

Does Conditioning Information Matter in Estimating Continuous Time Interest Rate Diffusions?

Abhay Abhyankar and Devraj Basu*

Abstract

We examine an important aspect of empirical estimation of term structure models; the role of conditioning information in dynamic term structure models. The use of both real world or simulated data implicitly incorporates conditioning information. We examine the bias created in estimating the drift by a specific form of conditioning, namely truncation. Using the theory of enlargement of filtrations we provide estimates of the extent of this truncation bias for commonly used short rate models. We find that this truncation bias causes the drift of these models to have a nonlinear structure.

I. Introduction

The short rate is a key building block in many continuous time models of the term structure of interest rates. For example, in time-homogenous single-factor models (Vasicek (1977), Cox, Ingersoll, and Ross (1985)), the state variable is taken to be the short rate. It is not surprising, therefore, that empirical research has focused on estimating the parameters of diffusion processes used to model the short rate. Recent research using nonparametric techniques (e.g., Ait-Sahalia (1996)) provides some evidence that the drift is highly nonlinear. This finding presents a considerable challenge to existing term structure models. Several authors (e.g., Chapman and Pearson (2000)) show that biases in the nonparametric estimators might lead to spurious nonlinearities being observed. Further, Bandi (2000), using estimation methods robust to nonstationarity, finds statistical evidence of martingale behavior over the lower range of interest rate values but is unable to draw strong conclusions about nonlinearity during episodes of very high interest rates in the 1973–1995 period. Jones (1999), using a Bayesian approach, concludes that the large negative drift for high interest rates reported in previous work can be confirmed only under prior distributions that are best described as

*Both authors, Warwick Business School, University of Warwick, Coventry CV4 7AL, U.K., Abhyankar (afinaa@wbs.warwick.ac.uk), and Basu (d.basu@warwick.ac.uk). We thank Timothy Crack and Saul Jacka for helpful comments. We would also like to thank Paul Malatesta (the editor) and especially David Chapman and Turan Bali (the referees) for suggestions that greatly improved the paper. All remaining errors are our own.

informative. He shows that the finding of nonlinear mean reversion may be generated solely by the finite sample bias of frequentist estimators such as maximum likelihood.

The main purpose of this paper is to show that the empirically observed nonlinear behavior in the drift term (for interest rate diffusions) may be neither spurious nor an artifact of the poor statistical properties of the estimators used. We rely on the intuition that empirical estimation of the parameters of diffusion processes necessarily involves the use of conditioning information whether using real world data or simulated data. In the case of empirical estimation, any real world data is implicitly conditioned by the fact that interest rates in whatever period considered have, in practice, some finite upper bound. Empirical research typically employs a variety of proxies for the short rate.¹ For example, Stanton (1997) uses the yield on the three-month Treasury bill, Chan, Karolyi, Longstaff, and Sanders (1992) use the one-month Treasury bill yield, while Ait-Sahalia (1996) and Bandi (2000) use the one-week Eurodollar rate. Since the stylized facts of short rate data are relevant to our main point, we now briefly describe them. The data set used, for example, in Ait-Sahalia (1996) covers the period 1973–1995. During this period, the short rate proxy has a mean value of 8.36% with a minimum value of 2.9% and a maximum value of 24%. Similar patterns are seen in the one-month Treasury bill rate over the 1960–1997 period: the mean rate is 5.69% and the minimum and the maximum values are 1.54% and 16.71%, respectively. Chapman and Pearson (2000) were the first to consider this phenomenon, which they term the truncation bias, and to analyze its impact on non-parametric and semi-nonparametric short rate estimators. They observe that the drift estimator in Stanton (1997), which uses only local information, is particularly susceptible to this bias and exhibits nonlinearities even when applied to sample paths simulated from a linear mean-reverting square root process. Relative to Chapman and Pearson (2000), this paper contributes explicit closed-form expressions for the drift of certain short rate processes under truncation bias. We do this using the theory of enlargement of filtrations and show that the drift term is highly nonlinear when the truncation bias is incorporated. We demonstrate the significance of this bias for specific interest rate processes. The first application of the theory of enlargement of filtrations in finance is Brown, Goetzmann, and Ross (1995), where they use the idea of conditioned diffusions to analyze the role of survivorship bias in the magnitude of the equity premium. Other examples that apply this method in finance include Grorud and Pontier (1998), Karatzas and Pikovsky (1996), and Taskin (1999).

