for R. This is explained by the fact that we use a two-stage optimization procedure. Thus, β_i's are estimated in the second stage, conditional on R. Whereas R is estimated from the first stage and is taken as a given quantity in the second minimization problem.

Figure 3.3 illustrates the estimated yield curve for the 31st of July 1989. It can be seen that the curve has an inverted hump. The CIR model fails to capture such a complicated shape. The EP models fits the term structure of that particular day using nine parameters plus the long rate. The RMSE for this day is around 16 cents which is slightly larger than the average RMSE for the whole sample. The short end is fitted as accurately as the long end.

Figure 3.4 represents the cross-sectional fitting obtained for the sample of the 11th of January 1990. The yield curve in early 1990 was essentially flat. For this reason, the estimation was achieved using only eight factors plus the long rate. Indeed, in such a term structure environment, fewer factors seem to be needed. The EP model has the advantage of accommodating various shapes of the term structure with the ability to add or reduce a factor. Using only 8 factors means starting the decreasing series from 2.75 and stopping at 0.01. Adding a factor corresponds to using the last term of the series which is $l_9 = 0.005$. The number of factors increases for "complicated" shapes and diminishes otherwise.

Figure 3.5 illustrates the evolution of the yield curve as estimated from the EP model for the sample of the 20th of December 1994. In late 1994, the U.S yield curve showed one of the steepest short-term slopes ever observed in the Treasury market. Using eight factors, the EP model fitted the term structure with a RMSE of around

10 cents. Thus the EP model seems to track satisfactorily the mean reversion of the short rate without any need for an additional factor. Moreover, it can be seen from the same figure, that the estimated term structure is bending and capturing very well the movement of the curve for the long maturities.

Figure 3.6 reports the estimated yield curve for the sample of the 22nd of July 1996. It can be seen that the curve is uniformly rising. For the long maturities, the curve is bending downward. This phenomenon is not specific to the EP model. It has been reported by many authors, without any reasonable explanation. Indeed, bond dates for long maturities are relatively sparse with an average of about five observations per sample for maturities greater than twenty years. Shea (1984) has showed that fitting the curve at long maturity with few observations can lead to some fitting anomalies. The EP model is not completely immune but does not encounter major problem. The estimated EP curve usually bends to capture all of the movement of term structure in the very long run. This advantage of the EP model is mainly the result of the flexibility provided by a variable number of factors. By being able to add another factor to the estimation, the long end of the term structure can be estimated with more accuracy than any model with a single state variable.

Through this list of examples, we examined the in-sample performance of the EP model. Our general conclusion is that the EP model captures very well the different shapes of term structure encountered in reality. However, this superior performance is sometimes achieved at the cost of adding another factor or having highly correlated factor estimates. In fact, the correlations between state factors sometimes exceed 90% (see Table 3.10). Brown and Dybvig and Brown and Schaefer reported collinearity between parameters of the CIR model. Jordan and Kuipers also noted this phenomenon when they tested the CIR model with STRIPS data. One way to reduce the collinearity is to reduce the number of state factors. We have tried to fit the term structure for the same period using six factors only but this reduction altered the quality of the in-sample results. Hence, we conclude that using between eight and nine state factors plus the long rate. is necessary for obtaining a reliable fit even if this is achieved at the expense of some collinearity in the state factors.

3.3.3.3 Pricing errors

Our concern here is to study the pricing errors of the EP model. Our analysis will be conducted with the objective of answering the following questions:

- How are the pricing errors distributed for each cross-sectional sample ?
- Are the pricing errors independent within cross-sections?

Figures 3.7 to 3.10 present the distributional patterns of errors for the same cross-sectional samples illustrated earlier. It appears from these figures that the residuals do not systematically follow any systematic pattern. However, as it is clear from several figures, the residuals are more variable at longer maturities and also exhibit some autocorrelation. The greater variability at longer maturity partly reflects the fitting of R to short maturities. A careful examination of the autocorrelation coefficients of the residuals allows us to have a better understanding of serial dependence across maturities. In Figures 3.11 to 3.14, we report the autocorrelation coefficients of the residuals estimated from the EP model. The two boundaries represent the upper and lower two standard deviation confidence bounds, based on the assumption that all autocorrelations are zero. The samples we report correspond to the ones presented previously. The residuals from the cross-sectional estimation of the 31st July 1989 as well as the 20th December 1994 show some autocorrelation. For the other samples, the autocorrelation is not significant. Among the 1805 daily cross-sectional samples, only 2.4% of them have one or two autocorrelation coefficients (out of 20 coefficients) outside the 95% bounds. This percentage can be considered as small. Thus, our assumption about the independence of the errors is tenable and the efficiency of our OLS estimator is assured.

In order to examine the independence of pricing errors, we group the pricing errors into one-year maturity categories, the available maturities at each cross-section are classified into 30 categories (i.e., less than one year, one to two years, ..., twenty nine to thirty years). For each cross-sectional estimation, the mean of pricing errors within each category of maturities is computed. This exercise is repeated for all the categories and for all cross-sectional samples. Thus, we obtain for each maturity category a series of mean pricing errors. We compute the mean and standard deviation of each series. These results are summarized in Table 3.6 and in Figure 3.15.

It appears from these results that the residual errors are not independent within cross-sections. Despite the fact that the average of the mean pricing errors as well as their standard deviation are small, one can see from Figure 3.15 that there is a curvilinear pattern in the pricing errors and pronounced departures from the zero line. Pricing errors in the short end are minor, because R is fitted by minimizing the SRMSE. However, the long end of the maturity range, with the exception of the 30-year maturity class, shows overpricing (positive errors). On the other hand, the mid-maturity range shows underpricing (negative errors). This same phenomenon was also reported by Brown and Schaefer when testing the CIR model with British bonds. They also reported underpricing in the intermediate ranges of bonds. Jordan and Kuipers observed similar pattern with CIR, Vasicek and Merton models when using STRIPS data. However, all of these authors concluded that the pattern was not related to the model but instead to the structure of the bond market itself.

The cross-sectional estimation of the 1805 samples from 27 July 1989 to 15 October 1996 allow us to answer the question raised at the beginning of this chapter. After examining all of the in-sample results, we can confirm that the EP model, for this time period and these types of securities, fits the observed term structure with the same exponential basis. Every term structure pattern encountered during this period can be adequately represented by the long rate R and eight or nine additional state factors.

3.4 Comparison between the EP model and other models in the literature

In this section, we compare our results in the previous sections with what has been reported for empirical tests of other term structure models in the literature. Our comparison is more qualitative than quantitative. In other words, we do not estimate other models using our U.S. bonds data and compare the results to what we obtained so far. Instead, we report what reported studies found with similar data sets and try to place our results in the context of this literature.

3.4.1 Some results from the theoretical models

In the last few years, some authors have become interested in comparing the empirical performance of the theoretical models of term structure as well as the curve fitting models of term structure. In general, the studies compare several models and report in-sample and out-of-sample results. Jordan and Kuipers (1997) compared three term structure models: Vasicek, CIR and Merton. Their study is based on Treasury coupon STRIPS prices and yields over the period 1990 through 1994. Some of their results are reported in Table 3.7. These results cannot be compared directly to our results since they are obtained from a completely different data set. However, they provide us with some indications of the range of pricing errors for common term structure models. The units of the price errors are dollars per \$100 in face value of STRIPS. The RMSE is the time-series mean of daily cross-sectional RMSE's. When testing the EP model, we reported a mean RMSE of 13.2 cents per \$100 of face value of Treasury bonds obtained with 1805 samples. This number is less than any RMSE of the three models reported in Table 3.7. Our results are obtained using coupon bonds prices instead of STRIPS prices. This fact would be expected to produce larger pricing errors. Indeed, most of the studies in the literature used T-bills or STRIPS to avoid complications arising from estimating pure discount rates from coupon bonds. The fact that the EP model yielded smaller pricing errors overall, in more trying circumstances, suggests it is a superior model. We can also claim that the EP model offers the advantage of having a "small" RMSE achieved using a simple linear estimation compared to the complexity required by a nonlinear estimation of CIR or Vasicek or Merton models.

3.4.2 Some results from the curve fitting models

The first curve-fitting model to be seriously tested is the cubic spline of McCulloch. Many drawbacks have been attributed to this model; in particular, the instability observed in the long forward rate. Although there is no economic theory imposing a restriction on the oscillation of the forward rate, practitioners prefer to have a model with stable forward rates.

Bliss has compared several curve-fitting models. He used the monthly CRSP bond file from 1970 to 1995. His main finding is that the Mean Absolute Error (MAE) for price errors is small in economic terms for almost all methods he examined, with the Unsmoothed Fama-Bliss method performing the best. It must be noted that this method, in particular, is nothing more than a modified version of the "bootstrap" technique (see Chapter 1), which cannot be considered as a model but simply a curve-fitting technique with a high number of parameters. In Table 3.8, we report a representative sample of his results. Using the EP model, we get an MAE of 4.18 cents per $100 par value. This number is smaller than any of the four models studied by Bliss. Moreover, it can be seen from Table 3.8 that even the smallest MAE among these models is obtained through a model having between 42 and 163 parameters! Instead, the MAE of the EP model is obtained using 9 to 10 state factors.

Bekdache and Baum (1995) used the monthly CRSP government bond file and U.S. Treasury STRIPS data to compare the in-sample and out-of-sample performance of six curve-fitting models. Their results indicate that the competitive models are very similar in terms of their in-sample performance, especially for T-bills prices. Table 3.9 presents the RMSE and MAE found with each model. Given that the RMSE and the MAE of the EP model are 13.2 cents and 4.18 cents per $100 par value respectively, we can conclude that the EP still offers the advantage of producing smaller pricing errors.

This completes our review of the empirical performance of some theoretical term structure models and curve fitting methods. The review brings the following points to our attention:

- The evidence indicates that the EP model tends to produce smaller RMSE than many other term structure models in the literature.
- The EP model achieves a high degree of accuracy in terms of in-sample results, even when using samples of heterogeneous bonds.

3.5 Eigen analysis

A model with eight or nine state factors and a long rate may be necessary for an accurate cross-sectional estimation of the term structure. However, when modeling interest contingent claims, a large number of state factors can be burdensome. For this reason, we would like to know whether the EP model can be reduced to a parsimonious model having fewer factors. For this determination, we first start by computing the variance-covariance matrix of the nine factors and the long rate, based on the time series derived in the previous section. Table 3.11 reports the variance-covariance matrix of the state factors and the long rate. Table 3.12 reports the eigenvalues as well as the eigenvectors. It appears that there are three eigenvalues that are significantly greater than the rest of the eigenvalues. The magnitude of the first eigenvalue is much larger than the other two. The first principal component explains 98.9% of the total variation, the second component explains 0.9% and the third component has a weight of 0.1% in the overall variation of term structure. These three components have an explanatory power of 99.9% for all the term structure movements. The remaining 0.1% is shared by the other seven components. Thus, the term structure state factors appear to have just three main principal components. The question is why are the remaining dimensions needed? It is the need for accuracy and goodness of fit that explains the presence of more than three factors in the EP model. The three factors are mainly needed to provide a "rough" representation of the shape of the term structure. The remaining factors, although not significant in terms of explaining the variance, take into account the subtleties of the term structure shape and help us to estimate an accurate yield curve. Zhang (1993) conducted a factor analysis to determine the number of factors behind the term structure evolution. He found that term structure is driven by three principal factors. Nelson and Siegel suggested an interpretation for the three main factors that drive the term structure: (1) the general interest rate level, (2) the slope of the yield curve and (3) the curvature of the yield curve. Recent developments in the term structure models seem to retain the two-and three-factor models as reasonable descriptions of reality [2]. The one-state factor models, such as those of CIR or Vasicek, although theoretically attractive, have very poor cross-sectional estimation

[2]See for instance Subrahmanyam (1996).

characteristics. Thus, the EP model can be a possible candidate as a three-factor model. Of course, the cross-sectional estimation properties of such a model must be investigated.

3.6 Empirical relation between the EP model and an HJM specification

In the previous chapter, we claimed that the forward function derived from the EP model can be approximated by a function which is no more than the re-indexed HJM model. In this section, we will empirically test this claim and present some results. In Chapter 2, we showed that the EP forward function can be written as

$$f(\beta,\tau) = R - \frac{\partial \ln(1 + \sum_{i=1}^{n} \beta_i(e^{-l_i\tau} - 1))}{\partial\tau}. \tag{3.14}$$

Using the fact that

$$f(\beta,\tau) = -\frac{\partial D/\partial\tau}{D},$$

the EP forward function can also be written as

$$f(\beta,\tau) = \frac{Re^{-R\tau} - R\sum_{i=1}^{n} \beta_i e^{-R\tau} - \sum_{i=1}^{n} \beta_i(l_i + R)e^{-(l_i+R)\tau}}{e^{-R\tau}[1 + \sum_{i=1}^{n} \beta_i(e^{-l_i\tau} - 1)]}. \tag{3.15}$$

On the other hand, we claimed that the EP forward function can be empirically approximated by a proxy forward function given by

$$\hat{f}(\varphi,\tau) = R + \sum_{i=1}^{n} \varphi_i e^{-l_i\tau}, \tag{3.16}$$

which is no more than a simple form of the re-indexed HJM model.

Now, our objective is to estimate the φ_i coefficients such that the sum of the squared differences between the EP forward function and the proxy forward function is minimized. The empirical existence of the φ_is implies the equivalence between the EP forward function and the re-indexed HJM model. The φ_i coefficients are estimated following these steps. Once we run the OLS regression of the cross-sectional estimation on equation 3.10, the set of estimates R and $\hat{\beta}$ are obtained. Then at each cross-sectional sample, the variables φ_i can be estimated from the following equation using the OLS regression

$$[g(\tau_j) - R] = \sum_{i=1}^{n} \varphi_i h_j^i + u_j \quad \text{for } j = 1, \ldots, N, \tag{3.17}$$

where $g(\tau_j) = f(\hat{\beta}, \tau_j)$ and $e^{-l_i \tau_j} = h_j^i$.

The forward errors can be defined as

$$g(\tau_j) - \hat{f}(\tau_j) = \hat{u}_j \quad \text{for } j = 1, \ldots, N, \tag{3.18}$$

where $\hat{f}(\tau_j) = \hat{f}(\hat{\varphi}_i, \tau_j)$.

Below we will present some examples of this close relation between the proxy forward function (the re-indexed HJM model) and the EP forward function. Our empirical investigation, using the same data set as before, shows that this difference is extremely small for all samples. As a result, we can claim that there exist coefficients $\varphi_i, i = 1 \ldots, n$, such that the EP forward function defined in 3.14 and the re-indexed HJM model coincide. In Figures 3.16 to 3.19, we report two distinct curves: the forward curve obtained from the EP model (solid line) and the proxy forward curve (plus signs) which is obtained from equation 3.16. The difference between the EP forward function and the proxy forward function is measured by the maximum difference between the two functions over the maturity range. We find the mean of the maximum difference between the two functions 3.67 basis points. This empirical finding is important because it confirms that the EP model and the re-indexed HJM model can be reconciled in an empirical framework.

3.7 Conclusion

In this chapter, we examined the cross-sectional performance of the EP model. The time series properties of the model and the state factors, in particular, remain to be developed further in later chapters.

The cross-sectional analysis, conducted on the U.S. Treasury bonds from 27 July 1989 to 15 October 1996, concluded that the EP model is successful as a model of the term structure of interest rates. This estimation of the EP parameters is achieved while keeping the exponential basis constant. The EP model is a model that lies "between" two categories, the theoretical models and curve fitting models. It offers the advantage of simplicity: linear estimation instead of non linear one. Moreover, the EP model is derived from the PDE of bond pricing and thus has a strong theoretical basis compared to the curve-fitting models which are atheoretical motivated solely by goodness of fit.

In Chapter 5 we will investigate the time series properties of the estimated state factors and to what extent they can be modeled by ARMA and GARCH processes. In this chapter, we had the opportunity to present the empirical evidence for the existence of a relationship between the EP model and a specific version of the re-indexed HJM model. This can be viewed as preliminary evidence that the theoretical term structure model as defined by HJM is inherently related to the EP model..

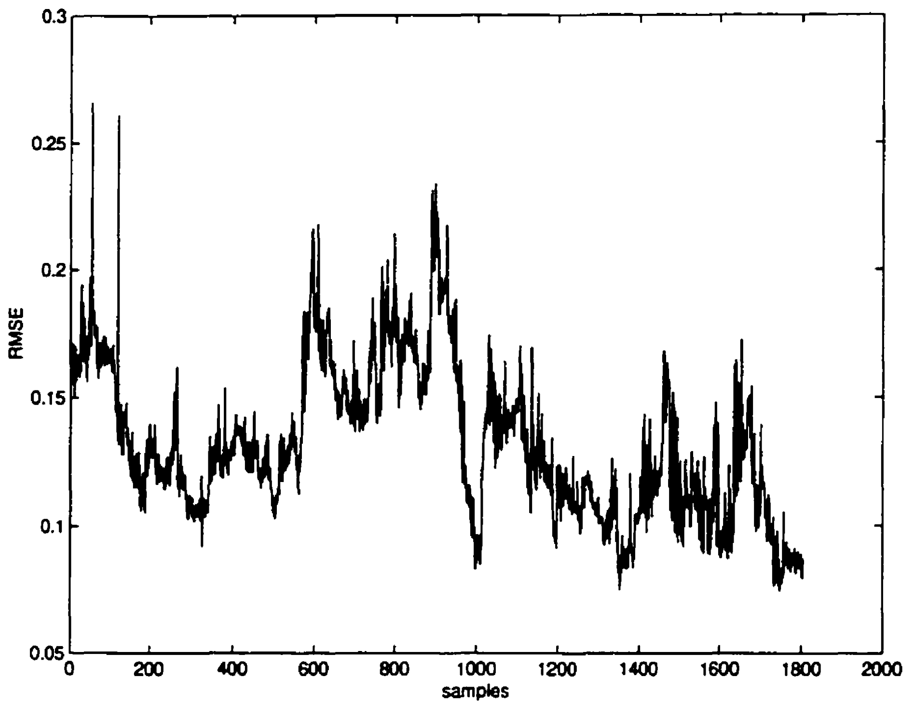

Figure 3.1: This figure represents the evolution of the daily $RMSE$ for 1805 cross-sections. The daily term structure is fitted using the EP model over the period 1989 through 1996. The RMSE is measured in dollars for a \$100 of face value.

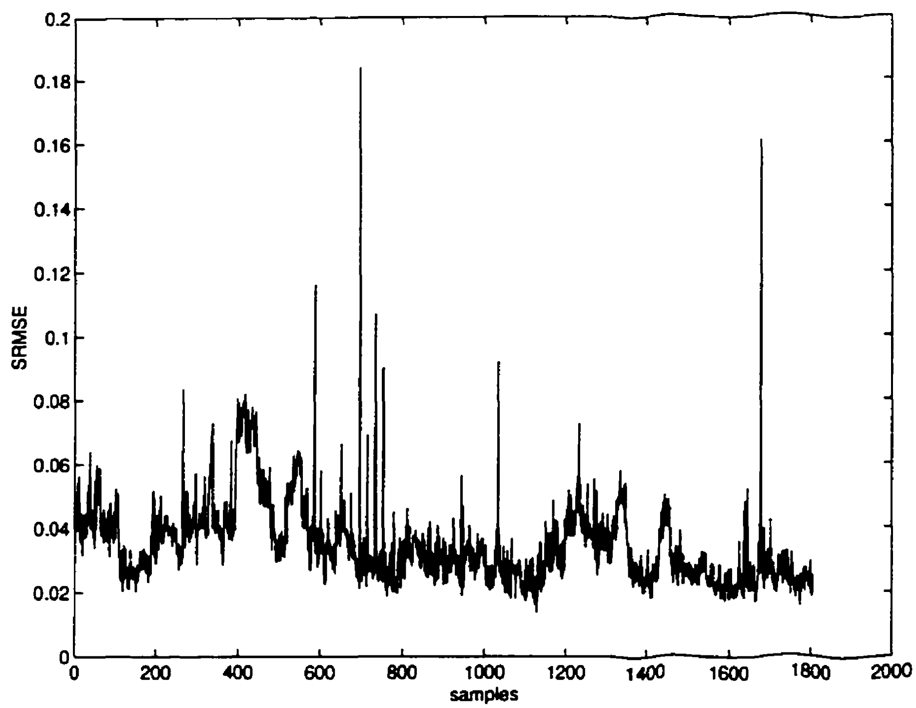

Figure 3.2: This figure represents the evolution of the daily *SRMSE* for 1805 cross-sections. The daily *SRMSE* reports the daily *RMSE* of U.S Treasury bills, notes and bonds with a maturity less than one year. The daily term structure is fitted using the EP model over the period 1989 through 1996. The *SRMSE* is measured in dollars for a $100 of face value.

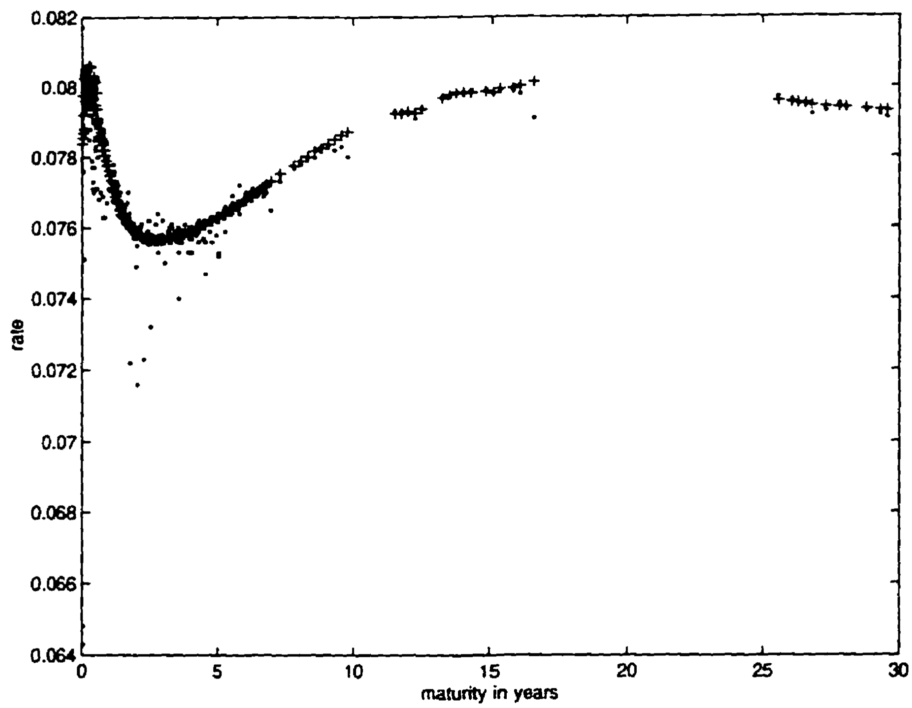

Figure 3.3: Cross-sectional estimation of the term structure for the 31st of July 1989. It shows the yield as a function of maturity. The crosses are the yields-to-maturity as computed from the EP model with state factors. The dots represent the yield-to-maturity as computed from the NYFRB. The (OLS) regression used in this cross-sectional estimation is given by equation 3.12.

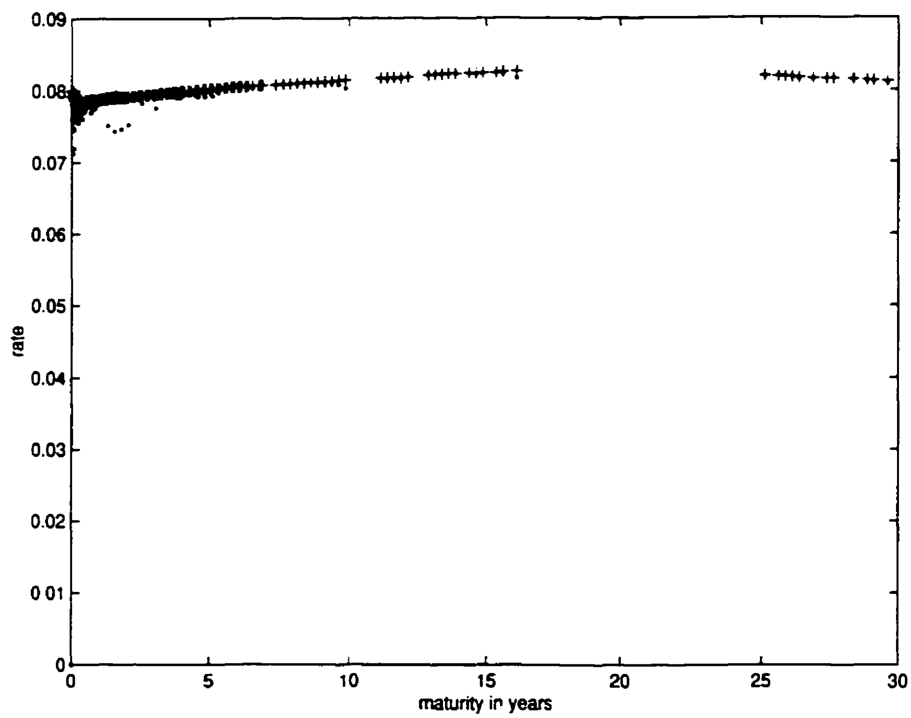

Figure 3.4: Cross-sectional estimation of the term structure for the 11th of January 1990. It shows the yield as a function of maturity. The crosses are the yields-to-maturity as computed from the EP model. The dots represent the yield-to-maturity as computed from the NYFRB. The (OLS) regression used in this cross-sectional estimation is given by equation 3.12.

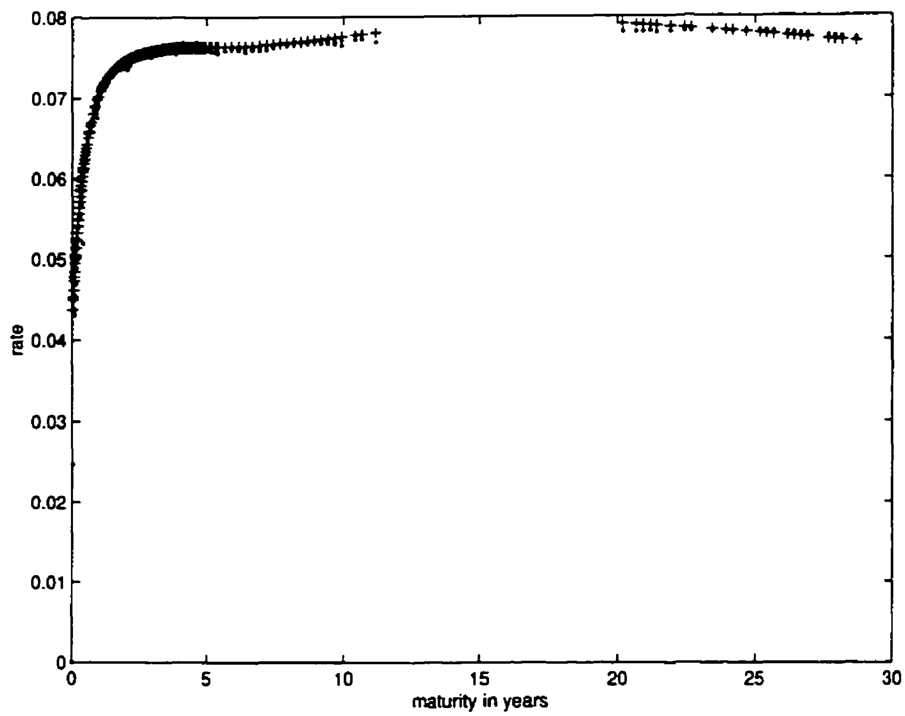

Figure 3.5: Cross-sectional estimation of the term structure for the 20th of December 1994. It shows the yield function in terms of the maturity. The crosses are the yields-to-maturity as computed from the EP model. The dots represent the yield-to-maturity as computed from the NYFRB. The (OLS) regression used in this cross-sectional estimation is given by equation 3.12.

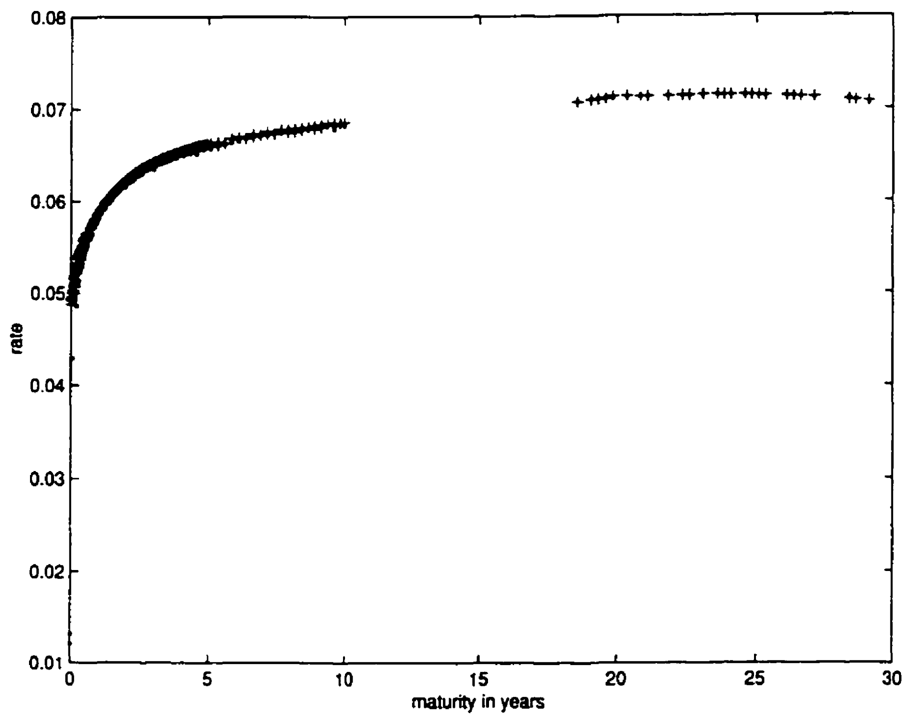

Figure 3.6: Cross-sectional estimation of the term structure for the 22nd of July 1996. It shows the yield function in terms of the maturity. The crosses are the yields-to-maturity as computed from the EP model. The dots represent the yield-to-maturity as computed from the NYFRB. The (OLS) regression used in this cross-sectional estimation is given by equation 3.12.

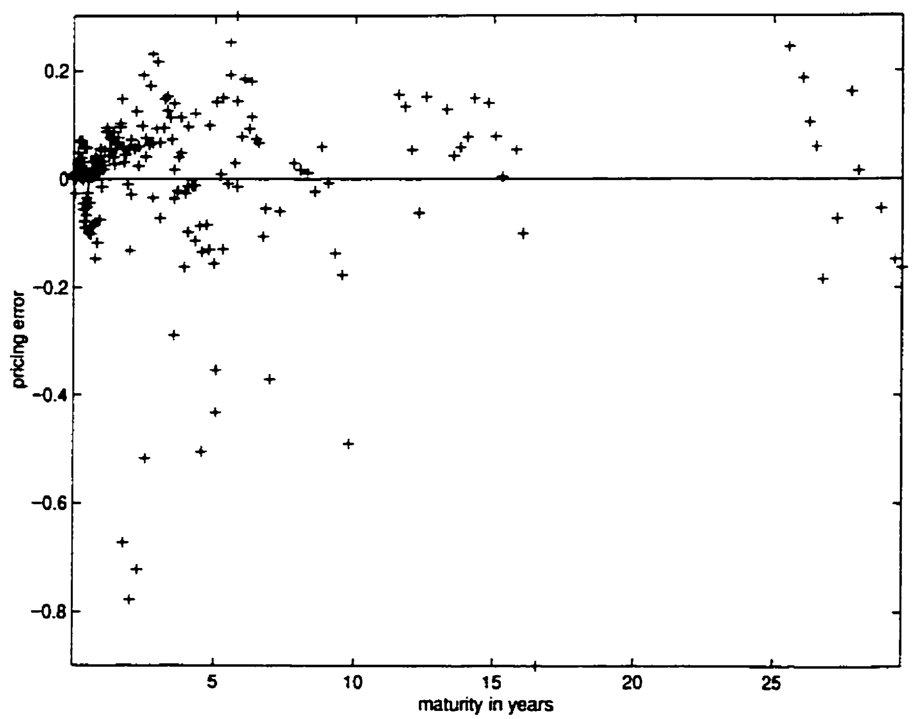

Figure 3.7: Pricing errors from the cross-sectional estimation of the term structure for the 31st of July 1989. The pricing errors are in dollars per $100 par value. They are defined as $\hat{\epsilon} = (\hat{P}_j - P_j^M)$, where P_j^M is the bond market price and $\hat{P}_j$ is the estimated bond price from the OLS regression 3.12.

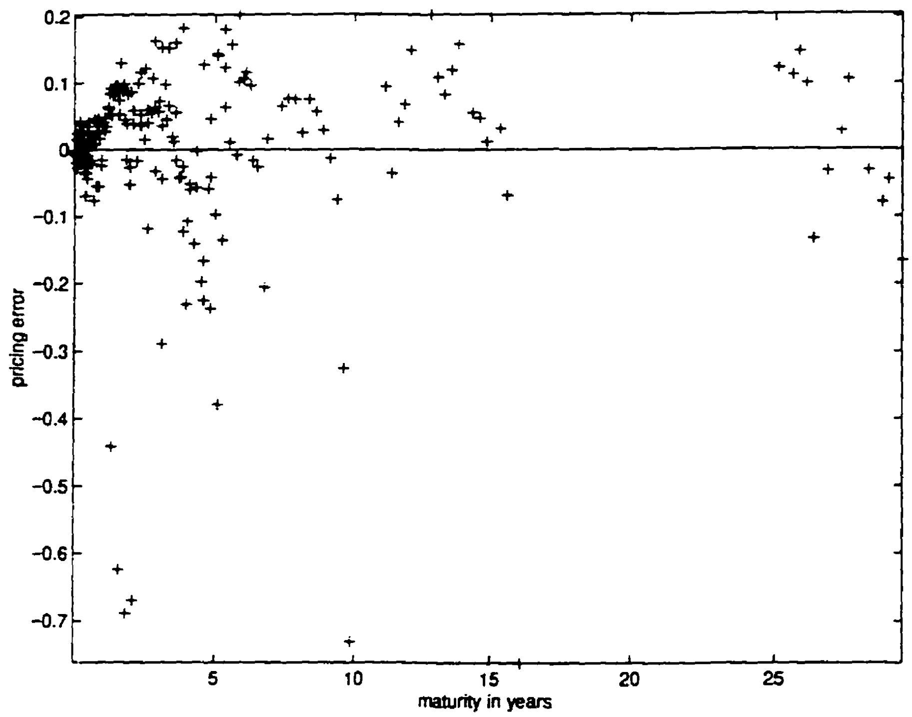

Figure 3.8: Pricing errors from the cross-sectional estimation of the term structure for the 11th of January 1990. The pricing errors are in dollars per \$100 par value. They are defined as $\hat{\epsilon} = (\hat{P}_j - P_j^M)$, where P_j^M is the bond market price and $\hat{P}_j$ is the estimated bond price from the OLS regression 3.12.

