Flexible Term Structure Estimation: Which Method is Preferred?

Andrew Jeffrey*, Oliver Linton†, and Thong Nguyen‡

January 10, 2001

Abstract

We show that the recently developed nonparametric procedure for fitting the term structure of interest rates developed by Linton, Mammen, Nielsen, and Tanggaard (2000) overall performs notably better than the highly flexible McCulloch (1975) cubic spline and Fama and Bliss (1987) bootstrap methods. However, if interest is limited to the Treasury bill region alone then the Fama-Bliss method demonstrates superior performance. We further show, via simulation, that using the estimated short rate from the Linton-Mammen-Nielsen-Tanggaard procedure as a proxy for the short rate has higher precision then the commonly used proxies of the one and three month Treasury bill rates. It is demonstrated that this precision is important when using proxies to estimate the stochastic process governing the evolution of the short rate.

1 Introduction

The term structure of interest rates is central to all models of fixed-income security pricing. Prime examples of continuous time models of the term structure include Vasicek (1977), Cox, Ingersoll and Ross (1985), Hull and White (1990), and Heath, Jarrow, and Morton (1992). In both the Vasicek and Cox-Ingersoll-Ross models the evolution of the short term interest rate and risk preferences are specified; this determines the term structure on any given day. Hull and White demonstrates how to extend both of these models so that they can be calibrated to an observed initial term structure.

*Andrew Jeffrey, Yale School of Management, PO Box 208200, New Haven CT. 06520-8200. Phone: (203) 432 6029. Fax: (203) 432 6970. Email: Andrew.Jeffrey@yale.edu

†Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. E-mail address: lintono@lse.ac.uk.

‡Bank of America Securities, San Francisco, CA 94111, U.S.A.

Heath-Jarrow-Morton deviate significantly from these paradigms by taking an initial term structure as given and models the evolution of the whole curve.

Unfortunately, at any point in time the whole term structure is not directly observable. Consider government bonds, which is a natural data set from which to obtain the term structure. There are several obstacles in place. The first, and most obvious, difficulty is that only a finite collection of bonds are observed as opposed to a desired continuum. Second, all bonds issued with maturities greater than one year are coupon bearing, except for bonds created using the recently introduced STRIPS program.[1] The existence of coupon payments is undesirable because theoretical term structure models, such as those mentioned above, always make reference to the zero-coupon bond term structure. Third, many theoretical models also assume that all bonds are default-free without features such as callability and/or special tax privileges. This unfortunately limits the data that can be used. Finally, liquidity problems are sometimes present. It has been well documented that Treasury notes and bonds with less than one year to maturity, and Treasury bills with less than one month to maturity, are illiquid (see Fama and Bliss (1987), Sarig and Warga (1989), Amihud and Mendelsohn (1991), Duffee (1996), Bliss (1997)).

Our interest in estimating the zero-coupon bond term structure is two-fold. First, estimation of the discount function each day should be considered as a vehicle to explore the intertemporal behavior of the term structure. In fact such extraction procedures are also a pre-requisite for intertemporal models requiring the specification of an initial term structure, such as Hull and White (1990) and Heath, Jarrow and Morton (1992). Second, of particular interest is the estimation of the very short end of the term structure since most models of the term structure's evolution have the short term interest rate as a state variable. The importance of the short end is highlighted by the fact that the evolution of the short-term interest rate under the risk-neutral probability measure is enough to characterize the whole term structure.

The first work dealing with the extraction of the (unobserved) zero-coupon bond term structure is accredited to McCulloch (1971, 1975) who proposed fitting the discount function with quadratic and cubic splines. Various alternative parametric methods followed: Chambers, Carleton, and Waldman (1984) use polynomials to estimate the yield curve, Vasicek and Fong (1982) use exponential splines to estimate the discount function, and Nelson and Siegel (1987) use a second-order constant-coefficient partial differential equation to fit the yield curve in a parsimonious fashion. Fisher, Nychka, Zervos

[1] "Stripped" notes and bonds are issues that have had their component cash flows traded separately. The Treasury does not sell the individual cash flows, this is done by dealers that first purchase a coupon bearing note or bond. In 1985 the Treasury permitted this resale of individual cash flows for a limited selection of notes and bonds - called the STRIPS program. As of September 1998 all Treasury notes and bonds issued on or after September 30, 1997 are all eligible for the STRIPS program.

