

Nonparameteric Estimation of Interest Rate Term Structure  
and Insurance Applications

Samuel H. Cox and Hal W. Pedersen

Department of Risk and Insurance  
Georgia State University  
Atlanta, Georgia USA 30302

Telephone: 404/651-4854

Facsimile: 404/651-1296

Email: samcox@gsu.edu

inshwp@panther.gsu.edu

**Abstract:** This paper describes nonparametric estimation of the drift and diffusion of a continuous-time short rate process, based on discrete observations. It also shows how to estimate the market price of risk nonparametrically based on discrete observations of bonds prices. We critique and confirm Stanton's results on nonparameteric estimators [17]. The nonparameteric approach overcomes some of Venter's criticisms of parametric models [20].

# NONPARAMETRIC ESTIMATION OF INTEREST RATE TERM STRUCTURE WITH INSURANCE APPLICATIONS

SAMUEL H. COX AND HAL W. PEDERSEN

ABSTRACT. This paper describes nonparametric estimation of the drift and diffusion of a continuous-time short rate process, based on discrete observations. It also shows how to estimate the market price of risk nonparametrically based on discrete observations of bonds prices. We critique and confirm Stanton's results on nonparametric estimators [17]. The nonparametric approach overcomes some of Venter's criticisms of parametric models [20].

## 1. INTRODUCTION

Models of insurance policies are used for pricing, reserving, allocation of capital, and risk management. The interest rate is a key ingredient in the model and therefore actuaries are very interested in interest rate models. Christiansen [7] remarks “[m]odels of possible future paths for interest rates are a key element of actuarial and other financial studies” and Tilley [18] comments, “[a] stochastic interest rate generator is a valuable actuarial tool”. It has been suggested that the most pressing questions in an actuary's mind when contemplating the use of interest-rate models are “given data, how do I get those parameters and how do I generate scenario paths?” At the 1998 ASTIN Colloquium in Glasgow, Gary Venter criticised some parametric models because they lacked properties observed in bond markets [20]. He offered some criteria for a yield curve model to be used in insurance models. Nonparametric models may meet his criteria.

This article discusses nonparametric estimation of interest rate model ingredients: drift, volatility, and market price of risk. The nonparametric approach seems promising for actuarial applications because it provides a consistent approach, keeping the continuous-time arbitrage theory and estimating the model with discrete-time data.

**1.1. Continuous-time model.** The model is based on an idealized bond market in which there are no transaction costs, traders can borrow as much as they like at default-free rates, all traders have the same information, and there is no arbitrage. Mathematically, the securities market is described in terms of a filtered probability space  $(\Omega, \mathcal{F}, F, \mathbb{P})$  where  $F = \{\mathcal{F}_t : 0 \leq t \leq T^*\}$ .

For general one-factor models, the *short rate*  $r$  is an Itô process

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dZ_t \tag{1.1}$$

---

*Date:* March 19, 1999.

1991 *Mathematics Subject Classification.* Primary: 90A09 ; Secondary: 62M10, 62P05, 90A20.

*Key words and phrases.* interest rates, term structure, interest rate volatility nonparametric estimation. The Actuarial Foundation supported this research with a grant.

where  $\{Z_t : 0 \leq t \leq T^*\}$  is a standard Brownian motion adapted to the market information. The state space of the interest rate process is  $\mathbb{R}_+ = (0, +\infty)$ ,  $\mathbb{R}$  denotes the real numbers, and the two functions  $\mu, \sigma : \mathbb{R}_+ \times [0, T^*] \rightarrow \mathbb{R}$  are to be estimated from the data. The functions  $\mu$  and  $\sigma$  are given functions of two variables, called the drift and diffusion, respectively. In financial modeling the diffusion function is often called the volatility.

The model is called *time-homogeneous* if the drift and volatility are independent of time. Stanton [17] offers support for the time-homogenous assumption as follows. Interest rate models such as Hull and White [11], Ho and Lee [10], Black, Derman and Toy [3], and Black and Karasinski [4] allow arbitrary time dependence in the parameters to match the current term structure exactly. Often the term “arbitrage-free” is used (inappropriately) to describe the fit to the current term structure. As we will see, the lack of arbitrage in the model has an entirely different meaning. Stanton also asserts that fitting the current term structure can introduce counterfactual behavior for future interest rates and that it requires reestimation of the model every time the term structure changes. Throughout this paper we will assume that the drift and diffusion are time-homogenous.

Most market models are based on an assumed parameterized form for the short rate process. The idea is that market data can be used to estimate the short rate process parameters and this should lead to model prices of bonds, interest rate derivatives, and other securities that depend on interest rates. If the model agrees well with actual bond prices and interest rate derivatives, then it would be natural to use it also to model over-the-counter securities (swaps, caps, floors, *etc.*) and insurance contracts (annuities, policy loan options, *etc.*). However, specifying the short rate is not enough information for the model to determine bond prices. The market price of risk is also required as we will discuss below. There are several popular time-homogeneous term structure models, three of which are

$$dr_t = \kappa(\alpha - r_t)dt + \sigma\sqrt{r_t}dZ_t \quad (\text{CIR})$$

$$dr_t = \kappa(\alpha - r_t)dt + \sigma dZ_t \quad (\text{Vasicek})$$

$$dr_t = \kappa(\alpha - r_t)dt + \sigma r_t^\gamma dZ_t \quad (\text{CKLS})$$

where  $\kappa$ ,  $\alpha$ ,  $\sigma$  and  $\gamma$  are constants. The drift  $\mu(r) = \kappa(\alpha - r)$  is said to be *mean reverting*. This is an intuitively appealing property and the data seems to support it. Mean reversion is equivalent to a linear drift, *i. e.*, it can be written in the form  $\mu(r) = a_0 + a_1 r$  for constants  $a_0$  and  $a_1$ . The intuition is illustrated in Figure 1.

Chan, Karolyi, Longstaff and Sanders [6] describe the most general of these with the diffusion coefficient  $\sigma(r) = \sigma r^\gamma$  for constants  $\sigma$  and  $\gamma$ . The Vasicek [19] model has a constant diffusion ( $\gamma = 0$ ). Cox, Ingersoll and Ross [8] specify a diffusion coefficient that is proportional to the square root of the interest rate level ( $\gamma = \frac{1}{2}$ ). The CKLS family of models with  $\gamma > 0$  has the desirable property that model interest rates are positive. This is desirable because we are modeling nominal rates, which are always positive. Indeed, any model for which the drift at  $r = 0$  is positive,  $\mu(0) > 0$ , and the diffusion at  $r = 0$  is zero,  $\sigma(0) = 0$ , must have positive interest rates. Intuitively, the reason is this: Since  $\mu(0) > 0$  and  $\mu(r)$  is continuous, then there is a range  $[0, \epsilon)$  over which  $\mu(r)$  is positive. Consider what happens to  $dr_t$  as  $r_t$  approaches 0. We can assume that  $0 < r_t < \epsilon < \alpha$  and that  $\epsilon$  is small enough

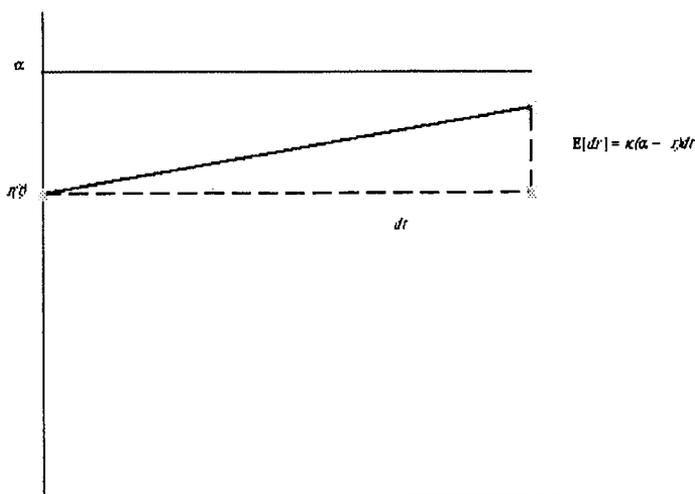


FIGURE 1. Mean Reversion. The time  $(t, t + dt)$  is on the horizontal axis and the short rate  $r(t)$  on the vertical axis. The parameter  $\alpha$  represents a rate to which the short rate “reverts.” The parameter  $\kappa$  represents the strength of the “pull” toward the preferred rate  $\alpha$ .

that  $\text{Var}(r(t + \epsilon) - r_t) = \sigma^2(r_t)\epsilon$  is negligible. This implies that the change in  $r_t$  is essentially deterministic and almost certainly equal to its expected value  $E[dr_t] = \kappa(\alpha - r_t) > 0$  so it is almost certain that  $r_{t+\epsilon} = r_t + \kappa(\alpha - r_t) > r_t$ . Thus the short rate process moves away from  $r = 0$  whenever it gets close. Therefore, since we start with a positive short rate, the model will never produce a negative interest rate.