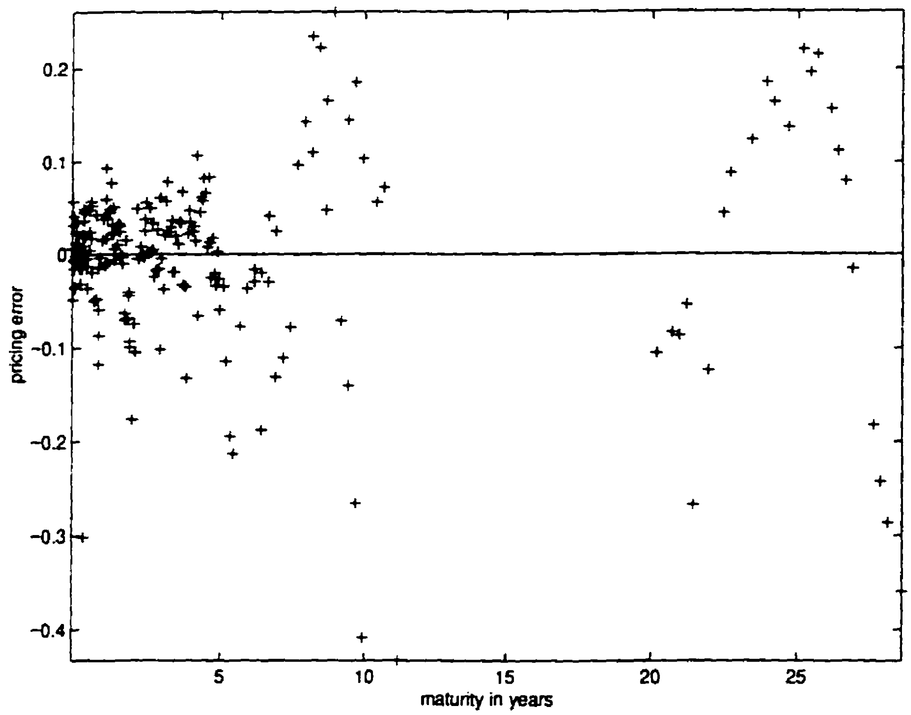

Figure 3.9: Pricing errors from the cross-sectional estimation of the term structure for the 20th of December 1994. The pricing errors are in dollars per \$100 par value. They are defined as $\hat{\epsilon} = (\hat{P}_j - P_j^M)$, where P_j^M is the bond market price and $\hat{P}_j$ is the estimated bond price from the OLS regression 3.12.

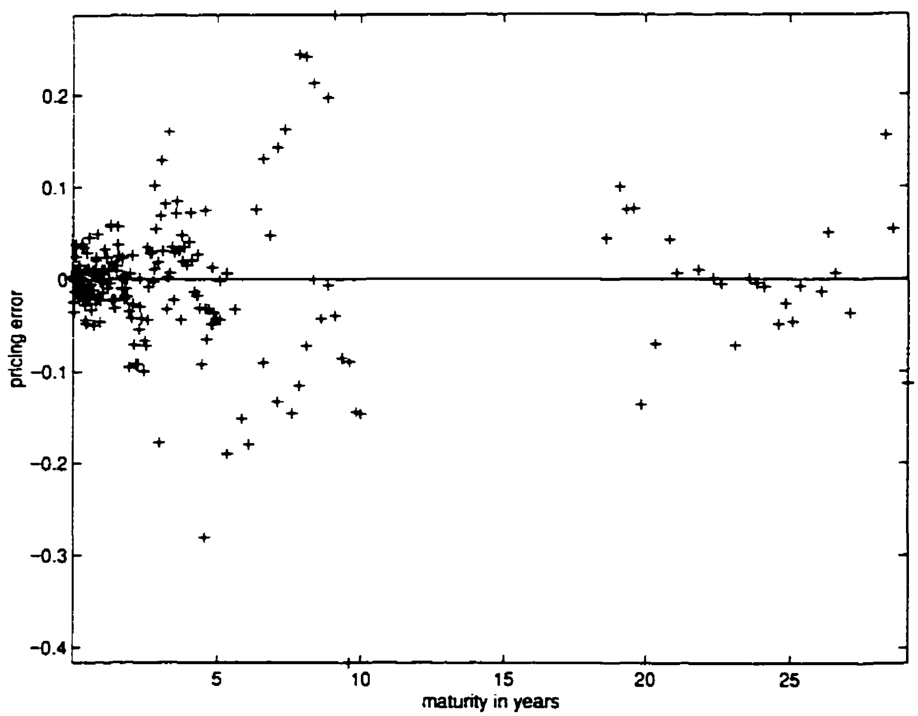

Figure 3.10: Pricing errors from the cross-sectional estimation of the term structure for the 22nd of July 1996. The pricing errors are in dollars per \$100 par value. They are defined as $\hat{\epsilon} = (\hat{P}_j - P_j^M)$, where P_j^M is the bond market price and $\hat{P}_j$ is the estimated bond price from the OLS regression 3.12.

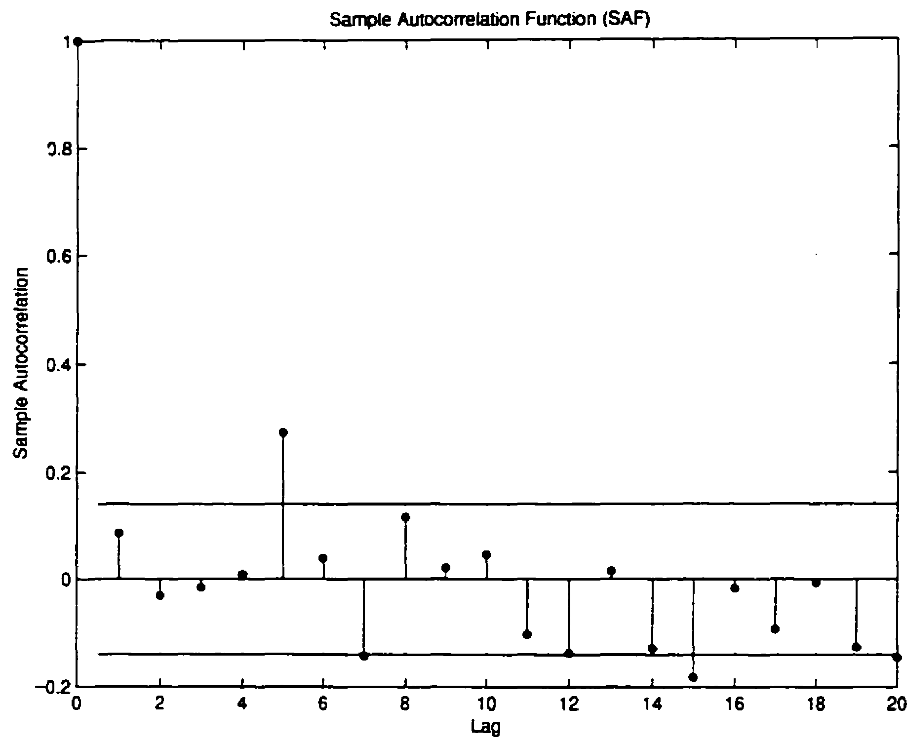

Figure 3.11: Autocorrelation of the residuals of the term structure estimation for the 31st of July 1989.

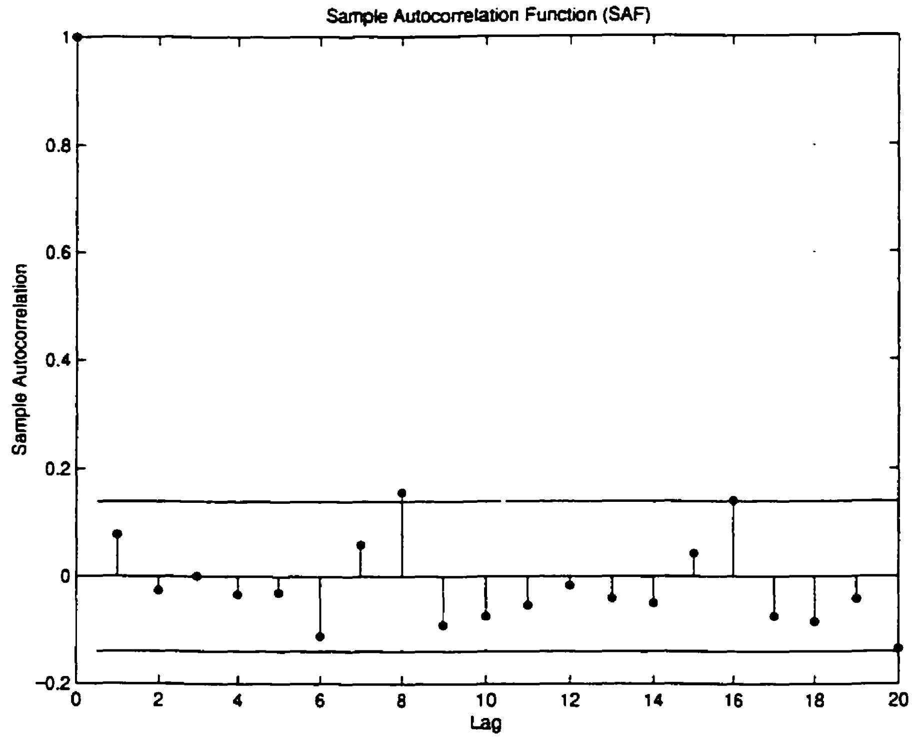

Figure 3.12: Autocorrelation of the residuals of the term structure estimation for the 11th of January 1990.

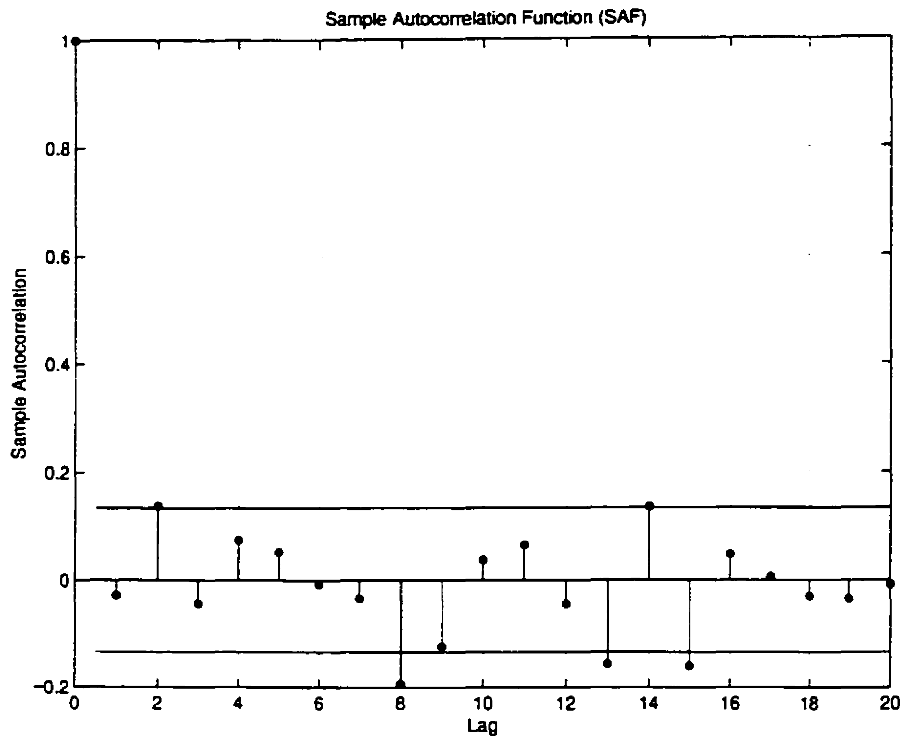

Figure 3.13: Autocorrelation of the residuals of the term structure estimation for the 20th of December 1994.

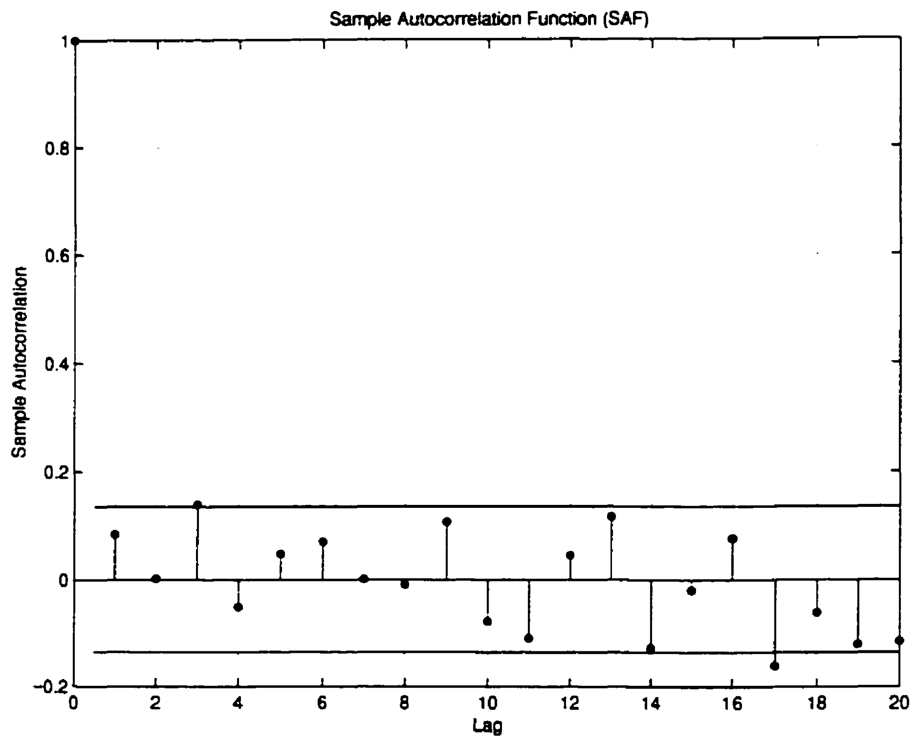

Figure 3.14: Autocorrelation of the residuals of the term structure estimation for the 22nd of July 1996.

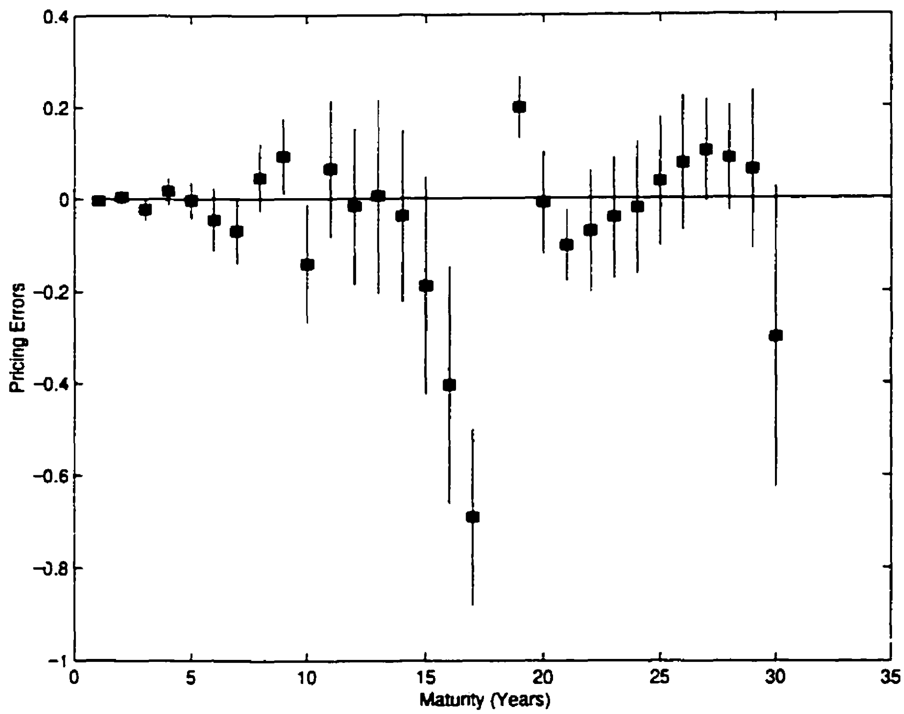

Figure 3.15: This figure represents the evolution of the pricing errors for 388,022 daily observations of U.S Treasury bills, notes and bonds for the period 1989-1996. The theoretical prices of bonds are estimated from the EP model. Errors are differences between estimated and actual prices. The units of the pricing errors are dollars per $100 in face value. The mean and standard deviation of series of daily mean pricing errors are computed for each maturity category. Center point for each category is mean; whiskers represent one standard deviation bounds. For the sample we used, there was no bond with maturity [17-18] years.

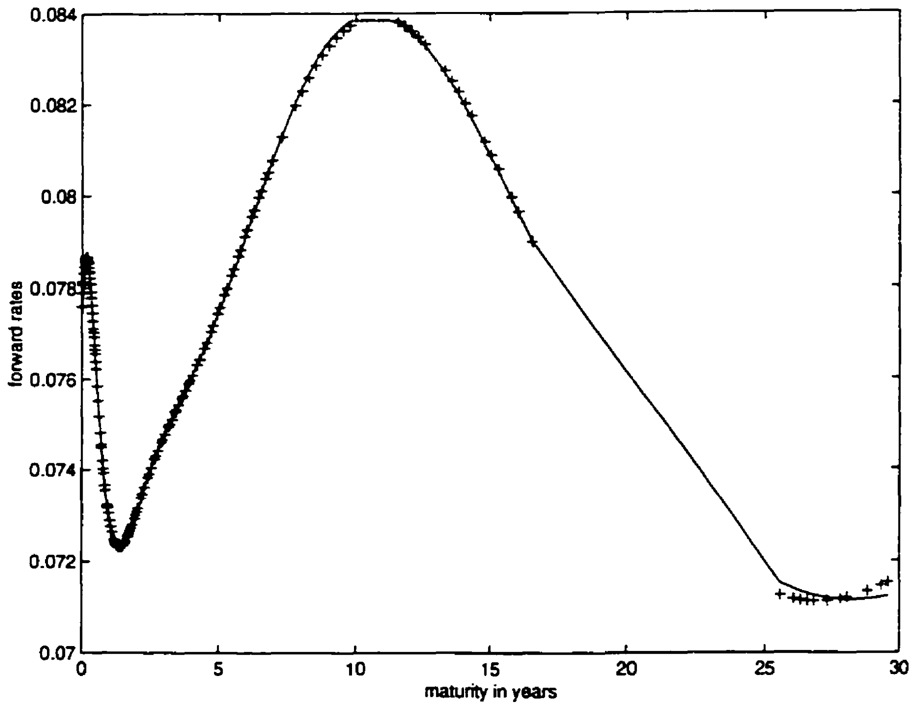

Figure 3.16: This figure is from the 31st of July 1989. It illustrates the EP forward curve (solid line) as described by equation (3.14) and the proxy forward curve (plus signs) described by equation (3.17). The proxy forward function is closely related to the HJM re-indexed model. In this case the maximum error between the two functions is $9.253\text{x}10^{-4}$.

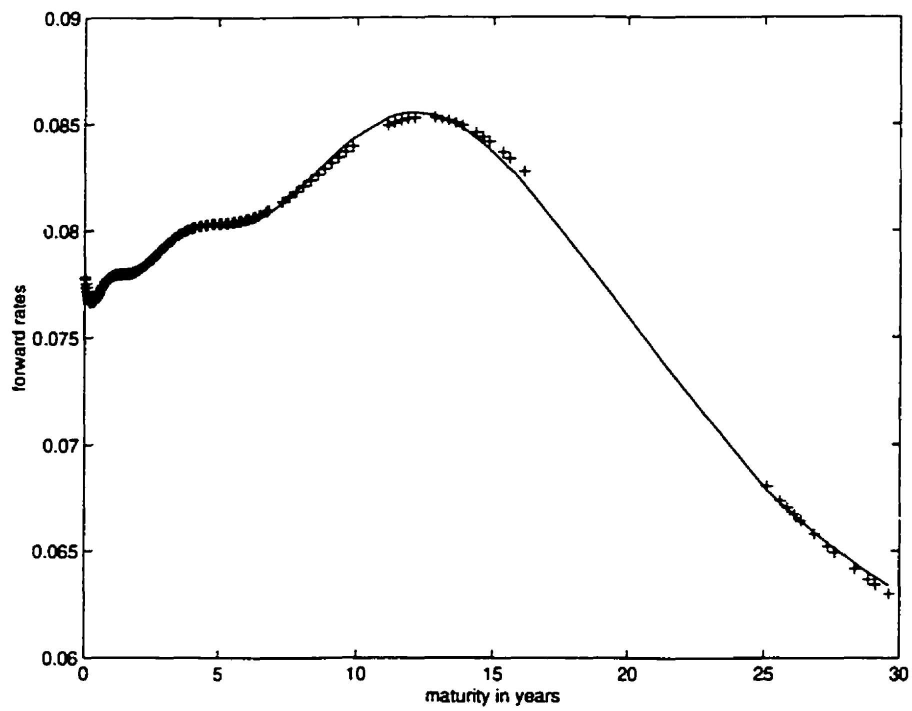

Figure 3.17: This figure is from the 11th of January 1990. It illustrates the EP forward curve (solid line) as described by equation (3.14) and the proxy forward curve (plus signs) described by equation (3.17). The proxy forward function is closely related to the HJM re-indexed model. In this case the maximum error between the two functions is 2.304x10^{-4}.

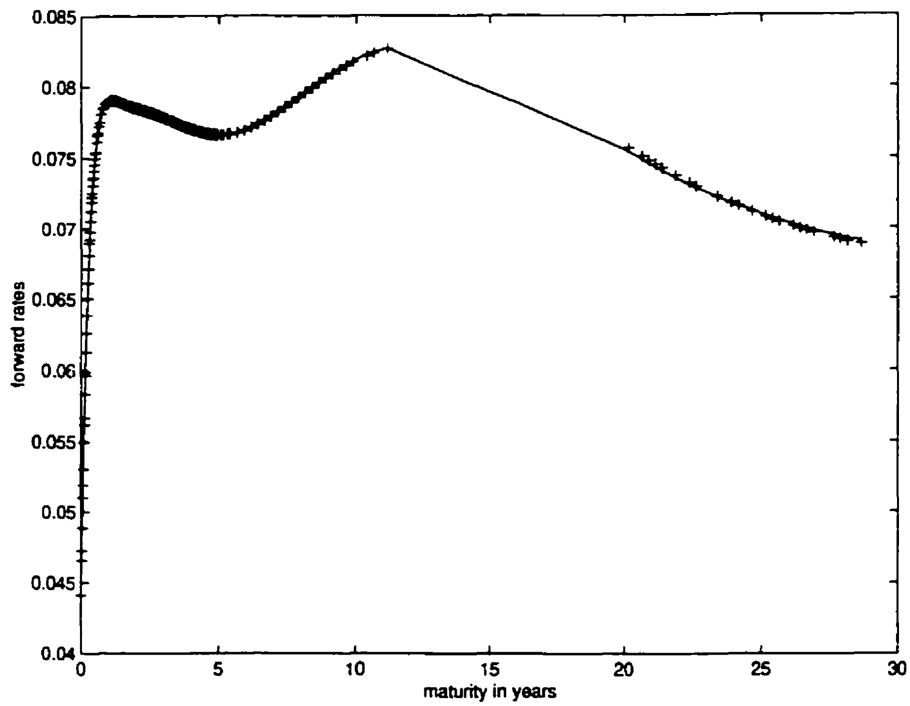

Figure 3.18: This figure is from the 20th of December 1994. It illustrates the EP forward curve (solid line) as described by equation (3.14) and the proxy forward curve (plus signs) described by equation (3.17). The proxy forward function is closely related to the HJM re-indexed model. In this case the maximum error between the two functions is 1.021×10^{-4}.

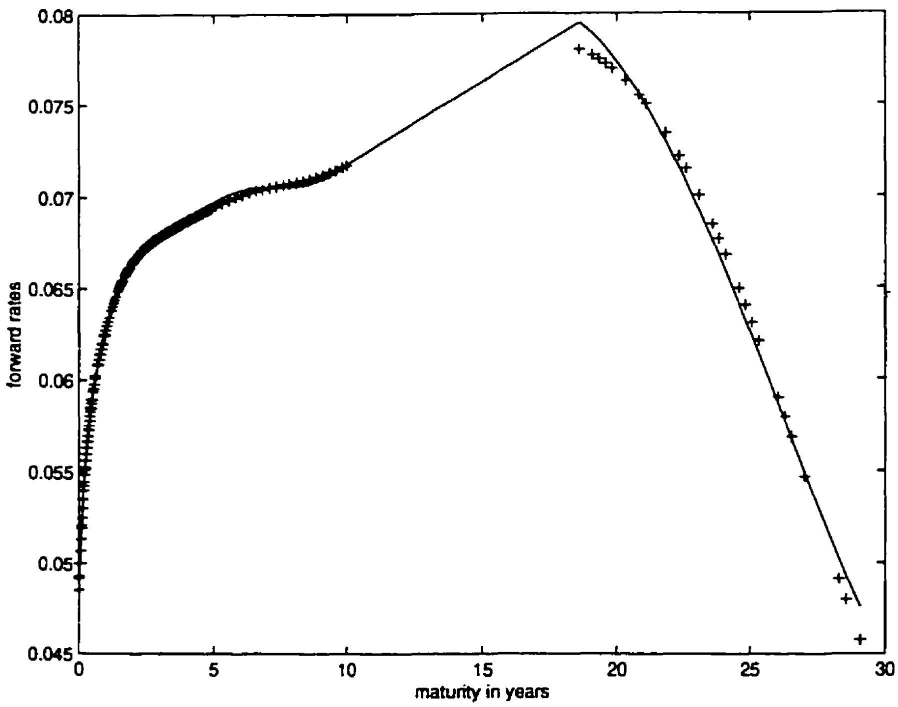

Figure 3.19: This figure is from the 22nd of July 1996. It illustrates the EP forward curve (solid line) as described by equation (3.14) and the proxy forward curve (plus signs) described by equation (3.17). The proxy forward function is closely related to the HJM re-indexed model. In this case the maximum error between the two functions is 2.553×10^{-4}.

$RMSE$	$SRMSE$	MAE	Mean
0.1322	0.0349	0.0418	-0.0046

Table 3.1: The price errors are from 1805 daily cross-sections of 388,022 daily U.S Treasury coupon bonds. The estimated prices are computed using the EP model. Errors are measured as differences between estimated and market prices. Errors are in dollars per $100 par value. RMSE is the time series mean of the daily cross-sectional RMSE. SRMSE is the time series mean of the daily cross-sectional SRMSE. MAE is the time series mean of the daily MAE. Mean is the time series mean of the daily mean pricing errors.

Cross-sectional Results	Values	Std. err
Date 7/31/1989	-	-
Long rate (R)	0.0963	-
Sample	204	-
β_1	-0.0283	0.0422
β_2	0.1292	0.2538
β_3	-0.2351	0.8801
β_4	0.2517	2.3726
β_5	-0.6525	5.5904
β_6	4.1192	10.4226
β_7	-16.0217	16.4522
β_8	21.0846	13.7779
β_9	-11.0262	4.3399
Short rate(r)	0.0774	-
$RMSE$	0.1618	-
$SRMSE$	0.0495	-
MAE	0.0983	-

Table 3.2: Cross-sectional estimation results of the sample for the 31st of July 1989.

Cross-sectional Results	Values	Std. err
Date 1/11/1990	-	-
Sample	205	-
Long rate (R)	0.0575	-
β_1	0.0265	0.0265
β_2	-0.1932	0.1352
β_3	0.7594	0.3699
β_4	-2.0587	0.7244
β_5	4.1142	1.1225
β_6	-5.2355	1.2636
β_7	3.5048	1.0098
β_8	-0.4574	0.4189
Short rate(r)	0.0780	-
$RMSE$	0.1446	-
$SRMSE$	0.0272	-
MAE	0.0843	-

Table 3.3: Cross-sectional estimation results of the sample for the 11th of January 1990.

Cross-sectional Results	Values	Std. err
Date 12/20/1994	-	-
Sample	213	-
Long rate (R)	0.0684	-
β_1	-0.0285	0.0184
β_2	0.0056	0.0933
β_3	0.2703	0.2519
β_4	-1.3009	0.4836
β_5	3.4946	0.7371
β_6	-5.1937	0.8329
β_7	4.2260	0.6857
β_8	-1.2764	0.2939
Short rate (r)	0.0434	-
$RMSE$	0.0981	-
$SRMSE$	0.0515	-
MAE	0.0423	-

Table 3.4: Cross-sectional estimation results of the sample for the 20th of December 1994.

Cross-sectional Results	Values	Std. err
Date 7/22/1996	-	-
Sample	218	-
Long rate (R)	0.035	
β_1	-0.0207	0.0145
β_2	0.1338	0.0784
β_3	-0.5967	0.2369
β_4	2.0100	0.5450
β_5	-5.8703	1.0976
β_6	12.8337	1.7980
β_7	-22.6783	2.5404
β_8	19.7181	1.9903
β_9	-5.3506	0.5816
Short rate (r)	0.0482	-
$RMSE$	0.0745	-
$SRMSE$	0.0216	-
MAE	0.0461	-

Table 3.5: Cross-sectional estimation results of the sample for the 22nd of July 1996.

Maturity Category	Mean	Std. dev.	Maturity category	Mean	Std.dev
0-1	-0.0038	0.0027	15-16	-0.1883	0.2366
1-2	0.0046	0.0144	16-17	-0.6911	0.1911
2-3	-0.0223	0.0237	17-18	na	na
3-4	0.0170	0.0309	18-19	0.1990	0.0688
4-5	-0.0028	0.0399	19-20	-0.0082	0.1121
5-6	-0.0451	0.0691	20-21	-0.1018	0.0780
6-7	-0.0689	0.0721	21-22	-0.0704	0.1329
7-8	0.0462	0.0755	22-23	-0.0412	0.1325
8-9	0.0930	0.0840	23-24	-0.0195	0.1444
9-10	-0.1407	0.1282	24-25	0.0376	0.1414
10-11	0.0660	0.1496	25-26	0.0773	0.1473
11-12	-0.0163	0.1700	26-27	0.1044	0.1120
12-13	0.0059	0.2116	27-28	0.0881	0.1155
13-14	-0.0368	0.1869	28-29	0.0636	0.1724
14-15	-0.4042	0.2576	29-30	-0.3002	0.3262

Table 3.6: This table reports statistics on the pricing errors of the EP model. These results are based on daily cross-sections over the period 1989-1996. The errors are differences between estimated and actual prices. They are in dollars per $100 par value. There is no bond with a [17-18] maturity during the sample period.

Model	Sample size	Mean	RMSE	Median
CIR(1985)	1250	0.020	0.203	-0.007
Vasicek(1977)	1250	0.011	0.183	-0.007
Merton(1973)	1250	0.046	0.328	0.018

Table 3.7: Results from Jordan and Kuipers(1997).

Model	Sample size	MAE	number of parameters
Unsmoothed Fama-Bliss	312	0.057	42 – 163
McCulloch	312	0.118	7 – 14
Extended Nelson and Siegel	312	0.181	5
Fisher, et al. cubic spline	312	0.101	2 – 33

Table 3.8: Results from Bliss (1997).

Model	Sample size	RMSE	MAE
Fisher, et al cubic spline	226-245	0.279	0.139
McCulloch	226-245	0.283	0.137
Nelson and Siegel	116-120	0.415	0.267

Table 3.9: Results from Bekdache and Baum (1997).

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	R
β_1	1.0000	-0.9865	0.9511	-0.8976	0.8356	-0.7850	0.7457	-0.7256	0.6908	0.3730
β_2	-0.9865	1.0000	-0.9875	0.9524	-0.9033	0.8580	-0.8189	0.7956	-0.7436	-0.4581
β_3	0.9511	-0.9875	1.0000	-0.9877	0.9555	-0.9197	0.8842	-0.8597	0.7931	0.5369
β_4	-0.8976	0.9524	-0.9877	1.0000	-0.9893	0.9673	-0.9397	0.9168	-0.8414	-0.5911
β_5	0.8356	-0.9033	0.9555	-0.9893	1.0000	-0.9934	0.9764	-0.9576	0.8797	0.6153
β_6	-0.7850	0.8580	-0.9197	0.9673	-0.9934	1.0000	-0.9941	0.9814	-0.9101	-0.6055
β_7	0.7457	-0.8189	0.8842	-0.9397	0.9764	-0.9941	1.0000	-0.9959	0.9418	0.5634
β_8	-0.7256	0.7956	-0.8597	0.9168	-0.9576	0.9814	-0.9959	1.0000	-0.9669	-0.5099
β_9	0.6908	-0.7436	0.7931	-0.8414	0.8797	-0.9101	0.9418	-0.9669	1.0000	0.3105
R	0.3730	-0.4581	0.5369	-0.5911	0.6153	-0.6055	0.5634	-0.5099	0.3105	1.0000

Table 3.10: Correlation matrix of the original series for the state factors β_i for $i = 1, \ldots, 9$ and the long rate R.

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	R
β_1	0.0026	-0.0173	0.0623	-0.1613	0.3387	-0.5437	0.7231	-0.5321	0.1354	0.0003
β_2	-0.0173	0.1176	-0.4346	1.1492	-2.4577	3.9895	-5.3306	3.9169	-0.9786	-0.0028
β_3	0.0623	-0.4346	1.6470	-4.4600	9.7299	-16.0041	21.5402	-15.8396	3.9063	0.0122
β_4	-0.1613	1.1492	-4.4600	12.3806	-27.6198	46.1491	-62.7639	46.3158	-11.3629	-0.0369
β_5	0.3387	-2.4577	9.7299	-27.6198	62.9574	-106.8784	147.0647	-109.0871	26.7900	0.0866
β_6	-0.5437	3.9895	-16.0041	46.1491	-106.8784	183.8552	-255.8702	191.0414	-47.3636	-0.1457
β_7	0.7231	-5.3306	21.5402	-62.7639	147.0647	-255.8702	360.3393	-271.4145	68.6183	0.1898
β_8	-0.5321	3.9169	-15.8396	46.3158	-109.0871	191.0414	-271.4145	206.1206	-53.2793	-0.1299
β_9	0.1354	-0.9786	3.9063	-11.3629	26.7900	-47.3636	68.6183	-53.2793	14.7302	0.0211
R	0.0003	-0.0028	0.0122	-0.0369	0.0866	-0.1457	0.1898	-0.1299	0.0211	0.0003

Table 3.11: Variance-covariance matrix of the original series for the state factors β_i for $i = 1, \ldots, 9$ and the long rate R.