(1995) and Waggoner (1997) modify the McCulloch cubic spline procedure by adding a function to penalize large variations in the estimated yield curve that can occur with over-fitting. Fama and Bliss (1987) approach the term structure estimation problem differently. Instead of providing a curve fitting procedure they use an iterative scheme, referred to as "bootstrapping", where a piece-wise constant forward rate curve is chosen to exactly price all bonds.[2] An excellent paper that compares a large subset of the above term structure extraction methods is provided by Bliss (1997). In short, based on out-of-sample tests, he concludes that the Fama-Bliss method performs better than all other methods and the McCulloch cubic spline is the better performer amongst the remaining curve fitting procedures.

Recently, Linton, Mammen, Nielsen, and Tanggaard (2000) (LMNT hereafter) have developed a nonparametric kernel smoothing procedure to fit the discount function. This approach is highly flexible with regard to the functional form of the estimated curve; the trade-off between under/over-fitting is controlled by the "bandwidth" that determines the quantity of nearby information used to estimate the yield at a particular maturity. LMNT (2000) mostly provides a description of the large sample theoretical properties of the discount function's estimate but the relative performance of this estimate compared with other term structure extraction methods has not been established on real data. This motivates the first part of the present paper: we consider a modified version of the LMNT procedure, provide the first order conditions necessary to solve the optimization problem in a timely fashion, and finally empirically compare this term structure estimation procedure to other flexible term structure extraction methods.[3] In our comparison we only consider the Fama-Bliss (1987) bootstrapping method and the McCulloch (1975) cubic spline method since: i) the LMNT method is not parsimonious and we wish to compare it to other non-parsimonious methods, and ii) Bliss (1997) has clearly demonstrated the superior performance of both the Fama-Bliss and McCulloch procedures relative to other existing methods.[4]

Using U.S. Treasury bills, notes and bonds with time-to-maturities out to ten years obtained from the CRSP bonds data set over the period January 1970 to December 1998 we conclude that the modified LMNT procedure demonstrates notable superior performance on average. In-sample results suggest that the LMNT method is preferred to McCulloch's cubic spline 70% to 85% of the time across the whole maturity spectrum.[5] Out-of-sample the LMNT procedure is preferred to

[2]Note that their implementation requires the elimination of suspicious quotes using a series of filters.

[3]The modification considered is to estimate the yield curve instead of the discount function. This was suggested in LMNT (2000) as an extension.

[4]We conducted some preliminary analysis using other term structure fitting procedures and our results conformed to the findings of Bliss (1997).

[5]In-sample comparison to the Fama-Bliss method is not considered as this method will, by construction, provide an almost perfect fit.

McCulloch's cubic spline approximately 75% of the time for securities with less than one year until maturity, 69% of the time for securities with maturities between one and three years, 52% of the time for securities with maturities between three to five years, and 63% of the time for securities with maturities greater than five years. With respect to the Fama-Bliss method we find that the LMNT procedure is only preferred a disappointing 18% of the time for securities with less than one year until maturity. However beyond one year the LMNT method is preferred 71%, 77% and 64% of the time on the one to three, three to five, and beyond five year maturity regions respectively. A similar pattern occurs when comparing the McCulloch method to the Fama-Bliss scheme in that the latter is only preferred when considering securities with less than one year until maturity. This suggests that the primary benefit of using the Fama-Bliss bootstrapping procedure is in the Treasury-bill region of the term structure however for longer maturity securities we are better off using a curve fitting procedure such as the LMNT method.