Aït-Sahalia [1] also provides support for the nonparametric approach. We will elaborate on his points. First, derivative prices are very sensitive to the precise form of the diffusion function. The Vasicek model has a closed form solution for options on bonds (or interest rates). In this model the diffusion is a constant and one can easily see that the option price is very sensitive to the size of the constant. So it is important to get the diffusion estimate as precisely as possible. Second, we cannot observe the diffusion (instantaneous interest rate volatility) so we have no *a priori* idea of what the diffusion should look like. Finally, we have huge data sets available. According to Aït-Sahalia, this is a “perfect setup” for the nonparametric approach. He estimates the volatility and Stanton estimates the drift, volatility, and market price of risk using the nonparametric approach. We will describe the nonparametric approach in detail and critique Stanton’s results.

**1.2. Description of the bond market.** The price  $P(r, t, T)$  of a default-free zero coupon bond is a function of the time  $t$ , the value of the short rate  $r$ , and the time  $T \leq T^*$  of maturity of the bond<sup>1</sup>. This is a fundamental assumption of the model. The particular form of the functional relationship is not specified, rather we assume only that the price is a differentiable function of the three parameters. We could also say that, in general, the price depends on the information  $\mathcal{F}$  as well.

By Itô's formula<sup>2</sup>, the process  $\{P(r_t, t, T) : 0 \leq t \leq T\}$  is also an Itô process and its differential equation is

$$dP = \left[ \frac{1}{2} \sigma^2 \frac{\partial^2 P}{\partial r^2} + \mu \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} \right] dt + \sigma \frac{\partial P}{\partial r} dZ$$

where it is understood that the functions are evaluated at  $(r, t) = (r_t, t)$ . The drift and volatility components are written in this special form:

$$\mu_P(r, t, T) = \frac{1}{P} \left[ \frac{1}{2} \sigma^2 \frac{\partial^2 P}{\partial r^2} + \mu \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} \right]$$

and

$$\sigma_P(r, t, T) = \frac{1}{P} \sigma \frac{\partial P}{\partial r}. \quad (1.2)$$

With this notational convention the stochastic differential equation for the bond price is usually written in this equivalent form:

$$\frac{dP}{P} = \mu_P dt + \sigma_P dZ \quad (1.3)$$

The notation  $P$ ,  $P(t, T)$ , and  $P(t, r, T)$  all denote the price, but it is customary to suppress the reference to  $r$ , or even  $t$  and  $T$ , when no confusion can arise.

**1.3. Trading and Portfolios.** Continuous default-free borrowing and lending at the short rate is available to all traders. It is usually described in terms of an account, like a savings account. The account values from a process  $\{B(t)\}$  satisfying  $B(0) = 1$  and  $dB(t) = r(t)B(t)dt$  for  $t \geq 0$ . The account has one dollar initially and grows by constantly earning and reinvesting the short rate. A trader who invests  $k$  dollars (buys the account) at time  $t_1$  receives a right to  $k/B(t_1)$  units of the account. At time  $t_2 > t_1$ , the trader can withdraw (sell units) with a value of  $B(t_2)k/B(t_1)$ .

The account value can be written explicitly

$$B(t) = \exp \left( \int_0^t r(u) du \right)$$

so an investment of  $k$  at time  $t_1$  accumulates to a value of

$$B(t_2)k/B(t_1) = k \exp \left( \int_{t_1}^{t_2} r(u) du \right)$$

<sup>1</sup>Readers familiar with the mathematics of bond markets may skim this section. Björk [2] is a good reference for this material.

<sup>2</sup>The appendix to this paper has a review of properties of Itô processes.

Let  $S_1$  and  $S_2$  denote two traded securities in this market with differential equations

$$dS_i(t) = \mu_i(t)dt + \sigma_i(t)dZ(t)$$

for  $i = 1, 2$ . A *portfolio* of  $S_1$  and  $S_2$  is a pair of processes  $(\phi_1(t), \phi_2(t))$  where  $\phi_i(t)$  denotes the number of units of the security  $i$  in the portfolio at time  $t$ . The number of units can be negative, indicating short selling (borrowing). The units of units need not be an integer; any real number is allowed for the number of units. The value of the portfolio at time  $t$  is

$$V(t) = \phi_1(t)S_1(t) + \phi_2(t)S_2(t).$$

The number of units in the portfolio have to be determined before the prices are revealed. This means that in setting a trading strategy a trader may base the values of  $(\phi_1(t), \phi_2(t))$  on information  $\mathcal{F}_s$  for  $s < t$ , but not on  $\mathcal{F}_t$ . In addition, the trading strategy must satisfy the technical condition: For  $i = 1, 2$ ,

$$\Pr \left[ \int_0^T \phi_i(t)^2 \sigma_i(t)^2 dt < \infty \right] = 1 \quad \text{for any } T < \infty$$

A portfolio is *self-financing* provided that the portfolio's value changes over an interval  $(t, t + dt)$  solely due to price changes, not due to changes in the number of units held. Mathematically this is expressed as follows:

$$dV(t) = \phi_1(t)dS_1(t) + \phi_2(t)dS_2(t) \tag{1.4}$$

By substitution we can find the Itô representation of  $V$ :

$$dV(t) = \phi_1(t)[\mu_1(t)dt + \sigma_1(t)dZ] + \phi_2(t)[\mu_2(t)dt + \sigma_2(t)dZ] \tag{1.5}$$

$$= [\phi_1(t)\mu_1(t) + \phi_2(t)\mu_2(t)]dt + [\phi_1(t)\sigma_1(t) + \phi_2(t)\sigma_2(t)]dZ \tag{1.6}$$

Self-financing portfolios are a key ingredient in pricing and hedging. In applying the concept we will usually arrange the strategy so that the portfolio's volatility term is zero and use the following additional important assumption.

**1.4. No arbitrage principle.** An *arbitrage opportunity* is a self-financing portfolio with zero initial value, having at some future time no possibility of a negative value and a positive probability of a positive value. In terms of the notation defined above, an arbitrage opportunity based on the securities  $S_1$  and  $S_2$  is a portfolio  $(\phi_1(t), \phi_2(t))$  for which the value process  $V(t) = \phi_1(t)S_1(t) + \phi_2(t)S_2(t)$  satisfies

- $dV = \phi_1 dS_1 + \phi_2 dS_2$
- $V(0) = 0$
- There is a  $T > 0$  for which

$$\Pr(V(T) < 0) = 0 \text{ and } \Pr(V(T) > 0) > 0.$$

The bond market is assumed to be arbitrage free, meaning there are no arbitrage opportunities. This natural assumption has some strong implications for modeling. A common application goes like this: If a traded security (or portfolio of traded securities) with price  $\{X_t\}$  has zero volatility, then its drift is  $r_t X_t$  where  $r_t$  is the short rate. By zero volatility we mean that, when written in its Itô representation,

$$dX = \mu_X(t)dt + \sigma_X(t)dZ,$$

the volatility term is identically zero with probability one:

$$\Pr[\sigma_X(t) = 0 \text{ for all } t] = 1$$

For any such security or portfolio we must have  $\mu_X(t) = r(t)X(t)$  for all  $t$  with probability one.

**1.5. Market Price of Risk.** Consider two zero coupon bonds denoted by  $S_1(t) = P(r, t, T_1)$  and  $S_2(t) = P(r, t, T_2)$  with  $T_2 < T_1$ . The zero coupon bonds have differential equations  $dS_i = \mu_i S_i dt + \sigma_i S_i dZ$ . Make a portfolio with one unit of  $S_1$  and  $\phi(t)$  units of  $S_2$  at time  $t$ , selected so that the resulting portfolio has zero drift:

$$V(t) = S_1(t) + \phi(t)S_2(t)$$

The differential equation of  $V(t)$  is

$$\begin{aligned} dV(t) &= [\mu_1(t)S_1(t) + \phi(t)\mu_2(t)S_2(t)]dt \\ &+ [\sigma_1(t)S_1(t) + \phi(t)\sigma_2(t)S_2(t)]dZ \end{aligned}$$

For a diffusion coefficient of zero,  $\phi(t) = -S_1(t)\sigma_1(t)/S_2(t)\sigma_2(t)$  for all  $t$ . Since there is no arbitrage and the portfolio has zero volatility, then its drift is  $r(t)V(t)$ . Substituting the values of the parameters and solving gives us the relation:

$$\begin{aligned} r(t)V(t) &= \mu_1(t)S_1(t) + \phi(t)\mu_2(t)S_2(t) \\ r(t)[S_1(t) + \phi(t)S_2(t)] &= \mu_1(t)S_1(t) - S_1(t)\mu_2(t)\sigma_1(t)/\sigma_2(t) \\ r(t)[S_1(t) - S_1(t)\sigma_1(t)/\sigma_2(t)] &= \mu_1(t)S_1(t) - S_1(t)\mu_2(t)\sigma_1(t)/\sigma_2(t) \\ r(t)[\sigma_2(t) - \sigma_1(t)] &= \mu_1(t)\sigma_2(t) - \mu_2(t)\sigma_1(t) \\ r(t) &= \frac{\mu_1(t)\sigma_2(t) - \mu_2(t)\sigma_1(t)}{\sigma_2(t) - \sigma_1(t)} \end{aligned}$$

This can be rearranged to obtain:

$$\frac{\mu_2(t) - r(t)}{\sigma_2(t)} = \frac{\mu_1(t) - r(t)}{\sigma_1(t)}$$

Thus the no arbitrage principle forces this relation: The excess of a bond's drift over the default free rate divided by its volatility is independent of the maturity of the bond. The common value is a characteristic of the market rather than the security. This quantity is called the *market price of risk*. Although it is not a price, but the name is widely used.