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}
	832.5474	8.2632	1.2730	0.0518	0.0113	0.0029	0.0011	0.0001	0.0000	0.0000
β_1	0.8762	0.0702	-0.4527	0.0464	-0.1258	0.0391	0.0492	-0.0199	0.0067	0.0013
β_2	0.4314	-0.0316	0.6144	-0.2832	0.5156	-0.1500	-0.2300	0.1084	-0.0454	-0.0099
β_3	0.1886	-0.0784	0.5753	0.0238	-0.4783	0.2098	0.4935	-0.2904	0.1565	0.0399
β_4	0.0867	-0.0692	0.2579	0.4443	-0.4620	-0.0262	-0.4204	0.4487	-0.3439	-0.1157
β_5	0.0424	-0.0477	0.0694	0.6078	0.1218	-0.3575	-0.2246	-0.2972	0.5172	0.2699
β_6	0.0214	-0.0446	0.0166	0.4913	0.3768	-0.1081	0.4235	-0.1696	-0.4183	-0.4680
β_7	0.0114	-0.0505	0.0106	0.2772	0.2960	0.4306	0.2641	0.3593	-0.1267	0.6576
β_8	0.0079	-0.0510	0.0110	0.1476	0.1727	0.6228	-0.1435	0.1807	0.5094	-0.4949
β_9	0.0045	-0.0400	0.0088	0.0344	0.0421	0.4637	-0.4518	-0.6506	-0.3704	0.1252
R	0.0233	-0.9860	-0.1213	-0.0955	-0.0038	-0.0581	0.0050	0.0075	0.0027	0.0003

Table 3.12: Eigenvalues and eigenvectors for the variance-covariance matrix of the variables shown in the rows. The eigenvalues are in the first row of the table. The eigenvectors are in the columns below the corresponding eigenvalues.

Chapter 4

Empirical Performance of the EP Model: Results for Canadian Bonds

In this chapter, we test the empirical performance of the EP model using a Canadian daily data set. We would like to answer the following question:

- Can the proposed EP model presented in Chapter 2, fit the observed term structures?

In Chapter 3, we conclude for the period 1989-1996, that the EP model fully describes the US term structures with eight or nine states factors plus the long rate. Is this finding confirmed by the Canadian data?

We will adopt the same notation as in Chapter 3. Using the EP model, we showed that the price of a coupon-bond can be written as a linear combination of some hypothetical prices:

$$P_j^M(s,t) + AI_j - P_j^0(s,t) = \sum_{i=1}^{n} \beta_i [P_j^i(t,s) - P_j^0(t,s)] + \epsilon_j, \qquad (4.1)$$

where P_j^M is a bond market price, AI_j is its accrued interest, P_j^0 and P_j^i are hypothetical prices and ϵ_j is an error term. Our objective in this chapter is to fit equation 4.1 to the Canadian data of bond prices. Equation 4.1 can be written as

$$y_j = \sum_{i=1}^{n} \beta_i x_j^i + \epsilon_j \text{ for all } j = 1, \ldots, N, \tag{4.2}$$

where

$$\begin{cases} y_j &= P_j^M + AI_j - P_j^0, \\ x_j^i &= P_j^i - P_j^0. \end{cases}$$

As mentioned in Chapter 3 the error terms ϵ_j are assumed to be independently distributed. Unlike Brown and Dybvig when estimating the CIR model, we do not assume that the error terms are normally distributed. Instead, the error distribution form will be subject of an empirical investigation.

The linear coefficients β_i of equation 4.2 will be estimated using ordinary least square method.

4.1 Cross-sectional analysis

4.1.1 Data

In this study we use Canadian data set from the Bank of Canada. There are nearly 60,667 daily observations on Canadian bills, notes and bonds over the period of 29th of June 1992 to 29th of May 1995, totaling 731 daily samples in all. The cross-sectional samples have an average of 79 observations (bills, notes and bonds). We excluded two types of bonds from the data set: bonds with options and bonds with special features. The time to maturity is computed as the difference between the maturity date indicated by the Bank of Canada and the settlement date. The basis year is 365 days. The yields to maturity (ytm), as mid point yields, provided by the Bank of Canada are used for comparison with the estimated yields to maturity from the EP model.

4.1.2 Procedure

In this cross-sectional estimation, we follow the same procedure as we did with the US data set. To compute the independent variables x_j^i of equation 4.2, we must specify the row vector $\lambda = (\lambda_1, \ldots, \lambda_n)$ as well as the long rate $\lambda_0 = R$. The parameters λ_i are chosen to make the basis exponential functions as distinct as possible in order to avoid extreme multicollinearity in the OLS regression. The length of this vector depends on the number of state factors chosen for the regression. With the US data, we found that eight to nine state factors plus the long rate are necessary to fit all the term structures in the sampie. This choice is motivated by the quality of cross-sectional fitting assessed by the RMSE or the SRMSE as well as the multicollinearity encountered among the distinct factors.

For the Canadian data, we started the estimation using nine factors. We noticed a high level of multicollinearity among the factors which translates to very large values of the state factors. After many trials, we conclude that using six factors plus the long rate brings a better result in terms of multicollinearity as well as cross-sectional fitting. However, the values chosen for the λ vector are different from the ones used with the US data. In the Canadian case, the values of l_i which enter into the computation of λ_i are set to

i :	1	2	3	4	5	6
l_i :	0.2	0.3	0.5	0.8	1.3	2.1

This difference in the choice of the λ vector can be explained by the existence of some structural differences between the Canadian and the US market. At the end of 1992, the US Treasury market was the largest issuer of debt in the world, with over $3 trillion of different type of bonds outstanding. This huge volume makes the US Treasury market the most liquid in the world. At the same period, the government of Canada issued $430 billions of debt as bills notes and bonds. Thus, compared to the US Treasury market, the Canadian market has less volume and is less liquid.

The long rate is chosen to minimize the SRMSE. Here, we defined the SRMSE as the RMSE for all bonds with maturity less than six months. Note that defining the SRMSE on a period of one year, as we did for US data, does not alter the quality of results. For the rest, we followed the same procedure as for US data. Below, we discuss our results.

4.1.3 Results

Our analysis of the Canadian cross-sectional results will be conducted from three viewpoints. First, we examine the pricing performance of the EP model. Second, we present some specific examples showing the fitting properties of the EP term structure model. We will emphasize on the difference in results between US and Canadian data. Finally, we examine and discuss the residuals of the EP model as obtained from Canadian data set.

4.1.3.1 EP pricing results

In Table 4.1, we report some statistics regarding the pricing results using the EP model with Canadian data. The average pricing errors is less than 3 Canadian cents on government bonds with \$100 of face value. With US data this number was less than 1 cent per \$100 par value. Thus, based on the average pricing errors criterion, the EP model provides a slightly better fitting of the US term structures than of the Canadian ones.

The average RMSE is 44.3 cents per \$100 of face value, and the average $SRMSE$ is 3.69 cents per \$100 par value. The equivalent numbers for the US data are 16 and 3.49 cents per \$100 par value, respectively. The $RMSE$ is slightly higher in the Canadian case than in the US case. Figure 4.1 represents the evolution of the RMSE for all daily cross-sectional Canadian samples. It can be noted that the $RMSE$ ranges between 7 to 95 cents per \$100 par value with an average near 44 cents. Moreover, we can see from the same figure that the RMSE is decreasing in magnitude over the study period. A sharp decrease, occurred at the beginning of 1994, is most likely attributed to progress in the quality of data collection. The new system has substantially improved the accuracy of Canadian bond price data. Several studies have considered the effect of quality of financial data on bond pricing. For instance, Elton and Green (1998) studied the impact of quality of financial data on tax and liquidity effects of term structure. They concluded that a significant portion of liquidity and tax effects found by previous authors appears to be no longer relevant because data problems influenced the calculations of the original estimates. Figure 4.2 depicts the evolution of the SRMSE. As we mentioned previously, the SRMSE is the RMSE for all bills, notes and bonds with a maturity less than six months. As can be seen from this figure, the SRMSE values range between less than 1 cent to 13.5 cents. This result is very similar to the evolution of the SRMSE of the US data. Thus, we conclude that the improvement of the Canadian bond data benefited the long term securities more than the short ones.

4.1.3.2 Examples of curve shapes

Almost all of the yield curves are upward sloping during the study period. Indeed, this shape was very common during the early 1990s. An inverted term structure shape was frequent during 1989 and early 1990; however, our data set does not cover this period of time. We now look at term structure estimation for a sample of four days. In Tables 4.2 to 4.5, we report for each sample, the long rate R obtained from the adjustment procedure, the short rate as computed from equation 2.23, the estimated β_i for $i = 1, \dots n$, the $RMSE$ and the $SRMSE$ for each sample. In general, the RMSE for all the examined samples are higher in magnitude compared to what we obtain for the US samples. These examples confirm our general discussion about the evolution of the RMSE.

In Figure 4.3. we report the estimation of the term structure for the 2nd of July 1992. The curve has a slight inverted hump. The CIR model fails to capture such a complicated shape. The EP model fits Canadian term structure of that particular day using six state factors plus the long rate. The RMSE for this day is around 84 cents which is larger than the average RMSE of all Canadian samples. We attribute this large value to the poor quality of the data collection system of Bank of Canada as well as to the lower liquidity of the Canadian bond market compared to its US counterpart.

Figure 4.4 represents the cross-sectional fitting obtained for the sample of the 30th of March 1993. The curve is upward. Using six factors plus the long rate, the EP model fits the term structure with a RMSE of 66 cents. Moreover, it can be seen from this figure that the estimated term structure is bending and capturing very well the movement of the curve for the long maturities. Shea (1984) showed that fitting the curve at long maturities with few observations can lead to some fitting anomalies. The EP model does not encounter such problems with US data. Here, with Canadian data, the estimated EP curve bends to capture all of the movements of term structure in the long run. As mentioned in Chapter 3, this advantage of the EP model is mainly the result of the flexibility provided by a variable number of factors.

Figure 4.5 reports the estimated curve of the 28th of September 1994. The curve is uniformly rising. The RMSE of this sample is 16 cents. It is a dramatic decrease compared to previous examples. As can be seen from Figure 4.1, a dramatic drop in the values of the RMSE is noticeable around the beginning of year 1994. Since, the average RMSE is around 20 cents per $100 of face value.

Figure 4.6 illustrates the evolution of the estimated yields to maturity for the sample of the 1st of February 1995. The Canadian curve showed a steep short-term slope. Using six factors plus the long rate, the EP model fits the term structure with a RMSE of near 11 cents. As for the US example in Figure 3.5 of Chapter 3, the EP model seems to track satisfactorily the reversion of the short rate without any need for an additional factor.

Through the four examples, we examined the in-sample performance of the EP model, using Canadian data. Our general conclusion is that the EP model captures very well the different shapes of term structure encountered in reality. This superior performance is achieved using only six state factors plus the long rate. With the US data, the same estimation procedure required eight to nine factors plus the long rate. This difference in the number of state factors is explained by the structural differences between the two markets. The US bond market is the largest in the world, it accounts a huge number of transactions per day. Thus, a large number of factors is necessary to capture all the movements of the US term structures. The Canadian market is smaller and less liquid. Six state factors plus the long rate are sufficient to fully describe the movement of the term structures.

4.1.3.3 Pricing errors

Here we study the distribution of residuals from the estimation of the EP model. Our analysis will be structured around two levels:

- The pricing errors for each cross-sectional sample.
- The pricing errors with cross-sections.

Figures 4.7 to 4.10 present residual plots of the four examples studied before. The residuals are defined as the difference between the EP model price and its market counterpart. The residual plots show no systematic pattern except for the larger scatter at maturities beyond about one year. To examine the residuals pattern more closely, we calculate the autocorrelation coefficients of the residual series for each cross-section. They are represented in Figures 4.11 to 4.14. The two boundaries represent the upper and lower standard deviation bounds, based on the assumption that all autocorrelation coefficients are zero. They are 95% bands for individual coefficients r_k. The residuals from the cross-sectional estimation of the 2nd of July 1992 and of the 30th of March 1993 show, for one or two lags (out of 20 lags), some autocorrelation. For the other samples, the autocorrelation is not significant. Only 3% of all the Canadian samples has autocorrelation outside the 95% bounds. This percentage is small and hence OLS estimation conducted in the above cross-sectional study appears to be adequate. For the US data, we find a percentage of 2.4% of all daily samples having autocorrelation coefficients outside 95% confidence interval. From this perspective, the EP model is yielding similar results with both data sets. Hence, we get the empirical confirmation that the assumption about independence of the errors is tenable and the efficiency of our OLS estimator is assured.

In order to examine the independency of pricing errors, we proceed as follows: First, the pricing errors are grouped into a one-year maturity categories, i.e., the available maturities at each cross-section are classified into 30 categories: less than one year, one to two years,..., twenty nine to thirty years. For each cross-sectional estimation, the mean of pricing errors within each category of maturity is computed. We repeat this exercise for all categories of maturities and all 731 samples. Thus, we obtain for each maturity category a series of mean pricing errors. In all, we have 30 series of pricing errors. Second, we compute the mean of these series and their standard deviations. The results are summarized in Figure 4.15 and Table 4.6. It appears from Figure 4.15 that the residuals errors are not independent with cross-sections. Despite the fact that the average of the mean pricing errors as well as their standard deviation are small, one can see from Figure 4.15 that there is a curvilinear pattern in the pricing errors and significant departures from the zero line. The mean pricing errors in the short end are close to zero. The long end shows some overpricing (positive errors), on average, but the bias is relatively small compared with what we found in US. data. The 21-year maturity range seems to be the most distorted range with large negative pricing errors on average. For the US. case, it was the 16-17 maturity range which suffered from underpricing. This empirical finding confirms our argument regarding the existence of some structural differences between the two markets. As for the US. market, we do not think that the existence of some pattern in the pricing errors for Canadian data is a sign of model inadequacy but rather a market-related issue.

4.2 Comparison between the EP model and other models

In this section, we report on some empirical studies of term structure which have used Canadian data. The purpose of this exercise is to compare the EP results obtained with the Canadian data to what other authors have found. Once again, this comparison is more qualitative than quantitative and is not intended to provide a definite ranking of models.

Brennan and Schwartz (1979) studied Canadian monthly data. They used Canadian Bankers' Acceptances and the average yields to maturity on Government of Canada bonds with maturities more than 10 years as proxies for the instantaneous interest rate and the long-term rate, respectively. Their data set covered a period from 1964 to 1976. They postulated a model with two state variables: the short rate and the long rate. However, they found evidence of the existence of a third unknown state variable. The predicted returns for bond portfolios of different maturities derived from their two-factor model explained only about 50% of the variation in actual returns within the sample period.

Bolder and Stréliski (1999) used a sample of 30 dates chosen to span the period from 1989 to 1999. The dates were selected to include 10 observations each from an upward-sloping, flat, and inverted term structure environments. Three distinct models were used:

- Nelson and Siegel (1987),
- Svensson (1994) which is an extension of the Nelson and Siegel model,
- Super-Bell model developed by Bell Canada in the 1960s.[1]

Their work is divided into two separate aspects: the estimation problem, i.e., the choice of the best yield curve model and the optimization of its parameters, and the data problem i.e., the selection of the appropriate set of market data. In the analysis of the estimation problem, three models were examined. Each of the studied alternative is summarized in terms of goodness of fit, speed of estimation and robustness of the results. According to their result, the Svensson model is the best. At a second step, they considered the data problem. Three alternative filtering settings were considered. Their final result encourages the filtering of data (e.g. elimination of short term bonds, inclusion of certain bonds or not) for the estimation of term structure. Unlike our estimation procedure, their estimation was conducted in terms of yields instead of prices. The Bolder and Stréliski study, although interesting, is restricted to only 30 dates. Moreover, the authors suggest the use a data filtering in order to improve the term structure fit. However, in our case, we use the EP model without filtering the data. As noted earlier in discussion about the in-sample results, the EP model seems to produce larger pricing errors for Canadian data compared to US. data. However, the errors are still, on average, within any bid-ask spread. One plausible explanation for this phenomenon is the relatively lower liquidity of the Canadian bond market in comparison to the US market.

[1]See Appendix A page 197 for more details about this model.

4.3 Eigen analysis

The EP model specification we used with Canadian data assumes the existence of six factors in addition to the long rate. However, we would like to know whether the EP model could be reduced to a more parsimonious model with a limited number of factors. Thus, we conduct an eigen analysis of the variance-covariance matrix of the variables ($\beta_i, i = 1, \ldots, n$ and R) given in Table 4.7. Table 4.8 presents the eigenvalues and eigenvectors. On the light of these results, we conclude that there are three eigenvalues which are significantly larger than the rest of eigenvalues. The magnitude of the first eigenvalue is much larger than the other two. The first principal component explains almost 84% of the total variation. The second component contributes with a portion of about 14%, and the third component explains 1.7% of the total variation of the term structure. These three components explains 99.7% of the total variation of the term structure, whereas only 0.3% is explained by the other components. Thus, the term structure state factors appear to possess just three principal components. A natural question is why are the remaining dimensions needed? It is the need for accuracy and goodness of fit that explains the presence of more than three factors in the EP model. In general, three factors are needed to capture a "rough" representation of the shape of the term structure. The other factors, although not significant, in terms of explaining the variance, take into account the subtleties of the term structure shape and provide us with an accurate estimation of the yield curve. This result confirms our findings with the US. data set. Once again, we are able to claim that the EP model can be simplified to a three-factor model. As for the US. data, we suggest as Neslon and Siegel claimed, that these three factors can have an intuitive interpretation: (1) the general interest rate level, (2) the slope of the yield curve and (3) the curvature of the yield curve. More and more term structure models seem to retain the two- and three-factor models as plausible descriptions of reality (i.e., see Balduzzi, Das, Foresi, and

Sundaram (1996)). The EP model is a possible candidate as a three-factor model. Further investigation about the time series properties of the EP model can confirm our claim.

4.4 Empirical relation between the EP model and an HJM specification

In this section, we present four examples which show the close relation between the proxy model (which is based on the EP model) and the re-indexed HJM model. Our empirical investigation shows that this difference is extremely small for all samples. The results support the earlier theoretical claim that there exist constants $\varphi_i, i = 1 \ldots, n$, such that the proxy yield function defined in Chapter 3 in equation (3.16) and the re-indexed HJM model nearly coincide. Figures (4.16) to (4.19) show two distinct curves: the forward curve obtained from the EP model (solid line) and the proxy forward curve (plus signs). The maximum difference between the proxy forward function and the forward function derived from the EP model is on average in the range of 3.63 basis points for the study period. This result is similar to the one obtained for the US. data set. Thus, we have the confirmation from both data sets that the EP model can be empirically related to a specific re-indexed HJM version.

4.5 Conclusion

In this chapter, we examined the cross-sectional performance of the EP model using a Canadian data set.

In general, the main results obtained with this data set are in line with what we found for US. data. The cross-sectional estimation conducted on a daily basis has brought accurate results in terms of $SRMSE$. The specification we used includes the long rate and six state factors. The results suggest that the exponential basis can be maintained at six factors during the estimation period. Finally, as with the US. data, we were able to show empirically that there exists a relationship between the EP model and the re-indexed model of HJM.

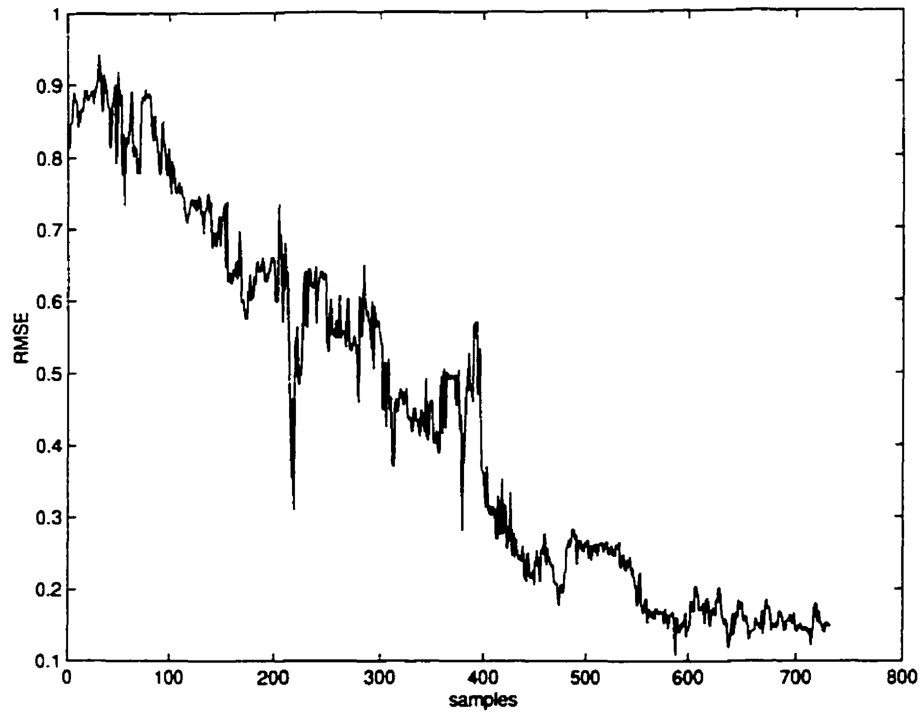

Figure 4.1: This figure represents the evolution of the daily RMSE for 731 cross-sections. The daily term structure is fitted using the EP model over the period 1992 through 1995. The RMSE is measured in dollars for a $100 of face value.

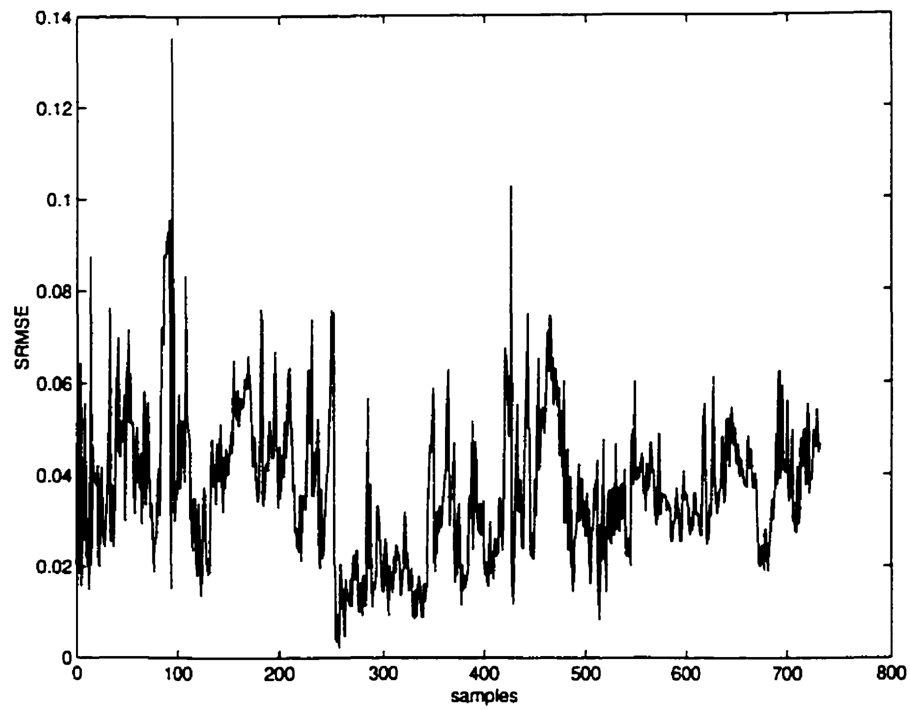

Figure 4.2: This figure represents the evolution of the daily SRMSE for 731 cross-sections. The daily SRMSE reports the daily RMSE of Canadian bills, notes and bonds with maturity less than six months. The daily term structure is fitted using the EP model over the period 1992 through 1995. The SRMSE is measured in dollars for a $100 of face value.

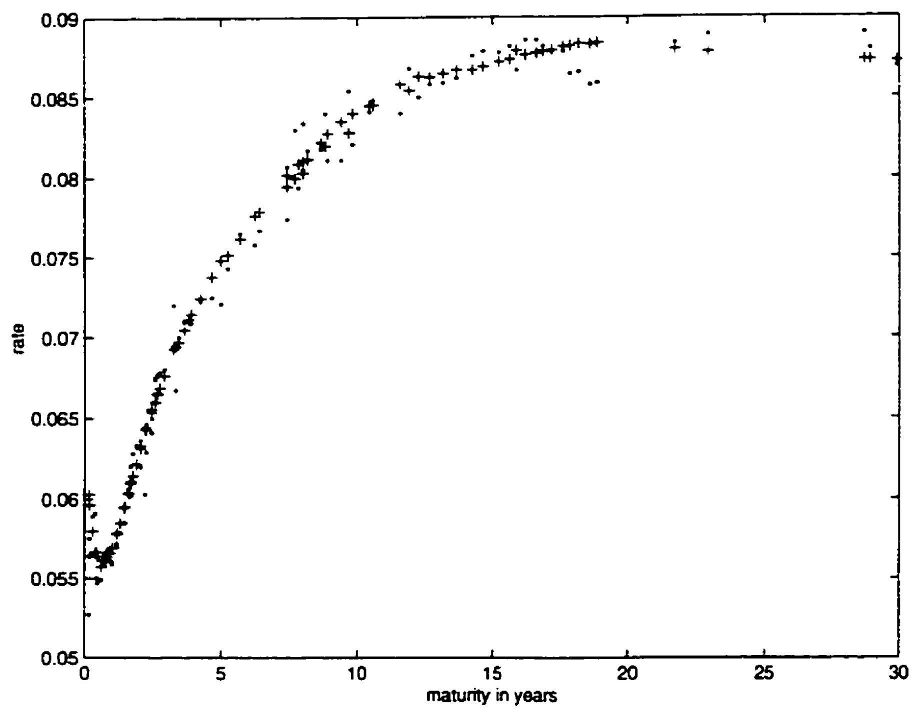

Figure 4.3: Cross-sectional estimation of term structure for the 2nd of July 1992. It shows the yields in terms of maturity. The crosses are the yields-to-maturity as computed from the EP model. The dots represent the yield-to-maturity as reported from the Bank of Canada files. The (OLS) regression used in this cross-sectional estimation is given by equation (3.12).

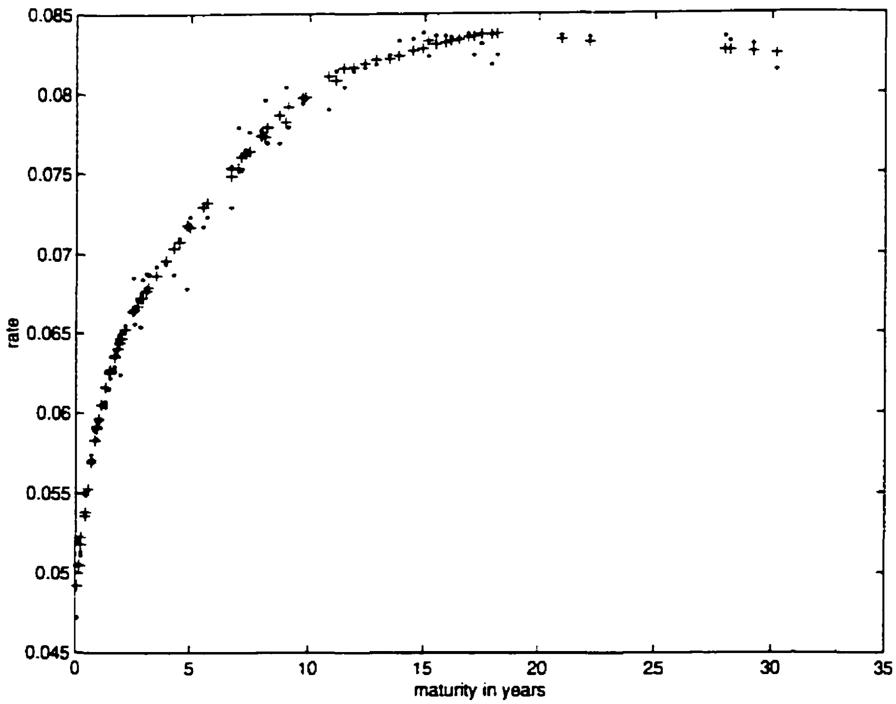

Figure 4.4: Cross-sectional estimation of the term structure for the 30th of March 1993. It shows the yields in terms of maturity. The crosses are the yields-to-maturity as computed from the EP model. The dots represent the yield-to-maturity as reported from the Bank of Canada files. The (OLS) regression used in this cross-sectional estimation is given by equation (3.12).

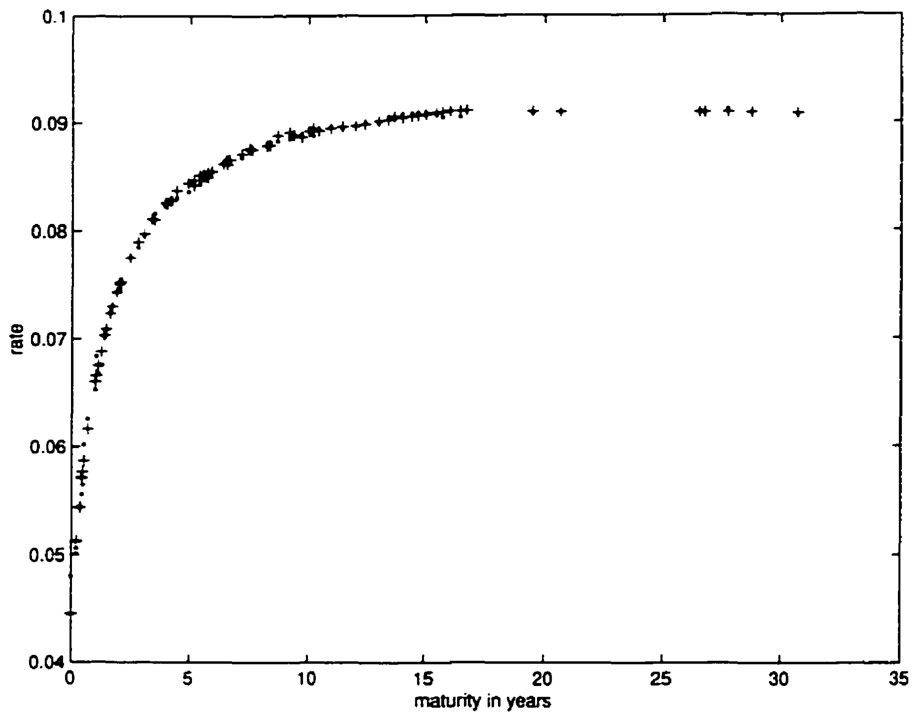

Figure 4.5: Cross-sectional estimation of the term structure for the 28th of September 1994. It shows the yields in terms of maturity. The crosses are the yields-to-maturity as computed from the EP model. The dots are the yields-to-maturity as reported from the Bank of Canada files. The (OLS) regression used in this cross-sectional estimation is given by equation (3.12).

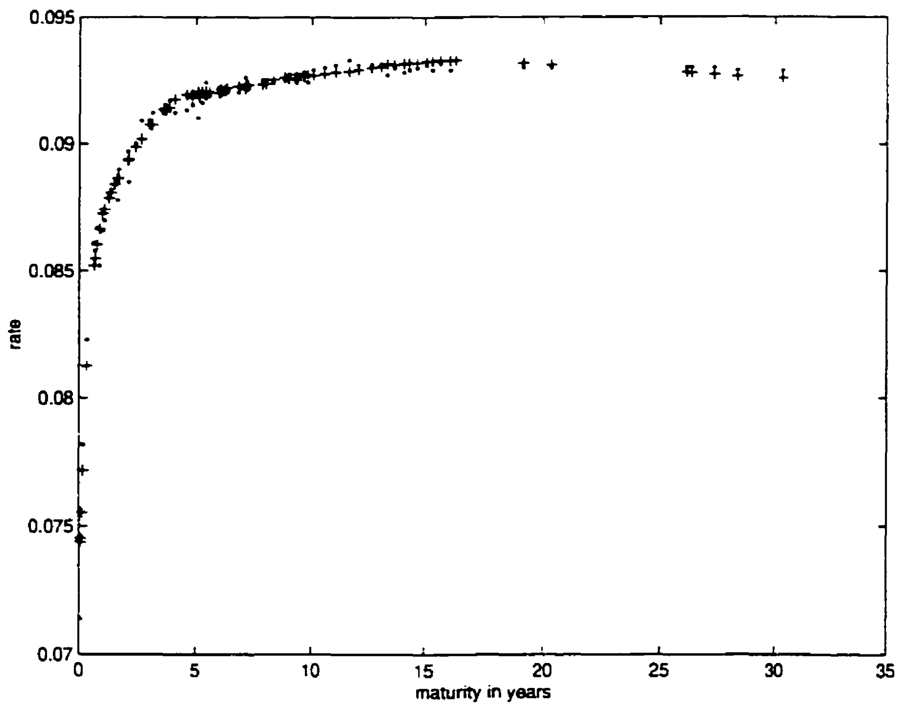

Figure 4.6: Cross-sectional estimation of the term structure for the 1st of February 1995. It shows the yields in terms of maturity. The crosses are the yields-to-maturity as computed from the EP model. The dots represent the yield-to-maturity as reported from the Bank of Canada files. The (OLS) regression used in this cross-sectional estimation is given by equation (3.12).

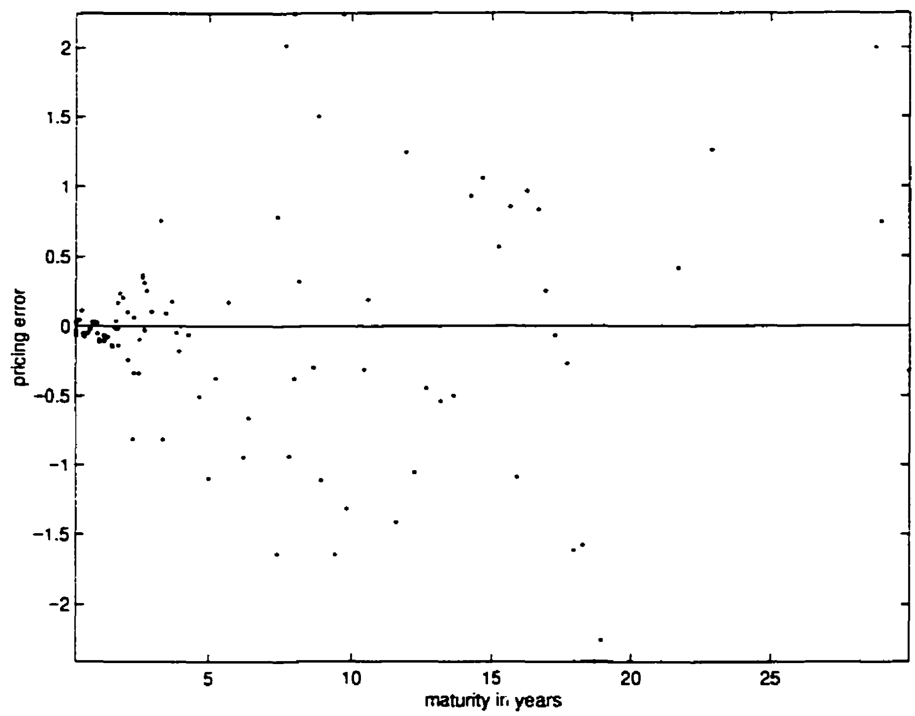

Figure 4.7: The pricing errors from the cross-sectional estimation of the term structure for the 2nd of July 1992. The pricing errors are in Candadian dollars per $100 of face value. They are defined as $\hat{\epsilon_j} = (\hat{P}_j - P_j^M)$, where P_j^M is the bond market price and $\hat{P}_j$ is the estimated bond price from the OLS regression (3.12).