The paper is organized as follows. In Section II, we provide the theoretical background and the derivation of drift corrections in the case of the Vasicek model, the Cox-Ingersoll-Ross model, and the case of a Brownian motion constrained to remain in an interval. The constrained Brownian motion process displays many of the characteristics of the empirical behavior of the short rate processes. Section III discusses the impact of the truncation bias on the drift and the estimation of the drift of these short rate processes. Section IV concludes.

¹Chapman, Long, and Pearson (1999) highlight the possible biases that might result from the use of these proxies for the unobservable short rate.

II. Theory

We now describe the effect of conditioning in the case of processes commonly used to model the short rate. These are i) a standard Ornstein-Uhlenbeck process first used in Vasicek (1977) and ii) the Cox, Ingersoll, and Ross (1985) model. In these cases, we enlarge the “natural” filtration, i.e., the filtration generated by the Brownian motions that drive the processes, by an event. This is equivalent to conditioning the diffusion processes on the event. The diffusion processes, with respect to the enlarged filtration, have a different drift term and the same diffusion coefficients as the original diffusions (Brown, Goetzmann, and Ross (1995), Jeulin and Yor (1985)). Assume that we have a diffusion process,

$$dr_s = \mu(s, r_s)ds + \sigma(s, r_s)dW_s,$$

with respect to the filtration $\{\mathcal{F}_s\}$. The new drift μ^* , with respect to the enlarged filtration $\{\mathcal{F}_s\} \vee \sigma(A)$, where we have conditioned on an event A with non-zero probability is

$$\mu^*(s, r_s) = \mu(s, r_s) + \frac{\pi_p}{\pi} \sigma^2(s, r_s),$$

where $\pi(s, r_s) = P(A \mid r(s) = \xi)$ and π_p denotes the derivative with respect to the spatial variable.

The volatility of the conditioned process remains the same, i.e., $\sigma^2(s, r_s) = \sigma^{*2}(s, r_s)$.² The expression for the drift of the conditioned process may be rewritten as (see Jeulin and Yor (1985))

$$\mu^*(s, r_s) = \left(\mu(s, r_s) + \frac{d < \pi(s, r_s), r_s > 1_A}{\pi(s, r_s)} \right) ds.$$

Example 1. We now derive the effect of conditioning information on the drift of the Vasicek model of the short rate. This model assumes that the short rate satisfies the diffusion equation,

$$dr_s = a(b - r_s)ds + \sigma dW_s.$$

In this case, the drift function is linear and given by $a(b - r_s)$. The new drift function with respect to the filtration enlarged by event A as above again leads to a non-linear drift function. We have

$$r_t = \exp(-a(t - s))r_s + (1 - \exp(-a(t - s)))b + \exp(-at) \int_s^t \exp(au)dW(u).$$

²See Lemma 1, on pp. 857–858 of Brown, Goetzmann, and Ross (1995) for a proof.

It follows from this representation that

$$\begin{aligned}
 M^s &= P(M_T \leq x | \mathcal{F}_s) \\
 &= 1 - \frac{2}{\sqrt{\frac{2\pi\sigma^2}{2a} (1 - e^{(-2a(T-s)})}}} \int_{x - e^{(-a(T-s))} r_s}^{\infty} \\
 &\quad \times \exp - \left(\frac{(y - (1 - e^{(-a(T-s)))} b)^2}{\frac{2\sigma^2}{2a} (1 - e^{(-2a(T-s)})}} \right) dy.
 \end{aligned}$$

It follows from Ito's Lemma that

$$\begin{aligned}
 d \langle M^s, r_s \rangle &= \\
 &- \left(\frac{2\sigma^2}{\sqrt{\frac{2\pi\sigma^2}{2a} (1 - e^{(-2a(T-s)})}}} \right) \\
 &\times \exp \left(- \frac{(z - (1 - e^{(-a(T-s)))} b)^2}{\frac{2\sigma^2}{2a} (1 - e^{(-2a(T-s)})}} \right) dy,
 \end{aligned}$$

where $z = x - (e^{(-a(T-s))} r_s$.