So far we do not have a formula for valuing zero coupon bonds (or any other security) in terms of the short rate. The missing ingredient is the market price of risk. Therefore, in addition to our other assumptions, we assume that the market price of risk, denoted  $\lambda(r, t)$ , is known. Its relation to the parameters of the zero coupon bond  $P(r, t, T)$  is

$$\begin{aligned} \lambda(r, t) &= \frac{P(t, r, T)\mu_P(r, t, T) - rP(t, r, T)}{\sigma_P(r, t, T)P(t, r, T)} \\ &= \frac{\mu_P(r, t, T) - r}{\sigma_P(r, t, T)} \end{aligned} \tag{1.7}$$

where

$$\frac{dP(r, t, T)}{P(r, t, T)} = \mu_P(r, t, T)dt + \sigma_P(r, t, T)dZ$$

for all zero coupon bond maturities  $T$ . There are important consequences of these assumptions. The equation becomes a partial differential equation with known coefficients when we (again) rewrite it but now in terms of the bond price function and its derivatives.

$$\begin{aligned} \lambda \sigma_P &= \mu_P - r \\ \lambda P \sigma_P &= (\mu_P - r)P \\ \lambda \sigma \frac{\partial P}{\partial r} &= \frac{1}{2} \sigma^2 \frac{\partial^2 P}{\partial r^2} + \mu \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} - rP \end{aligned}$$

Therefore we can determine zero coupon bond prices in terms of the short rate drift, diffusion, and the market price of risk by solving this partial differential equation

$$\frac{1}{2} \sigma^2 \frac{\partial^2 P}{\partial r^2} + (\mu - \lambda \sigma) \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} - rP = 0 \tag{1.8}$$

for  $0 < r < \infty$ ,  $0 < t < T$ , subject to the boundary conditions:  $P(T, T) = 1$  and  $r(0) = r_0$  is the observed value of the short rate at  $t = 0$ .

We also rewrite the bond's stochastic differential equation incorporating the market price of risk:

$$\begin{aligned} \frac{dP}{P} &= \mu_P dt + \sigma_P dZ \\ &= (r + \lambda \sigma_P) dt + \sigma_P dZ \end{aligned} \tag{1.9}$$

The market price of risk may be zero, but this is an empirical issue. Stanton shows that for the US T-bill data January 1965 – July 1995 the market price of risk is significantly different from zero [17]. A nonzero market price of risk can be embedded in the model parameters. The mathematical details are described by Duffie [9], Musiela and Rutkowski [14], and others. Replace the original process  $\{Z(t)\}$  with  $\{Z^*(t)\}$  where  $Z^*(t) = Z(t) + \int_0^t \lambda(u) du$  or, in differential notation,  $dZ^* = dZ + \lambda dt$ . The original short rate dynamics

$$dr = \mu dt + \sigma dZ$$

becomes

$$\begin{aligned} dr &= (\mu - \lambda \sigma) dt + \sigma dZ^* \\ &= \mu^* dt + \sigma dZ^*. \end{aligned}$$

There is a new measure  $\mathbb{P}^*$  on the same probability space,  $(\Omega, \mathcal{F})$ , equivalent to  $\mathbb{P}$  in the sense that  $\mathbb{P}(A) = 0$  if and only if  $\mathbb{P}^*(A) = 0$ , for which the new process  $\{Z^*(t)\}$  is a standard Brownian motion adapted to the original filtration. Often at this point one simply works with the new process. The original process  $\{Z(t)\}$  is frequently called the *real* process or the *physical* process. The new process is called the *risk neutral process*. For the new dynamics,

zero coupon bond prices follow the dynamics like (1.9), but  $\{Z^*(t)\}$  is the driving process:

$$\frac{dP(r, t, T)}{P(r, t, T)} = \mu_P^*(r, t, T)dt + \sigma_P^*(r, t, T)dZ^*$$

where

$$\begin{aligned} \sigma_P^*(r, t, T) &= \sigma \frac{\partial P(r, t, T)}{\partial r} = \sigma_P(r, t, T) \\ \mu_P^*(r, t, T) &= \mu_P(r, t, T) - \lambda(r, t)\sigma_P^*(r, t, T) \\ &= \mu_P(r, t, T) - \lambda(r, t)\sigma_P(r, t, T) = r(t) \end{aligned}$$

Therefore

$$\frac{dP(r, t, T)}{P(r, t, T)} = rdt + \sigma_P(r, t, T)dZ^* \quad (\text{Use the } \mathbb{P}^* \text{ measure.})$$

or

$$\frac{dP(r, t, T)}{P(r, t, T)} = [r + \lambda\sigma_P]dt + \sigma_P(r, t, T)dZ \quad (\text{Use the } \mathbb{P} \text{ measure.})$$

The Feynman-Kač formula (we review this in the appendix) shows that zero coupon bond prices are expected discounted values, using the new measure:

$$P(t, r, T) = \mathbf{E}_{\mathbb{P}^*} \left[ e^{-\int_t^T r(u)du} | r_t = r \right]$$

This can be generalized. Consider a security with a single payment to be made at time  $T$  which is a known function  $H(r)$  of the short rate. The price at time  $t$  of the security is

$$\mathbf{E}_{\mathbb{P}^*} \left[ e^{-\int_t^T r(u)du} H(r(T)) | r_t = r \right].$$

This equation is the basis for valuation by Monte Carlo simulation and equation (1.8) is the basis for solution by numerical analysis techniques. If we have functions  $\sigma$  and  $\mu^* = \mu - \lambda\sigma$ , then for the Monte Carlo approach we can choose a random sample of interest rate paths  $\{r_j(u) : 0 \leq u \leq T, j = 1, \dots, N\}$  by using the following recursion  $N$  times. For a suitable, fixed value of  $k$ , let  $\Delta u = T/k$ ,  $u_0 = 0$  and  $u_i = iT/k$ . Each sample path starts with  $r(0)$  set to the current value of the short rate and is constant over each subinterval  $[u_{i-1}, u_i]$ . The successive values satisfy

$$r(u_i + \Delta u) = r(u_{i-1}) + \mu^*(r(u_{i-1}), u_{i-1})\Delta u + \sigma(r(u_{i-1}), u_{i-1})\epsilon_i$$

where  $\{\epsilon_i\}$  are independent normal random variables with mean zero and variance  $\Delta u$ . The average of the sample values

$$e^{-\int_t^T r_j(u)du} H(r_j(T)) = \exp \left( -\sum_{i=1}^k r_j(u_{i-1})\Delta u \right) H(r_j(T))$$

is an approximation to the price. For large enough samples and small enough time increments, the approximation is as good as it needs to be. Changing probability measures does not ameliorate the estimation problem. In order to get the risk neutral process drift  $\mu^*$  and volatility  $\sigma$  we have to estimate the physical process parameters (observations of the short

rate are sufficient for this) and the market price of risk (for which observations of a price are required).

## 2. NONPARAMETRIC ESTIMATION

**2.1. Basic idea.** Nonparametric estimation allows us to determine the drift and diffusion terms in a continuous time process, even though we observed it at discrete intervals. The estimation procedure is called nonparametric because no parametric assumption is made with regard to the drift, volatility and market price of risk functions. They are functions of  $r$  only, but nothing more is assumed. Actually, Stanton uses the daily (and for some calculations weekly and monthly) observations of the secondary market yields on 90-day US Treasury bills as a substitute for the short rate. The rates were converted from discount yield to an equivalent interest rate. The market price of risk estimates use the 180-day rate as well as the 90-day rate. The daily data consists of 7,262 observations. The observations are on average  $u = 30.5/7,262 = 0.004$  years apart. This corresponds to  $1/(0.004) = 250$  observations per year which seems correct since there are no observations on weekends or holidays. No adjustments were made for weekends and holidays – we simply assumed the observations are equally spaced. We followed Stanton and others in using the annualized yield to maturity corresponding to observations of the 90-day bond price as a surrogate for the short rate. Chan, Karolyi, Longstaff, Sanders and Ait-Sahalia use the same data and assumptions.

The stochastic differential equation for the short rate is  $dr = \mu(r_t)dt + \sigma(r_t)dZ_t$ . We have sample interest rates earned on 90-day zero coupon bond prices. The observed 90-day bond price  $P_i$  at time  $t_i$  is converted to yield-to-maturity by the formula

$$r_i = -\frac{1}{0.25} \log P(t_i, t_i + 0.25).$$

We take these to be observations of  $r$  denoted  $r_1, r_2, \dots, r_n$ , at times  $t_1, \dots, t_n$ . The observations are assumed to be equally spaced with  $u = t_{i+1} - t_i$ . Now we see how the methods described above are applied to obtain non-parametric estimators.

We begin with estimation of the probability density function  $f(r)$  of the short rate  $r$  and then describe the drift, volatility and market price of risk estimators.