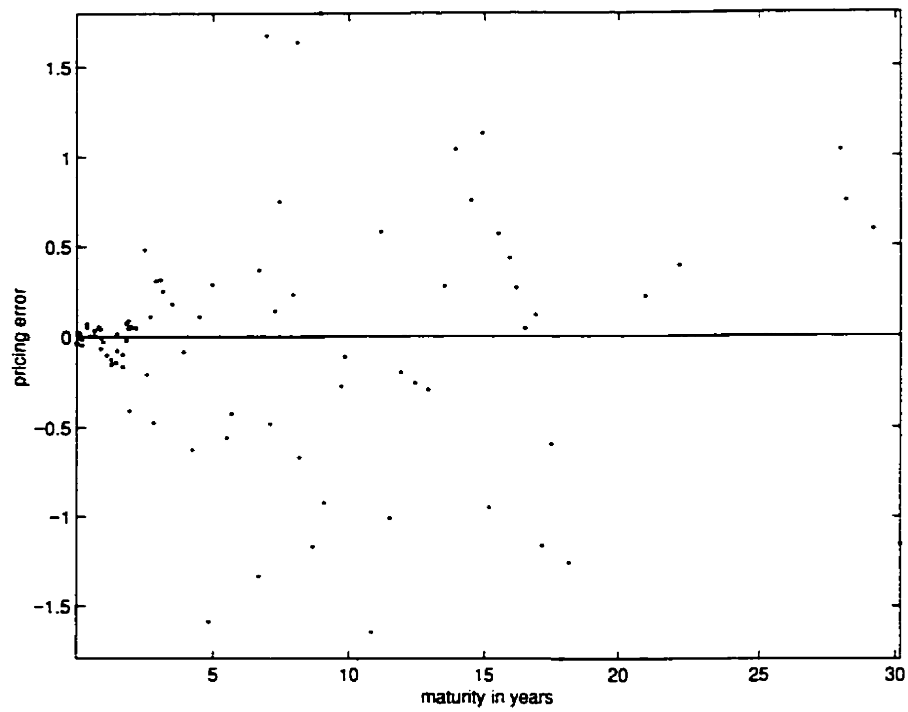

Figure 4.8: The pricing errors from the cross-sectional estimation of the term structure for the 30th of March 1993. The pricing errors are in Canadian dollars per \$100 of face value. They are defined as $\hat{\epsilon_j} = (\hat{P}_j - P_j^M)$, where P_j^M is the bond market price and $\hat{P}_j$ is the estimated bond price from the OLS regression (3.12).

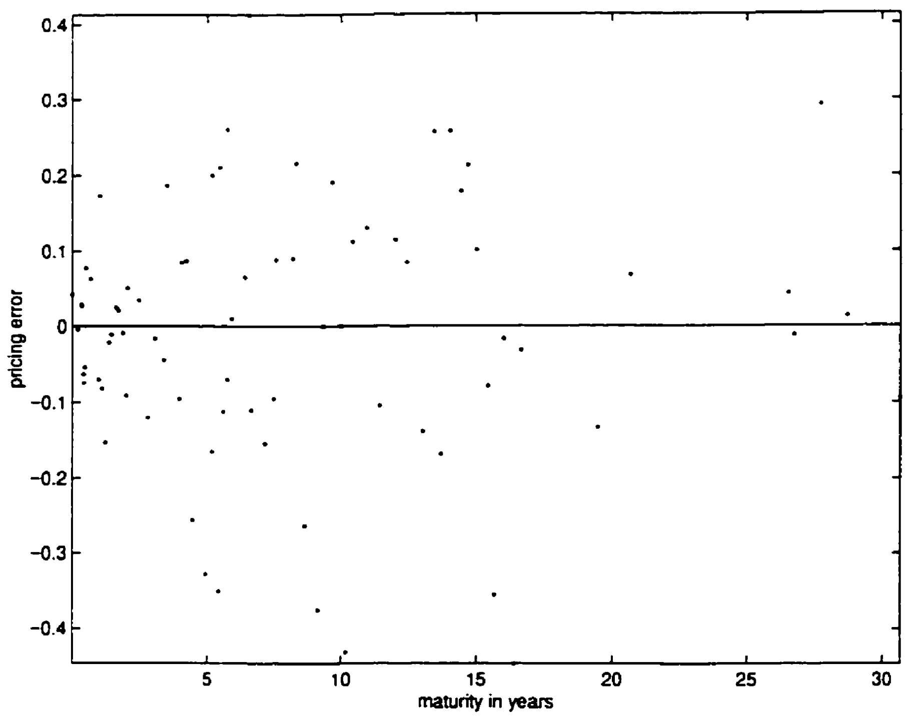

Figure 4.9: The pricing errors from the cross-sectional estimation of the term structure for the 28th of September 1994. The pricing errors are in Canadian dollars per $100 of face value. They are defined as $\hat{\epsilon_j} = (\hat{P}_j - P_j^M)$, where P_j^M is the bond market price and $\hat{P}_j$ is the estimated bond price from the OLS regression (3.12).

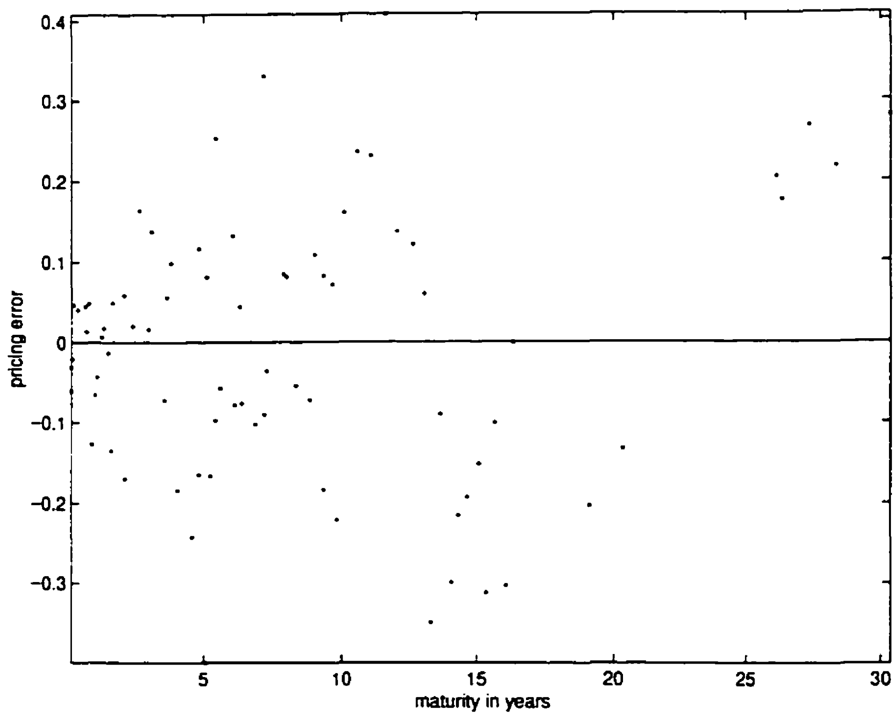

Figure 4.10: The pricing errors from the cross-sectional estimation of the term structure for the 1st of February 1995. The pricing errors are in Canadian dollars per $100 of face value. They are defined as $\hat{\epsilon_j} = (\hat{P}_j - P_j^M)$, where P_j^M is the bond market price and $\hat{P}_j$ is the estimated bond price from the OLS regression (3.12).

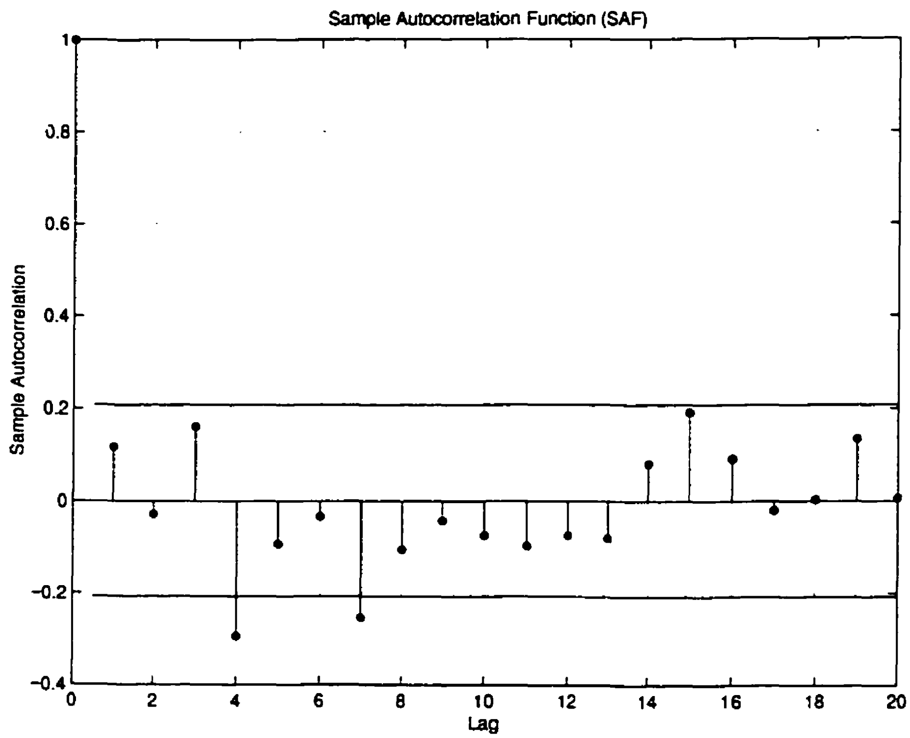

Figure 4.11: Autocorrelation of the residuals of the Canadian term structure estimation for the 2nd of July 1992.

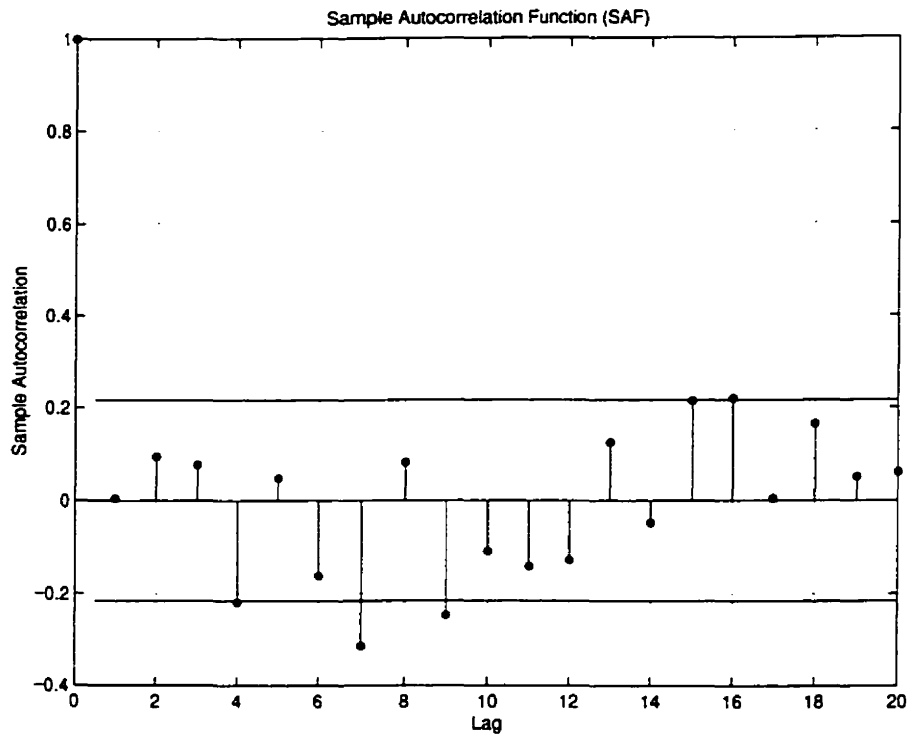

Figure 4.12: Autocorrelation of the residuals of the Canadian term structure estimation for the 30th of March 1993.

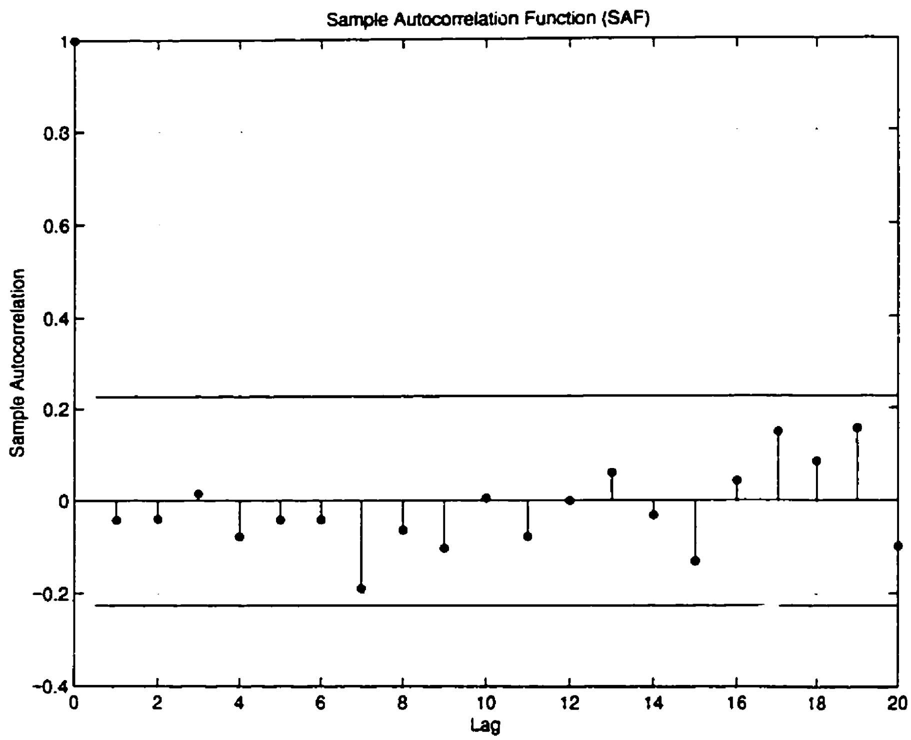

Figure 4.13: Autocorrelation of the residuals of the Canadian term structure estimation for the 28th of September 1994.

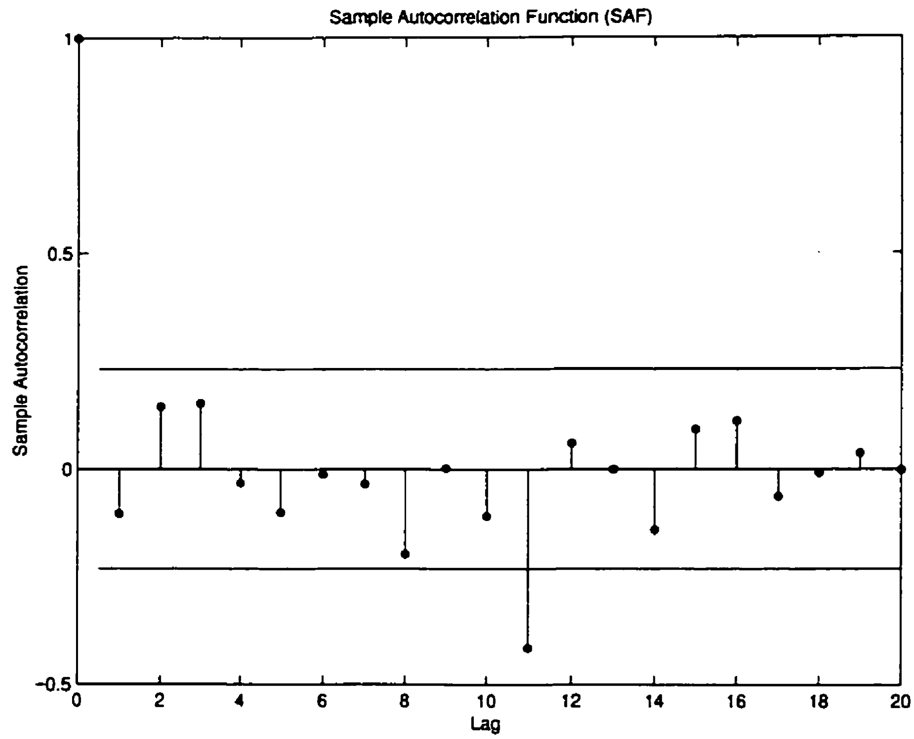

Figure 4.14: Autocorrelation of the residuals of the Canadian term structure estimation for the 1st of February 1995.

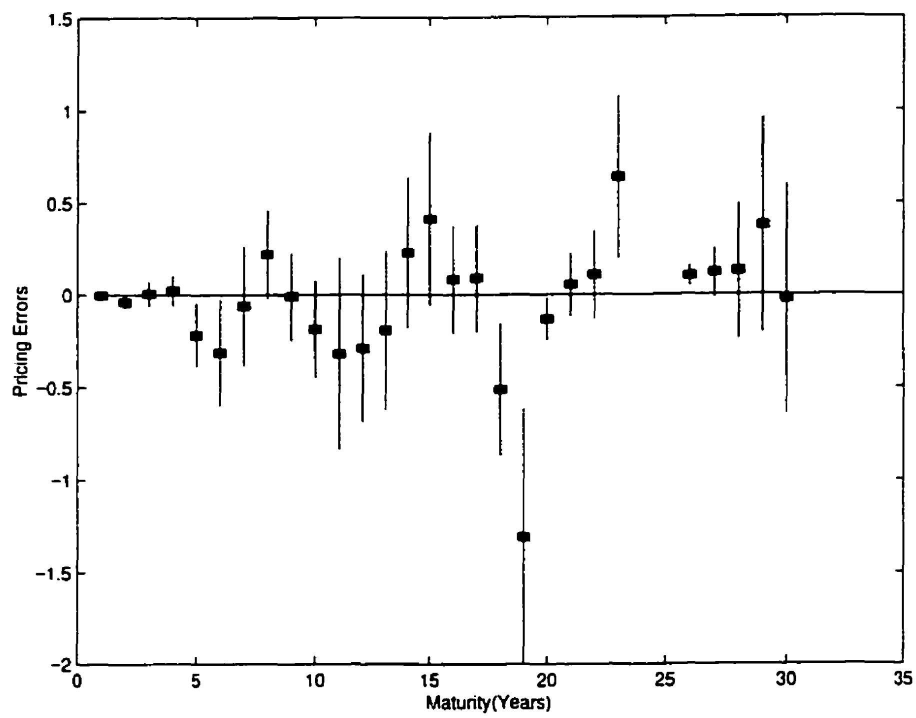

Figure 4.15: This figure shows the variation in pricing errors with maturity for 60,667 daily observations of bills, notes and bonds of Canadian governement for the period 1992-1995. The theoretical prices of bonds are estimated from the EP model. Errors are differences between estimated and actual prices. The units of the pricing errors are dollars per $100 in face value. The whiskers represent one standard deviation bounds on the time series means of the pricing errors within the respective maturity class. For the sample we used, there was no bond with maturities [23-24] years and [24-25] years.

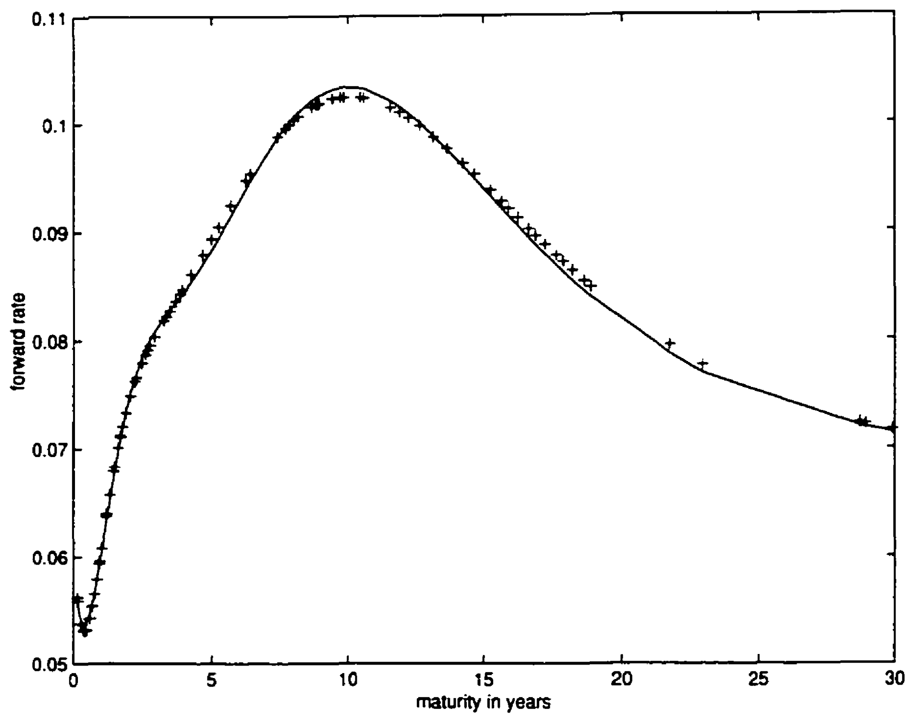

Figure 4.16: This figure is of the 2nd of July 1992. It illustrates the EP forward curve (solid line) as described by equation (3.14) and the proxy forward curve (plus signs) described by equation (3.17). The proxy forward function is closely related to the HJM re-indexed model. In this case the maximum error between the two functions is 18×10^{-4}.

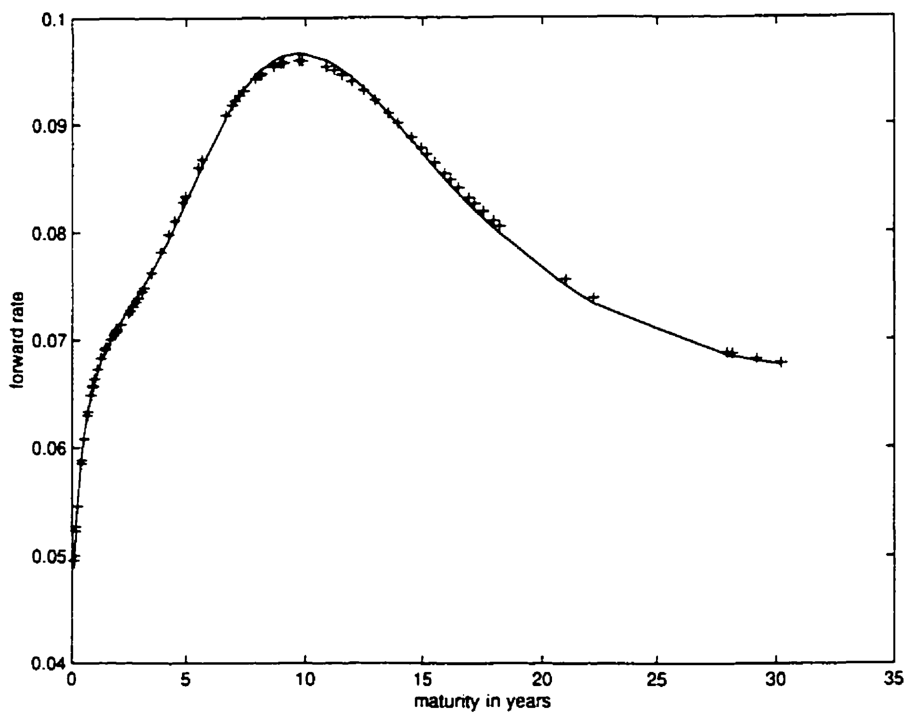

Figure 4.17: This figure is of the 30th of March 1993. It illustrates the EP forward curve (solid line) as described by equation (3.14) and the proxy forward curve (plus signs) described by equation (3.17). The proxy forward function is closely related to the HJM re-indexed model. In this case the maximum error between the two functions is 12×10^{-4}.

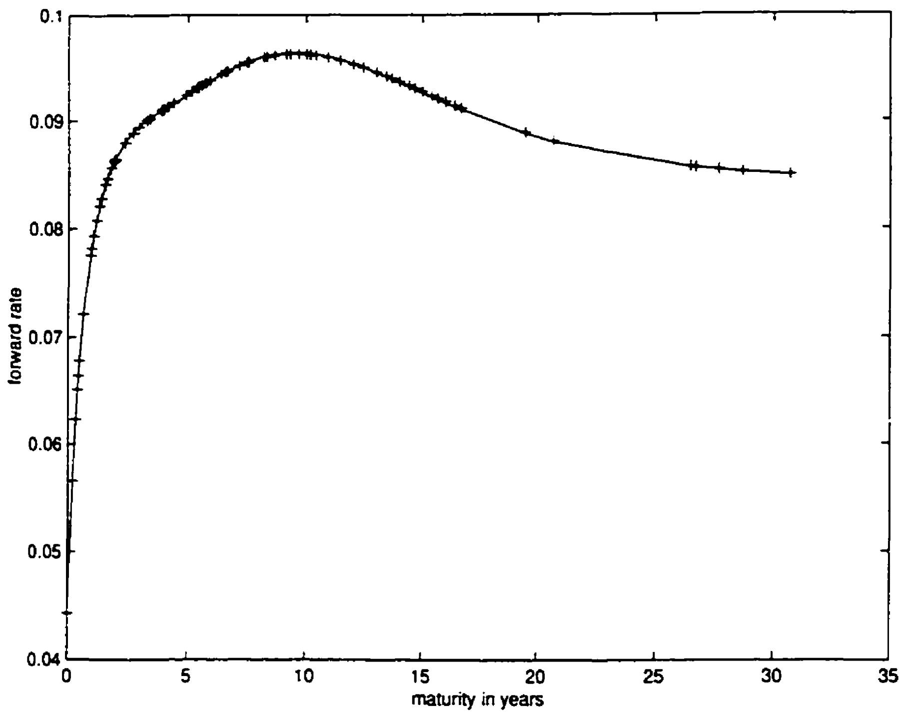

Figure 4.18: This figure is of the 28th of September 1994. It illustrates the EP forward curve (solid line) as described by equation (3.14) and the proxy forward curve (plus signs) described by equation (3.17). The proxy forward function is closely related to the HJM re-indexed model. In this case the maximum error between the two functions is 3.04×10^{-4}.

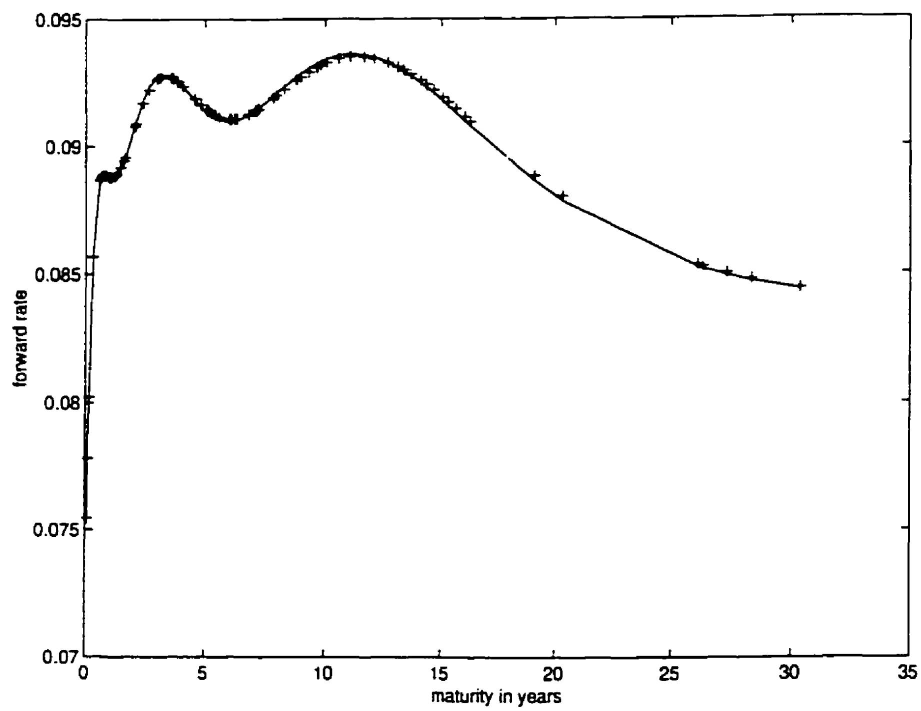

Figure 4.19: This figure is of the 1st of February 1995. It illustrates the EP forward curve (solid line) as described by equation (3.14) and the proxy forward curve (plus signs) described by equation (3.17). The proxy forward function is closely related to the HJM re-indexed model. In this case the maximum error between the two functions is 1.65×10^{-4}.

$RMSE$	$SRMSE$	MAE	Mean
0.4431	0.0369	0.3016	-0.0258

Table 4.1: The price errors are from 731 daily cross-sections of 60,667 daily prices for Canadian coupon bonds. The estimated prices are computed using the EP model. Errors are measured as differences between estimated and market prices. RMSE is the time series mean of the daily cross-sectional RMSE. SRMSE is the time series mean of the daily cross-sectional SRMSE. MAE is the time series mean of the daily MAE. Mean is the time series mean of the daily mean pricing errors. Measurement unit is dollars per $100 of face value.

Cross-sectional Results	Values	Std. err
Date 7/2/1992	-	-
Sample	93	-
Long rate (R)	0.0692	-
β_1	3.3398	0.4633
β_2	-4.7029	0.1.2005
β_3	2.4048	1.6094
β_4	-0.6771	1.3818
β_5	-0.0269	0.6976
β_6	0.0530	0.1724
Short rate(r)	0.0635	-
$RMSE$	0.8452	-
$SRMSE$	0.0644	-
MAE	0.5704	-

Table 4.2: Cross-sectional estimation results of the sample for the 2nd of July 1992.

Cross-sectional Results	Values	Std. err
Date 3/30/1993	-	-
Sample	86	-
Long rate (R)	0.0659	-
β_1	2.8763	0.3476
β_2	-3.8745	0.9061
β_3	1.7021	1.2197
β_4	-0.3368	1.0630
β_5	0.0147	0.5461
β_6	0.0025	0.1366
Short rate(r)	0.0467	-
$RMSE$	0.6610	-
$SRMSE$	0.0412	-
MAE	0.4437	-

Table 4.3: Cross-sectional estimation results of the sample for the 30th of March 1993.

Cross-sectional Results	Values	Std. err
Date 9/28/1994	-	-
Sample	78	-
Long rate (R)	0.0843	-
β_1	1.4292	0.1114
β_2	2.0972	0.2836
β_3	1.2284	0.3769
β_4	-0.4866	0.3290
β_5	0.0955	0.1681
β_6	-0.0219	0.0409
Short rate (r)	0.0441	-
$RMSE$	0.1668	-
$SRMSE$	0.0411	-
MAE	0.1250	-

Table 4.4: Cross-sectional estimation results of the sample for the 28th of September 1994.

Cross-sectional Results	Values	Std. err
Date 2/1/1995	-	-
Sample	76	-
Long rate (R)	0.0836	
β_1	1.6664	0.1070
β_2	-2.9762	0.2722
β_3	2.5894	0.3622
β_4	-1.5820	0.3190
β_5	0.5836	0.1675
β_6	-0.1151	0.0429
Short rate (r)	0.0700	-
$RMSE$	0.1638	-
$SRMSE$	0.0423	-
MAE	0.1318	-

Table 4.5: Cross-sectional estimation results of the sample for the 1st of February 1995.

Maturity Category	Mean	Std. dev	Maturity Category	Mean	Std. dev
0-1	-0.0027	0.0095	15-16	0.0821	0.2919
1-2	-0.0373	0.0379	16-17	0.0890	0.2924
2-3	0.0059	0.0681	17-18	-0.5111	0.3568
3-4	0.0230	0.0842	18-19	-1.3066	0.6912
4-5	-0.2158	0.1721	19-20	-0.1322	0.1149
5-6	-0.3114	0.2881	20-21	0.0541	0.1723
6-7	-0.0590	0.3224	21-22	0.1067	0.2415
7-8	0.2201	0.2413	22-23	0.6375	0.4424
8-9	-0.0070	0.2405	23-24	na	na
9-10	-0.1849	0.2612	24-25	na	na
10-11	-0.3189	0.5184	25-26	0.1007	0.0549
11-12	-0.2889	0.4002	26-27	0.1178	0.1351
12-13	-0.1906	0.4344	27-28	0.1281	0.3667
13-14	0.2282	0.4076	28-29	0.3763	0.5830
14-15	0.4130	0.4698	29-30	-0.0229	0.6220

Table 4.6: This table reports statistics on the pricing errors of the EP model. These results are based on Canadian daily cross-sections over the period 1992-1995. The errors are differences between estimated and actual prices. Units of pricing errors are in Canadian dollars per $100 of face value. There is no bond with maturities [23-24] and [24-25] years, during the sample period.

	β_1	β_2	β_3	β_4	β_5	β_6	R
β_1	0.6769	-0.9721	0.5845	-0.2599	0.0776	-0.0101	-0.0057
β_2	-0.9721	1.6098	-1.2727	0.7341	-0.2605	0.0429	0.0071
β_3	0.5845	-1.2727	1.4017	-0.9792	0.3819	-0.0693	-0.0025
β_4	-0.2599	0.7341	-0.9792	0.7422	-0.3015	0.0566	0.0002
β_5	0.0776	-0.2605	0.3819	-0.3015	0.1260	-0.0243	0.0001
β_6	-0.0101	0.0429	-0.0693	0.0566	-0.0243	0.0049	-0.0001
R	-0.0057	0.0071	-0.0025	0.0002	0.0001	-0.0001	0.0001

Table 4.7: Variance-covariance matrix of the original series for the state factors β_i for $i = 1, \ldots, 6$ and the long rate R for Canadian data.

	e_1	e_2	e_3	e_4	e_5	e_6	e_7
	3.8264	0.6336	0.0801	0.0013	0.0003	0.0000	0.0000
β_1	0.0270	0.0010	0.0213	0.5614	-0.4822	0.5823	-0.3347
β_2	0.0804	0.0055	0.1188	0.6084	-0.2180	-0.4174	0.6225
β_3	0.1963	-0.0181	0.3678	0.3945	0.4345	-0.3777	-0.5821
β_4	0.3431	-0.0795	0.5666	-0.0467	0.3657	0.5272	0.3755
β_5	0.5094	-0.1730	0.3518	-0.3773	-0.6010	-0.2514	-0.1417
β_6	0.6924	-0.2562	-0.6344	0.1138	0.1898	0.0533	0.0249
R	-0.3123	-0.9475	0.0534	0.0386	0.0177	-0.0073	0.0020

Table 4.8: The eigenvalues are in the second row of the table. The eigenvectors are in the columns below the corresponding eigenvalues. These results are for Canadian data. We used the covariance matrix of the state factors β_i and R to compute the eigenvalues as well as the eigenvectors.

Chapter 5

Time Series Analysis

5.1 Introduction

In the cross-sectional study of the EP model in Chapter 3, we have extracted the time series of ten state factors including the long rate R. All of them drive the term structure of interest rates. These unobservable state factors were obtained as estimated coefficients of a linear regression. The state factors in the EP model were assumed to follow the following process:

$$d\beta = \alpha(\beta, t)dt + \sigma(\beta, t)dz_t. \tag{5.1}$$

Unlike the majority of term structure models, the linear solution of the EP model does not impose any restriction on the drift of these factors or on their diffusion. Instead, the solution of the EP model was derived through an arbitrage argument similar to the APT approach. At this stage, we would like to study the time series properties of these state factors that were estimated in Chapter 3. Indeed, we showed that the estimated state factors, the β coefficients, were obtained from the following equation

$$\beta_t = (X'X)^{-1}X'y_t. \tag{5.2}$$

Equation 5.2 represents the connection between the estimated state factor vector β_t and the vector of bond prices y_t. $X = (x_j^i)$ is an $(N \times n)$ matrix of present values of bonds P_j discounted at a rate λ_i. Notice that in this equation, we introduced a time-dependent notation. Our objective is to study the time-series properties of the β coefficients. This investigation is interesting for two reasons. First, it will allow us to deepen our understanding of the processes behind the term structure models. Second, it will allow us to investigate risky arbitrage strategies using the EP model.