The new drift with respect to the enlarged filtration $\mathcal{F}_s \vee \sigma(A)$ is given by

$$\left(a(b - r_s) + \frac{d \langle M^s, r_s \rangle 1_A}{M^s} \right) ds,$$

again using the result in Jeulin and Yor (1985). The diffusion coefficient in this case also remains unchanged. It is clear that this drift function is highly nonlinear.

We thus see that incorporating conditioning information has the effect of changing the linear drift to a highly nonlinear drift structure.

Example 2. The CIR model of interest rates specifies the short rate as

$$dr_s = \kappa(\theta - r_s)ds + \sigma\sqrt{r_s}dW_s.$$

We condition this diffusion process over the time interval $[T_1, T_2]$ on event A , i.e., the short rate r_t on $[T_1, T_2]$ lies in the interval $(0, b)$, which introduces an *upper* truncation bias,

$$A = \{ (r_s)_s \in [T_1, T_2] \in (0, b) \}.$$

The drift of the conditioned process with respect to the enlarged filtration $\{\mathcal{F}_s\} \vee \sigma(A)$ is

$$\left(\kappa(\theta - r_s) + \frac{\pi_r(s, r_s)}{\pi(s, r_s)} \sigma^2 r_s \right) ds,$$

where $\pi(s, r_s) = P(A \mid \mathcal{F}_s)$. To compute this probability, we need to solve the Kolmogorov backward equation subject to suitable boundary conditions. The solution involves the confluent hypergeometric function and is extremely complicated. To simplify the analysis, we assume that the process has reached its stationary distribution, which is a gamma distribution with parameters $((2\kappa/\sigma^2), (2\kappa\theta/\sigma^2))$ (Cox, Ingersoll, and Ross (1985)). With this assumption,

$$\pi(s, r_s) = \pi(r) = c \int_r^b e^{-\frac{2\kappa}{\sigma^2}x} x^{\frac{2\kappa\theta}{\sigma^2}-1} dx.$$

Hence, the new drift with respect to the enlarged filtration is

$$\kappa(\theta - r) - \frac{e^{-\frac{2\kappa}{\sigma^2}r} r^{\frac{2\kappa\theta}{\sigma^2}-1}}{\int_r^b e^{-\frac{2\kappa}{\sigma^2}x} x^{\frac{2\kappa\theta}{\sigma^2}-1} dx} \sigma^2 r.$$

We note that this new drift term is highly nonlinear in r . We use this new drift to calculate the quantity $E(r_{t+\Delta} - r_t \mid r_t, A)$, which is analyzed in Chapman and Pearson ((2000), Section 4), assuming the process has reached a steady state distribution as

$$E(r_{t+\Delta} - r_t \mid r_t, A) = \mu(r_t, A) \Delta + o(\Delta),$$

where the second term can be ignored for small Δ .

The true conditional expectation for the unconditional process $E(r_{t+\Delta} - r_t \mid r_t)$ is given by

$$E(r_{t+\Delta} - r_t \mid r_t) = \theta(1 - e^{-\kappa\Delta}) - (1 - e^{-\kappa\Delta})r_t,$$

which is equal to $\kappa(\theta - r_t)\Delta$ for small Δ .³

The conditional expectation incorporating this truncation bias is different from the conditional expectation of the unconditional process. This is illustrated in Chapman and Pearson (2000), Figure 9, where they run an ordinary least squares (OLS) regression of $(r_{t+\Delta} - r_t)$ on r_t for $r_t > 0.14$ that is obtained from a single simulated sample of a square root process. They find that the slope of the OLS