**2.2. Estimating the marginal density.** The marginal density estimator is

$$\hat{f}(r) = \frac{1}{nh} \sum_{i=1}^n K((r - r_i)h^{-1}) \tag{2.1}$$

$$\tag{2.2}$$

where  $n = 7,626$  is the number of observations in the 30.5 years from January 1, 1965 to July 31, 1995.

This is a nonparametric estimator of the drift and requires only that we specify the function  $K(z)$ , called the *kernel density*, and the *bandwidth*  $h$ . Stanton uses the Gaussian kernel density defined by the standard normal probability density function:  $K(z) = (1/\sqrt{2\pi}) \exp(-z^2/2)$ . Simonoff [16, Chapter 3] and Campbell, Lo and MacKinley [5, page 500] discuss this and are good references for kernel density estimation. Simonoff suggests a

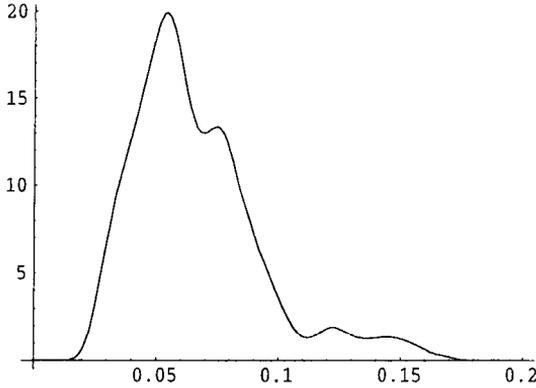


FIGURE 2. Estimated marginal density  $\hat{f}(r)$  of the short rate  $r$ , based on US T-bill data January 1965 to July 1995.

bandwidth of  $h = 1.059\sigma n^{-1/5} = 0.0047$ . We used the rounded value  $h = 0.005$ . Our graph of the marginal density is presented in Figure 2. It looks very much like Stanton's [17, Figure 3]. The estimator  $\hat{f}(r)$  is a mixture of  $n$  equally weighted densities

$$f_i(r) = \frac{1}{h\sqrt{2\pi}} \exp(-(r - r_i)^2/2h^2)$$

which are normal with mean equal to the observed value  $r_i$  and variance  $h^2$ . Simonoff describes other kernel densities and techniques to improve the estimator at boundaries. Since interest rates are bounded below at zero, these other techniques should be investigated.

**2.3. Estimation of the drift.** Stanton describes estimation of the drift and volatility in general terms as follows. Consider an Itô process

$$dX_t = \mu(X_t)dt + \sigma(X_t)dZ_t$$

with time-homogeneous drift and volatility. Stanton uses results on the *infinitesimal generator* of  $\{X_t\}$  ([15, page 115]). However, all of the first order approximations follow from more elementary analysis.

Consider the problem of estimating the drift and diffusion functions of  $\{X_t\}$  based on a sample of observations  $x_1, x_2, \dots, x_n$  at times  $t_1 < t_2 < \dots < t_n$ . For a first order approximation we consider an interval  $(t, t + u)$  and the corresponding change in  $X_t$ , using the integral representation:

$$X_{t+u} - X_t = \int_t^{t+u} \mu(X_s)ds + \int_t^{t+u} \sigma(X_s)dZ_s$$

The expectations are

$$E[X_{t+u} - X_t | X_t = x] = \int_t^{t+u} E[\mu(X_s) | X_t = x] ds$$

and therefore

$$E[\Delta_u X_t | X_t = x] \approx u\mu(x).$$

This gives us the first order approximation:

$$\mu(x) \approx \frac{1}{u} E[\Delta_u X_t | X_t = x]$$

Now use the sample values to estimate  $E[\Delta_u X_t | X_t = x]$  as follows:

$$E[\Delta_u X_t | X_t = x] \approx \frac{1}{c(x)} \sum_{i=1}^{n-1} (x_{i+1} - x_i) K[(x - x_i)h^{-1}].$$

We assign probability  $K((x - x_i)h^{-1})/c(x)$  to the increment  $x_{i+1} - x_i$  that was observed when  $t = t_i$  and  $X_t = x_i$ . The weight  $c(x)$  is selected so that

$$\sum_{i=1}^{n-1} K[(x - x_i)h^{-1}]/c(x) = 1$$

and

$$\sum_{i=1}^{n-1} K[(x - x_i)h^{-1}] = c(x).$$

Finally, we have the formula for the drift estimator:

$$\hat{\mu}(x) = \frac{1}{u c(x)} \sum_{i=1}^{n-1} (x_{i+1} - x_i) K[(x - x_i)h^{-1}] \tag{2.3}$$

where  $u = \Delta t_i$  is the length of the interval between observations. The estimator is a smooth function of  $x$ . The bandwidth  $h$  controls the smoothing effect of the estimator. We began with Simonoff's recommendation, which is essentially the same as the value Stanton suggests (although he does not give the value explicitly):  $h = 1.059\sigma n^{-1/5}$  where  $\sigma$  is the standard deviation of the sample  $\{r_i\}$  and  $n$  is the sample size. For the data set at hand, we computed a value of  $h = 0.0047$ . For a bandwidth of  $h = 0.005$ , our estimate of the marginal density  $f(r)$  for the observed short rate is essentially the same as Stanton's. However, to get our estimates of the short rate drift and diffusion to be as smooth as his we had to use a bandwidth of  $h = 0.02$ . Our Figure 3 is essentially the graph as Stanton's [17, Figure 4].

**2.4. Estimation of the diffusion.** For the estimator of the diffusion of an Itô process  $dX_t = \mu(X_t)dt + \sigma(X_t)dZ_t$ , consider again the integral representation:

$$X_{t+u} - X_t = \int_t^{t+u} \mu(X_s)ds + \int_t^{t+u} \sigma(X_s)dZ_s$$

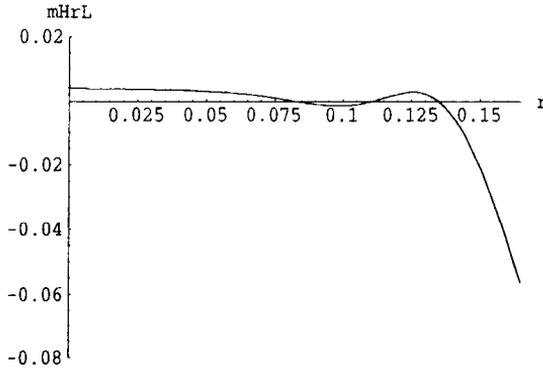


FIGURE 3. Estimated drift of the short rate process, based on US T-bill data January 1965 to July 1995.

For a very brief interval  $u$ , we can use this approximation to the increment, conditional on  $X_t = x$ :

$$(X_{t+u} - X_t | X_t = x) = \mu(x)u + \sigma(x)\Delta_u Z_t$$

Therefore, the conditional variance may be approximated as

$$\text{Var}(X_{t+u} - X_t | X_t = x) = \sigma^2(x)\text{Var}(\Delta_u Z_t) = \sigma^2(x)u.$$

This leads to the following estimator of the square of the diffusion function:

$$\begin{aligned} \hat{\sigma}^2(x) &= \frac{1}{u}\text{Var}(X_{t+u} - X_t | X_t = x) \\ &= \frac{1}{u}\text{E}[(X_{t+u} - \text{E}[X_{t+u}])^2 | X_t = x] \\ &= \frac{1}{u}\text{E}[(X_{t+u} - \hat{\mu}(x)u)^2 | X_t = x] \\ &= \frac{1}{uc(x)} \sum_{i=1}^{n-1} (x_{i+1} - u\hat{\mu}(x_i))^2 K((x - x_i)h^{-1}). \end{aligned} \tag{2.4}$$

where  $u = \Delta t_i$ . The graph of our estimator  $\hat{\sigma}(\tau)$  of the diffusion of the short rate process is shown in Figure 4. It essentially the same as Stanton's [17, Figure 5].

If the drift satisfies  $\sigma(0) = 0$  and  $\mu(0) > 0$  then the model interest rates will be positive with probability one. This is a desirable characteristic and it can be required as follows. Estimate the function  $g(x) = \sigma^2(x)/x$  using the formula

$$\hat{g}(x) = \frac{1}{uc(x)} \sum_{i=1}^{n-1} \frac{(x_{i+1} - u\hat{\mu}(x_i))^2}{x_i} K((x - x_i)h^{-1}).$$

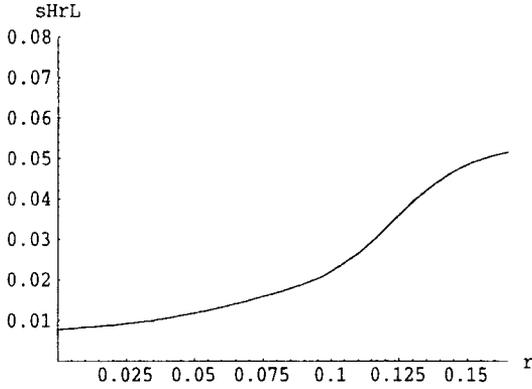


FIGURE 4. Estimated diffusion of the short rate process, based on US T-bill data January 1965 to July 1995.