5.2 Statistical description of the estimated time series

In this section, we summarize the main statistical findings relative to the estimated state factors. Our data set consists of daily estimation of ten state factors from July 27, 1989 to October 15, 1996, for a total of 1805 observations. Moreover, we can see that the state factor, β_9, is not included in all the cross-sectional estimations. Indeed, the presence of this factor was necessary only in a certain environment. We relate this factor to a more turbulent macroeconomic environment. Table 5.1 reports the mean, standard deviation, skewness and kurtosis of the state factors. Skewness measures the symmetry of the distribution around its mean and kurtosis is a measure of how outlier-prone a distribution is, i.e., the tail thickness of the distribution. Figures 5.1, 5.2 and 5.3 show histograms of the values of the state factors with superimposed normal density functions. It appears from these figures that the state factors have different behaviors with different degrees of asymmetry around the mean. After examining these histograms, the skewness and the kurtosis coefficients of all the estimated factors, it is obvious that the assumption of conditionally normally distributed factors does not hold. Moreover, it can be observed from the same figures, that the distributions of the state factors β_i for $i = 7, \ldots, 9$

have a high concentration of values at zero. For β_7 and β_8, this is explained by the fact that the role of these state factors, in the estimation procedure, is to capture all the fine subtleties of the term structure movements. Thus for relatively "stable" economical environments, their role is barely needed and thereby their values are close to zero. For β_9, this is explained by the fact that this factor is not included in all the cross-sections. Thus, its value is confined to zero each time we use only nine factors instead of ten in the regression. For R, we restricted the lower boundary value of this state factor to 3.5%, this is motivated by some economic significance.

Figures 5.4 to 5.6 report the time series evolution of all the state factors. It seems from these figures that the evolution of all the state factors is governed by some regular shifts in volatility. This causes each series to change dramatically at some points in time. This phenomenon can also be observed in Figures 5.1, 5.2 and 5.3. Indeed, the distributions of these state factors, especially for higher order factors, i.e., β_i, for $i = 4, \ldots, 9$ and R, appear to be bimodal. The occurrence of the shifts and the bimodal distributions are the result of the same phenomenon. It is related to our estimation procedure described in Chapter 3. We believe that it is a "mathematical artifact" induced by high multicollinearity among the components, especially of higher order β_i for $i = 4, \ldots, 9$, of the exponential basis of the EP model. Thus, studying the original series of the state factors can be misleading. Indeed, we might be inclined to model a shift which is nothing more than an "artifact" induced by our previous estimation procedure. Instead, we suggest to study the distribution of the eigen principal components of the state factors in the EP model. The principal components are by definition orthogonal linear combinations of the state factors. In this context, orthogonality implies uncorrelated principal components. In Chapter 3, we conducted an eigen analysis from which we concluded that the EP model is driven by three principal components. The relationship between the state factors and their centred principal components is defined as follows

$$F_c = E'(B - Mean), \tag{5.3}$$

where B is a (10×1805) matrix containing the 10 state factors, $Mean$ is a (10×1805) matrix. Each row of $Mean$ represents the mean of a state factor. E is the (10×10) matrix of column eigenvectors ordered by the eigenvalues of B and F_c is a (10×1805) matrix containing the principal components of the state factors. An apostrophe denotes a matrix transpose. Each row of matrix F_c represents the series for a principal component, denoted by f_i for $i = 1, \ldots, 10$. The uncentred principal components of the state factors are defined as follows

$$F = E'B. \tag{5.4}$$

Centred and uncentred principal components are connected by the following relation

$$F = F_c + E'Mean. \tag{5.5}$$

The row elements of the second term in 5.5 are constant mean values that do not influence the time series properties of the principal components and are in fact small. Thus, we use the uncentred principal components instead of the centred principal components. Now, our objective is to study the time series properties of the first three uncentred principal components denoted by f_1, f_2 and f_3[1], and attempt to model them.

Figure 5.7 presents the time series of the three principal components. It is obvious that these series are smoother and less erratic than the times series of the original state factors. Moreover, their distributions do not exhibit the bimodal property observed with the original state factors. Figure 5.8 plots the corresponding three histograms with superimposed normal density functions. It is clear from this

[1] In the remaining text, the uncentred principal components will be referred to as principal components for short.

figure that the principal components have different distributions than the original state factors. First, the two-peak phenomenon is no longer present. Second, their kurtosis and skewness are not similar to those of the original state factors. Table 5.2 reports some descriptive statistics for the principal components series. The three components have heavy tails relative to a normal distribution. For instance f_1 has a kurtosis of 3.84 which is larger than 3. We study the times series of the three principal components in the next section.

5.3 Stationarity of the principal components

Traditionally, the examination of stationarity of a time series starts by computing the sample autocorrelation function (SAF) as well as the sample partial autocorrelation function (SPAF). The SAFs are computed as follows

$$\hat{\rho}_k(x) = \frac{\sum_{t=k+1}^{T} (x_t - \bar{x})(x_{t-k} - \bar{x})}{Ts^2}, \quad k = 1, 2...20, \tag{5.6}$$

where x is a general variable, T is the sample size, $\bar{x}$ is the mean of x, and s the sample standard deviation of x. If the autocorrelation coefficients are null for $k > q$, the variance of $\hat{\rho}_k(.)$ is

$$V(\hat{\rho}_k) = T^{-1}(1 + 2\hat{\rho}_1^2 + \ldots + 2\hat{\rho}_q^2). \tag{5.7}$$

The SPAF is calculated by fitting autoregressive models of increasing orders: the estimate of the last coefficient in each model is the sample partial autocorrelation, $\hat{\alpha}_{kk}$. Figure 5.9 illustrates the evolution of the SAF for all the three principal components, while Figure 5.10 presents the evolution of their SPAF. The shapes of the SAF lead us to conclude that all three principal components have SAF that are "infinite in extent", which is a behavior compatible with autoregressive models. The SPAF behavior seems to reach zero for certain lags in some principal components but reappears later on. However, in general we are more inclined to consider the SPAF to be finite in extent. The behavior of these two functions suggest the use of an autoregressive moving average ARMA model as a representation of these principal components. However, the fitting procedure cannot be applied until the stationarity of the principal components is examined. Indeed, before suggesting any model to fit the principal components, we start by checking whether the series are stationary or not, thereby suggesting an order of integration. The stationarity of a series implies the existence of one or multiple unit roots. The theory and practice of testing unit roots have been reviewed by many authors [2]. Here, we will present two types of tests:

- The Phillips and Perron (1988) test for unit roots.
- The Augmented Dickey and Fuller (1979) test.

Consider the following OLS estimation of the following regression function:

$$f_t = \alpha + \rho f_{t-1} + u_t, \tag{5.8}$$

where u_t is assumed to be normally distributed. The objective is to investigate whether the series described by f_t is stationary or not. Thus, we test the null hypothesis: $H_0 : \rho = 1$ against the stationary alternative $H_a : \rho < 1$. An obvious test statistic is the usual 't-ratio' of the estimate of $(\rho - 1)$ to its estimated standard error. Dickey and Fuller showed that this statistic does not have a Student's t distribution. Instead the distribution, which is denoted by τ_μ, has a specific distribution determined by Fuller (1976). Phillips and Perron (1988) generalize the unit root test when the errors u_t are assumed to be white noise to the case when u_t is serially correlated and possibly heteroskedastic as well. Hence

$$f_t - f_{t-1} = \psi(L)\epsilon_t, \tag{5.9}$$

[2] See for instance Hamilton (1994) and Harvey (1993) for a review of these tests.

where

$$\psi(L)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j},$$

with $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma^2$ and $E(\epsilon_t \epsilon_\tau) = 0$ for $t \neq \tau$. Phillips and Perron proposed estimating equation 5.8 by OLS even when u_t is serially correlated and then modifying the statistic to take into account the serial autocorrelation and potential heteroskedasticity in the disturbances [3]. Dickey and Fuller presented an alternative approach which accounts for serial correlation by including higher-order autoregressive terms in the regression function:

$$\Delta f_t = \alpha + \phi_1 f_{t-1} + \sum_{i=1}^{p-1} \phi_{i+1} \Delta f_{t-i} + \epsilon_t. \quad (5.10)$$

The augmented Dickey-Fuller test statistic is computed as $\tau_\mu = \frac{\hat{\phi}_1}{std(\hat{\phi}_1)}$. At this stage, it must be mentioned that the augmented Dickey-Fuller test is, in theory, only valid if the underlying process is indeed a finite autoregression. However, Said and Dickey (1984) showed that the augmented Dickey-Fuller test could still be justified on asymptotic grounds. Descriptive statistics such as the mean, standard deviation and selected autocorrelation coefficients as well as stationarity test statistics of the original and first difference series are reported in Table 5.3. We can see that the autocorrelation coefficients in the original time series of f_1 decay very slowly. Those of the day-to-day change are generally small, except for $\hat{\rho}_1$, and are not consistently positive or negative. The results of the formal augmented Dickey-Fuller non-stationarity test with $p = 4$, as well as the Phillips-Perron test indicate a strong rejection of the null hypothesis at the 5% significance level. Note that both tests have the same critical value. These results imply that the stationarity of the series is very likely. The same results are reported for the other two principal components in Tables 5.4 and 5.5.

[3] See Hamilton (1994) section 17.6 for an explicit formula of the adjusted statistics.

5.4 Modelling the principal components of the EP model

The results obtained in the previous section allow us to consider the three principal components as stationary processes. Thus, we choose to model them by ARMA(p,q) models.

5.4.1 Fitting an ARMA(p,q) model

In this section, we fit an ARMA(p,q) model to the principal component f_1. We will follow the same estimation procedure for all three principal components. An ARMA(p,q) process includes both autoregressive and moving average terms:

$$f_t = c + \phi_1 f_{t-1} + \ldots + \phi_p f_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \epsilon \theta_{t-q}. \tag{5.11}$$

In a more compact way, equation 5.11 can be rewritten as

$$(1 - \phi_1 L - \ldots - \phi_p L^p) f_t = c + (1 + \theta_1 L + \ldots + \theta_q L^q) \epsilon_t, \tag{5.12}$$

where L is the lag operator. The stationarity condition of this process requires that the roots of the characteristic equation

$$x^p - \phi_1 x^{p-1} - \ldots - \phi_p = 0, \tag{5.13}$$

are less than one in absolute value, i.e., they lie within the unit circle. An alternative way of expressing this condition is in terms of the lag operator

$$1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p = 0, \tag{5.14}$$

where the lag operator is simply replaced here with the scalar z. The stationarity condition is that the roots of 5.14 should all lie outside the unit circle. In other words, the absolute values of the roots of this equation must be strictly greater than one. If the stationarity condition is satisfied, the ARMA process turns out to be covariance stationary and both side of 5.12 can be divided by $(1 - \phi_1 L - \ldots - \phi_p L^p)$ to obtain

$$f_t = \mu + \psi(L)\epsilon_t, \tag{5.15}$$

where

$$\begin{cases} \psi(L) & = \dfrac{(1 + \theta_1 L + \ldots + \theta_q L^q)}{(1 - \phi_1 L - \ldots - \phi_p L^p)}, \\ \sum_{j=0}^{\infty} |\psi_j| & \leq \infty, \\ \mu & = c/(1 - \phi_1 - \phi_2 - \ldots - \phi_p). \end{cases}$$

Hence, the stationarity of an ARMA process depends entirely on the autoregressive parameters $(\phi_1, \phi_2, \ldots, \phi_p)$ and not on the moving average parameters $(\theta_1, \theta_2, \ldots, \theta_q)$. The estimation of these ARMA(p,q) coefficients is obtained through the minimization of the likelihood function using a Gauss-Newton algorithm. The results of the estimation are summarized in Table 5.6. Since the orders p and q are unknown, we suggest that the series can be modelled by some ARMA process of reasonably low order. Thus, we consider that $p = 0, \ldots, 2$ and that $q = 0, \ldots, 2$. This choice of orders may seem *ad hoc*, but for ARMA models fitted with p and q higher than 2, we find that the added coefficients to be statistically insignificant. Table 5.6 reports all the estimated parameters of the fitted ARMA(p,q) models. The estimation of all the models seems to imply reasonable values for the coefficients. For models with autoregressive order $p = 1$, the estimates of the autoregressive coefficients ϕ_1 are close to unity. For the models with $p = 2$, the sum of both parameters ϕ_1 and ϕ_2 is close to unity. This phenomenon is reported by many financial time series. For instance, for the AR(1) model, the estimate ϕ_1 is less than 1, yet, the series is very close to a random walk, i.e., $\phi_1 = 1$.

There are different selection criteria that may be used to choose a model of appropriate order. The most popular ones are the Akaike (1974) Information Criteria (AIC) and Schwarz (1978) criterion (BIC). BIC is strongly consistent since it determines the true model asymptotically, whereas AIC favors overparametrized models. Usually, the criteria are used such that

$$AIC(p^*, q^*) = \min AIC(p, q) \text{ for } p = 1, \dots P, \text{ and } q = 1, \dots Q.$$

From Table 5.6, we can see that both AIC and BIC select the pair of orders (1,1) and (2,1). We select the ARMA(1,1) model because it is more parsimonious than the ARMA(2,1) model. Now, the question arises: Can the selected model, i.e., ARMA(1,1), pass diagnostic checks on the residuals? To answer this question, we examine the properties of the residuals $\hat{\epsilon}_t$ of the estimated model. Figure 5.11 displays the sample SAF of the residuals for ARMA(1,1) with upper and lower 95% confidence bounds that are based on the assumption that all autocorrelations are zero beyond lag zero. Note here that the significance of the correlation coefficients are being tested individually and not simultaneously. About one in 20 coefficients would be expected to lie outside the bounds by chance under the hypothesis of no autocorrelation. As two coefficients lie outside the bounds, it appears that the residuals of the ARMA(1,1) model exhibit little or no correlation. Thus, we can consider this model as an adequate working model. Table 5.7 reports the results of the ARMA(p,q) models estimation for the principal components f_2 and f_3. After suggesting different orders for p and q, we conclude that the ARMA(1,1) model is suitable for both principal components. The autoregressive coefficient ϕ_1 is 0.98 for f_2 and 0.97 for f_3.

5.4.2 Testing for the presence of ARCH errors

Figure 5.12 presents a time series plot of the first differences of the three principal components. It can be seen from the figure that the daily estimated values are not homoskedastic. They are rather characterized by periods of tranquility followed by periods of more turbulent movements (a phenomenon known as volatility clustering). We have also seen previously that there may be a little serial dependence in the residuals of the ARMA(1,1) model. Indeed, the Ljung and Box (1978) portmanteau test for up to the twentieth order serial correlation of $\hat{\epsilon}$ for the series of f_1 equals 44.39, whereas the same test for the twentieth order serial correlation in the squared errors equals 221.13. Under the null hypothesis of identically and independently distributed principal components, both test statistics are asymptotically the realization of a chi-square distribution with twenty degrees of freedom (χ^2_{20}). However, it must be noted that, in the presence of ARCH, the portmanteau test for serial correlation in $\hat{\epsilon}_t$ tends to over-reject. The presence of ARCH can lead to serious model mispecification if it is ignored. Weiss (1984) showed that ignoring ARCH will lead to the identification of ARMA models that are overparametrized. Therefore, before deciding on the adequacy of the fitted model, we test for the presence of ARCH in the residuals of the model. Engle (1982) suggested a test based on the Lagrangian multiplier (LM) in which the null hypothesis is that ϵ_t possesses a constant conditional variance against the alternative that the latter is given by an ARCH(p) process.

This test is described as follows: First, the estimated residuals of the model are saved and then $\hat{\epsilon}_t^2$ is regressed on a constant and k of its lagged values:

$$\hat{\epsilon}_t^2 = a_0 + a_1\hat{\epsilon}_{t-1}^2 + \ldots + \hat{\epsilon}_{t-k}^2 + e_t, \quad \text{for} \quad t = 1, \ldots, T. \tag{5.16}$$

The statistics $T \times R^2$, from the regression of 5.16, converges in distribution to a χ^2 under the null hypothesis. It must be mentioned that some complications may arise when the alternative is a GARCH(l,k) process. However, in practice, this test is more useful in testing the squared residuals than determining whether the residuals follow an ARCH or a GARCH process. The results of the LM test as well as the Ljung-Box statistics of the squared residuals are reported in Table 5.8. According to those results, the residuals of the ARMA(1,1) model are heteroskedastic. This is a first indication that ARMA(1.1) is insufficient to capture the behavior of the principal component f_1. We applied the same test procedure on the principal components f_2 and f_3. We found that ARMA(1.1) model is insufficient to fully capture the behavior of these two principal components.

5.4.3 ARMA(p,q)/GARCH(l,k) model

In order to include the heteroskedastic aspect of the errors, we consider the following model:

$$ARMA\,(1,1): \quad f_1 = c + \phi_1 f_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}, \tag{5.17}$$

$$GARCH(2,1): \quad h_t = K + \omega_1 \epsilon_{t-1}^2 + \alpha_1 h_{t-1} + \alpha_2 h_{t-2}, \tag{5.18}$$

where h_t is the conditional variance of ϵ_t. With this model, we still assume that the mean of the process follows an ARMA(p.q) model however we allow the error term ϵ_t to have some form of heteroskedasticity. Thus. the conditional volatility of ϵ_t. as described by equation 5.18. is a GARCH(2.1) process. This model has been used by various authors for different financial series. The results of the GARCH estimation. for the principal component f_1, are presented in Table 5.9. The coefficients of the ARMA(1.1)/GARCH(1.1) indicate that α_1 and ω_1 are both highly significant at the conventional 5% level. Moreover. we note that the estimate of $(\omega_1 + \alpha_1)$ turns out to be very close to unity. The same remark is also valid for the ARMA(1.1)/GARCH(2.1) where $(\omega_1 + \alpha_1 + \alpha_2)$ is close to one. This result is very similar to other results using higher frequency financial data. Engle and Bollerslev (1986) called such a process an integrated GARCH(l,k), IGARCH(l,k) or GARCH(l.d.k), where $d = 1$ is the order of integration. In this class of models, the autoregressive polynomial equation modelling the variance in 5.18, has a unit root and consequently a shock to the conditional variance is persistent in the sense that it remains important for future forecasts of all horizons. As noted in Bollerslev, Engle, and Nelson (1994), the notion of "persistence" of a shock to volatility within the ARCH class of models is even more complicated than the corresponding concept of persistence (integration) in the mean for linear models (ARMA). However, the same authors claim that the IGARCH models are strictly stationary and ergodic. Nevertheless, they suspect that the apparent persistence of shocks may be driven by thick-tailed distributions of the errors rather than inherent non-stationarity.

In order to determine whether we have an integrated series or not, we computed the t-statistics for the null hypothesis that $(\omega_1 + \alpha_1 + \alpha_2) = 1$ as well as for $(\omega_1 + \alpha_1) = 1$. We found that the null hypothesis is rejected at the 5% significance level for both cases. Thus, we cannot accept the hypothesis that f_1 follows an IGARCH(1,k)) model. In our case, the GARCH(2,1) model seems to have a better explanatory power compared to the GARCH(1,1). The likelihood ratio (LR) test, which assesses the adequacy of a model relative to another nested model, is equal to 26.88 which is greater than χ_1^2 at the 5% significance level. Thus, it seems that an ARMA(1,1)/GARCH(2,1) is a good description of the empirical properties of the first principal component f_1. This result is confirmed by the results obtained from the Ljung-Box portmanteau test on the squared residuals of the estimated parameters as well as the LM test. Those results are reported in Table 5.10. The upper plot in Figure 5.13 shows the evolution of the standardized residuals (the residuals divided by their conditional standard deviation). They appear generally stable with little clustering. The lower plot of the same figure shows the SAF of the squared standardized residuals from the same ARMA/GARCH model, with 95 % confidence bands. It is clear from this figure that the SAF here shows no autocorrelation. For the remaining two principal components, we follow the same procedure as described in the present section. Table 5.11 reports our findings. All three principal components seem to be well modelled by an ARMA(1,1)/GARCH(2,1).

In summary, the ARMA/GARCH models suggested for all three principal components do capture the serial temporal dependencies in the volatility of these factors.

Thus, the highly significant portmanteau test for serial correlation of the squared residuals of the ARMA(p,q) models drops dramatically for the squared standardized residuals of all the ARMA/GARCH specification.

5.5 The properties of the estimated short rate

In Chapter 2, the short rate from the EP model was defined as a linear combination of all the state factors of the model. Since there is a large body of empirical literature devoted to examining the time series properties of the short rate data, we would like to study the process of the short rate inferred from the EP model. Usually, the empirical research on the time series properties of any unobservable economic variable, such as the short rate, requires the specification of a proxy. Figure 5.14 shows the evolution of the short rate implied by the estimated EP model and the evolution of the three-month T-bills yields. One can clearly see from this figure that both series have very similar evolution. In general, there is not a consensus among authors regarding the choice of a proxy for the short rate. As Chapman, Long, and Pearson (1999) concluded in their paper "...the proxy problem is economically significant...". In our case the proxy will be the implied short rate from the EP model. Many authors have tried to answer the following question: "What is the process followed by the short rate?" Chan, Karolyi, Longstaff, and Sanders (1992), Aït-Sahalia (1996), Duan and Jacobs (1998), Conley, Hansen, Luttmer, and Scheinkman (1997), Jiang (1998), and Ahn and Gao (1999) are some of the studies related to the subject. The empirical findings regarding the process of the short rate deals with two different aspects: the drift of the short rate and the diffusion of the short rate. Regardless of the proxy used, almost all studies conclude that the conditional volatility of the short rate exhibits heteroskedasticity. On the other hand, the estimation of the drift remains imprecise and still ambiguous. Aït-Sahalia

(1996) claims that the linearity of the drift imposed in the literature, especially in the affine models, is the main source of misspecification. However, Ahn and Gao claim that the drift function of some proxies of the short rate are consistent with the assumption of linear drift. The EP model did not impose any restriction on the drift or on the diffusion of the short rate. Our objective is to find a "good" model which fits the EP implied short rate.

5.5.1 The stationarity of the short rate

In section 5.3, using the Phillips-Perron and the augmented Dickey-Fuller tests, we concluded that the first three principal components of the EP model are stationary. The EP implied short rate is a linear combinations of the state factors. It is also (approximately) a linear combination of the three orthogonal principal components studied above. Given these findings, we claim that the EP implied short rate must follow a stationary process. In Table 5.12, we report some descriptive statistics for the implied short rate. The results are slightly different from what we obtained so far for the principal components. The null hypothesis of non-stationarity for the EP implied short rate is rejected at the 10% significance level. For the principal components, we were able to reject the non-stationarity hypothesis at the 5% significance level. This result is interesting in the sense that it can be compared to the findings of other authors. Jiang (1998) and Aït-Sahalia (1996) tested the stationarity of two different proxies of the short rate: the three-month Treasury bills rate and the 7-day Eurodollar rate. Both studies found a slight rejection of the null hypothesis. This result explains the reason why the short rate, especially in macroeconomics, is usually modeled as having a unit root and hence a non-stationary process, as pointed out by Conley, Hansen, Luttmer, and Scheinkman (1997)

5.5.2 A dynamic for the implied short rate

Vasicek (1977) assumed that the short rate can be written, in a discrete framework, as an AR(1) process

$$r_t = \delta(1-\phi) + \phi r_{t-1} + \sigma\epsilon_t, \tag{5.19}$$

where $\epsilon_t \sim NID(0, \sigma^2)$. This model is easy to implement. However, empirical evidence on the diffusion process of the interest rate goes against the assumption of a constant volatility of the short rate. Sun (1992) gave a discrete time version of the Cox, Ingersoll, and Ross (1985) model of bond pricing. He suggested that the square root process can be approximated by

$$r_t = \delta(1-\phi) + \phi r_{t-1} + \lambda r_{t-1}^2 \epsilon_t. \tag{5.20}$$

Despite the unusual form of the innovation, this relation is still considered as an AR(1). Gibbons and Ramaswamy (1993) estimated the ϕ parameter of equation 5.20. The parameter $\bar{\phi}$ will be interpreted as the implied first order autocorrelation coefficient. They found this value to be very small, 0.37, compared to the autocorrelation coefficient of 0.95, computed from a series of US Treasury bills rates. Moreover, Backus and Zin (1994) have showed that using an ARMA(p,q) model for the short rate can better accommodate the observable dynamic of the term structure. Figure 5.15 shows the evolution of the SAF and the SPAF of the EP implied short rate. The SAF decays at a very slow pace. However, the SPAF has a behavior similar to what is usually produced by an AR(p) model. Here, we will fit an ARMA(p,q) model to the EP implied short rate and analyze the results. Table 5.13 indicates that either an ARMA(2,1) or an ARMA(2,2) is a good description of the process of the short rate. The fitted models seem to be stationary. Our findings seem to head in the same direction as those found by Backus and Zin. We claim that previous term structure models imposed some simplifying assumptions on the process followed by the short rate. Those assumptions do not necessarily capture all the movements occurring in the bond markets. A good specification of the short rate process is a first step towards a better understanding of the term structure subject. Both Vasicek and CIR models assumed an AR(1) model for the short rate process. AR(1) is a special case of ARMA(p,q). Hence, fitting the EP implied short rate with an ARMA(p,q) model assures *a priori* a better description of the short rate. Nevertheless, we cannot conclude on the adequacy of the ARMA models unless the residuals of these models are checked. Figure 5.16 shows the evolution of the SAF of the residuals of the ARMA(2,1) and the ARMA(2,2) models fitted for the EP implied short rate. It appears that five autocorrelation coefficients lie outside the bounds for both models. This clearly violates the hypothesis of no autocorrelation of the residuals of the models. Moreover, the residuals of the ARMA models must be checked to determine whether they exhibit heteroskedasticity. Results from Table 5.14 imply that the two selected ARMA models may not be adequate for the short rate. In fact, the residuals of both models support the hypothesis of ARCH effects in the residuals as well as autocorrelation. This finding leads to considering an ARMA/GARCH specification for the short rate. Table 5.15 summarizes the results for a range of specifications. Given the values of some coefficients, it is clear that the ARMA(2,1) and ARMA(2,2) are not adequate as models of the mean of the short rate . Indeed, these two specifications imply non stationary processes. The sum of their autoregressive coefficients is very close to one. For instance, the model ARMA(2,1)/GARCH(1,1) implies a sum $\phi_1 + \phi_2 = 0.9991$. The hypothesis that this sum is statistically different from one is rejected at the 5% level of significance. This confirms the claim of Weiss (1984) that the existence of ARCH effects may lead to overparameterized ARMA models. On the contrary, ARMA(1,1)/GARCH(1,1) as well as ARMA(1,1)/GARCH(2,1) seem to be plausible descriptions of the implied short rate. Based on our estimation results, we find that the short rate exhibits heteroskedasticity and some persistence of volatility shocks. The result is in line with that of Duan and Jacobs (1998) who concluded that the short rate follows a GARCH(1,d,1) process. In this notation, "*d*" is the order of differencing of the series. Much of the analysis of financial time series considers the case when the order of differencing is either 0, i.e., the series is stationary, or 1, i.e., the series is integrated of order one. However, if "*d*" is a non-integer, the series is said to be fractionally integrated. Duan and Jacobs found that "*d*" is statistically different from 0 and 1. They reported that the series is fractionally integrated and thus concluded that the short rate has a long memory component. In our case, the results in Table 5.16 support the adequacy of the ARMA(1,1)/GARCH(2,1) model. Indeed, the residuals do not seem to be correlated or heteroskedastic. In Figure 5.17, we show a plot of the SAF of the squared standardized residuals of the ARMA(1,1)/GARCH(1,1) and ARMA(1,1)/GARCH(2,1) models. We notice that both models are adequate. However, on the basis of the LR test, we consider the ARMA(1,1)/GARCH(2,1) as a better description of the series. We tested whether the GARCH(2,1) model for the volatility of the short rate is integrated or not. The hypothesis is slightly rejected. This result implies that the order of integration is not equal to one. In a future study one must estimate the value of d to know whether the short rate is fractionally integrated or not.

Thus, the EP implied short rate and the three principal components are described by the same ARMA/GARCH model. The result for the EP implied short rate is not surprising as it is largely determined by the three main principal components.

In general, the findings here are encouraging and in line with the very recent development in the term structure literature. Indeed, more and more authors (see for instance Subrahmanyam (1996)) suggest the use of multi-factor models. Balduzzi, Bertola, and Foresi (1996) suggest a three-factor model of term structure: the short rate, the long-run mean of the short rate and the volatility of the short rate. They suggest that the volatility factor follows a GARCH process.

In Chapter 3, we gave empirical proof that the term structure is driven by more than one factor. In our case ten state factors were necessary for a good fit. However, we showed through an eigen analysis that only three principal components are needed to explain a large proportion of the term structure dynamic. In the time series study, we found that all three principal components can be modelled by an ARMA(1,1)/GARCH(2,1) process. Our results confirm the idea that any term structure model must allow for a rich dynamic which captures better the actual evolution of the structure. The simplifying assumptions made by previous term structure models are clearly no longer tenable.

5.6 The economic significance of the principal components

The empirical investigation of Chapter 3 found that daily US term structures over almost seven years can be estimated accurately by nine to ten state factors. We showed that these state factors have three principal components. We concluded that the EP model is largely determined by three principal components. One question rises: " Are the three principal components of the EP model, economically meaningful?" State variables in other term structure models have been assumed to be the short rate or the long rate or the volatility of the short rate. Our suspicion is that the principal components are correlated with macroeconomic variables which influence the term structure movements. Our objective in this section is to present empirical evidence which confirms our suspicion.

In the literature, many authors have empirically showed that some macroeconomic variables have an impact on interest rates and thereby on term structure. Urich and Wachtel (1984), McQueen and Roley (1993) reported that the Producer

Price Index (PPI) has an impact on interest rates. Hardouvelis (1987) found significant effects for the Consumer Price Index, the Trade Balance and the Unemployment Rates. Balduzzi, Elton, and Green (1996) found that announcements about Durable Goods Orders, Initial Jobless Claims and Nonfarm Payrolls affect the three-month bills prices. They also found that prices of medium and long maturity bonds are affected by the CPI, Durable Goods Orders, Housing Starts, Initial Jobless Claims, Nonfarm Payrolls, PPI, Consumer Confidence, National Association of Purchasing Managers (NAPM) Index, New Home Sales and M2 median. Our objective is to check whether the three principal components are correlated with macroeconomic variables. Of course, we do not expect to find perfect correlation. Nevertheless, the presence of correlation would be evidence of the connection between the principal components and the real macroeconomic environment.

First, we start by choosing a basket of nine macroeconomic variables. We include Housing Starts (Housing), Retail Sales (Retail). Monetary Aggregates M1 and M2, Yield of the Longest Bond (Long), the Federal Discount Rate (Fed), the Producer Price Index (PPI), the Consumer Price Index (CPI) and Durable Goods (DuraGoods). The data on these variables are published by the Federal Reserve Bank of Saint-Louis through their web site. Since most of these variables are published monthly, we select from the three principal components the values that correspond to the end of each month for the period between 1989 to 1996 (87 months). We construct two matrices: an (87×3) matrix containing the monthly principal components, an (87×9) matrix containing the monthly data on macroeconomic variables.

5.6.1 Procedure

Our objective is to examine the relationship between the three principal components of the state factors, inferred from the EP model, and the pool of macroeconomic variables using two procedures: simple correlation analysis and canonical correlation analysis.

5.6.1.1 Correlation analysis

We compute the simple correlations between the logarithms of the macroeconomic variables and:

1. the three principal components f_i for $i = 1, 2, 3$,
2. the short rate r,
3. the long rate R.

Let f_i denote the ith estimated state factor, and L_j the logarithm of the jth macroeconomic variable. Logarithms are used to reduce the scale of the raw data. The simple correlation coefficient between any pair of these two types of variables is computed using the following standard formula:

$$r_{ij} = \frac{s(f_i, L_j)}{s(f_i)s(L_j)}, \quad \text{for } i = 1, 2, 3, \text{ and } j = 1, \ldots, 9,$$

where

$$\begin{aligned} s^2(f_i) &= \frac{\sum_{l=1}^{n}(f_{l,i} - \bar{f}_i)^2}{n-1}, \\ s^2(L_j) &= \frac{\sum_{l=1}^{n}(L_{l,j} - \bar{L}_j)^2}{n-1}, \\ s(f_i, L_j) &= \frac{\sum_{l=1}^{n}(f_{l,i} - \bar{f}_i)(L_{l,j} - \bar{L}_j)}{n-1}, \quad \text{and } n = 87. \end{aligned}$$

Table 5.17 reports the correlation matrix of the macroeconomic variables L_j with the three principal components, the short rate r and the long rate R. f_1 has a relatively strong correlation with M1, PPI, and CPI. The principal components f_2 and f_3 as well as the long rate R are highly correlated with Fed. The EP implied short rate is highly correlated with Fed as well with M1 and M2. Moreover, the EP implied short rate is moderately related to all other macroeconomic variables. This last result is perfectly predictable since the short rate is a state variable intended to capture the evolution of the economy.

Based on this table, we cannot conclude if any macroeconomic variable is influencing one principal component in particular. Indeed, most of the correlation coefficients with the principal components are of the same order, therefore, big differences among the principal components are not evident. In order to gain better insight into the nature of the correlation between the principal components and the macroeconomic variables, we conduct a canonical analysis.

5.6.1.2 Canonical analysis

Gittins (1984) defines the canonical analysis in the following terms: "by canonical correlation analysis, we mean a technique of multivariate analysis which seeks linear functions of two sets of variables with special properties in terms of correlation irrespective of the nature of the variables comprising either set." Thus, the aim here is to clarify the relationship between what is called, in the language of canonical analysis, the two domains: the principal components and the pool of macroeconomic variables listed above. From the analysis we conduct, we will examine the significance of some coefficients and try to infer conclusions regarding the relationship between the principal components and the macroeconomic variables.

First, we must mention that the analysis is performed on standardized data (mean 0 and standard deviation 1). Second, u_k and v_k for $k = 1, 2, 3$, are defined as linear combinations of f_i and L_j variables, respectively. Finally, canonical pairs (u_k, v_k) are constructed to have maximum correlation between u_k and v_k, while distinct pairs are constrained to be mutually uncorrelated.

Table 5.18 reports the correlation coefficients r_k $(k = 1, 2, 3)$. It is evident that the correlation between u_1 and v_1 is strong since $r_1 = 0.86$. This suggests the existence of a linear relationship between the principal components and macroeconomic variables. The remaining r_k ($k > 1$) decline until the smallest coefficient r_3 attains 0.48. The strengths of the first three canonical correlations confirm the presence of several linear relationships between the two domains.