Given the importance of the short-end of the term structure the second component of this paper studies the short rate estimate obtained by using the above LMNT procedure. Proxies for the short rate vary from study to study: Chan, Karolyi, Longstaff, and Sanders (1992) use the one month Treasury bill rate; Aït-Sahalia (1996b) uses the seven day Eurodollar deposit spot rate; Conley, Hansen, Luttmer, and Scheinkman (1997) use the Federal funds rate; Stanton (1997) uses the three month Treasury bill rate. At first glance all of these proxies seem reasonable however some questions regarding their appropriateness have recently been raised. In particular Duffee (1996) finds idiosyncratic behavior in very short term U.S. Treasury bills and suggests using the three month Treasury bill rate as opposed to a shorter maturity rate. Chapman-Long-Pearson (1999) demonstrate that if the term structure of interest rates is driven by a non-affine term structure model then an economically significant difference can arise between the use of one month and three month spot rate proxies for the short rate. However, one issue that is not considered in the above studies is the impact of observation noise in prices and it is this issue that we wish to address.

One method of reducing the influence of observation noise is to use a curve-fitting procedure to estimate the very short term end of the term structure. The resulting proxy for the short rate should exhibit less error for the obvious reason that the curve-fitting procedure "averages the errors" present in observed prices. We demonstrate this in the paper via a simulation study: coupon-bond prices are generated with error, the term structure is then extracted using the modified LMNT methodology, and a comparison between several short term interest rate proxies is conducted. The comparison involves studying: i) each proxy's deviation from the true short rate generated from the simulation, and ii) the use of each proxy in the estimation of the drift and diffusion coefficients of the stochastic process characterizing the evolution of the short rate. As expected, the LMNT-based estimate of the short rate has higher precision relative to the use of the one/three month Treasury bill rates. We also recover the result of Duffee (1996) in that the three month Treasury bill rate provides a better proxy to the short rate than the three month Treasury bill rate. This is because the error in yield resulting from the noise in prices is amplified when the security's time-to-maturity decreases. The importance of using a short rate proxy with low noise is highlighted when using it to estimate drift and diffusion coefficients from interest rate changes. The simulation demonstrates that the bias in a diffusion estimate and the ability to estimate the slope of the drift depends heavily on the ability to obtain a more precise estimate.

The remainder of this paper proceeds as follows. In Section 2 the modified LMNT (2000), Fama-Bliss (1987) and McCulloch (1975) term structure extraction procedures are described. Section 3 contains the comparison methodology, data description, and results. A simulation study demonstrating a benefit of using a curve fitting procedure to extract a proxy for the short term interest rate is provided in section 4. Section 5 summaries and concludes.

2 Term Structure Extraction Methods

In this section we provide our implementation of the LMNT (2000) procedure and describe both the McCulloch (1975) and Fama and Bliss (1987) methods for extracting the term structure of interest rates. All of these procedures estimate the term structure at time t with bond data observed at that time only and hence they ignore the intertemporal aspects of the term structure's evolution. All bonds are idealistically default-free and provide a stream of non-random cash flows at known times in the future. In ensuing sections we take as given N bond prices all observed at the same point in time. The i^{th} observed bond price is denoted P^i and this bond provides known cash flows b^i_j at times τ^i_j in the future for $j = 1, \ldots, m^i$. The discount function, denoted $d(\cdot)$, is extracted from these observed bond prices by imposing the static no-arbitrage condition, which has the interpretation that the price of any bond is the sum of all its discounted cash flows, that is $P^i = PV^i$, where $PV^i = \sum_{j=1}^{m^i} b^i_j d(\tau^i_j)$ is the discounted present value in which $d(\tau)$ is the discounted value of one dollar to be received at time τ. In practice, we can expect there to be small errors in this relation when applied to actual data. First, we do not observe the actual trade price, instead we observed quoted bid and ask prices. Further these prices may not have been quoted at exactly the same time so small deviations may result from non-synchronous trading. In addition to this there are other complications such as liquidity and taxes that have been proposed as reasons to expect further small violations of the arbitrage condition. We shall therefore suppose that $P^i = PV^i + \varepsilon^i$, where ε^i is a random error term.