³An analysis for the CIR process conditioned on the event B that $r \geq l$ for some positive constant l , i.e., introducing a lower truncation bias shows that, for values of r close to the lower limits of the data, local estimators would also find a high degree of mean reversion. The drift of this conditioned process is given by

$$\kappa(\theta - r) + \frac{e^{-\frac{2\kappa}{\sigma^2}r} r^{\frac{2\kappa\theta}{\sigma^2}-1}}{\int_u^r e^{-\frac{2\kappa}{\sigma^2}x} x^{\frac{2\kappa\theta}{\sigma^2}-1} dx} \sigma^2 r.$$

For values of r close to u , we can use L'Hospital's Rule to evaluate the second term in the above expression and we obtain the following expression for the drift term,

$$\kappa(\theta - r)\Delta - 2\kappa\Delta r + (2\kappa r)\Delta.$$

This shows that for values of r close to the lower limit of the data, the lower truncation bias causes drift estimators that use local information, such as that of Stanton (1997), to find a higher degree of mean reversion.

regression is more negatively sloped than would be expected from the theoretical slope in the above equation. Their regression illustrates clearly the effect of the truncation bias on the conditional expectation. As the quantity $E(r_{t+\Delta} - r_t | r_t) / \Delta$ is essentially the Stanton (1997) estimator of the drift process, the discussion in Chapman and Pearson ((2000), Section 4), shows that the drift estimator is significantly affected by the truncation bias near the edges of the data.

The conditional expectation incorporating the truncation bias is given by

$$\left(\kappa(\theta - r) - \frac{e^{-\frac{2\kappa}{\sigma^2}r} r^{\frac{2\kappa\theta}{\sigma^2}-1}}{\int_r^b e^{-\frac{2\kappa}{\sigma^2}x} x^{\frac{2\kappa\theta}{\sigma^2}-1} dx} \sigma^2 r \right) \Delta.$$

In the above expression, we now take $b = 0.1583$, which is the upper limit of the single simulated sample used in Chapman and Pearson (2000). For values of r close to b , we can use L'Hospital's Rule to evaluate the second term in the above expression and we obtain the following expression for the drift term,

$$\kappa(\theta - r)\Delta - (2\kappa r)\Delta + (2\kappa\theta)\Delta.$$

We now take the parameters, as in Chapman and Pearson, to be $\kappa = 0.85837$, $\sigma = 0.15660$, $\Delta = 1/250$, and find that the coefficient of r is -0.0102969 , which is more negatively sloped than would be expected from the theoretical slope for the conditional expectation of the unconditioned process.

The above numerical example provides theoretical evidence for the conclusion in Chapman and Pearson ((2000), Section 4, Figure 9), where it is shown that the Stanton (1997) estimator of the drift does exhibit significant bias due to truncation effects particularly near the extremes of the data.

Example 3. We now provide another example that further confirms Chapman and Pearson's (2000) observation that the nonlinearity of the short rate drift is not a robust stylized fact.

Since many papers have found that there is no mean reversion in the heart of the density of the short rate data, we consider the effect of conditioning information when the short rate follows a drift-free Brownian motion. Let us consider the process given by⁴

$$dr_s = dW_s,$$

conditioned to remain in an interval (a, b) up to time T . We assume that T is large so that this process has converged to its steady state diffusion, which is given by

$$dr = \frac{\pi}{(b-a)} \frac{\cos \frac{\pi(r-a)}{b-a}}{\sin \frac{\pi(r-a)}{b-a}} dt + dW.$$

⁴This, of course, is a rather unappealing model of the short rate as the process can take on negative values and does not capture the stylized facts of interest rate processes such as mean reversion.

This follows from Pinsky ((1985), p. 375) by replacing the interval $(-c, c)$ with the interval (a, b) . The drift of this process is zero when $r = (b - a)/2$, which is the mid-point of the interval (a, b) . As r tends toward either a or b , the process mean reverts in a highly nonlinear fashion, i.e., the drift displays nonlinear behavior. We now choose plausible values of a and b in terms of the historical range of the short rate specifically by taking $a = 1.9\%$ and $b = 24.5\%$. Figure 1 illustrates the behavior of the drift for these values of a and b . It is clear from Figure 1 that the above model exhibits many of the features of the short rate noted by Ait-Sahalia (1996). Specifically, near its mean (which, in our example, is the middle of the range (a, b)), the spot rate behaves like a random walk and then mean reverts nonlinearly when far away from the mean. Our Figure 1 is remarkably similar to Ait-Sahalia's (1996) Figure 4b, p. 408, which depicts a nonlinear parametric drift function that is not rejected by his data and methodology.