The squared correlation coefficients r_k^2 express the proportion of the variance of the kth canonical variate u_k, that is explained by its conjugate v_k, or vice versa. From Table 5.18, we find that $r_1^2 = 0.7452$, which means that 74.52% of the variation in the linear combination of the principal components specified by u_1 is attributable to the variation in that particular linear combination of the macroeconomic variable specified by v_1. From the magnitudes of r_1^2, r_2^2 and r_3^2, one can deduce that the overall relationship between the two domains is reasonably strong. We rely on the percentage attributable to the kth root, r_k^2, as a way to determine the dimensionality. From Table 5.18, it can be deduced that the roots r_k^2, for $k = 1, 2$, account for 84% of the predictable variance. The inclusion of the third root necessarily increases this percentage to 100% because the effective dimensionality of the linear association between the principal components and the macroeconomic variables is limited to the smaller number of variables in each domain (three in this case). This result implies that three linear relations may be considered to fully describe the effective dimensionality of the linear association between the principal components and the macroeconomic variables.

Table 5.19 reports different types of correlation of the principal components with the canonical variates u_k and of the macroeconomic variables with v_k. The intraset correlations (see the lower part of Table 5.19) measure the correlation of the macroeconomic variables with the canonical variates v_k. For instance, we note that the canonical variate v_1 is characterized principally by the macroeconomic variables, Fed (0.70), M1 (-0.47) and CPI (-0.43). v_2 is essentially a combination of Fed (0.36), Long (0.30) and Dura (0.26). v_3 is a contrast of percentage of M1 (-0.19), Ret (-0.17) and CPI (-0.15). Of the three canonical variates, v_1 is the strongest one, absorbing about 16% (the variance extracted = 0.1582) of the total variance associated with the macroeconomic variables. The second canonical variate accounts for only 4% of the total variance. The third canonical variate explains little of the total variance associated with the macroeconomic variables. Collectively, the v_k account for about 22% of the total variance of the macroeconomic variables.

The interset correlations are reported in the upper part of Table 5.20. They are the correlations of the estimated state factors with the canonical variates v_k. The magnitude of the correlations of the principal components with v_1 shows that all principal components are at least moderately related to v_1. Particularly, the first canonical variate of the macroeconomic variables domain is characterized by a positive weight to all the principal components. Since, we conclude previously that v_1 is highly related to Fed, we can suggest that the principal components f_1, f_2 and f_3 tend to be positively associated with Fed. v_2 is positively correlated with f_2 (0.65), while negatively correlated with f_1 (-0.23) and f_3 (-0.16). Knowing that v_2 is largely related to Fed and Long, we can affirm that f_2 is positively related to both Fed and Long. The canonical variate v_3 is positively related to f_1 (0.32) and negatively related to f_3 (-0.34) whereas it is poorly related to f_2. Since v_3 is negatively related to $M1$, then we can claim that f_1 and f_3 are positively related to $M1$.

Our findings show a strong link between the principal components of the state factors of the EP model and the macroeconomic environment, as expected. Moreover, as suggested by our canonical analysis, it seems that some principal components are particularly related to the Federal Discount Rate, to the Yield of the Longest Bond and to M1. Part of these results is plausible based on previous work in the literature. Indeed, Brennan and Schwartz (1979) developed a term structure model where one of the state variables is the return on the longest bond. Thus, our results confirm the relevancy of such variable but point to the presence of other factors that influence the term structure of interest rates. We think that a continuation of this investigation could include a wider choice of macroeconomic variables and will bring a better confirmation of our preliminary results.

5.7 Conclusion

This chapter studies the time series properties of the principal components estimated from the EP model. Our preliminary investigation revealed a mathematical artifact in the distribution of the estimated state factors. This artifact is induced by the high multicollinearity of the exponential components of the EP model. Thus, the analysis focused on the first three principal components of the state factors. They are orthogonal linear combinations of the state factors and explain up to 99.9% of the total variation in the EP model. After conducting a time-series analysis on these principal components, we concluded that all of them are stationary processes and can be described by a common ARMA(1,1)/GARCH(2,1) model. This result is interesting for several reasons:

1. It confirms recent findings in the literature about the processes followed by some state variables of term structure models. See for instance Balduzzi, Das, Foresi, and Sundaram (1996). They use a three-factor term structure model. One of the state variables, the volatility term, is described by a GARCH process.

2. The discovery that a common ARMA/GARCH model describes the principal components of the state variables opens the door for a better understanding of the term structure. It suggests a possible use of the EP model in arbitrage strategies. This idea, in particular, will be investigated in the next chapter.

3. The short rate, defined as a linear combination of all the state factors of the

EP model, is found to follow the same ARMA/GARCH process. This result confirms the widely held idea reported in the literature that the short rate process is heteroskedastic.

4. Finally, in our attempt to attribute some specific economic significance to the principal components, we discovered that the three principal components are related to macroeconomic variables such as the Yield of the Longest Bond, the monetary aggregate M1 and the Federal Discount Rate. This result is very promising because it shows that there is a relationship between the estimated principal components of the EP model and macroeconomic indicators related to the Federal reserve policy (Federal Discount Rate) and to monetary policy (M1). We suggest this as a subject of a future research.

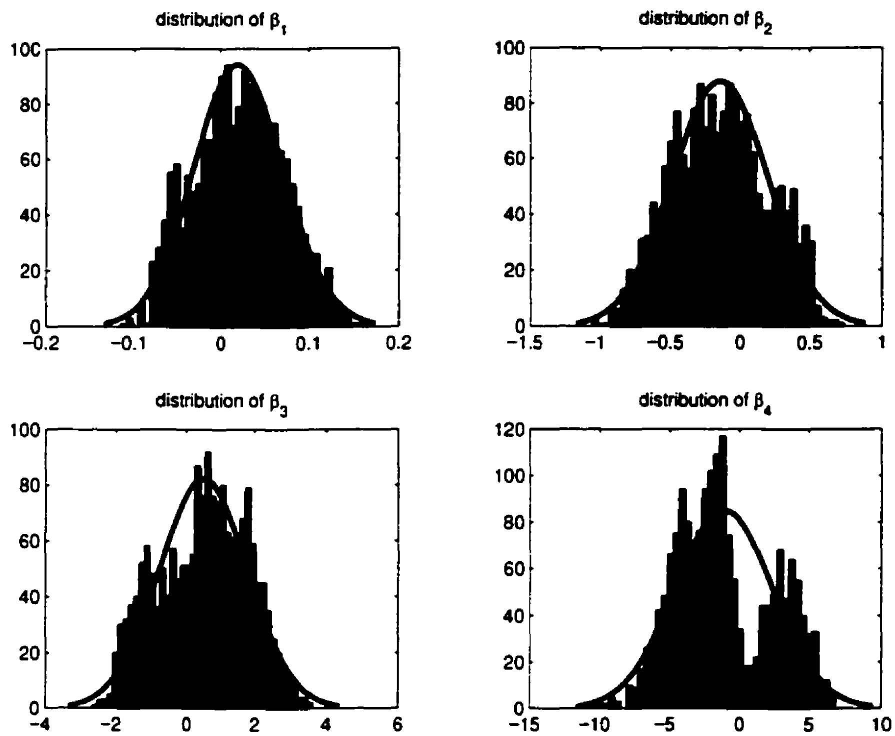

Figure 5.1: These figures depict the distibution of the state factors with a superimposed normal distribution for β_i for $i = 1 \ldots, 4$ estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period from 1989 to 1996.

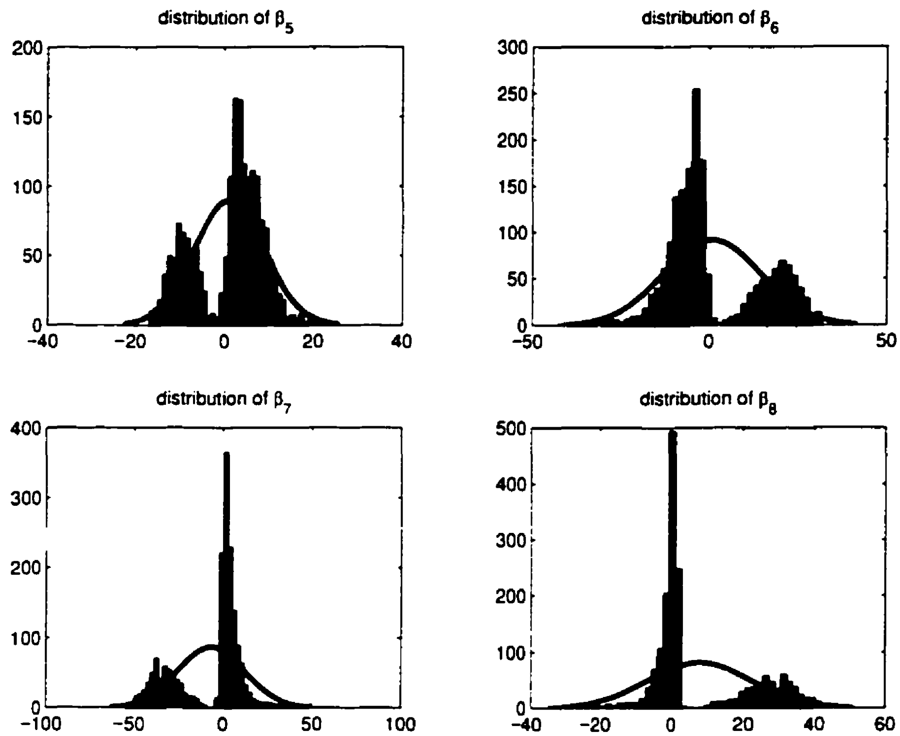

Figure 5.2: These figures depict the distribution of the state factors with a superimposed normal distribution for β_i for $i = 5 \ldots, 8$ estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period from 1989 to 1996.

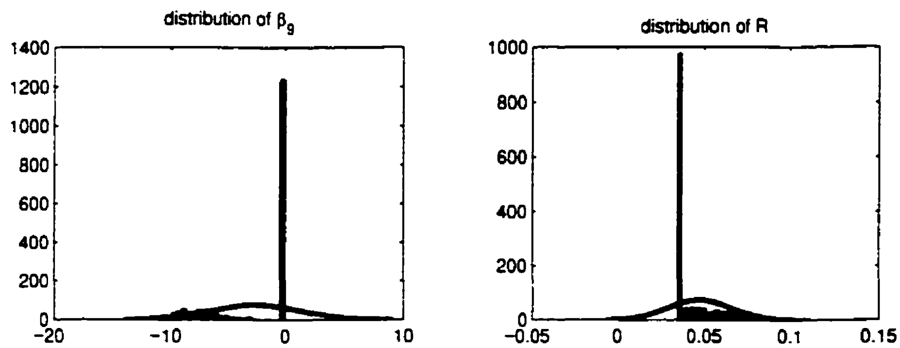

Figure 5.3: These figures depict the distribution of the state factors with a superimposed normal distribution. for β_9 and R estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period 1989 to 1996.

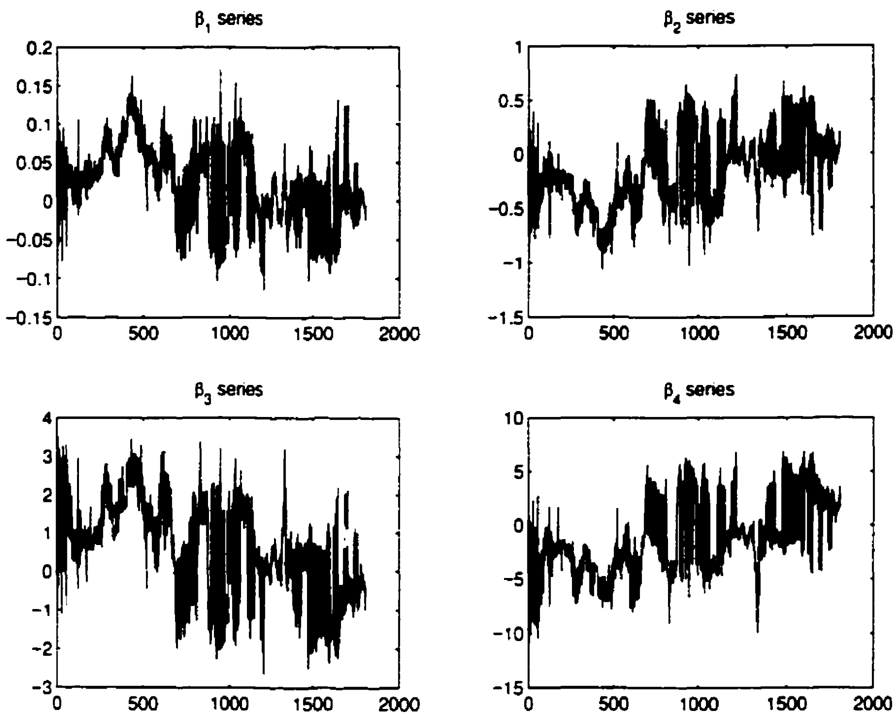

Figure 5.4: These figures depict the evolution of the state factors β_i for $i = 1 \ldots, 4$ estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period from 1989 to 1996.

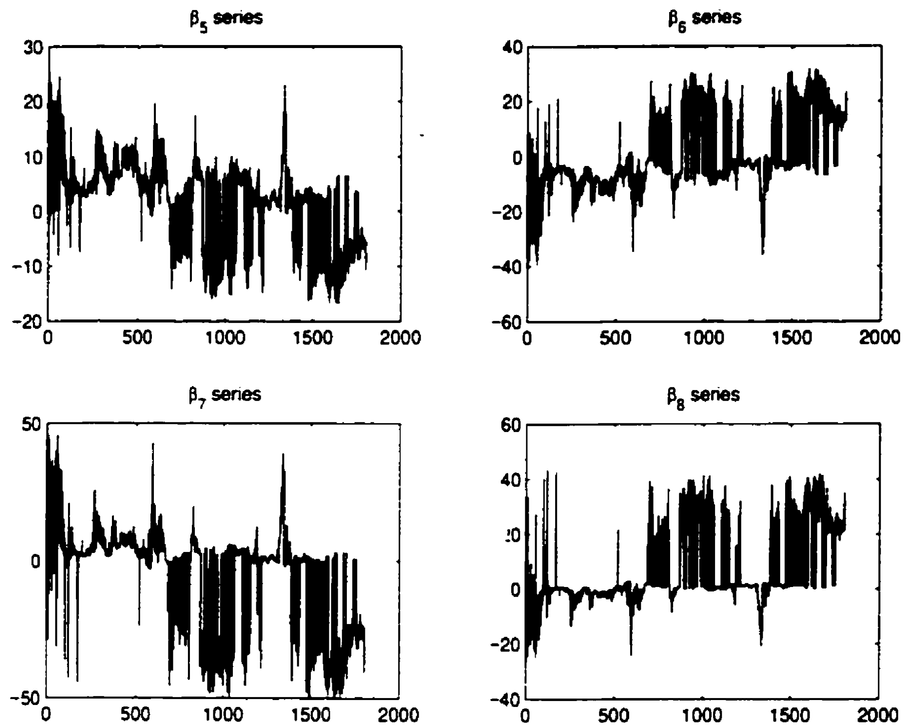

Figure 5.5: These figures depict the evolution of the state factors β_i for $i = 5 \ldots, 8$ estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period from 1989 to 1996.

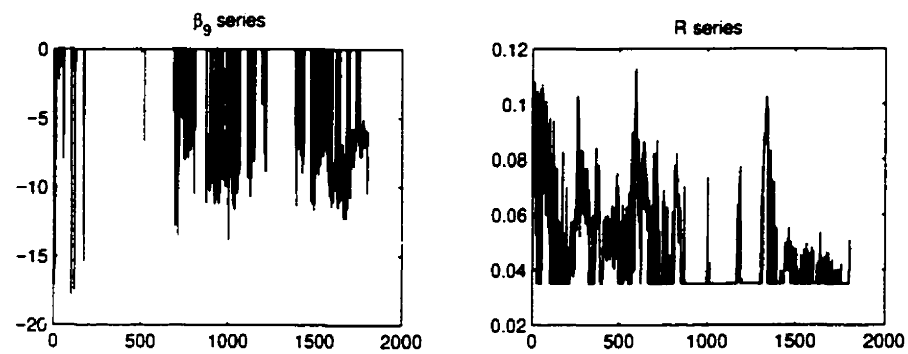

Figure 5.6: These figures depict the evolution of the state factors β_9 and R estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period 1989 to 1996.

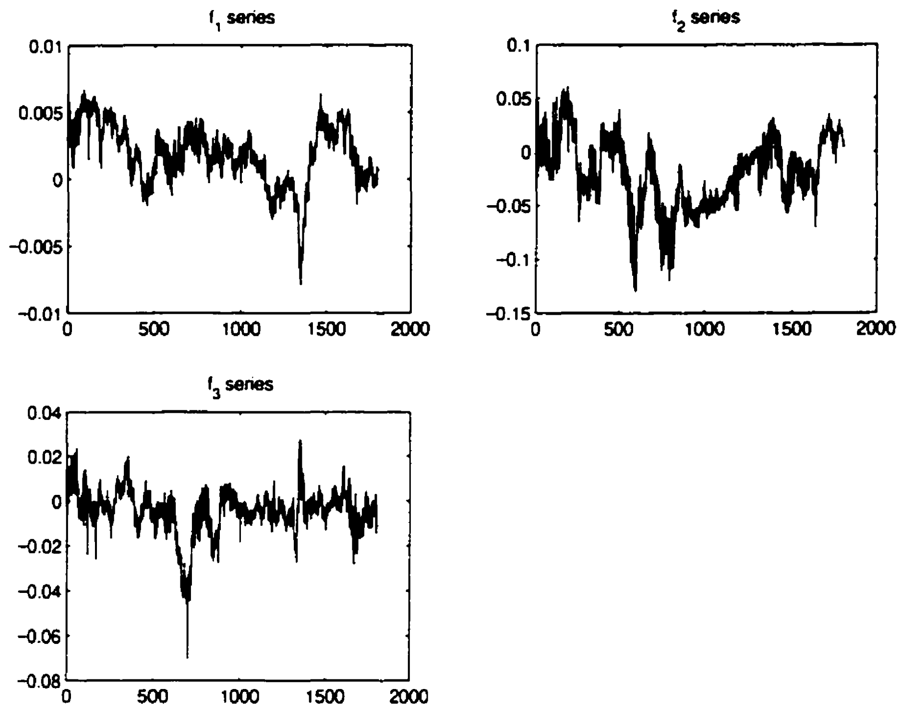

Figure 5.7: These figures depict the evolution of the principal components of the state factors f_i for $i = 1 \ldots, 3$ estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period from 1989 to 1996.

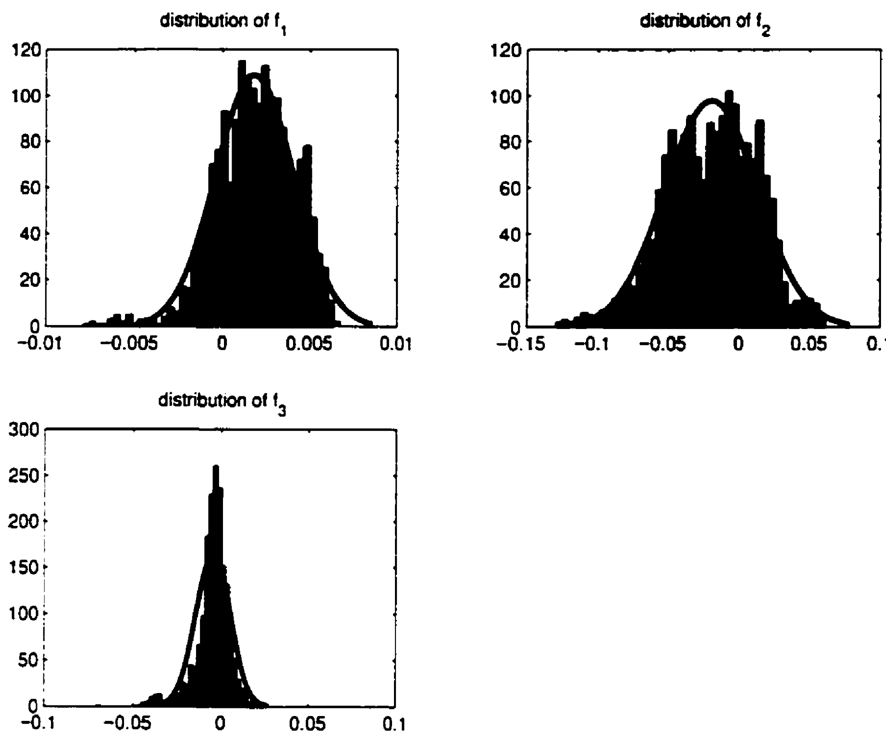

Figure 5.8: These figures depict the distribution of the principal components of the state factors f_i for $i = 1 \ldots, 3$ estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period from 1989 to 1996.

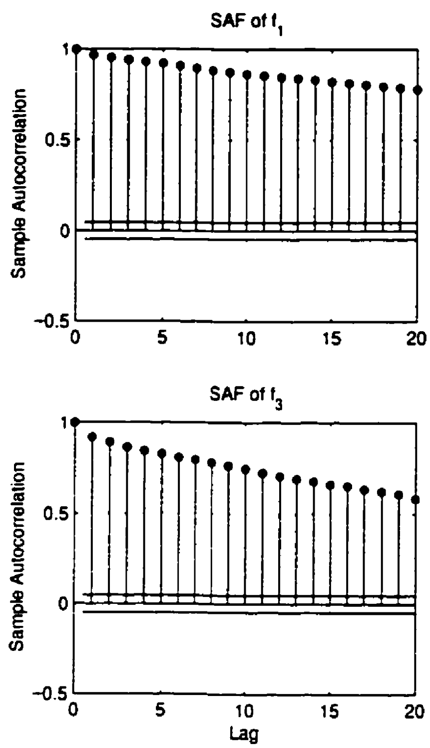

Figure 5.9: These figures depict the evolution of the SAF of the principal components f_i for $i = 1, \ldots, 3$ estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period 1989 to 1996. The two bands represent the upper and lower two standard deviation 95% confidence bounds, based on the assumption that all autocorrelations are zero.

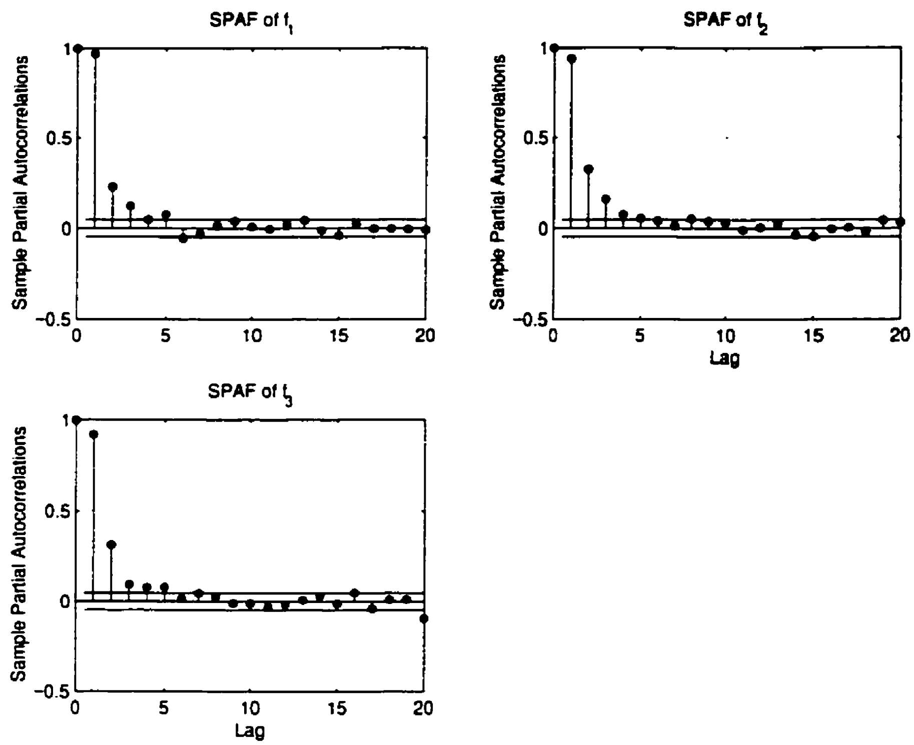

Figure 5.10: These figures depict the evolution of the SPAF of the principal components of the state factors f_i for $i = 1 \ldots, 3$ estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period 1989 through 1996. The two bands represent the upper and lower two standard deviation 95% confidence bounds, based on the assumption that all autocorrelations are zero.

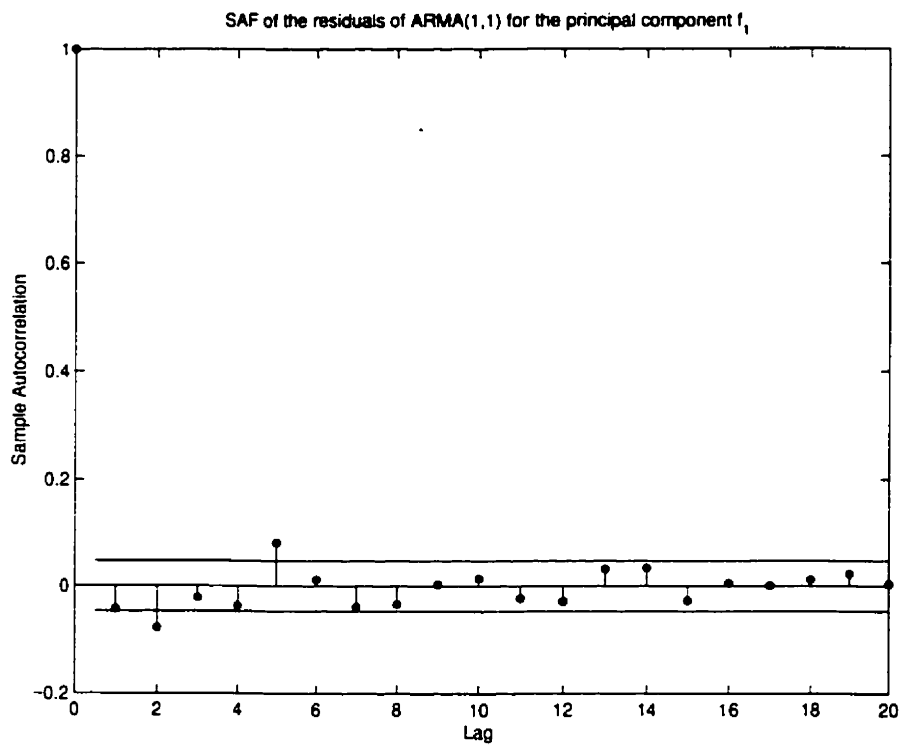

Figure 5.11: This figure represents the evolution of the SAF of the residuals of the ARMA(1,1) model for the principal component f_1.

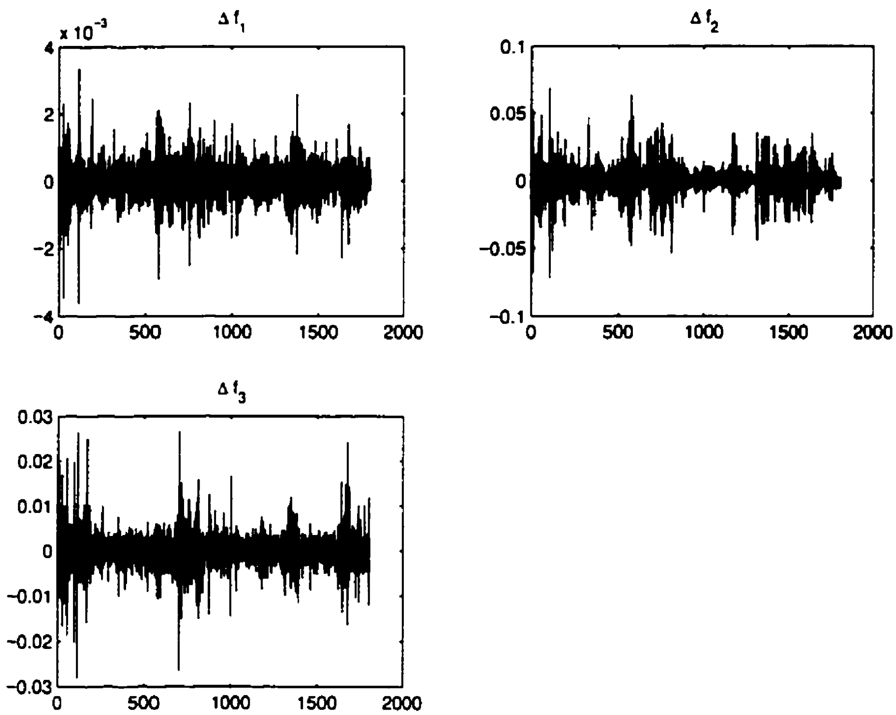

Figure 5.12: These figures depict the evolution of first differences of the principal components f_i for $i = 1 \ldots, 3$ of the state factors estimated from 1805 cross-sections of US Treasury bills, notes and bonds fit to the EP model for the period 1989 to 1996.

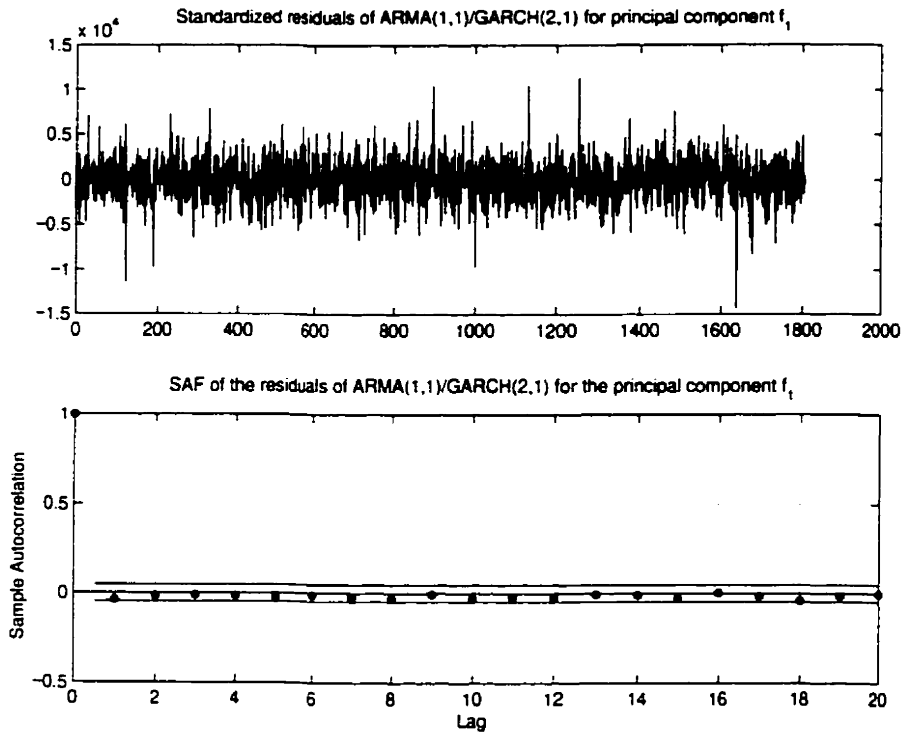

Figure 5.13: These figures illustrate the evolution of the standardized residuals and the SAF of the squared standardized residuals of the ARMA/GARCH model for the principal component f_1.

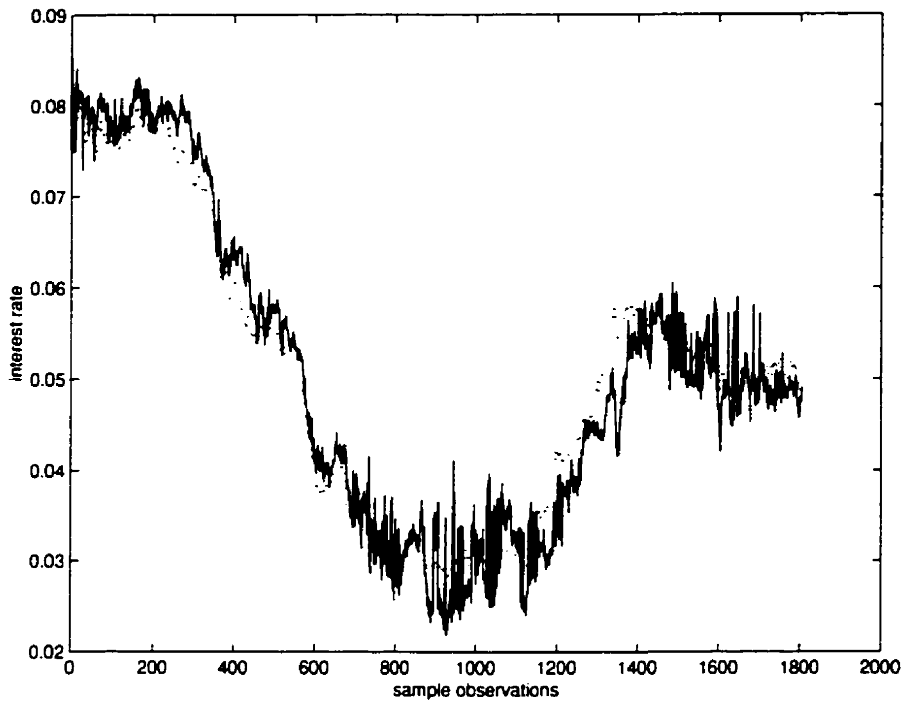

Figure 5.14: This figure presents the evolution of two time-series: the implied short rate (solid line) for the EP model estimated from 1805 daily cross-sections of U.S Treasury coupons bonds over the period of 1989-1996; the daily yields of three-months U.S T-bills (dots).

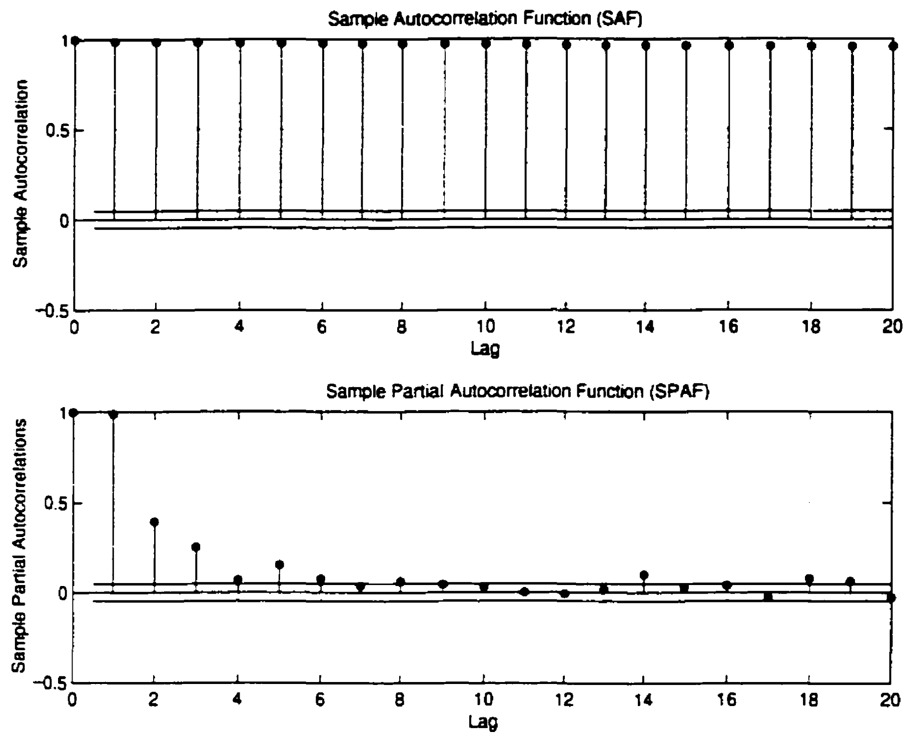

Figure 5.15: The SAF and the SPAF of the EP implied short rate.