2.1 The LMNT Method

LMNT (2000) suggest various kernel smoothing estimators for the discount function. In particular two estimators are studied in depth, the "local constant" and "local linear" methods. In brief these estimators are, respectively, based on locally approximating the discount function as a constant and a linear function of maturity. The LMNT implementation that we adopt modifies the local linear method by applying it to the yield curve instead of the discount function; the yield curve $y(\cdot)$ is defined by $y(\tau) = -\frac{1}{\tau}\ln(d(\tau))$ or equivalently $d(\tau) = \exp(-\tau y(\tau))$. The motivation for considering this version of the LMNT procedure is threefold: i) the discount function at the origin is guaranteed to be one, ii) the discount function will be strictly positive for finite maturities, and iii) the discount function is "closer" to being log-linear than linear.

This version of the LMNT procedure provides estimates for both the yield curve and its first derivative (denoted $y'(\cdot)$). The idea is that for any point τ close to a point v we can approximate $y(\tau)$ by the linear function $y(v) + (\tau - v)y'(v)$. Therefore, we can approximate the present value of a given bond by

$$\widehat{PV}^i = \sum_{j=1}^{m^i} b_j^i e^{-\left(y(v_j)+(\tau_j^i - v_j)y'(v_j)\right)\tau_j^i},$$

where v_j are points close to τ_j^i. The estimation method is based on minimizing the sum of squared pricing errors $P^i - \widehat{PV}^i$. Specifically, we find the functions $\widehat{y}(\cdot)$ and $\widehat{y'}(\cdot)$ that minimize the following criterion function with respect to the functions $y(\cdot)$ and $y'(\cdot)$:

$$Q_N(y, y') = \sum_{i=1}^{N} \int \cdots \int \left(P^i - \sum_{j=1}^{m^i} b_j^i e^{-\left(y(v_j)+(\tau_j^i - v_j)y'(v_j)\right)\tau_j^i} \right)^2 \prod_{k=1}^{m^i} \left\{ K_h(v_k - \tau_k^i) dv_k \right\}, \tag{1}$$

where K is the kernel function, h is the bandwidth, and $K_h(\cdot) = K(\cdot/h)/h$. Here, the integrals are taken over the support of the kernel. An interpretation for this criterion is that it is a kernel smoothed version of the sample least squared errors criterion; the kernel weighting scheme measures the distance between bonds by the distance between their respective vectors of cash flow payment times.[6]

[6] The objective function (1) has the same probability limit as

$$\sum_{i=1}^{N} \left(P^i - \sum_{j=1}^{m^i} b^i(\tau_j^i) e^{-\tau_j^i y(s_j^i)} \right)^2,$$

which can be seen by letting $h \downarrow 0^+$, $N \to \infty$ and presuming that the bonds ensure the sequence $\{\tau_j^i\}$ becomes dense in a compact support.

The above minimization problem generates the following first order conditions for $y(.), y'(.)$ for all points v in the support of the payment times:

$$\begin{aligned} \sum_{i=1}^{N}\sum_{k=1}^{m^i} X_k^i(v; y(.), y'(.)) &= 0 \\ \sum_{i=1}^{N}\sum_{k=1}^{m^i} X_k^i(v; y(.), y'(.)) \left(v - \tau_k^i\right) &= 0, \end{aligned} \tag{2}$$

where:

$$\begin{aligned} X_k^i(v; y, y') &= \left(K_h(v-\tau_k^i) b_k^i \tau_k^i d_k^i(v)\right)\left(P^i - b_k^i d_k^i(v) - \sum_{\substack{j=1 \\ j\neq k}}^{m^i}\left(\int K_h(x-\tau_j^i) b_j^i d_j^i(x)dx\right)\right), \\ d_k^i(v) &= \exp\left\{-\left(y(v) + (\tau_k^i - v)y'(v)\right)\tau_k^i\right\}. \end{aligned}$$

Numerically it is easier to solve these first order conditions as opposed to the minimization problem (1) due to the high dimensionality of the multiple integrals present in (1).[7]

To implement the above LMNT procedure choices must be made regarding: i) the kernel, ii) the bandwidth, and iii) the set of yields to solve for which is defined by the choice of v's considered in the first order conditions.[8] Our choices are as follows. A common kernel choice in the applied nonparametric literature is the Gaussian kernel, mostly because it produces very smooth curves and it is also well attuned for estimation of the derivatives.[9] Consequently we adhere to this choice. For the bandwidth our choice is based on the observation that the cash flows generated by all bonds becomes increasingly sparse as time to maturity increases. Hence, when computing the present value of the n^{th} cash flow from bond i, occurring at time τ_n^i, it is desirable for the bandwidth h to be an increasing function of τ_n^i. We choose h to be of the form $h(\tau_n^i) = a + b\tau_n^i$ where a and b are arbitrarily chosen so that $h(0) = 2/12$ and $h(10) = 1$.[10] This bandwidth choice ensures that the

[7] These conditions are derived in a similar manner to the first order conditions for the local linear estimator of the discount function in LMNT (2000).