FIGURE 1

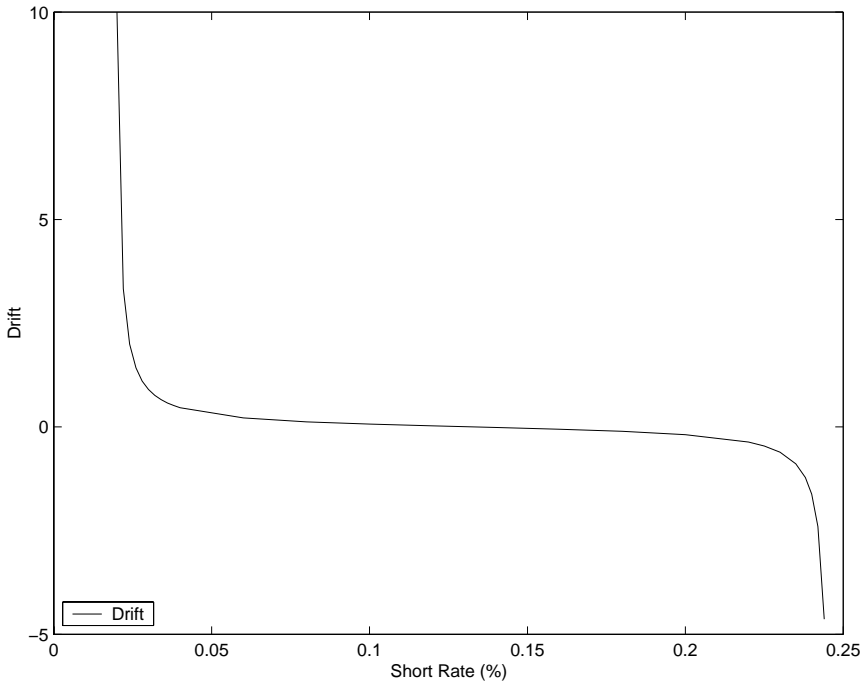


Figure 1 shows the drift $\mu(r)$ of the process

$$dr = \frac{\pi}{(b - a)} \frac{\cos \frac{\pi (r - a)}{b - a}}{\sin \frac{\pi (r - a)}{b - a}} dt + dW$$

for various values of r and $a = 1.9$ and $b = 24.5$.

Our analysis shows that the truncation bias could cause local estimators of the drift to find nonlinear behavior near the edges of the data even when the true

drift is actually zero. This also illustrates that the truncation bias is an important specification bias as Ross (1987) observes.

III. Significance of the Bias

We now consider the case of the Vasicek model conditioned on event A that the maximum over a time period is less than a fixed constant x . The new drift process with respect to the enlarged filtration $\mathcal{F}_s V\sigma(A)$ is given by

$$\left(a(b - r_s) + \frac{d < M^s, r_s > 1_A}{M^s} \right) ds,$$

where

$$d < M^s, r_s > = - \left(\frac{2\sigma^2}{\sqrt{\frac{2\pi\sigma^2}{2a} (1 - e^{-2a(T-s)})}} \exp \left(- \frac{(z - (1 - e^{-a(T-s)}) b)^2}{\frac{2\sigma^2}{2a} (1 - e^{-2a(T-s)})} \right) dy \right)$$

as obtained in Section III earlier where $z = x - (e^{-a(T-s)})r_s$.