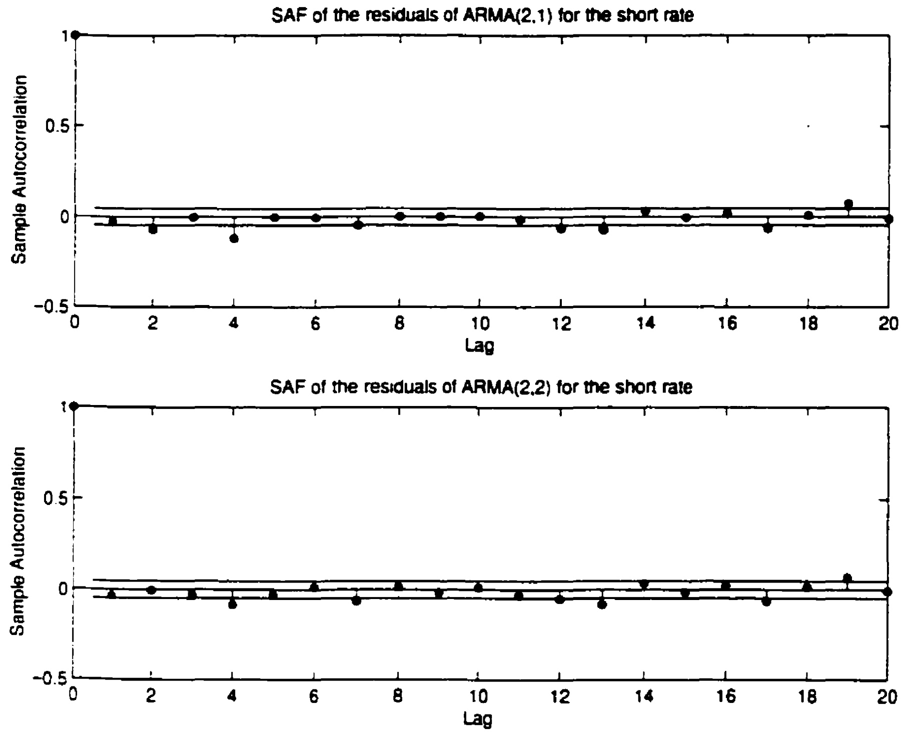

Figure 5.16: The SAF of the residuals of the ARMA(2,1) and ARMA(2,2) models for the EP implied short rate.

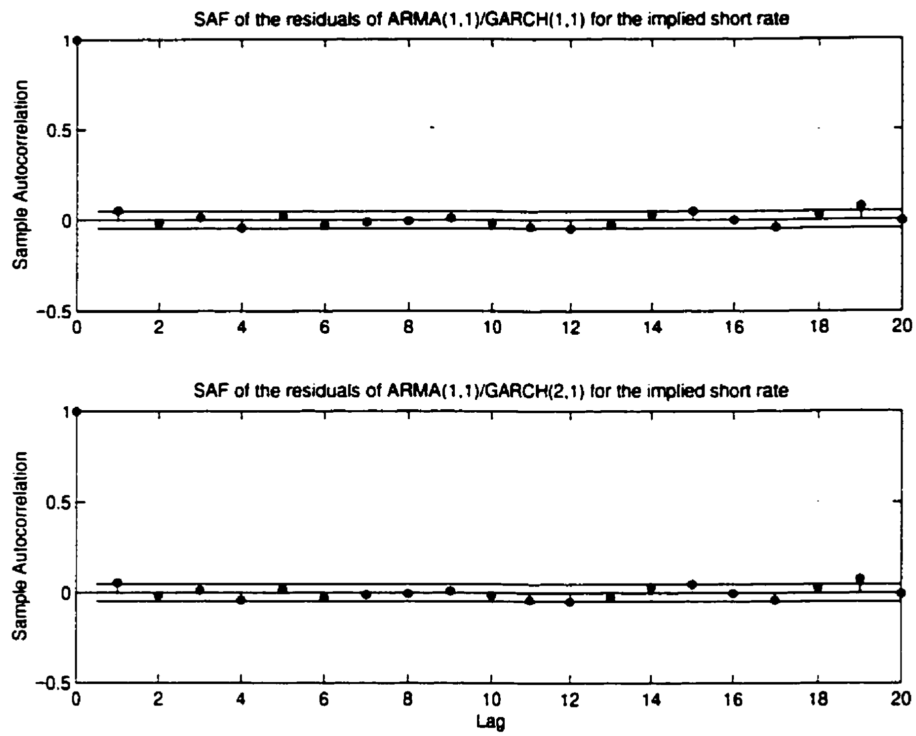

Figure 5.17: The SAF of the squared standardized residuals of two ARMA/GARCH models for the EP implied short rate.

factor	mean	std	skewness	kurtosis
β_1	0.0196	0.0510	0.0212	2.4136
β_2	-0.1435	0.3429	0.0535	2.3024
β_3	0.5097	1.2833	-0.1370	2.1994
β_4	-1.1045	3.5186	0.2397	2.2617
β_5	1.1424	7.9346	-0.3315	2.4347
β_6	0.8851	13.5593	0.4397	2.5403
β_7	-6.4032	18.9826	-0.6249	2.4935
β_8	7.8010	14.3568	0.8136	2.4099
β_9	-2.4429	3.8379	-1.1501	2.8346
R	0.0472	0.0177	1.4394	4.1407

Table 5.1: Descriptive statistics for series of 1805 observations on the estimated state factors β_i for $i = 1, \ldots, 9$ and R.

Component	mean	std	skewness	kurtosis
f_1	0.0018	0.0022	-0.5701	3.8413
f_2	-0.0197	0.0327	-0.2710	2.9321
f_3	-0.0044	0.0099	-1.1487	6.6558

Table 5.2: Descriptive statistics for series of 1805 on the first three principal components f_i for $i = 1, \ldots, 3$.

Variables	N	Mean	Std.Dev.	$\hat{\rho}_1$	$\hat{\rho}_3$	$\hat{\rho}_7$	$\hat{\rho}_{11}$	$\hat{\rho}_{15}$	$\hat{\rho}_{19}$
$f_{1,t}$	1805	0.0018	0.0022	0.9695	0.9419	0.8946	0.8544	0.8202	0.7869
$f_{1,t+1} - f_{1,t}$	1804	-0.00006	0.0368	-0.2423	0.0005	-0.0363	-0.0220	-0.0380 -0.0296	0.0163
Augmented daily Dickey-Fuller		-3.2151							
Phillips-Perron statistics		-86.0466							
H_0: Nonstationary		reject at 95% (critical value=-2.86)							

Table 5.3: Summary statistics of the first principal component f_1.

Variables	N	Mean	Std.Dev.	$\hat{\rho}_1$	$\hat{\rho}_3$	$\hat{\rho}_7$	$\hat{\rho}_{11}$	$\hat{\rho}_{15}$	$\hat{\rho}_{19}$
$f_{2,t}$	1805	-0.0197	0.0327	0.9407	0.9083	0.8580	0.8243	0.7827	0.7449
$f_{2,t+1} - f_{2,t}$	1804	1.6135×10^{-5}	0.0112	-0.3471	0.0006	-0.0407	-0.0133	-0.0277	0.0062
Augmented daily Dickey-Fuller		-3.7338							
Phillips-Perron statistics		-284.1847							
H_0: Nonstationary		reject at 95% (critical value=-2.86)							

Table 5.4: Summary statistics of the second principal component f_2.

Variables	N	Mean	Std.Dev.	$\hat{\rho}_1$	$\hat{\rho}_3$	$\hat{\rho}_7$	$\hat{\rho}_{11}$	$\hat{\rho}_{15}$	$\hat{\rho}_{19}$
$f_{3,t}$	1805	-0.0044	0.0099	0.9192	0.8656	0.7974	0.7272	0.6633	0.6102
$f_{3,t+1} - f_{3,t}$	1804	-8.9677×10^{-6}	0.0039	-0.3401	-0.0507	0.0067	-0.0098	-0.0516	0.0828
Augmented daily Dickey-Fuller		-4.6674							
Phillips-Perron statistics		-312.2564							
H_0: Nonstationary		reject at 95% (critical value=-2.86)							

Table 5.5: Summary statistics of the third principal component f_3.

Model	ϕ_1	ϕ_2	θ_1	θ_2	c	AIC	BIC
ARMA(1.0)	0.9760 (0.0056)				0.00004 (0.00001)	-15.0745	-15.0714
ARMA(2.0)	0.8057 (0.0230)	0.1745 (0.0229)			0.00003 (0.00001)	-15.1248	-15.1187
ARMA(0.1)			0.8340 (0.01299)		0.00189 (0.00005)	-13.2439	-13.2409
ARMA(0.2)			1.0682 (0.01804)	0.6309 (0.01804)	0.0018 (0.00006)	-13.7839	-13.7778
ARMA(1.1)	0.9858 (0.0042)		-0.2288 (0.0237)		0.00002 (0.00001)	-15.1434	-15.1343
ARMA(1.2)	0.9885 (0.0039)		-0.2164 (0.0238)	-0.0824 (0.0238)	0.00001 (0.00001)	-15.1391	-15.1330
ARMA(2.1)	1.2739 (0.0806)	-0.2817 (0.0792)	-0.5069 (0.07473)		0.00001 (0.000008)	-15.1432	-15.1341
ARMA(2.2)	0.7591 (0.2534)	0.2264 (0.2500)	0.0132 (0.2520)	-0.1348 (0.0592)	0.00002 (0.00001)	-15.1419	-15.1297

Table 5.6: Parameter estimates for the principal component f_1 from 1805 observations. The standard deviation of the parameters are enclosed in parentheses.AIC is the Akaike Information Criterion and BIC is the Schwarz criterion.

Component	Model	ϕ_1	θ_1	c	AIC	BIC
f_2	ARMA(1.1)	0.9841 (0.0052)	-0.3221 (0.0239)	0.0002 (0.0001)	-9.1433	-9.1372
f_3	ARMA(1.1)	0.9713 (0.0064)	-0.3053 (0.0242)	0.0001 (0.00006)	-11.2001	-11.1940

Table 5.7: Parameter estimates of ARMA models for the principal component f_2 and f_3. The standard deviation of the parameters are enclosed in parentheses.

order	LM test	Ljung-Box test	Critical Values
$k = 10$	189.7554	197.0073	18.3070
$k = 15$	195.3712	203.1289	24.9957
$k = 20$	201.1191	217.7110	31.4104

Table 5.8: Test for the presence of ARCH effects in the residuals of the ARMA(1,1) model for the principal component f_1.

Model	α_1	α_2	ω_1	$K \times 10^{-6}$	ϕ_1	θ_1	c	$LogLikelihood$
GARCH(1,1)	0.9209 (0.0014)		0.0600 (0.0021)	1.5391 (0.5492)	0.9725 (0.0012)	-0.2187 (0.0135)	0.0002 (0.00011)	11190.2316
GARCH(2,1)	0.4977 (0.0098)	0.2479 (0.0119)	0.1756 (0.0149)	2.2321 (0.3323)	0.9877 (0.0036)	-0.2438 (0.0236)	0.00002 (0.00001)	11203.6765

Table 5.9: Parameter estimates of the ARMA/GARCH model for the first principal component f_1.

order	LM test	Ljung-Box test	Critical Values
$k = 10$	4.7548	4.6256	18.3070
$k = 15$	6.0609	5.7149	24.9957
$k = 20$	8.8663	8.7130	31.4104

Table 5.10: Test for the presence of ARCH effects in the residuals of the ARMA(1,1)/GARCH(2,1) model for the principal component f_1.

Component	ARMA(p,q)/ GARCH(l,k)	α_1	α_2	ω_1	$K \times 10^{-6}$	ϕ_1	θ_1	c	$LogLikelihood$
f_2	(1,1)/(2,1)	0.3313 (0.1093)	0.5427 (0.1017)	0.1135 (0.0149)	1.6778 (0.1897)	0.9863 (0.0041)	-0.3829 (0.0231)	-0.0002 (0.00012)	5935.1037
f_3	(1,1)/(2,1)	0.3630 (0.1096)	0.4662 (0.1018)	0.1437 (0.0144)	0.4807 (0.07644)	0.9629 (0.0051)	-0.3114 (0.0233)	-0.0001 (0.00005)	7722.0964

Table 5.11: Parameter estimates of the ARMA/GARCH model for the principal components f_i, for $i = 2, 3$.

Variables	N	Mean	Std.Dev.	$\hat{\rho}_1$	$\hat{\rho}_3$	$\hat{\rho}_7$	$\hat{\rho}_{11}$	$\hat{\rho}_{15}$	$\hat{\rho}_{19}$
r_t	1805	0.0503	0.0171	0.9886	0.9837	0.9759	0.9698	0.9635	0.9589
$r_{(t+1)} - r_t$	1804	-2.011E-005	0.0024	-0.3977	0.0744	-0.0453	0.0102	-0.0237	0.0744
Augmented daily Dickey-Fuller		-2.61							
Phillips-Perron statistics		-67.24							
H_0: Nonstationary		reject at 90% (critical value=-2.57)							

Table 5.12: Summary statistics of the EP implied short rate.

Model	ϕ_1	ϕ_2	θ_1	θ_2	c	AIC	BIC
ARMA(1,0)	0.9885				0.0005	-12.0295	-12.0264
	(0.0033)				(0.0001)		
ARMA(2,0)	0.5979	0.3942			0.0004	-12.1972	-12.1911
	(0.0216)	(0.0216)			(0.0002)		
ARMA(0,1)			0.8974		0.0503	-9.2887	-9.2856
			(0.0095)		(0.00043)		
ARMA(0,2)			1.2148	0.7087	0.0504	-9.9632	-9.9571
			(0.015)	(0.0158)	(0.0004)		
ARMA(1,1)	0.9992		-0.3477		0.000008	-12.2311	-12.2251
	(0.0019)		(0.0237)		(0.0001)		
ARMA(2,1)	0.9903	0.0085	-0.5048		0.000019	-12.2763	-12.2671
	(0.0320)	(0.0319)	(0.0321)		(0.00007)		
ARMA(2,2)	0.1875	0.8106	0.3011	-0.4828	0.00002	-12.2836	-12.2714
	(0.0244)	(0.0245)	(0.0309)	(0.0233)	(0.0001)		

Table 5.13: Parameter estimates of the ARMA model fitted to the EP implied short rate r (original series) from 1805 cross-sectional samples.

Model	Order	LM test	Ljung-Box test	Critical Values
ARMA(2,1)	$k = 10$	280.3327	602.4590	18.3070
	$k = 15$	297.6179	725.9694	24.9957
	$k = 20$	341.0450	978.8884	31.4104
ARMA(2,2)	$k = 10$	266.3816	569.0741	18.3070
	$k = 15$	286.4189	705.8384	24.9957
	$k = 20$	329.3882	954.7532	31.4104

Table 5.14: Test for the presence of ARCH effects in the residuals of the models ARMA(2,1) and ARMA(2,2) fitted to the EP implied short rate.

ARMA(p,q)/ GARCH(1,k)	α_1	α_2	ω_1	$K \times 10^{-6}$	ϕ_1	ϕ_2	θ_1	θ_2	c	*LogLikelihood*
(1,1)/(1,1)	0.9296		0.0606	0.0174	0.9957				0.0002	8943.2221
	(0.0029)		(0.0049)	(0.0003)	(0.0008)				(0.00004)	
(1,1)/(2,1)	0.3928	0.4980	0.0909	0.0239	0.99585		-0.4529		0.00023	8946.7807
		(0.0025)	(0.0073)	(0.0005)	(0.0007)		(0.0205)		(0.00003)	
(2,1)/(1,1)	0.9261		0.0639	0.0137	1.2443	-0.2452	-0.6477		0.000147	8949.0487
	(0.0030)		(0.0050)	(0.0034)	(0.0041)	(0.0041)	(0.0128)		(0.00002)	
(2,1)/(2,1)	0.3685	0.5156	0.1058	0.0242	1.2361	-0.2368	-0.64834		0.00014	8953.0557
	(0.0035)	(0.0043)	(0.0079)	(0.0056)	(0.0496)	(0.0493)	(0.0338)		(0.00003)	
(2,2)/(1,1)	0.9224		0.0675	0.0166	1.3913	-0.3924	-0.7987	0.0703	0.00011	8950.8727
	(0.0032)		(0.0054)	(0.0037)	(0.0110)	(0.0110)	(0.0135)	(0.0194)	(0.00002)	
(2,2)/(2,1)	0.4025	0.4829	0.1044	0.02367	1.7462	-0.7470	-1.1740	0.2673	0.00004	8958.0422
	(0.0019)	(0.0026)	(0.0078)	(0.0056)	(0.0116)	(0.0115)	(0.0109)	(0.0160)	(0.0001)	

Table 5.15: Parameter estimates of the ARMA/GARCH models fitted to the EP implied short rate (original series) from 1805 cross-sectional samples.

order	LM test	Ljung-Box test	Critical Values
$k = 10$	7.5511	8.0044	18.3070
$k = 15$	10.6334	10.6554	24.9957
$k = 20$	15.3612	15.6247	31.4104

Table 5.16: Test for the presence of ARCH effects in the residuals of the ARMA(1,1)/GARCH(2,1) model for the EP implied short rate.

Variable	Housing	Retail	M1	M2	Long	Fed	PPI	CPI	Dura-Goods
f_1	-0.1096	-0.3565	-0.4460	-0.3929	0.0822	0.3551	-0.4058	-0.4232	-0.2498
f_2	0.1151	0.0734	-0.1710	-0.0526	0.3558	0.5799	-0.0780	-0.0739	0.2095
f_3	-0.0507	-0.0938	-0.1707	-0.2604	0.0123	0.3537	-0.1930	-0.1970	-0.0383
r	0.1978	-0.4436	-0.6868	-0.6237	0.6333	0.9178	-0.6095	-0.6234	-0.2588
R	-0.2367	-0.4620	-0.4973	-0.4934	0.5223	0.3514	-0.4974	-0.4954	-0.4037

Table 5.17: Simple correlation coefficients between the principal components, the short rate r and the long rate R and selected macroeconomic variables based on monthly data 1989-1996.

k	r_k	r_k^2	$\%tr$	$C\%$
1	0.8632	0.7452	50.80	50.80
2	0.7004	0.4906	33.45	84.25
3	0.4808	0.2312	15.75	100.00
Total	-	1.4669	100.00	-

Table 5.18: Canonical correlation coefficients. Relationships between the first three principal components and nine macroeconomic variables.

Canonical Variate	u_1	u_2	u_3
f_1	0.6567	-0.3381	0.6741
f_2	0.3399	0.9395	0.0415
f_3	0.6565	-0.2412	-0.7147
Var.Ext	0.3260	0.3517	0.3223
Canonical Variate	v_1	v_2	v_3
Hou	-0.0590	0.1453	-0.0322
Ret	-0.2677	0.1828	-0.1764
M1	-0.4761	-0.0010	-0.1983
M2	-0.4512	0.1111	-0.0876
Long	0.2068	0.3007	0.0713
Fed	0.7005	0.3626	0.0263
PPI	-0.4262	0.0778	-0.1477
CPI	-0.4387	0.0870	-0.1568
Dura	-0.1057	0.2672	-0.1336
Var.Ext	0.1582	0.0415	0.0166

Table 5.19: Canonical analysis of principal components-macroeconomic variables. Correlation between the original variables and the canonical variates.

Canonical Variate	v_1	v_2	v_3
f_1	0.5669	-0.2368	0.3241
f_2	0.2934	0.6581	0.0200
f_3	0.5667	-0.1689	-0.3436
Var.Ext	0.2429	0.1726	0.0745
Canonical Variate	u_1	u_2	u_3
Hou	-0.0683	0.2075	-0.0670
Ret	-0.3101	0.2609	-0.3669
M1	-0.5516	-0.0014	-0.4124
M2	-0.5227	0.1587	-0.1822
Long	0.2396	0.4293	0.1483
Fed	0.8115	0.5177	0.0546
PPI	-0.4938	0.1110	-0.3072
CPI	-0.5082	0.1243	-0.3261
Dura	-0.1224	0.3814	-0.2780
Var.Ext	0.2124	0.0846	0.0717

Table 5.20: Canonical analysis of principal components-macroeconomic variables. Correlation between the original variables and the canonical variates. Dual of Table 5.19.

Chapter 6

Out-of-Sample Results and Arbitrage

In the previous chapter, we concluded that the principal components of the state factors, inferred from the EP model, follow GARCH processes that we have fully identified. In theory, the dependence implied by a GARCH model, means prediction is possible. Our final objective is to investigate in practical terms, if useful prediction is possible in economic terms.

First, we will construct two types of out-of-sample GARCH forecasts by using past information on the principal components, available at the time the forecasts are made. Then, we judge the forecasting performance of each series. Second, the forecasting series will be used to determine whether we can exploit the predictive power of the GARCH forecasts. In other words, we assess whether profits can be generated from an arbitrage strategy involving the out-of-sample GARCH forecasts of the three principal components of the EP model.

6.1 Forecasting using GARCH processes

In chapter 5, we found that the principal components f_i, for $i = 1, \ldots, 3$, possess the following type of ARMA/GARCH process:

$$f_{t,i} = c + \phi_1 f_{t-1,i} + \epsilon_t + \theta_1 \epsilon_{t-1,i}, \tag{6.1}$$

$$h_{t,i} = K + \omega_1 \epsilon_{t-1,i}^2 + \alpha_1 h_{t-1,i} + \alpha_2 h_{t-2,i}. \tag{6.2}$$

Since, we are dealing with models where time-dependent conditional heteroskedasticity is present, it is not obvious how to write the expression of the optimal forecasts for the mean of such processes.

Baillie and Bollerslev (1992) determined the general expression of the optimal forecasts in such a context. Applying their general formula to our case and going through algebraic manipulations, we recover the expression for the optimal s-step-ahead predictor of $f_{t,i}$ as

$$E_t(f_{t+s,i}) = c(1 + \phi_1 + \ldots + \phi_1^{s-1}) + \phi_1^s f_{t,i} + \theta_1 \phi_1^{s-1} \epsilon_{t,i}. \tag{6.3}$$

See Appendix D page 202 for the relevant technical development. Given this expression for the mean of the GARCH forecasts, we use the GARCH specification identified in the previous chapter for each principal component and compute the corresponding GARCH forecasts. We estimate both rolling and updating GARCH forecasts. The rolling forecast uses a constant sample size of 1604 observations for each principal component. We start the estimation procedure at January 1, 1996. After making the forecasts for date $(t + 1)$ using observations from $(t - 1604)$ to t, all the GARCH parameters are reestimated by adding the observation on day $(t + 1)$ and deleting the observation of day $(t - 1604)$. Thus, from January 1 1996 to October 14 1996, we re-estimate the GARCH process everyday using this rolling procedure. In all, we obtain 200 daily out-of-sample GARCH forecasts.

The updating procedure simply adds information as time progresses to construct a forecast. Thus, we start the estimation procedure at time t, which corresponds to

January 1 1996. Then, we keep updating our sample with the most recent observations on the principal component. The last estimation of the GARCH parameters is based on 1804 observations.

The out-of-sample GARCH forecasts thus obtained are compared to the actual principal components for the same 200-day period calculated from the principal components as previously explained in chapter 5. We will try to answer the following question: "How well do the projected future principal components predict their actual values?" Specifically, are they efficient or biased estimates? It is obvious that an ideal forecast model would produce estimates of the future principal components that closely approximate the actual in-sample values. Figures 6.1 and 6.2 show the evolution of the rolling GARCH forecasts and the updating GARCH forecasts, respectively. Both types of forecasts are compared to the evolution of the in-sample estimates of the principal components. In general, the GARCH forecasts follow the same evolution as their in-sample counterparts. This result confirms our findings in chapter 5 about the relevance of the GARCH processes followed by the three principal components.

Forecasting performance of the two types of GARCH series is judged by comparing the ability of the forecasts to predict the in-sample values of the principal components. Comparisons are based on the mean error, defined as actual minus forecast, (ME), the mean absolute error (MAE) and the RMSE. The results of this test are contained in Table 6.1. It is obvious from this table that:

- The updating GARCH forecasts slightly outperform the rolling GARCH forecasts for all three principal components under the ME criterion.
- The RMSE and MAE criteria both indicate that updating and rolling GARCH forecasts are quite similar for all three principal components.

Moreover, based on the ME criterion, it seems that both types of GARCH forecasts of the three principal components are slightly downward biased for f_1 and f_2 and upward biased for f_3. However, given the relative size of the ME, we can consider such bias as negligible. Fair and Shiller (1990) noted that comparing out-of-sample forecasts using the RMSE criterion has some limitations. Thus, further insights into the difference between the two types of forecasts can be obtained by regressing the in-sample estimates on the out-of-sample forecasts:

$$f_t = a + b f_{p,t} + u_t, \tag{6.4}$$

where $f_{p,t}$ is alternatively the rolling and the updating GARCH forecast. This type of regression is often called "encompassing regression". In particular, Hendry and Richard (1982) and Fair and Shiller (1990) developed a rich literature about the subject. The idea is the following: if a forecast equals the true expected value of f_t, then regressing in-sample values of the principal components on their expectations should produce regression estimates of 0 and 1 for a and b, respectively. Any deviation from those values is interpreted as evidence for bias and inefficiency in the forecast. Equation 6.4 is fitted with OLS. In this type of equation, OLS is a consistent estimator of the regression coefficients. Moreover, the forecast horizon of one day coincides with the frequency of the sample we used. This will rule out any dependency in the errors of the regression. Therefore, standard errors will not be underestimated when computed by OLS.

6.1.1 Interpretation of results

The results for the "encompassing regression" on both forecasts are reported in Tables 6.2 and 6.3, respectively. All of the estimated values of b are positive and are significantly different from zero and from one at the 5% significance level. However, the estimated values of b are very close to unity. For f_1, f_2 and f_3, the slope coefficients of the updating GARCH are 0.96, 0.99 and 0.88, respectively. This indicates that weights of the GARCH forecasts, even though different from one, are very significant and may represent a valuable source for predicting the future values of the principal components of the EP model. For both types of GARCH forecasts, the intercept is statistically significant for all principal components. The t-statistics of the intercept terms for f_1, f_2 and f_3 using the rolling GARCH forecast are 3.5, 2.25 and 2.5, respectively. Nevertheless, none is large enough to be material. Thus, there exists a bias in the GARCH forecasts but it is statistically small. Finally, the adequacy of both GARCH forecasts is confirmed by the R^2 statistics. For all three principal components f_i, $i = 1 \ldots, 3$, the multiple determination coefficients R^2 for the rolling forecasts are 0.51, 0.65 and 0.52, respectively. These values are quite large and indicate that the explanatory power of the GARCH model is satisfactory for the three principal components. More than 50% of the variability in the actual values is explained by the forecasts for each component. From previous results, we can conclude that both types of forecasts are useful in conveying a large amount of information about the future evolution of the principal components. Now, we would like to know whether this statistical predictability can be exploited to generate significant economic profits.

6.2 Arbitrage

Our previous results indicate that the principal components changes are predictable in a statistical sense. Of course, an accurate forecast is desirable, but GARCH forecasts of the principal components might be used for trading even if the forecasts are not totally accurate. Thus, the final question to be answered here is whether the reported predictability is large and persistent in order to be economically significant. In other words: "Can the GARCH forecasts be exploited to make a material trading profit?"

In Chapter 5, we mentioned that the principal components of the EP model are related to the n state factors through the following relation:

$$F = E'B, \tag{6.5}$$

where B is an $(n \times M)$ matrix of values for n state factors for M days. E is an $(n \times n)$ orthogonal matrix in which each column represents an eigenvector and F is an $(n \times M)$ matrix containing the resultant principal components. The above relation is very important for it allows us to easily switch from the state factors of the EP model to the principal components. We also showed in Chapter 5 that the following matrix equation connects the estimated matrix of state factors B and the matrix of bond prices Y

$$B = (X'X)^{-1}X'Y. \tag{6.6}$$

where $Y = (y_1, \ldots, y_M)$ is a matrix of bond prices. y_i is a vector of bond prices of the ith cross-sectional sample. Equations 6.5 and 6.6 imply that

$$F = E'(X'X)^{-1}X'Y. \tag{6.7}$$

Therefore, bond prices can be mapped into principal components. Since we concluded in Chapter 3 that only three principal components are needed to describe the term structure, equation 6.7 can be written for a principal component

$$f_t = E'(X'X)^{-1}X'y_t. \tag{6.8}$$

As the mapping of y_t to f_t is many to one, it follows that a subspace of y_t maps into any given f_t. Equation 6.8 is very interesting since it implies that any principal component f_t can be considered as a linear combination of the vector of bond prices y_t (i.e., portfolio of bonds). Buying(selling) this portfolio is mathematically equivalent to buying(selling) the corresponding principal components. It is also clear from equation 6.8 how an individual bond portfolio would provide a derivative portfolio of estimated principal components f_t. Moreover, equation 6.8 shows that any derivative portfolio of the principal components can be connected to the real bond market. Thus, from a theoretical construct, i.e., f_t, we can switch to a real portfolio of bonds. Indeed,

$$y_t = P_{M,t} - P_{0,t},$$

where $P_{M,t}$ is the average market price of a bond j plus the accrued interest at time t. $P_{0,t}$ is the cash flows of the same bond discounted at the long rate R at time t. Thus, any portfolio of principal components can be related back to the set of available bonds.

Moreover, when we studied the time series properties of the principal components of the EP model, we found that f_t is well described by a GARCH model of the following form

$$f_t = c + \phi_1 f_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}, \tag{6.9}$$

$$h_t = K + \omega_1 \epsilon_{t-1}^2 + \alpha_1 h_{t-1} + \alpha_2 h_{t-2}. \tag{6.10}$$

Equation 6.9 is useful in predicting f_t from f_{t-1}. Given this relationship between f_t and f_{t-1} and knowing that f_t can ultimately be related to the bond market, one can think of many ways to generate profits. As an illustration, we consider a derivative portfolio of the three principal components with weight vector w_{t-1} at time $t-1$, where each component weight w_{t-1}, for $i = 1, 2, 3$ is the number of units of the corresponding principal component that is purchased for the portfolio at price $f_{t-1,i}$. In order to be considered as an arbitrage portfolio, the weights are selected as follows

$$w_{t-1,1} f_{t-1,1} + w_{t-1,2} f_{t-1,2} + w_{t-1,3} f_{t-1,3} = 0. \tag{6.11}$$

This arbitrage portfolio requires no capital outlay. At time t, the value of this portfolio will be

$$w_{t-1,1} f_{t,1} + w_{t-1,2} f_{t,2} + w_{t-1,3} f_{t,3} = w_{t-1,1} df_{t,1} + w_{t-1,2} df_{t,2} + w_{t-1,3} df_{t,3}, \tag{6.12}$$

where $df_{t,i} = f_{t,i} - f_{t-1,i}$. We want to investigate empirically whether such a strategy will, on average, generate profits.

One possible choice for the arbitrage portfolio is to select a weight vector such that:

$$w_{t-1} = f_{p,t-1} - f_{t-1},$$

where $f_{p,t-1}$ is the GARCH forecast made at time $t-1$ for the principal component f_t on the next day. Since we showed that the GARCH forecasts contain substantial information in predicting future principal components, this portfolio will probably be a good candidate to generate long-run profits. As shown later, this choice of weights very nearly satisfies condition 6.11 because the GARCH forecasts, even if slightly biased, contain a substantial amount of information about the actual principal components.

Below, we implement this choice of weights. The trading strategy we use to assess potential arbitrage profits is based on the updating forecast of the principal component. At date $(t-1)$, GARCH forecasts of the principal compouent for day t are formed. If a principal component is predicted to increase (decrease) from day $(t-1)$ to day t, the derivative portfolio is purchased (sold). At time $(t-1)$, the investment outlay of the arbitrage portfolio must be close to zero. At time t, the position is closed and the profit is computed. The prices of the derivative portfolio will be determined from the actual principal component as estimated from the term structure at time t. The trading profit is computed as

$$\pi_t = w_{t-1,1} f_{t,1} + w_{t-1,2} f_{t,2} + w_{t-1,3} f_{t,3}, \tag{6.13}$$

where the $f_{t,i}$ for $i = 1, \ldots, 3$, are inferred from the term structure at day t, through the EP model.

Table 6.4 contains a summary of the trading strategy results. In this table, we report the mean and the standard deviation of the daily portfolio cost. Also, we report the mean and standard deviation of the arbitrage profits realized. Note that this strategy is repeated every day in the 200 samples of GARCH forecasts, between January 2, 1996 to October 15, 1996. The mean cost of the portfolio is \$8.648 × 10^{-5}, which is extremely small, with a standard deviation of \$0.00071. The minimum and maximum define a small range close to zero. This indicates that the arbitrage portfolio involved little outlay of capital at almost no risk. The mean profit for this sampling period is \$0.26 . The standard deviation of this mean value, \$2.4604, indicates that the portfolio profits are far from certain. Nonetheless, mean profit is relatively large. The worst loss realized during this period is \$9.47 and the best profit is \$23.97. Apparently, trading on the basis of the principal components predictions cannot produce consistent positive profits. However, we have empirical evidence that, on average, the profits are significantly greater than zero, excluding transaction costs, for the one day trading horizon.

However, to implement this arbitrage strategy on the actual bond market, we must derive the principal components from the bond market. In theory, this is possible through relation 6.7. Second, the principal components f_t must be "constructed" in the real bond market. Indeed, the state factors are a linear transformation of bond prices y_t, i.e., portfolio of bond prices y_t. In Chapter 3, we defined y_t as:

$$y_t = P_{M,t} - P_{0,t},$$

where $P_{M,t}$ is the average market price of a bond j plus the accrued interest. $P_{0,t}$ is the cash flows of the bond j discounted at the long rate R at time t. $P_{M,t}$ can be easily traded in the bond market as it is the average of the bid and ask prices of bond j. However, $P_{0,t}$ is not available on the bond market and one must look for a bond having similar characteristics. Thus, the trading strategy we described above is subject to some limitations imposed by reality and availability of bonds.

Moreover, one can wonder whether the inclusion of transaction costs would eliminate the arbitrage profits? Knowing that an average commission for individual investors is about 35 cents, we have good reasons to suspect that the profits will be reduced sharply by transaction costs with this one-day trading strategy. Based on this assumption, we cannot reject the efficiency hypothesis of the US bond market. From a statistical standpoint, the GARCH forecasts appeared to be powerful predictors for the future principal components values. The magnitudes and accuracy of the predicted changes were large enough to generate, on average, substantial profits. However, given that the average size of these profits is about 26 cents, they will tend to vanish after transaction costs are considered. Therefore the result does not challenge the hypothesis that the bond market is informationally efficient.