[8] It is theoretically possible to compute every point on the yield curve from the first order conditions once the kernel and bandwidth have been chosen. However since this is impossible in practice we choose a finite set of specific maturities to form a "reference set of yields" and interpolate between them in a manner that will ben discussed shortly.

[9] The yield curve estimator will inherit the differentiability properties of the kernel implying that the resulting yield curve estimator will be infinitely differentiable on the entire real line.

[10] Ad hoc motivation for these choices is obtained from the following properties of our data set. We do not use any Government securities with a time-to-maturity of less than one month however we do wish infer an estimate of the short rate. We consider a bandwidth choice of two months to be adaquate to capture a reasonable amount of data at the short end of the term structure for such an inference. At the long end, we do not use any Government securities with a time-to-maturity of more than ten years. Using a one year bandwidth will "tie together" the last few cash flow

estimated yield curve is relatively more flexible at the short end where more information is available. Finally, we choose a finite number of maturity dates $v_1, \ldots, v_k$ for which to compute $\widehat{y}(v_i)$ and $\widehat{y'}(v_i)$ for $i = 1, \ldots, k$. These estimates can be viewed as a set of parameters summarizing the entire yield curve and an interpolation algorithm is then used to obtain an estimate of any desired yield. We choose $v_1, \ldots, v_k$ by setting $v_1 = 0$ and generate the remaining v_i's by $v_i = v_{i-1} + \frac{1}{2}h(v_{i-1})$. This structure is consistent with the notion that less yield measurements are required when the bandwidth increases and adjacent v_i's are "close" to each other relative to the bandwidth in that neighborhood.

The chosen interpolation algorithm is motivated from the original objective function (1) as follows. Suppose we have computed $\widehat{y}(\cdot)$ and $\widehat{y'}(\cdot)$. Now consider a pure discount bond with time to maturity τ and price $d(\tau)$. The yield on this zero coupon bond can be computed from the relationship $d(\tau) = e^{-y(\tau)\times\tau}$. An estimate of the yield $y(\tau)$ is obtained from the value $\widehat{Y}(\tau)$ that minimize the original objective function (1) with respect to $y(\tau)$ given $\widehat{y}(\cdot)$ and $\widehat{y'}(\cdot)$; that is

$$\min_{y(\tau)} \int \left(e^{-y(\tau)\times\tau} - e^{-\left(\widehat{y}(v)+(\tau-v)\widehat{y'}(v)\right)\tau}\right)^2 K_h(v-\tau)dv.$$

The solution to this minimization problem is

$$Y(\tau) = -\frac{1}{\tau}\ln\left(\int \left(e^{-\left(\widehat{y}(v)+(\tau-v)\widehat{y'}(v)\right)\tau}\right) K_h(v-\tau)dv\right).$$

Since we only have $\widehat{y}(v_i)$ and $\widehat{y'}(v_i)$ for $i = 1, \ldots, k$, we can approximate the above with

$$\widehat{Y}(\tau) = -\frac{1}{\tau}\ln\left(\frac{\sum_{i=1}^{k} K_h(v_i-\tau)\left(e^{-\left(\widehat{y}(v_i)+(\tau-v_i)\widehat{y'}(v_i)\right)\tau}\right)}{\sum_{i=1}^{k} K_h(v_i-\tau)}\right). \quad (3)$$