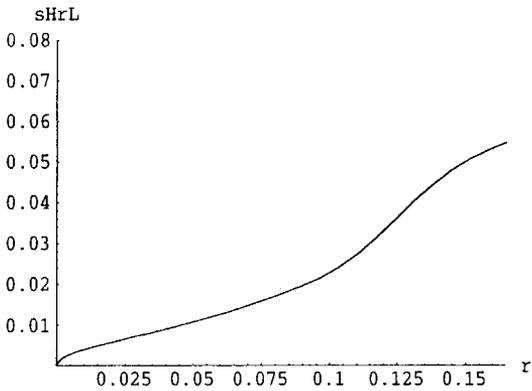


FIGURE 5. Estimated constrained diffusion of the short rate process, based on US T-bill data January 1965 to July 1995. The constraint  $\sigma(0) = 0$  guarantees positive interest rates.

Now use the product of  $\widehat{g}(x)$  and  $x$  as a new estimator:

$$\widehat{\sigma}_0^2(x) = x\widehat{\sigma}^2(x)$$

and

$$\widehat{\sigma}_0(x) = \sqrt{\widehat{\sigma}_0^2(x)} \tag{2.5}$$

Figure 5 is the graph of the constrained diffusion function. Again this is essentially the same as Stanton's estimate.

2.5. **Estimating the Market Price of Risk.** Let  $P^{(1)}(t_i)$  denote the price of the 180-day zero coupon bond, observed at time  $t_i$ . Actually the data provides us with  $y_i = \frac{1}{0.5} \log P^{(1)}(t_i)$ . Consider the yield over the period  $(t_i, t_i + 0.25)$  obtained by buying the 180-day bond at  $t_i$  for  $P_i^{(1)}$  and selling it at  $t_i + 0.25$ . When sold, it is a 90-day bond and its price is  $P^{(1)}(t_i + 0.25)$ . We can calculate the observed yield on 180-day bonds over the 90-day period  $(t_i, t_i + 0.25)$  as

$$\begin{aligned} R^{(1)}(t_i) &= \frac{1}{0.25} \log \frac{P^{(1)}(t_i + 0.25)}{P^{(1)}(t_i)} \\ &= \frac{1}{0.25} (-0.25r_{t_i+0.25} + 0.5y_i) \\ &= 2y_i - r_{t_i+0.25} \approx \frac{1}{0.25} \frac{P^{(1)}(t_i + 0.25) - P^{(1)}(t_i)}{P^{(1)}(t_i)} \end{aligned}$$

Of course this is only “observed” at those times  $t_i$  for which there are observations at  $t_i + 0.25$  also.

Let  $R^{(2)}(t_i)$  be the annualized return on the 90-day bond over the same 90-day period  $(t_i, t_i + 0.25)$ . Evidently,  $R^{(2)}(t_i) = r_i$ , but the following method will produce an estimate of the market price of risk for returns on any two zero coupon bonds so we will write it out in general terms.

Consider two zero coupon bonds with prices  $P(t, r, T_j)$  for  $j = 1, 2$ . The instantaneous return  $\frac{dP^{(j)}}{P^{(j)}}$  satisfies the stochastic differential equation (1.9), which we rewrite with slightly different notation:

$$\begin{aligned} \frac{dP^{(j)}}{P^{(j)}} &= \mu^{(j)}(r_t, t)dt + \sigma^{(j)}(r_t, t)dZ_t \\ &= [r(t) + \lambda(r_t)\sigma^{(j)}(r_t, t)] dt + \sigma^{(j)}(r_t, t)dZ_t \end{aligned}$$

Let  $q = 0.25$  be the length of the interval  $(t, t + q)$  over which the returns are defined. As a rough approximation, we can write the effective return as follows:

$$\begin{aligned} qR^{(j)} &\approx \frac{P^{(1)}(t_i + q) - P^{(1)}(t_i)}{P^{(1)}(t_i)} \approx \mu^{(j)}(r_t, t)q + \sigma^{(j)}(r_t, t)\Delta_q Z_t \\ &= [r(t) + \lambda(r_t)\sigma^{(j)}(r_t, t)] q + \sigma^{(j)}(r_t, t)\Delta_q Z_t \end{aligned}$$

This justifies using the sample values  $R_i^{(j)}$  to estimate  $\mu^{(j)}(r_t, t)$  and  $\sigma^{(j)}(r_t, t)$ . Note that the bond drift and volatility depend on  $t$  in general, but we are looking at the same part of the lifetime of each observation of  $P^{(j)}$  in the sample, so all depend on  $t$  in the same way. We can drop the dependence on  $t$ . Just keep in mind that we are estimating the drift and diffusion in the first 90-days of a 180-day bond. Write the drift of as  $\mu^{(j)}(r)$  and estimate it as follows:

$$\hat{\mu}^{(j)}(r) = \frac{1}{c(r)} \sum_{i=1}^n R_i^{(j)} K[(r - r_i)h^{-1}]$$

with  $h = 0.02$  and

$$c(r) = \sum_{i=1}^n K[(r - r_i)h^{-1}].$$

Estimate the volatility as follows:

$$\begin{aligned} (\hat{\sigma}^{(j)}(r))^2 q &= \frac{1}{c(r)} \sum_{i=1}^n [qR_i^{(j)} - q\hat{\mu}^{(j)}(r_i)]^2 K[(r - r_i)h^{-1}] \\ &= \frac{q^2}{c(r)} \sum_{i=1}^n [R_i^{(j)} - \hat{\mu}^{(j)}(r_i)]^2 K[(r - r_i)h^{-1}] \end{aligned}$$

Therefore, we have

$$(\hat{\sigma}^{(j)}(r))^2 = \frac{q}{c(r)} \sum_{i=1}^n [R_i^{(j)} - \hat{\mu}^{(j)}(r_i)]^2 K[(r - r_i)h^{-1}]$$

and

$$\hat{\sigma}^{(j)}(r) = - \left( \frac{q}{c(r)} \sum_{i=1}^n [R_i^{(j)} - \hat{\mu}^{(j)}(r_i)]^2 K[(r - r_i)h^{-1}] \right)^{1/2}.$$

Now calculate the excess return of bond 1 over bond 2.

$$\begin{aligned} \frac{dP_t^{(1)}}{P_t^{(1)}} - \frac{dP_t^{(2)}}{P_t^{(2)}} &= [\mu^{(1)}(r_t, t) - \mu^{(2)}(r_t, t)] dt + [\sigma^{(1)}(r_t, t) - \sigma^{(2)}(r_t, t)] dZ_t \\ &= \lambda(r_t) [\sigma^{(1)}(r_t, t) - \sigma^{(2)}(r_t, t)] dt + [\sigma^{(1)}(r_t, t) - \sigma^{(2)}(r_t, t)] dZ \end{aligned}$$

from which we find that

$$qR_i^{(1)} - qR_i^{(2)} \approx \lambda(r_i) [\sigma^{(1)}(r_i, t) - \sigma^{(2)}(r_i, t)] q + [\sigma^{(1)}(r_i, t_i) - \sigma^{(2)}(r_i, t_i)] \Delta_q Z_t.$$

Now cancel the  $q$  and use the estimators for the drift and volatility terms. This gives us an estimator for the market price of risk:

$$\hat{\lambda}(r) = \frac{\hat{\mu}^{(1)}(r) - \hat{\mu}^{(2)}(r)}{\hat{\sigma}^{(1)}(r) - \hat{\sigma}^{(2)}(r)} \tag{2.6}$$

Figures 6 and 7 show the graphs of the observations of pairs  $(r_i, R_i^{(j)})$  and the estimator  $\hat{\mu}^{(j)}(r)$ . This differs from Stanton's estimator in two ways. First, we are using annualized returns  $R_i^{(j)}$  and Stanton used effective returns so he has a factor of  $q$  in the denominator, but ours is included in the formula for the estimator of the volatility. Second, Stanton defines the volatility differently. His definition is  $\Lambda(r) = \sigma_0(r)\lambda(r)$ , where  $\lambda(r)$  is our market price of risk and  $\sigma_0(r)$  is the short rate volatility, constrained so that  $\sigma_0(0) = 0$ .

The difference  $\hat{\mu}^{(1)}(r) - \hat{\mu}^{(2)}(r)$  is the numerator of the market price of risk estimator. Their graphs are given in Figure 8. The difference is rather small. This is consistent with the sample statistics from the observations of returns shown in Table 1.

NONPARAMETRIC ESTIMATION OF TERM STRUCTURE

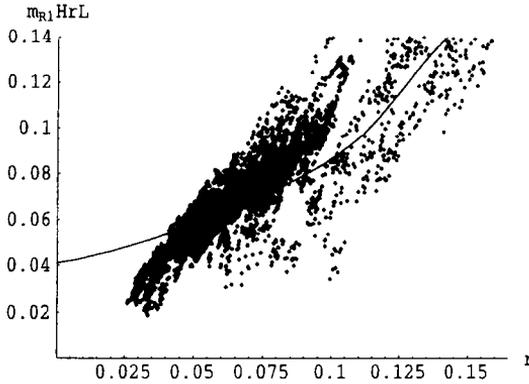


FIGURE 6. Return on 180-day bond over 90 days, based on US T-bill data January 1965 to July 1995 and the drift  $\mu^{(1)}(r)$ .