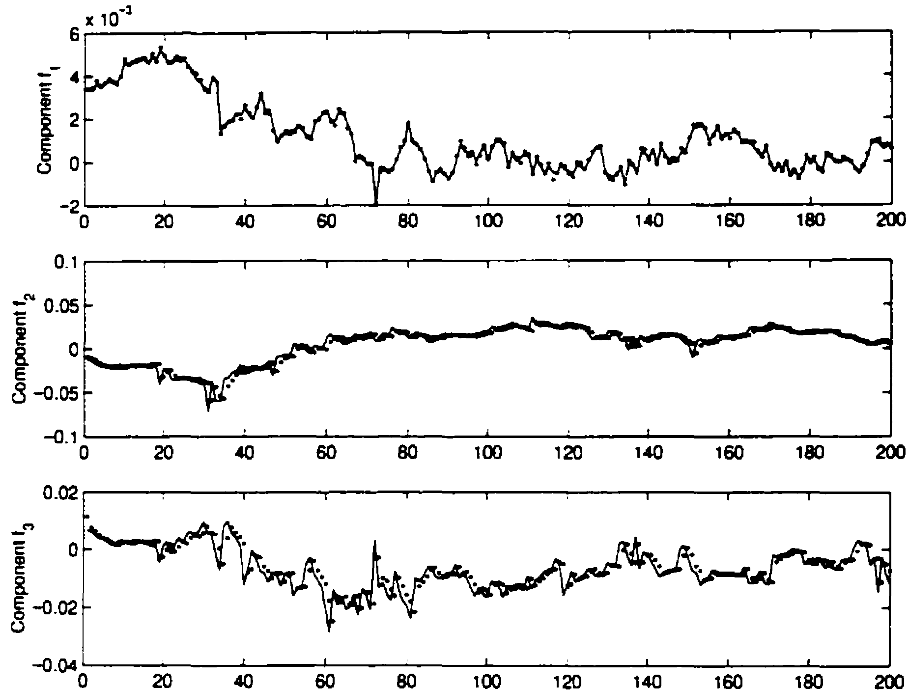

Figure 6.1: These figures depict the evolution of the principal components f_i, for $i = 1 \ldots 3$, estimated using the rolling forecast procedure. The dots are the GARCH forecasts and the solid lines are the actual principal components from the EP model for the period from January 2 1996 to October 15 1996.

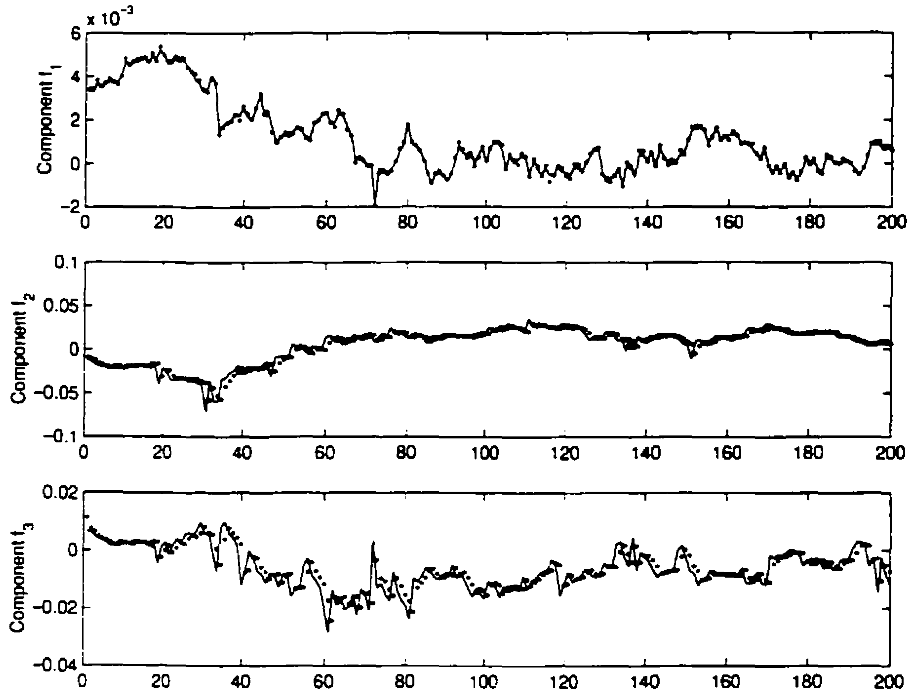

Figure 6.2: These figures depict the evolution of the principal components f_i, for $i = 1 \ldots, 3$, estimated using the updating forecast procedure. The dots are the GARCH forecasts and the solid lines are the actual principal components from the EP model for the period from January 2 1996 to October 15 1996.

Principal Component	ME ($\times 10^{-4}$)	MAE	RMSE
f_1(rolling)	5.4138	0.0023	0.0029
f_1(updating)	5.3750	0.0023	0.0031
f_2	8.8919	0.0042	0.0062
	7.9311	0.0042	0.0062
f_3	-2.3241	0.0030	0.0043
	-2.4860	0.0030	0.0043

Table 6.1: Comparisons between out-of-sample GARCH forecasts of the principal components.

Statistics	f_1	f_2	f_3
a	0.0007	0.0009	-0.0010
SE	0.0002	0.0004	0.0004
b	0.9742	0.9915	0.8701
SE	0.0379	0.0228	0.0459
R^2	0.5133	0.6534	0.5182

Table 6.2: In-sample estimates of the principal components regressed on rolling GARCH forecasts. The equation used is (6.4): $f_t = a + bf_{p,t} + u_t$.

Statistics	f_1	f_2	f_3
a	0.0006	0.0008	-0.0009
SE	0.0002	0.0004	0.0004
b	0.9611	0.9885	0.8772
SE	0.0337	0.0226	0.0461
R^2	0.5145	0.6506	0.5091

Table 6.3: In-sample estimates of the principal components regressed on updating GARCH forecasts. The equation used is (6.4): $f_t = a + bf_{p,t} + u_t$.

Arbitrage Strategy	Mean	Std. dev.	Minimal value	Maximum
Cost(in $)	8.648×10^{-5}	0.00071	0.00001	0.0317
Profit(in $)	0.2597	2.4604	-9.4752	23.9706

Table 6.4: The trading strategy is formed on the basis of the updating out-of-sample GARCH forecasts for the principal components of the EP model. The prediction of principal component for day t is based on the GARCH forecast of day $t-1$ for the same principal component. The weights of the arbitrage portfolio are $w_{t-1} = f_{p,t-1} - f_{t-1}$.

Chapter 7

Conclusion

Since the appearance of the article of Durand in 1942, which suggested a hand fitting of the term structure, many empirical and theoretical models have been proposed for term structure estimation. McCulloch (1975) and Vasicek and Fong (1982) emphasized empirical estimation of the term structure. In other words, they came up with a useful technique that attempts to replicate the shape of the term structure. However, Langetieg (1981) and Cox, Ingersoll, and Ross (1985) have derived a no-arbitrage condition known as the fundamental partial differential equation (PDE) for bond pricing. Through their model, they gave a solid theoretical understanding of term structure. However, a wide gap still exists between theory and empirical techniques for the estimation of term structure.

Guo (1998) poses the following question:"Does the PDE admit any linear solution?" He proceeds to show that the EP model, reproduced in Chapter 2, is the only discount model to be consistent with no-arbitrage. To a certain extent, this model is equivalent to the Exponential Spline model of Vasicek and Fong (1982) but without splines. Its solution can be written as a linear combination of exponential functions. It has the main advantage of being based on an arbitrage argument and thus satisfying the fundamental PDE of bond pricing. The EP model can be stated as follows:

$$D(\beta, s, t) = (1 - \sum_{i=1}^{n} \beta_i)e^{-R(s-t)} + \sum_{i=1}^{n} \beta_i e^{-(R+l_i)(s-t)}. \tag{7.1}$$

where $D(\beta, s, t)$ is the discount function at time t of one monetary unit to be received at time $s > t$. R represents the long rate, $(s - t)$ is the time to maturity and l_i are a selected series of decreasing (or increasing) values. Variables $\beta_1, \ldots, \beta_n$ are state factors of the EP model. The fitting of the term structure is always motivated by the problem that there are few long-term treasury securities available for fitting. The models of McCulloch (1975) and Vasicek and Fong (1982) suggest spline fitting techniques to represent this characteristic of the bond market. As is clear from equation 7.1, this feature is imbedded in the EP model, which allows the discount function to decay at different rates. The EP model does not divide the maturity range into subintervals for spline fitting. Instead, the component exponential functions with larger exponential parameters decay faster than those with smaller parameters. Unlike curve fitting models, such as McCulloch (1975), Bliss (1997) and others, the EP model is not an *ad hoc* function designed to fit the term structure but, more interesting, is a theoretical model derived from the fundamental PDE of bond pricing. Thus, the EP model has the advantages of both approaches; its simplicity is very similar to empirical techniques presented in the literature and its theoretical form is consistent with the PDE of bond pricing. Is it the right candidate for describing term structure? This is the question being answered in this thesis.

The research objective in this thesis is to conduct a cross-sectional and time-series investigation of the EP model. The purpose is to understand its cross-sectional fitting properties as well as the time-series properties of its state factors, which drive term structure.

In Chapter 3 and 4, we studied the cross-sectional properties of the EP model. Two different data sets were examined. One is provided by the NYFRB and consists of daily US T-bills, notes and bonds prices. The second is from the Bank of Canada and contains daily prices of Canadian T-bills, notes and bonds. Our first objective was to estimate the coefficients β_i of the following regression model:

$$P_M^j + AI^j - P_0^j = \sum_{i=1}^{n} \beta_i (P_i^j - P_0^j) + \epsilon_j \quad \text{for } j = 1 \ldots, N, \tag{7.2}$$

where P_M^j is the jth bond market price and AI^j is the accrued interest corresponding to the jth bond. P_i^j and P_0^j are the present values of bond j discounted at rate $(R+l_i)$ and R, respectively. N is the number of treasury securities in the cross-sectional sample. In this equation, the $\beta_i, i = 1, \ldots, n$ are n state factors to be estimated. The long rate R considered as a state factor is also estimated for each cross-sectional sample. Both data sets provided solid evidence that, over the periods under study, US and Canadian term structures could be accurately estimated by the constant exponential basis of the EP model. The fitting performance of the EP model was found to accommodate different shapes of the term structure curve, which is an essential feature of every term structure model. Since that the exponential basis is kept constant, the linear coefficients of the EP model can be considered as state factors that reflect the varying economic conditions that determine term structure. Both data sets confirmed the relevancy of the EP model. However, a different exponential basis was used for each data set. This is perfectly reasonable because US and Canadian bond markets do not share exactly the same sets of state factors. Nevertheless, through an eigen analysis of the n state factors, we found using both data sets, that term structure variability is largely determined by three principal components of the state factors. This result is consistent with the current tendency of modeling the yield curve as being driven by at least three sources of uncertainty; for instance Balduzzi, Das, Foresi, and Sundaram (1996) have presented a model in this direction. Moreover, Subrahmanyam (1996), in his review of the literature, urges the academicians to think about models with more than two state variables.

In Chapter 5, we studied the time-series properties of the estimated state factors. We were also interested to know whether any economic significance could be attributed to these factors. First, we found that the state factors cannot be directly studied because they exhibit in their evolution some regular dramatic shifts induced by high multicollinearity among the components of the exponential basis of the EP model. Hence, we found it more useful to study the time series of the principal components of the state factors. Our results show that the three principal components are stationary and are adequately described by ARMA/GARCH processes. This result clearly indicates that the principal components inferred from the EP model exhibit a heteroskedasticity pattern frequently observed in other financial and economic time series. Unlike many previous models, the EP model does not constrain the state variables to follow any process in particular. Instead, it allows a wide range of processes. This is, most probably, the principal reason behind its accuracy in capturing all the subtle movements of term structure. Through a correlation and canonical analysis, we show that the principal components are linearly related to macroeconomic variables such as the Federal Discount Rate, the Yield of the Longest Bond and the monetary aggregate M1. This result is very similar to those of Balduzzi, Elton, and Green (1996) who concluded that the term structure of interest rates is affected by announcements for macroeconomic variables in the US economy. We also find that the process of the short rate, a linear combination of all the state factors of the EP model, is well described by an ARMA/GARCH model. This result is in line with Duan and Jacobs (1998), who showed through an equilibrium argument, that the short rate follows a GARCH model. More recently, Ball and Torous (1999) presented empirical evidence that the volatility of the short-term interest rates of several foreign countries including the US, is volatile. In this sense, our results also confirm these findings.

The findings of Chapter 5, naturally led us to investigate the predictive power of the GARCH forecasts of the principal components of the EP model. We assessed the information content of the GARCH forecasts in predicting future values of the principal components. We formed two types of forecasts: rolling forecasts and updating ones. We assessed the performance of these GARCH forecasts using encompassing regression. We find that for the three principal factors, the GARCH forecasts are almost unbiased and highly predictive. Finally, we tested whether the reported prediction is economically significant. For this, we created a trading strategy that attempts to take advantage of the informational content of the GARCH forecasts. We showed that the profits generated by this trading strategy, over a horizon of 200 days, are on average significantly different from zero. However, given the relatively small size of the average profit, we have good reason to suspect that it would vanish in the presence of transaction costs.

Given all of our findings, we believe that the EP model is a promising candidate as a linear model for term structure. Its accuracy in estimating the cross-sectional samples of two major data sets is strong evidence in this direction. Moreover, the EP model is found to be largely determined by three principal components that are predictable. The time dependency of the three principal components is useful in predicting the shape of the future term structure. This feature is highly important, essential to every term structure intended for practical use.

Appendices

Appendix A

Notes on Three Models of Term Structure

A.1 Vasicek model

In the Vasicek model (1977), the instantaneous spot risk-free rate, $r(t)$, follows an Ornstein-Uhlenbeck process:

$$dr(t) = \kappa(\theta - r(t))dt + \sigma dz(t). \tag{A.1}$$

$\kappa > 0$ is the mean reversion parameter. θ is the long-run average of the instantaneous interest rate and σ is the volatility parameter. $z(t)$ a standard Brownian motion. Vasicek derived the equilibrium yield to maturity $y(r, t, \tau)$ for a zero-coupon bond as

$$y(r, t, \tau) = A(\tau) + B(\tau)r(t), \tag{A.2}$$

$$A(\tau) = l(1 - B(\tau)) + \frac{(1 - e^{-\kappa\tau})^2}{4\kappa^3\tau}\sigma^2, \tag{A.3}$$

$$B(\tau) = \frac{1 - e^{-\kappa\tau}}{\kappa\tau}. \tag{A.4}$$

with l the yield on the bond with maturity $\tau \to \infty$

$$\lim_{\tau\to\infty} y(r, t, \tau) = l = (\theta + \lambda) - \frac{\sigma^2}{2\kappa^2},$$

with λ the market price of risk.

A.2 CIR model

In the CIR model (1985), the short rate follows a square-root process:

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dz(t), \tag{A.5}$$

where κ, $\theta > 0$ and $\sigma > 0$ are constants.

CIR showed that the current price at time t of a zero-coupon bond which matures at $t + \tau$ and has a face value of one unit is given by the following expressions

$$P(r, t, \tau) = A(\tau)e^{-B(\tau)r(t)}, \tag{A.6}$$

$$A(\tau) = \left[\frac{2\gamma e^{(\gamma+\kappa+\lambda)\tau/2}}{(\gamma+\kappa+\lambda)(e^{\gamma\tau}-1)+2\gamma}\right]^{2\kappa\theta/\sigma^2}. \tag{A.7}$$

$$B(\tau) = \frac{2(e^{\gamma\tau}-1)}{(\gamma+\kappa+\lambda)(e^{\gamma\tau}-1)+2\gamma}. \tag{A.8}$$

$$\gamma = \left((\kappa+\lambda)^2 + 2\sigma^2\right)^{1/2}. \tag{A.9}$$

$$l = \frac{2\kappa\theta}{\gamma+\kappa+\lambda}. \tag{A.10}$$

with l defined as the long rate and λ the market price of risk.

A.3 The Super-Bell model

The Super-Bell model is based on the paper of Bolder and Stréliski (1999). It was developped by Bell Canada Limited in the 1960s. First, a par yield curve is derived. A par yield curve is a series of yields that would be observed if the sample of bonds were all trading at par value. The following regression is conducted:

$$yield_\tau = \alpha_0 + \alpha_1(\tau) + \alpha_2(\tau^2) + \alpha_3(\tau^3) + \alpha_4(\tau^{0.5}) + \alpha_5(\log\tau) + \alpha_6(C) + \alpha_7(C\tau) \tag{A.11}$$

This regression defines yield to maturity $yield_\tau$ as a function of term to maturity τ and the coupon rate C. Once the coefficients of equation A.11 are estimated, a vector of par yield estimates is obtained through the following algebraic arrangement:

$$\hat{yield}_\tau = \frac{\hat{\alpha}_0 + \hat{\alpha}_1(\tau) + \hat{\alpha}_2(\tau^2) + \hat{\alpha}_3(\tau^3) + \hat{\alpha}_4(\tau^{0.5}) + \hat{\alpha}_5(\log\tau)}{1 - \hat{\alpha}_6 - \hat{\alpha}_7(\tau)}. \tag{A.12}$$

Using the vector of estimated par yields, $\hat{yield}_\tau$, an additional regression is run to yield the fitted equation

$$\hat{\hat{yield}}_\tau = \hat{\hat{\alpha}}_0 + \hat{\hat{\alpha}}_1(\tau) + \hat{\hat{\alpha}}_2(\tau^2) + \hat{\hat{\alpha}}_3(\tau^3) + \hat{\hat{\alpha}}_4(\tau^{0.5}) + \hat{\hat{\alpha}}_5(\log\tau). \tag{A.13}$$

This last step is conducted in order to "smooth" the par yield curve, obtained in equation A.11.

Appendix B

Theoretical Prices of the EP Model

These details are adapted from Guo (1993). Here, we only present the closed-form formula of the present value of a hypothetical price, under the exponential base $e^{-\lambda_0}$. In the text we denoted this present value by P_j^0.

First, consider a standard semi-annual coupon bond with face value F, maturity $T-t$, and annual coupon interest payment C_j. We denote by p the total number of half-years

$$p = integer(2(T-t)).$$

Let t_c be the fraction of half-year to the nearest coupon interest payment

$$t_c = 2(T-t) - p.$$

The bond will have a total of $(p+1)$ cash flows at the following future time points (in years)

$$t_c/2, t_c/2 + 1/2, t_c/2 + 1, \ldots, t_c/2 + p/2,$$

at which the corresponding discount factors are

$$e^{-\lambda_0 t_c}, e^{-\lambda_0(t_c/2+1/2)}, e^{-\lambda_0(t_c/2+1)}, \ldots, e^{-\lambda_0(t_c/2+p/2)}.$$

For example. if $T - t$=15.3 years. then p = 30 half-years. t_c = 0.6 half year, and there will be a total of 31 cash flows. The present value of the bond is

$$P_j^0 = \frac{C_j}{2} e^{-\lambda_0 t_c/2} [\frac{1 - e^{-(p+1)\lambda_0/2}}{1 - e^{-\lambda_0/2}}] + F e^{-\lambda_0(T-t)}. \quad \text{(B.1)}$$

To separate the accrued interest. equation (B.1) can be re-written as

$$P_j^0 = \frac{C_j/2}{e^{\lambda_0/2} - 1}(1 - e^{-\lambda_0(T-t)}) + F e^{-\lambda_0(T-t)} + \frac{C_j}{2} \frac{(e^{(1-t_c)\lambda_0/2} - 1)}{e^{\lambda_0/2} - 1}. \quad \text{(B.2)}$$

If we take the first order Taylor series approximation that (i.e., $e^x - 1 \simeq x$ for small $|x|$). the last term in equation (B.2) simplifies to

$$\frac{C_j(1 - t_c)}{2}.$$

which is the accrued interest AI_j defined in the text.

Appendix C

GARCH Forecasts

Here we present a derivation of GARCH forecasts for an ARMA(1,1) process, y_t, as follows

$$y_t = \mu + \phi_1 y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t. \tag{C.1}$$

Using matrix notation, equation C.1 can be re-written as

$$Y_t = \mu e_1 + \Phi Y_{t-1} + (e_1 + e_2)\epsilon_t, \tag{C.2}$$

where,

- $Y_t = \begin{pmatrix} y_t \\ \epsilon_t \end{pmatrix}$,
- $\Phi = \begin{pmatrix} \phi_1 & \theta_1 \\ 0 & 0 \end{pmatrix}$,
- e_i is a 2×1 vector with 1 in the ith element and 0 elsewhere.

Equation C.2 is set up in this manner so it is a particular case of equation 13 of Baillie and Bollerslev (1992). Thus, following these authors, the optimal s-step ahead predictor of y_{t+s} is equal to

$$E_t(y_{t+s}) = \iota_s + \sum_{i=0}^{k-1} \tau_{i,s} y_{t-i} + \sum_{i=0}^{l-1} \lambda_{i,s}\epsilon_{t-i}, \tag{C.3}$$

where k and l are the order of AR and MA, respectively. For an ARMA(1,1), the previous equation can be written as

$$E_t(y_{t+s}) = \iota_s + \tau_{0,s} y_t + \lambda_{0,s}\epsilon_t. \tag{C.4}$$

Baillie and Bollerslev showed that

$$\iota_s = e_1'(I + \Phi + \ldots + \Phi^{s-1})e_1\mu, \tag{C.5}$$

$$\tau_{0,s} = e_1'\Phi^s e_1, \tag{C.6}$$

$$\lambda_{0,s} = e_1'\Phi^s e_2. \tag{C.7}$$

Given the structure of Φ, we can compute Φ^s as

$$\Phi^s = \begin{pmatrix} \phi_1^s & \theta_1\phi_1^{s-1} \\ 0 & 0 \end{pmatrix}.$$

Substituting the expression of Φ^s in ι, τ and λ, we get

$$\iota_s = \mu[1 + \phi_1 + \ldots + \phi_1^{s-1}], \tag{C.8}$$

$$\tau_{0,s} = \phi_1^s, \tag{C.9}$$

$$\lambda_{0,s} = \theta_1\phi_1^{s-1}. \tag{C.10}$$

Now, replacing these new expressions in equation C.4

$$E_t(y_{t+s}) = \mu[1 + \phi_1 + \ldots + \phi_1^{s-1}] + \phi_1^s y_t + \theta_1\phi_1^{s-1}\epsilon_t, \tag{C.11}$$

which corresponds to equation 6.3 in Chapter 6, with a different notation.

Bibliography

Ahn, Dong-Hyun, and Bin Gao, 1999, A parametric nonlinear model of term structure dynamics, *Review of Financial Studies* 12, 721–762.

Aït-Sahalia, Yacine, 1996, Testing continuous-time models of the spot interest rate, *Review of Financial Studies* 9, 385–426.

Akaike, Hirotsugu, 1974, A new look at the statistical model identification, *IEEE, Transactions on Automatic Control* AC-19, 716–723.

Amihud, Yacov, and Haim Mendelson, 1991, Liquidity, maturity, and the yield on Treasury securities, *Journal of Finance* 46, 1411–1426.

Backus, David K., and Stanley E. Zin, 1994, Reverse engineering the yield curve, Working paper, NBER.

Baillie, Richard T., and Tim Bollerslev, 1992, Prediction in dynamic models with time-dependent conditional variances, *Journal of Econometrics* 52, 91–113.

Balduzzi, Pierluigi, Giuseppe Bertola, and Silverio Foresi, 1996, A model of target changes and the term structure of interest rates, *Journal of Monetary Economics* 39, 223–249.

Balduzzi, Pierluigi, Sanjiv Ranjan Das, Silverio Foresi, and Rangarajan Sundaram, 1996, A Simple approach to three-factor affine term structure models, *Journal of Fixed Income* 6, 43–53.

Balduzzi. Pierluigi. Edwin J. Elton. and Clifton T. Green. 1996. Economic news and the yield curve: Evidence from the U.S. Treasury market. Working paper. Stern School of Business. New York.

Ball. Clifford A., and Walter N. Torous, 1999, The stochastic volatility of short-term interest rates: some international evidence, *Journal of Finance* 6, 2339–2359.

Barone. Emilio. Domenico Cuoco, and Emerico Zautzik, 1991, Term structure estimation using the Cox, Ingersoll, and Ross model: the case of Italian Treasury bonds. *Journal of Fixed Income* 1, 87–95.

Beim. David O.. 1992. Term structure and the non-cash value in bonds, Working paper. Columbia University.

Bekdache. Besma. and Christopher F. Baum. 1995. Comparing alternative no-arbitrage models of the term structure of interest rates. Working paper. Wayne State University.

Bliss. Robert R.. 1997. Testing term structure estimation methods. *Advances in Futures and Options Research* 9. 197–231.

Bolder. David. and David Stréliski. 1999. Yield curve modelling at the Bank of Canada. Working paper, 84 Bank of Canada.

Bollerslev. Tim, Robert F. Engle, and David Nelson, 1994, ARCH models, *Handbook of econometric* 4, 2959–3031.

Brace, Alan, and Mark Musiela, 1994, A multifactor gauss-markovian implementation of Heath Jarrow and Morton, *Mathematical Finance* 2, 259–283.

Brennan. Michael J., and Eduardo S. Schwartz. 1979. A continuous time approach to the pricing of bonds, *Journal of Banking and Finance* 3, 133–155.

Brown. Roger H.. and Stephen M. Schafer. 1993. Interest rate volatility and the shape of the term structure. *Philosophical Transactions of the Royal Society: Physical Sciences and Engineering* 347, 449–598.

Brown, Stephen J.. and Philip H. Dybvig, 1986, The empirical implications of the Cox, Ingersoll and Ross theory of the term structure of interest rates, *Journal of Finance* 41, 617–630.

Chacko, George. 1998. Continuous-time estimation of affine term structure models: A general approach. Working paper. Harvard University.

Chan, K. C., Andrew Karolyi, Francis A. Longstaff. and Anthony B. Sanders, 1992, An empirical comparison of alternative models of the short term interest rates, *Journal of Finance* 47, 1209–1227.

Chapman, David A., John B. Long, and Neil D. Pearson, 1999, Using proxies for the short rate: when are the three months like instant?, *Review of Financial Studies* 12, 763–806.

Chen, Ren-Raw, and Louis Scott, 1992, Pricing interest rate options in a two-factor Cox, Ingersoll and Ross model of the term structure, *Review of the Financial Studies* 5, 613–636.

Coleman, Thomas S., Lawrence Fisher, and Roger G. Ibbotson, 1992, Estimating the term structure of interest rates from data that include the prices of coupon bonds, *Journal of Fixed Income* 2, 85–116.

Conley, Timothy G., Lars P. Hansen, Erzo G.J. Luttmer, and José A. Scheinkman, 1997, Short-term interest rates as subordinate diffusions, *Review of Financial Studies* 10, 525–577.

Constantinides, George, and Jonathan E. Ingersoll, 1984, Optimal bond trading with personal taxes, *Journal of Financial Economics* 13, 229–335.

Constantinides, George M., 1992, A theory of the nominal term structure of interest rates, *Review of Financial Studies* 5, 531–552.

Cox, John, Jonathan Ingersoll, and Stephen Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–467.

Daves, Phillip R., and Michael C. Ehrhardt, 1992, Liquidity, reconstruction, and the value of U.S. treasury STRIPS, Working paper, University of Tennessee.

de Munnik, Jeroen F.J., and Peter C. Schotman, 1994, Cross-sectional versus time series estimation of term structure models: empirical results for Dutch bond market, *Journal of Banking and Finance* 18, 997–1025.

Dickey, David A., and Wayne A. Fuller, 1979, Distribution of the estimators for autoregressive time series with unit root, *Journal of American Statistical Association* 74, 427–431.

Duan, Jin-Chuan, and Kris Jacobs, 1998, Equilibrium interest rate dynamics: Short and long memory, Working paper, McGill University.

Duffie, Darell, 1992, *Dynamic Asset Pricing Portfolio*, (Princeton University Press Princeton, N.J).

Duffie, Darell, 1996, Special repo rates, *Journal of Finance* 51, 493–526.

Duffie, Darell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.

Durand, David, 1942, Basic yields of corporate bonds 1900-1942, Working paper, 3 NBER.

Elton, Edwin J., and Clifton T. Green, 1998, Tax and liquidity effects in pricing government bonds, *Journal of Finance* 53, 1533–1561.

Engle, Robert F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica* 50, 987–1008.

Engle, Robert F., and Tim Bollerslev, 1986, Modelling the persistence of conditional variances, *Econometric Reviews* 5, 1–50.

Fair, Ray C., and Robert J. Shiller, 1990, Comparing information in forecasts from econometric models, *American Economic Review* 80, 375–389.

Fama, Eugene F., and Robert R. Bliss, 1987, The information in long maturity forward rates, *American Economic Review* 77, 680–692.

Ferguson, Robert, and Steven Raymar, 1998, A comparative analysis of several popular term structure estimation models, *Journal of Fixed Income* 3, 17–33.

Fisher, Mark E., Douglas Nychka, and David Zervos, 1995, Fitting the term structure of interest rates with smoothing splines, Working paper, 95-1 Finance and Economics Discussion Paper Series, Federal Reserve Board.

Flesaker, Bjorn, and Lane P. Hughston, 1995, Positive interest, Merrill Lynch Preprint.

Fuller, William A., 1976, *Introduction to Statistical Time Series*, (Wiley New York).

Gibbons, Michael R., and Krishna Ramaswamy, 1993, A test of the Cox, Ingersoll, and Ross model of the term structure, *Review of Financial Studies* 6, 619–658.

Gittins, Robert, 1984, *Canonical Analysis, A Review with Application in Ecology*, (Springer-Verlag Berlin Heidelberg).

Guo, Chen, 1993. An exponential-quadrinomial model of the term structure of interest rates, Working paper, University of Ottawa.

Guo, Chen, 1997, Re-indexing the Heath Jarrow and Morton term structure model, Working paper, University of Ottawa.

Guo, Chen, 1998, A linear solution to multivariate term structure model of Langetieg, Working paper, University of Ottawa.

Hamilton, James D., 1994, *Time Series Analysis*. (Princeton University Press Princeton, N.J).

Hardouvelis, Gikas A., 1987, Macroeconomic information and stock prices, *Journal of Economics and Business* 39, 131–140.

Harrison, Michael J., and David Kreps, 1979, Martingale and arbitrage in multiperiod securities markets, *Journal of Economic Theory* 20, 381–408.

Harrison, Michael J., and Stanley Pliska, 1981, Martingales and stochastic integrals in the theory of continuous trading, *Stochastic Processes and Their Application* 11, 215–260.

Harvey, Andrew C., 1993, *Time Series Models*. (The MIT Press Cambridge).

Heath, David, Robert Jarrow, and Andrew Morton, 1992, Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica* 60, 77–105.

Hendry, David F., and Jean-François Richard, 1982, On the formulation of empirical models in dynamic economics, *Journal of Econometrics* 20, 3–33.

Ho, Thomas S.Y., and Sang-Bin Lee, 1986, Term structure movements and pricing interest rate contingent claims, *Journal of Finance* 41, 1011–1029.

Jamashidian, Farshid, 1988, Notes on Heath Jarrow and Morton, Unpublished Manuscript, Merril Lynch.

Jiang, George J., 1998, Nonparametric modelling 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.

Jin, Yan, and Paul Glasserman, 1998, Equilibrium positive interest rates: A unified approach, Working paper, Columbia Business School.

Jordan, Bradford D., and Susan D. Jordan, 1991, Tax options and the pricing of Treasury bond triplets, *Journal of Financial Economics* 30, 135–164.

Jordan, Bradford D., and David R. Kuipers, 1997, On the performance of affine term structure models: Evidence from the U.S. Treasury STRIPS market, Working paper, University of Houston.

Kamara, Avraham, 1994, Liquidity, taxes, and short-term Treasury yields, *Journal of Financial and Quantitative Analysis* 29, 403–417.

Kaushik, Amin I., and Andrew Morton, 1994, Implied volatility functions in arbitrage-free term structure models, *Journal of Financial Economics* 35, 141–180.

Langetieg, Terence C., 1981, A multivariate model of the term structure, *Journal of Finance* 35, 71–97.

Ljung, Greta M., and George E.P. Box, 1978, On a measure of lack of fit in time series models, *Biometrika* 66, 67–72.

Longstaff, Francis, and Eduardo S. Schwartz, 1992, Interest rate volatility and the term structure: A two-factor general equilibrium model, *Journal of Finance* 57, 1259–1281.

McCulloch, Huston J., 1975, The tax-adjusted yield curve, *Journal of Finance* 30, 811–829.

McQueen, Grant, and Vance V. Roley, 1993, Stock prices, news, and business conditions, *Review of Financial Studies* 3, 683–707.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–888.

Musiela, Mark, 1994, Nominal annual rates and lognormal volatility structure, Working paper, the University of NSW.

Nelson, Charles R., and Andrew F. Siegel, 1987, Parsimonious modeling of yield curves, *Journal of Business* 60, 473–489.

Pearson, Neil, and Tong-Sheng Sun, 1994, Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model, *Journal of Finance* 49, 1279–1304.

Phillips, Peter C., and Pierre Perron, 1988, Testing for unit root in time series regression, *Biometrika* 75, 335–346.

Rogers, Chris L. G., 1997, The potential approach to the term structure of interest rates and foreign exchange rates, *Mathematical Finance* 7, 157–176.

Ronn, Ehud I., 1987, A new linear programming approach to bond portfolio management, *Journal of Financial and Quantitative Analysis* 22, 439–466.

Ross, Stephen, 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Said, Said E., and David A. Dickey, 1984, Testing for unit roots in autoregressive moving-average models with unknown order, *Biometrika* 71, 369–374.

Schwarz, Gideon, 1978. Estimating the dimension of a model, *Annals of Statistics* 6, 461–464.

Shea, Gary S., 1984. Pitfalls in smoothing interest rate term structure data: Equilibrium models and spline approximation, *Journal of Financial and Quantitative Analysis* 19, 253–269.

Shea, Gary S., 1985. Interest rate term structure estimation with exponential splines: A note, *Journal of Finance* 40, 319–325.

Subrahmanyam, Marti G., 1996. The term structure of interest rates: Alternative approaches and their implication for the valuation of contingent claims, *The Geneva Papers on Risk and Insurance Theory.*

Sun, Tong-Sheng, 1992. Real and nominal interest rates: A discrete-time model and its continuous-time limit, *Review of Financial Studies* 5, 581–611.

Svensson, Lars E., 1994. Estimating and interpreting forward interest rates: Sweden 1992-1994, Working paper, International Monetary Fund.

Urich, Thomas, and Paul Wachtel, 1984. The effects of inflation and money supply announcements on interest rates, *Journal of Finance* 39, 1177–1188.

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

Vasicek, Oldrich, and Gifford Fong, 1982, Term structure modelling using exponential splines, *Journal of Finance* 37, 339–348.

Weiss, Andrew A., 1984. ARMA models with ARCH errors, *Journal of Time Series Analysis* 5, 129–143.

Zhang, Hua, 1993. The dynamic behaviour of the term structure of interest rates and its implications for interest-rate derivatives, Ph.D. thesis McGill University.