$\widehat{Y}(\tau)$ can be interpreted as the yield on the τ maturity pure discount bond whose price is obtained by the Nadaraya-Watson kernel smoothing estimate of pure discount bond prices computed from yield curve estimates $\widehat{y}(v_i)$ and $\widehat{y'}(v_i)$ for $i = 1, \ldots, k$. Since $\widehat{Y}(\tau)$ is obtained from other yield curve estimates we should interpret this purely as an interpolation scheme.[11]

2.2 The Fama-Bliss Method

The Fama-Bliss (1987) bootstrapping procedure considers expressing the term structure in terms of the forward rate curve $f(\cdot)$, which is defined by the relation $d(\tau) = \exp\left(-\int_0^\tau f(v)dv\right)$, and further the

payments of the longest maturity bond (cash flows from US Treasury notes/bonds occur semi-annually) in the sense that they will be discounted at similar rates.We also conducted some preliminary experiments with various choices of b, namely $b = 0.75$, 1.25 and 1.5. The results did not vary dramatically.

[11]Note that the first order conditions (2) require computation of the integral $\int K_h(x-\tau_j^i)e^{-\left(y(x)+(\tau_j^i-x)y'(x)\right)\tau_j^i}dx$. Using the above interpolation scheme the integral is approximated by $\exp\left\{-\tau_j^i \times \widehat{Y}(\tau_j^i)\right\}$.

forward rate curve is presumed constant between successive observed bond maturities. More specifically let the sequence of observed bonds $\{P^i\}_{i=1}^N$ be ordered from the shortest maturity to longest maturity and let τ^i denote the time-to-maturity of the i^{th} bond. Let F^i denote the constant forward rate on the time-to-maturity interval $(\tau^{i-1}, \tau^i]$ where $\tau^0 = 0$, that is $f(\tau) = F^i$ for $\tau \in (\tau^{i-1}, \tau^i]$. The discount function now takes the form $d(\tau) = \exp\left(-F^K(\tau - \tau^{K-1}) - \sum_{k=1}^{K-1} F^k(\tau^k - \tau^{k-1})\right)$ where K is chosen so that $\tau \in (\tau^{K-1}, \tau^K]$.

To extract (bootstrap) the forward rate curve proceed as follows. First determine F^1 by considering the shortest maturity instrument and solve for F^1 in $P^1 = \sum_{j=1}^{m^1} b_j^1 \exp\left(-F^1 \times \tau_j^1\right)$. Now consider the second shortest instrument and solve for F^2 in $P^2 = \sum_{j=1}^{m^2} b_j^2 d(\tau_j^2)$ given F^1, and so on. In general, to bootstrap F^i use the i^{th} observed bond and find the F^i that solves $P^i = \sum_{j=1}^{m^i} b_j^i d(\tau_j^i)$ where the sequence $\{F^j\}_{j=1}^{i-1}$ has been computed from previous bonds in the same fashion.

By construction the above procedure exactly prices all in-sample bonds. It consequently is subject to spurious behavior if some "mis-priced" bonds are in the sample. To lessen the impact of this problem Fama-Bliss propose the following filters for the data: i) only fully taxable, non-callable and non-flower bonds are used, ii) Treasury notes and bonds are excluded from the sample if their time-to-maturity is less than one year, iii) an instrument is included if either its yield-to-maturity is within 0.2% absolute difference of the yield-to-maturities of surrounding instruments or in between them, and iv) an instrument is included if the resulting yield curve when the instrument is included does not exhibit large yield reversals (adjacent changes that greater that 0.2% in absolute value and in opposite directions).[12]

2.3 The McCulloch Cubic Spline Method

McCulloch (1975) uses a cubic spline procedure to estimate the discount function. That is, the estimated discount function is of the form

$$\widehat{d}(\tau) = g_i(\tau) \text{ on the interval } [\tau_i, \tau_{i+1}] \text{ for } i = 1, \ldots, k-1, \tag{4}$$

where:

$\tau_1, \ldots, \tau_k$ are a pre-specified set of knot points where $\tau_1 = 0$,

$g_i(\tau) = a_i(\tau - \tau_i)^3 + b_i(\tau - \tau_i)^2 + c_i(\tau - \tau_i) + d_i$,

$g_i^{(n)}(\tau_{i+1}) = g_{i+1}^{(n)}(\tau_{i+1})$ for $i = 1, \ldots, k-2$ and $n = 0, 1, 2$ (n represents the n^{th} derivative),

[12] The above is only a brief description of the Fama-Bilss filter and does not do it justice. A detailed description of the filtering procedure can be found in the CRSP monthly bond file manual from the Center for Research in Security Prices.

and $g_1(0) = 1$.