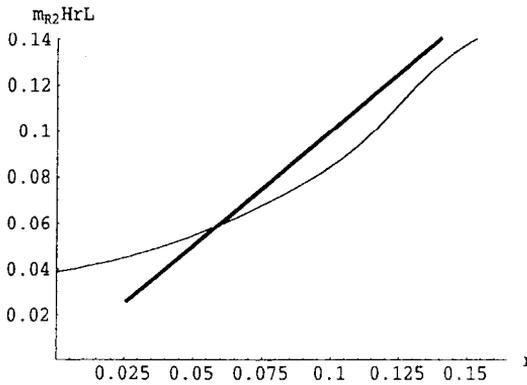


FIGURE 7. Return on 90-day bond over 90 days, based on US T-bill data January 1965 to July 1995 and the drift  $\mu_2(r)$ .

We estimated the diffusion functions as follows. First we calculated

$$\hat{\sigma}_j^2(r) = \frac{1}{\Delta c(r)} \sum_{i=1}^{n-1} (R_{j,i+1} - R_{j,i})^2 K((r - r_i)h^{-1}) \quad (2.7)$$

with  $h = 0.018$ . Now the estimators are given by the negative square root:

$$\hat{\sigma}_j(r) = -\sqrt{\hat{\sigma}_j^2(r)} \quad (2.8)$$

The negative square root is required for consistency with formula (1.2), which describes the relation between the diffusion of the bond price and the diffusion of the short rate:

$$\sigma_P = \sigma \frac{1}{P} \frac{\partial P}{\partial r}$$

Bond	Mean Over 90-days $R^{(j)}$	Yield First	Standard Deviation
180-day ( $j = 1$ )	0.06957		0.03012
90-day ( $j = 2$ )	0.06637		0.02658
Difference	0.00320		0.00354

TABLE 1. Sample statistics of observed 90-day returns of 180-day bonds and 90-day bonds from January 1, 1965 to July 31, 1995.

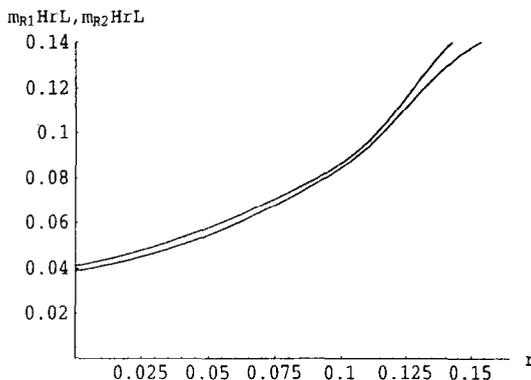


FIGURE 8. The estimators  $\hat{\mu}_1(r)$  and  $\hat{\mu}_2(r)$  of the returns of 180-day bonds (1) and 90-day bonds (2), based on US T-bill data January 1965 to July 1995.

Since the price  $P$  decreases as  $r$  increases, then the partial derivative is negative. The diffusion of the price has the opposite sign as the diffusion of the drift. The short rate diffusion coefficient  $\sigma(r)$  is positive, so we use the negative square root for the diffusion coefficient of the bond price. Figure 9 shows the graphs of  $\hat{\sigma}_1(r)$  and  $\hat{\sigma}_2(r)$ . The difference is very small. This could be a problem since it is the denominator of the estimator of the market price of risk.

The sample excess return statistics in Table 1 indicate that numerator should have an order of magnitude of 0.00320. The denominator should have an order of magnitude about  $-(0.00354)$ . We should expect the market price of risk to be around -1. The graph of our estimator is shown in Figure 10. The graph ranges from -2.5 to -0.5, which is consistent with the sample statistics.

Our definition of the market price of risk is more or less standard [5, page 434], [13, page 320], although Ingersoll defines the market price the same way Stanton defines it [12]. The relation between Stanton's definition, which we denote  $\Lambda(r)$ , and ours is  $\Lambda(r) = \sigma_0(r)\lambda(r)$ . In other words his estimator is the (constrained) short rate diffusion multiplied by our estimator of the market price of risk. We get essentially the same values for  $\Lambda(r)$ . The graph is shown in Figure 11. In summary, we have essentially replicated Stanton's nonparametric estimates of

## NONPARAMETRIC ESTIMATION OF TERM STRUCTURE

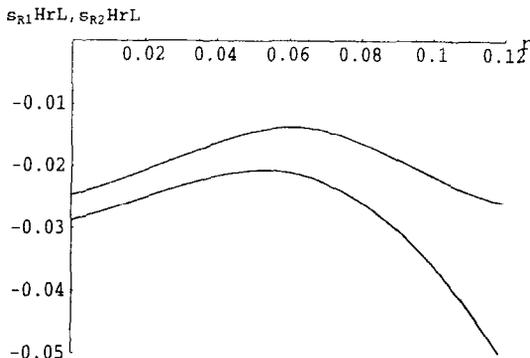


FIGURE 9. The diffusion estimators,  $\hat{\sigma}_1(r)$  and  $\hat{\sigma}_2(r)$ , of 90-day returns of 180-day bonds (1, dashed graph) and 90-day bonds (2), based on US T-bill data January 1965 to July 1995.

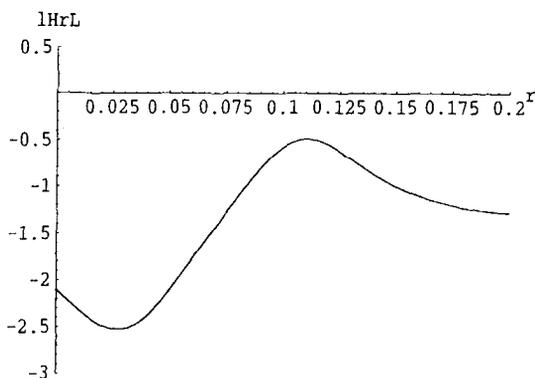


FIGURE 10. The market price of risk estimator,  $\hat{\lambda}(r)$  based on excess 90-day returns of 180-day bonds (1) and 90-day bonds (2) for US T-bill data January 1965 to July 1995.

the drift and diffusion of the short rate and the market price of risk. Of course the definition of the market price of risk is not critical providing it is interpreted properly.

### 3. ASSET MODELING

At the 1998 ASTIN Colloquium in Glasgow, Gary G. Venter presented criterion yield curve models should satisfy in order to be used for stochastic modeling of property and liability risks [20]. This is called dynamic financial analysis (DFA) in the United States. Venter suggests that the model should meet these general criteria:

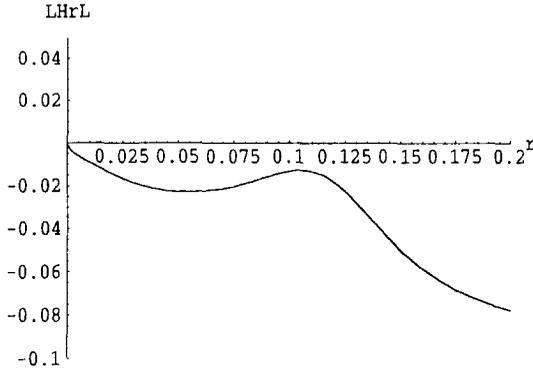


FIGURE 11. Stanton's market price of risk estimator,  $\hat{\Lambda}(r) = \hat{\sigma}_0(r)\hat{\lambda}(r)$ .

- (1) It should closely approximate the current yield curve.
- (2) It should produce patterns of change in the short-term rate that match those produced historically.
- (3) Over longer simulations, the ultimate distributions of yield curve shapes it produces, given any short-term rate, should match historical results.

The first criteria, we believe, is not so important especially for long term models. The nonparametric estimators based on historical data will necessarily satisfy the second two criteria. This suggests that nonparametric estimators may be useful in DFA models.

#### 4. CONCLUSION

We reviewed and verified recent work of Stanton [17] in estimation of the drift and diffusion of the short rate and the market price of risk in a bond market. The nonparametric approach should be of interest to actuaries because it seems to be a convenient way reflecting historical interest rate properties in the model. There is a good bit of work to be done in the future. Basing the market price of risk on only two bond prices is not adequate. In the future we expect to obtain observations of the yield curve as a function of time, with longer term bonds included in the observations. The data set used here has only 90-day and 180-day bonds; we would like to include 1-year, 5-year, 10-year and 20-year prices as well. This seems to be especially important in estimating the market price of risk. Beyond the estimation problem, one needs to develop methods of efficiently incorporating nonparametric estimators in insurance models. We expect to report on both problems in the future.

#### REFERENCES

1. Y. Ait-Sahalia, *Nonparametric pricing of interest rate derivative securities*, Working paper, Graduate School of Business, University of Chicago, 1992.
2. Tomas Björk, *Arbitrage theory in continuous time*, Oxford University Press, Oxford, 1998.
3. F. Black, E. Derman, and W. Toy, *A one-factor model of interest rates and its application to treasury bond options*, Financial Analysts Journal (1990), no. January-February, 33-39.