We can now obtain, based on the above, numerical estimates of the size of the drift correction for various short rate values. The parameter values we use, based on Ait-Sahalia (1996), are as follows: $a = 0.85837$, $b = 0.089102$, and $\sigma^2 = 0.0021854$. We assume that $T - s =$ one year and $0.x = 0.24$. Figure 2 depicts, for values ranging from a low of 2% to a high of 22%, the drifts of the unconditioned and conditioned processes. Figure 2 illustrates the dramatic influence of the upper truncation bias for large values of the short rate, i.e., near the upper edges of the data. The nonlinear nature of the drift of the conditioned process is also clearly apparent for large values of the short rate. We also see that the size of the *upper* truncation bias is small for low values of the short rate. Of course, there will be a *lower* truncation bias associated with the smallest realized value in the sample.

We next discuss the effect of truncation bias on the estimators such as that used in Stanton (1997) for the Vasicek process. Again, for the sake of tractability, we assume that the process has reached its stationary distribution, which is normal with mean b and variance $\sigma^2/2a$. As in the previous section, we calculate the quantity $E(r_{t+\Delta} - r_t | r_t, A)$, which is approximately equal to $\mu(r_t, A)\Delta$, where A is the event that $r_t \leq u$ for some fixed constant u . A calculation similar to that in Section II shows

$$\mu(r_t, A) = a(b - r_t) - \frac{e^{-\frac{(r_t - b)^2}{2\frac{\sigma^2}{2a}}}}{\int_{r_t}^u e^{-\frac{(y-b)^2}{2\frac{\sigma^2}{2a}}} dy} \sigma^2.$$

For r close to u , we may use L'Hospital's Rule to obtain

$$\mu(r_t, A) = a(b - r_t) - 2a(r_t - b).$$

FIGURE 2

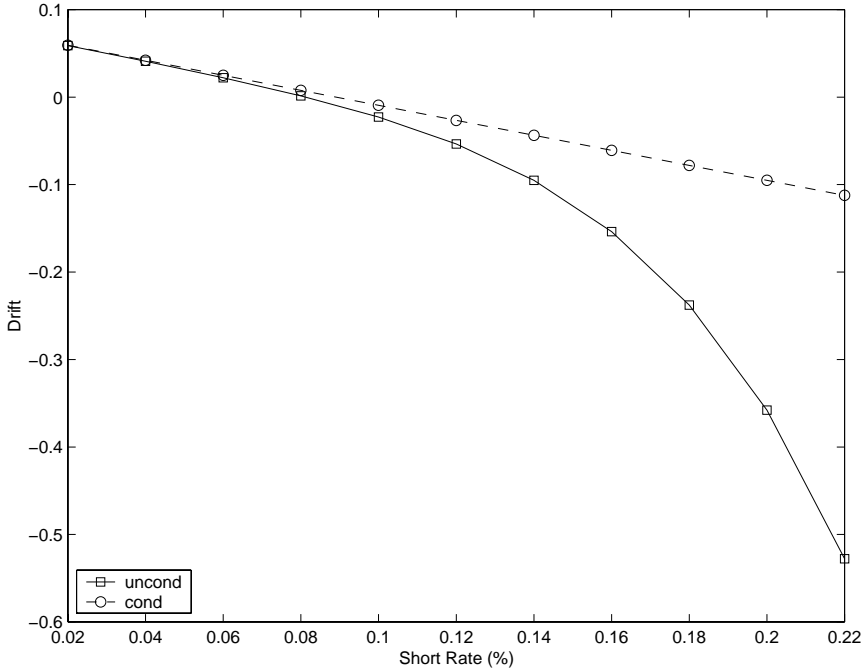


Figure 2 shows the drift correction for various values of the short rate r_s . The drift correction C_s for the Vasicek process conditioned on the event that the short rate is less than 24% is given by the expression $((d < M^s, r_s > 1_A)/M^s)$ and is depicted by the unbroken line. The drift of the unconditioned process $a(b - r_s)$ is depicted by the broken line. Note that the drift correction and, hence, the truncation bias for driftless Brownian motion $dr = \sigma dW$ (an uncorrelated process), conditioned on the process being less than a fixed value x , is $-\sigma^2/(x - r)$, which for values of r much smaller than x , is close to zero. This shows that the truncation bias is not always and everywhere equally important.