NONPARAMETRIC ESTIMATION OF TERM STRUCTURE

4. F. Black and P. Karasinski, *Bond and option pricing when short rates are lognormal*, Financial Analysts Journal (1991), no. July, 52-59.
5. J. Campbell, A. Lo, and A. MacKinlay, *The econometrics of financial markets*, Princeton University Press, Princeton, New Jersey, 1997.
6. K. Chan, G. Karolyi, F. Longstaff, and A. Saunders, *An empirical comparison of alternative models of the short-term interest rate*, Journal of Finance **47** (1992), 1209-1227.
7. S. Christiansen, *A practical guide to interest rate generators for c-3 risk*, Transactions of the Society of Actuaries **XLIV** (1992), 101-134.
8. J. C. Cox, Jr. J. E. Ingersoll, and S. A. Ross, *A theory of the term structure of interest rates*, Econometrica **53** (1985), 363-384.
9. Darrell Duffie, *Dynamic asset pricing theory*, 2nd ed., Princeton University Press, Princeton NJ 08540, 1996.
10. T.S.Y. Ho and S.-B. Lee, *Term structure movements and pricing interest rate contingent claims*, Journal of Finance **41** (1986), no. December 1986, 1011-1029.
11. John Hull and Alan White, *Pricing interest-rate-derivative securities*, Review of Financial Studies **3** (1990), no. 4, 573-592.
12. J. Ingersoll, *Theory of financial decision making*, Rowman & Littlefield, Totowa, NJ, 1987.
13. Y. K. Kwok, *Mathematical models of financial derivatives*, Springer, New York, 1998.
14. M. Musiela and M. Rutkowski, *Martingale methods in financial modelling*, Springer-Verlag, New York, 1997.
15. Brent Øksendal, *Stochastic differential equations*, fifth ed., Springer-Verlag, Berlin Heidelberg, 1998.
16. Jeffery S. Simonoff, *Smoothing methods in statistics*, Springer, New York, 1996.
17. Richard Stanton, *A nonparameteric model of term structure dynamics and the market price of risk*, JF (1997), 1973-2002.
18. J. Tilley, *An actuarial layman's guide to building stochastic interest rate generators*, Transactions of the Society of Actuaries **XLIV** (1992), 509-538.
19. Oldrich A. Vasicek, *An equilibrium characterization of the term structure*, Journal of Financial Economics **5** (1977), 177-188.
20. Gary G. Venter, *Asset modelling - empirical tests of yield curve generators*, 1998 General Insurance Convention and ASTIN Colloquium, 1998, pp. 177-189.

APPENDIX A. REVIEW OF ITÔ PROCESSES

**A.1. Brownian Motion - The Basis of Short Rate Models.** This is a brief, intuitive, non-rigorous discussion of Itô processes. We recommend study of one of the well-known texts for those interested in interest rate models. A stochastic process is an indexed family of random variables  $\{X_t : 0 \leq t \leq T^*\}$  defined on a probability space  $(\Omega, \mathbb{P}, \mathcal{F})$ . For each element  $\omega \in \Omega$ , the values  $X_t(\omega)$  define a deterministic (nonrandom) function of time  $t$ , called the sample path corresponding to  $\omega$ . For all times  $t$  and real numbers  $a, b$  we must be able to determine the probability of the event

$$\{\omega \in \Omega | a < X_t(\omega) \leq b\}$$

so each such event must be a member of  $\mathcal{F}$ , the set of events on which the probability measure is defined.

A standard Brownian motion  $\{Z_t : 0 \leq t \leq T^*\}$  is a stochastic process, defined on a probability space  $(\Omega, \mathbb{P}, \mathcal{F})$ , satisfying:

1.  $Z$  has independent increments.
2.  $Z(t + \Delta t) - Z(t)$  has a normal distribution with mean zero and variance  $\Delta t$
3.  $Z$  has continuous sample paths.
4.  $Z(0) = 0$

Higher dimensional standard Brownian motion is a vector valued process with components which are (one dimensional) standard Brownian motion. A  $k$ -factor interest rate model is based on a  $k$ -dimensional standard Brownian motion. Throughout this paper we are concerned with one-factor short rate models, so we will describe only one dimensional Itô processes.

In general (a one-factor) Itô process  $\{X_t\}$  is one that can be written in the form

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dZ_t$$

where  $\{Z_t : 0 \leq t \leq T^*\}$  is a standard Brownian motion. This is merely a symbolic description of an operational definition in terms of two integrals:

$$X_{t+h} - X_t = \int_t^{t+h} \mu(X_u, u)du + \int_t^{t+h} \sigma(X_u, u)dZ_u$$

The first integral (with respect to time) is simply the usual integral of calculus. The result is random because the path  $\{X_u | t \leq u \leq t+h\}$  is random, but given the path we calculate the integral by integrating over time. There is nothing new here. This component is sometimes called the path integral to distinguish it from the second component, the Itô integral.

The Itô integral is defined as a limit of finite sums. Given a partition  $t = t_0 < t_1 < t_2 < \dots < t_n = t+h$  of the interval of integration, let  $Y_n$  denote the sum

$$Y_n = \sum_{i=0}^{n-1} \sigma(X(t_i), t_i)\Delta Z(t_i)$$

where  $\Delta Z(t_i) = Z(t_{i+1}) - Z(t_i)$  is the increment<sup>3</sup> in  $Z$  over  $[t_i, t_{i+1})$ . The values of  $X(t_i)$  are calculated recursively, in terms of the path of  $Z$ . Starting with the current time  $t$ :

$$\begin{aligned} X(t_0) &= X(t) \\ X(t_{i+1}) &= X(t_i) + \mu(X(t_i), t_i)\Delta t_i + \sigma(X(t_i), t_i)\Delta Z(t_i) \end{aligned}$$

which can be written

$$\Delta X(t_i) = \mu(X(t_i), t_i)\Delta t_i + \sigma(X(t_i), t_i)\Delta Z(t_i)$$

Given  $Z(t_i)$ , we can calculate the conditional moments of  $\Delta X(t_i)$ :

$$\begin{aligned} E[\Delta X(t_i)|Z(t_i)] &= \mu(X(t_i), t_i)\Delta t_i + \sigma(X(t_i), t_i)E[\Delta Z(t_i)|Z(t_i)] \\ &= \mu(X(t_i), t_i)\Delta t_i \end{aligned}$$

and

$$\begin{aligned} \text{Var}[\Delta X(t_i)|Z(t_i)] &= \sigma^2(X(t_i), t_i)\text{Var}[\Delta Z(t_i)|Z(t_i)] \\ &= \sigma^2(X(t_i), t_i)\Delta t_i \end{aligned}$$

Given the entire path of  $Z$  over the partition,  $\{Z(t_i)\}$ , we can calculate the conditional moments of  $Y_n$ :

$$\begin{aligned} E[Y_n|\{Z(t_i)\}] &= \sum_{i=0}^{n-1} E[\sigma(X(t_i), t_i)E[\Delta Z(t_i)]] \\ &= 0 \\ \text{Var}[Y_n|\{Z(t_i)\}] &= \sum_{i=0}^{n-1} \sigma^2(X(t_i), t_i)\text{Var}[\Delta Z(t_i)] \\ &= \sum_{i=0}^{n-1} \sigma^2(X(t_i), t_i)\Delta t_i \\ &\approx \int_t^{t+h} \sigma^2(X_u, u)du \end{aligned}$$

The conditional means are zero, so the unconditional mean is too,  $E[Y_n] = 0$ . The conditional variance tends to a path integral and its mean is the variance of  $Y_n$ .

This suggests (and one can prove along the suggested lines) that as the partitions become finer and finer, the sums  $Y_n$  converge to a random variable, denoted

$$\int_t^{t+h} \sigma(X_u, u)dZ_u,$$

with mean zero

$$E\left[\int_t^{t+h} \sigma(X_u, u)dZ_u\right] = 0$$

---

<sup>3</sup>The notation  $Z(t_i)$  means the same thing as  $Z_{t_i}$ . We use it to avoid double subscripts.

and variance

$$\text{Var} \left[ \int_t^{t+h} \sigma(X_u, u) dZ_u \right] = \text{E} \left[ \int_t^{t+h} \sigma^2(X_u, u) du \right].$$

In the most popular term structure models, the *short rate*  $r$  is an Itô process:

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dZ_t$$

where  $\{Z_t : 0 \leq t \leq T^*\}$  is a standard Brownian motion. The drift and diffusion functions  $\mu(r, t)$  and  $\sigma(r, t)$  are determined from observations of the market. The Itô formula is the fundamental tool in describing the bond market.

**A.2. Itô's Formula.** If a process  $\{X_t\}$  has an Itô representation  $dX = \mu dt + \sigma dZ$  and  $f(x, t)$  is a smooth function defined on the state space of  $\{X(t)\}$ , then the process defined by  $Y_t = f(X_t, t)$  for all  $t$  is also an Itô process and its representation is

$$dY_t = \mu_Y dt + \sigma_Y dZ_t$$

where

$$\mu_Y(x, t) = \frac{1}{2}\sigma^2(x, t)\frac{\partial^2 f(x, t)}{\partial x^2} + \mu(x, t)\frac{\partial f(x, t)}{\partial x} + \frac{\partial f(x, t)}{\partial t}$$

and

$$\sigma_Y(x, t) = \sigma(x, t)\frac{\partial f(x, t)}{\partial x}$$

Here is a way to remember the formula. Write the two variable Taylor series expansion of  $Y(t + dt)$  and use the Itô differential multiplication rules:  $(dZ)^2 = dt$ ,  $dZdt = 0$  and  $(dt)^2 = 0$ . The rules imply that  $(dZ)^k = dt(dZ)^{k-2} = 0$  for  $k \geq 3$ . Apply the rules to  $dX$  to obtain  $(dX)^2 = \sigma^2 dt$  and  $dt dX = 0$ . So only a few terms of the expansion are nonzero:

$$Y(t + dt) = f(X_t, t) + \frac{\partial f(X_t, t)}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f(X_t, t)}{\partial x^2} (dX_t)^2 + \frac{\partial f(X_t, t)}{\partial t} dt$$

and so

$$\begin{aligned} dY_t &= Y(t + dt) - Y_t \\ &= \frac{\partial f(X_t, t)}{\partial x} (\mu dt + \sigma dZ) + \frac{1}{2} \frac{\partial^2 f(X_t, t)}{\partial x^2} \sigma^2 dt + \frac{\partial f(X_t, t)}{\partial t} dt \\ &= \left( \frac{1}{2} \sigma^2(x, t) \frac{\partial^2 f(x, t)}{\partial x^2} + \mu(x, t) \frac{\partial f(x, t)}{\partial x} + \frac{\partial f(x, t)}{\partial t} \right) dt + \frac{\partial f(X_t, t)}{\partial x} \sigma dZ \end{aligned}$$

The Itô rule  $(dZ)^2 = dt$  is the differential form of the intergral equation

$$\int_0^t (dZ_u)^2 = t.$$

The integral is defined as the limit of the sums

$$Y_n = \sum_{i=0}^{n-1} (\Delta Z(t_i))^2$$

where

$$\Delta Z(t_i) = Z(t_{i+1}) - Z(t_i)$$

and  $t_i = i \frac{t}{n}$ . Since  $\Delta Z(t_i)$  is normal with mean zero and variance  $\Delta t = t/n$ , then  $(\Delta Z(t_i))^2/\Delta t$  is  $\chi$ -square with one degree of freedom. Therefore, we can compute the moments of  $(\Delta Z(t_i))^2$  as follows:

$$E[(\Delta Z(t_i))^2] = \Delta t E \left[ \frac{(\Delta Z(t_i))^2}{\Delta t} \right] = \Delta t$$

and

$$\text{Var}[(\Delta Z(t_i))^2] = (\Delta t)^2 \text{Var} \left[ \frac{(\Delta Z(t_i))^2}{\Delta t} \right] = (\Delta t)^2 2$$

Since the increments of  $Z$  are independent the squares of the increments are independent too. So we can now determine the moments of  $Y_n$ :

$$E[Y_n] = \sum_{i=0}^{n-1} E[(\Delta Z(t_i))^2] = n\Delta t = t$$

and

$$\text{Var}[Y_n] = \sum_{i=0}^{n-1} \text{Var}[(\Delta Z(t_i))^2] = 2n(\Delta t)^2 = \frac{2t^2}{n}$$

Therefore in the limit as  $n$  tends to infinity, the limiting distribution of  $Y_n$  has zero variance. Thus  $Y_n$  simply tends to its mean value and we have

$$\int_0^t (dZ_u)^2 = t \quad \text{or in differential notation} \quad (dZ)^2 = dt.$$

**A.3. Feynmann-Kač Formula.** The price at time  $t$  is a function of  $t$  and the current value of the short rate  $r_t = r$ . The price of a zero coupon bond is the expected discounted value of its future cash payment. In general, the price of a security that pays its owner a single payment of  $H(r_T)$  at time  $T$  is also the expected discounted value of its cash payment. The expectation is taken over the paths described by the differential equation

$$dr_t = [\mu(r_t, t) - \lambda(r_t, t)\sigma(r_t, t)]dt + \sigma(r_t, t)dZ_t$$

subject to  $r_t = r$ . This follows from the Feynmann-Kač formula.

*Feynmann-Kač Formula.* Let  $\{X_t\}$  denote an Itô process with equation  $dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dZ_t$ . Let  $H(x)$  be a given function for the states of  $X_T$ . Define the function:

$$V(x, t) = E[H(X_T) \exp(-\int_t^T X_u du) | X_t = x]$$

Then  $V(x, t)$  satisfies the partial differential equation

$$\frac{\partial V(x, t)}{\partial t} + \mu(x, t) \frac{\partial V(x, t)}{\partial x} + \frac{1}{2} \sigma^2(x, t) \frac{\partial^2 V(x, t)}{\partial x^2} - xV(x, t) = 0$$

subject to  $V(x, T) = H(x)$  for all  $x$ .

This is an intuitive justification for the formula. Let  $Y_t = \exp(-\int_t^T X_u du)$  for  $t \leq T$ . The derivative of the path integral  $-\int_t^T X_u du$  with respect to the lower limit  $t$  is  $X_t dt$ . So by Itô's formula with  $f(x, t) = e^x$  the differential of  $Y_t$  is  $dY_t = Y_t X_t dt$ .

Another application of the Itô formula with  $f(x, t) = V(x, t)$  shows that the differential of  $Y_t = V(X_t, t)$  is

$$dY_t = \left[ \frac{\partial V(X_t, t)}{\partial t} + \mu(X_t, t) \frac{\partial V(X_t, t)}{\partial x} + \frac{1}{2} \sigma^2(X_t, t) \frac{\partial^2 V(X_t, t)}{\partial x^2} \right] dt + \sigma(X_t, t) \frac{\partial V(X_t, t)}{\partial x} dZ_t.$$

Now equate the two conditional expectations. The first representation gives

$$E[dY_t | X_t = x] = E[X_t Y_t dt | X_t = x] = x dt E[Y_t | X_t = x] = xV(x, t) dt.$$

The second gives

$$\begin{aligned} E[dY_t | X_t = x] &= E \left[ \frac{\partial V(X_t, t)}{\partial t} + \mu(X_t, t) \frac{\partial V(X_t, t)}{\partial x} + \frac{1}{2} \sigma^2(X_t, t) \frac{\partial^2 V(X_t, t)}{\partial x^2} \middle| X_t = x \right] dt \\ &\quad + E \left[ \sigma(X_t, t) \frac{\partial V(X_t, t)}{\partial x} dZ_t \middle| X_t = x \right] \\ &= \left[ \frac{\partial V(x, t)}{\partial t} + \mu(x, t) \frac{\partial V(x, t)}{\partial x} + \frac{1}{2} \sigma^2(x, t) \frac{\partial^2 V(x, t)}{\partial x^2} \right] dt + \sigma(x, t) \frac{\partial V(x, t)}{\partial x} E[dZ_t] \\ &= \left[ \frac{\partial V(x, t)}{\partial t} + \mu(x, t) \frac{\partial V(x, t)}{\partial x} + \frac{1}{2} \sigma^2(x, t) \frac{\partial^2 V(x, t)}{\partial x^2} \right] dt \end{aligned}$$

Equating the two expressions yields the Feynmann-Kač formula. The boundary condition follows from the fact that  $Y_T = H(X_T)$ .

This is applied to bond markets with the short rate process  $\{r_t\}$  in place of  $\{X_t\}$ . The physical short rate process is

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dZ_t.$$

The price of a short rate derivative at time  $t$  with current short rate  $r$  is denoted  $V(r, t)$ . It pays its owner a single payment of  $H(r_T)$  at time  $T$ . As before, it satisfies the partial differential equation

$$\frac{\partial V(r, t)}{\partial t} + [\mu(r, t) - \lambda(r_t, t)\sigma(r_t, t)] \frac{\partial V(r, t)}{\partial r} + \frac{1}{2} \sigma^2(r, t) \frac{\partial^2 V(r, t)}{\partial r^2} - rV(r, t) = 0$$

subject to  $V(r, T) = H(r)$  for all  $r$ .

By the Feynmann-Kač formula the function

$$V^*(r, t) = E^* \left[ H(r_T) \exp \left( - \int_t^T r_u du \right) \middle| r_t = r \right]$$

where short rate process is driven by

$$dr_t = [\mu(r_t, t) - \lambda(r_t, t)\sigma(r_t, t)]dt + \sigma(r_t, t)dZ_t^*$$

satisfies the same partial differential equation and boundary condition. Hence,  $V^*(r, t) = V(r, t)$ . We have two methods of numerically calculating the derivative price. We can solve the partial differential equation numerically or we can calculate the expected discounted value by simulating the short rate process, being careful to use the drift  $\mu - \lambda\sigma$ , and evaluate the value  $H(r_T) \exp \left( - \int_t^T r_u du \right)$  many times the average of the simulated values approximates the price.

(Samuel H. Cox and Hal W. Pedersen) ACTUARIAL SCIENCE PROGRAM, DEPARTMENT OF RISK MANAGEMENT AND INSURANCE, GEORGIA STATE UNIVERSITY, ATLANTA, GEORGIA, 30302 USA

*E-mail address*, Samuel H. Cox: [samcox@gsu.edu](mailto:samcox@gsu.edu)

*E-mail address*, Hal W. Pedersen: [inshwp@panther.gsu.edu](mailto:inshwp@panther.gsu.edu)