This shows that for values of r close to the upper limit of the data, the truncation bias causes drift estimators that use local information, such as that of Stanton (1997), to find a higher degree of mean reversion.

IV. Conclusions

A number of empirical studies have noted that the estimated drift function shows considerable evidence of nonlinearity. However, theoretical term structure models have linear drift specifications. Research has, therefore, re-examined the robustness of these findings and has raised issues about the biases in the estimators used (see, for example, Chapman and Pearson (2000) and Pritsker (1998)). In the case of empirical estimation, any real world data is implicitly conditioned by the fact that interest rates in whatever period considered have, in practice, some finite upper bound. Chapman and Pearson (2000) were the first to consider this truncation bias and to analyze its impact on non-parametric and semi-nonparametric short rate estimators. They observe, for example, that the drift estimator in Stanton (1997) is particularly susceptible to this bias. We show, using the theory of

enlargement of filtrations, that, in the presence of conditioning information, the drift has a nonlinear structure in the case of diffusion processes commonly used to model the short rate. We also find that the presence of conditioning creates a bias in the estimation of the drift for which we provide some estimates. Our contribution, relative to Chapman and Pearson (2000), is that we provide explicit closed-form expressions for the drift of certain short rate processes under truncation bias. The techniques of this paper may be relevant for studying other price processes, for example, the behavior of exchange rates in target zones.

References

- Ait-Sahalia, Y. "Testing Continuous-Time Models of the Spot Interest Rate." *Review of Financial Studies*, 2 (1996), 385–426.
- Bandi, F. "Short-Term Interest Rate Dynamics: A Spatial Approach." Working Paper, Graduate School of Business, Univ. of Chicago (2000).
- Brown, S.; W. Goetzmann; and S. Ross. "Survival." *Journal of Finance*, 50 (1995), 853–873.
- Chan, K. C.; A. Karolyi; F. Longstaff; and A. Sanders. "An Empirical Comparison of Alternative Models of the Short Term Interest Rate." *Journal of Finance*, 47 (1992), 1209–1227.
- Chapman, D.; N. D. Pearson; and J. Long. "Using Proxies for the Short Rate: When is Three Months Like an Instant?" *Review of Financial Studies*, 12 (1999), 763–806.
- Chapman D., and N. D. Pearson. "Is the Short Rate Drift Actually Nonlinear?" *Journal of Finance*, 55 (2000), 355–388.
- Cochrane, J. *Asset Pricing*. Princeton, NJ: Princeton Univ. Press (2001).
- Cox, J. C.; J. E. Ingersoll; and S. A. Ross. "A Theory of the Term Structure of Interest Rates." *Econometrica*, 53 (1985), 321–346.
- Gorud, A., and M. Pontier. "Insider Trading in a Continuous Time Market Model." *International Journal of Theoretical and Applied Finance*, 3 (1998), 331–347.
- Jeulin, T., and M. Yor. "Grossissements de Filtrations: Exemples et Applications." *Lecture Notes in Mathematics*. New York, NY: Springer Verlag (1985), 1118.
- Jones, C. "Nonlinear Mean Reversion in the Short Term Interest Rate." Working Paper, Simon School of Business, Univ. of Rochester (2000).
- Karatzas, I., and S. Shreve. *Brownian Motion and Stochastic Calculus*, 2nd ed. New York, NY: Springer-Verlag (1991).
- Pinsky, R. "On the Convergence of Diffusion Processes Conditioned to Remain in a Bounded Region for Large Time to Limiting Positive Recurrent Diffusion Processes." *Annals of Probability*, 13 (1985), 363–378.
- Pritsker, M. "Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models." *Review of Financial Studies*, 11 (1998), 449–487.
- Ross, S. "Regression to the Max." Working Paper, Yale Univ. (1987).
- Stanton, R. "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk." *Journal of Finance*, 52 (1997), 1973–2002.
- Taskin, A. "Application of Enlargement of Filtrations to the Term Structure of Interest Rates." Working Paper, Department of Economics, Univ. of Chicago (1999).
- Vasicek, O. "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics*, 5 (1977), 177–188.