The above structure implies that for k knot points there are $k+1$ free parameters that can be solved for using the standard ordinary least squares objective function; that is $\min_\theta \sum_{i=1}^{N} \varepsilon_i^2$ where θ is the set of free parameters in the cubic spline.

McCulloch chooses the spacing of the knot points so that there are an equal number of bond observations between successive knot points. The motivation for this choice is that having more issues in a particular maturity region of the term structure allows us to determine a more detailed depiction of the discount function in that region. Since more issues are observed at the short end of the term structure the McCulloch scheme allows for greater flexibility in that region. Furthermore McCulloch suggests setting the number of knot points equal to the square root of the number of observed bonds $\sqrt{N}$. Our first implementation of the McCulloch scheme adheres to this choice however we also consider doubling the number of knot points to $2\sqrt{N}$ in order to study the impact of a more flexible cubic spline.

3 Comparison of Extraction Procedures

3.1 Data

The data used in this study is obtained from the Center for Research in Security Prices (CRSP) and consists of U.S. Treasury bill, note and bond bid and ask prices recorded on the last day of each month over the period 1970 to 1998. To preserve homogeneity in the data we only consider fully-taxable, non-callable, and non-flower U.S. issues. This eliminates pricing complications apparent in some issues that have either a special tax-privileged status or option like features. To alleviate well known illiquidity problems associated with issues that are approaching their maturity date, Treasury bills with a time-to-maturity of less than one month and Treasury notes and bonds with a time-to-maturity of less than one year are excluded. As an additional ad-hoc illiquidity filter, issues that have a relatively large bid-ask spread compared to the average bid-ask spread of instruments with a similar time to maturity are also excluded.[13] The data remaining after imposing the above filters exhibits sparsity and notable gaps in maturity for bond issues with a time to maturity beyond

[13] To be more specific, on a given day the instruments that remain after invoking the previous filters are grouped by time-to-maturity into one year intervals. Denote the average bid-ask spread for the i^{th} interval as $\overline{s_i}$. If the bid-ask spread s of an instrument that has a time to maturity falling within the i^{th} interal is greater than the average spread by 50%, that is $s > 1.5 \times \overline{s_i}$, then this instrument is excluded from the sample. The particular filter only eliminates approximately three to six instruments each month.

ten years. This hinders the ability of all term structure extraction procedures introduced in section 2 to provide a reasonable depiction of the yield curve at the long term end of the term structure. Consequently we restrict attention to issues that have at most a maturity of ten years. As a result of the above filtering process the average sample size for a day in the 1970's, 1980's, and 1990's, is 64, 113 and 140 respectively.

3.2 Methodology

To compare the performance of the different term structure extraction procedures we consider four criteria, three of which were used by Bliss (1997).[14] First the mean absolute pricing error $MAPE = \sum_{i=1}^{N} \left|e_i^P\right|$, where e_i^P denotes the pricing error of the i^{th} reconstructed bond, provides the average dollar error where all bonds have a face value of \$100. As evidenced in Tables 1 and 2 to follow, and noted in Bliss (1997), pricing errors for longer maturity bonds tend to be larger. This motivates the use of the weighted mean absolute pricing error measure $WMAPE = \sum_{i=1}^{N} w_i \left|e_i^P\right|$, where Bliss (1997) suggests the weighting scheme $w_i = (D^i)^{-1} / \sum_{i=1}^{N} (D^i)^{-1}$, where D^i denotes the Macaulay duration of the i^{th} bond. As an alternative means of standardizing pricing error behavior we consider the mean absolute yield error $MAYE = \sum_{i=1}^{N} \left|e_i^Y\right|$, where e_i^Y denotes the error in yield-to-maturity of the i^{th} reconstructed bond. This standardization is motivated by the observation that an error in yield will have a greater impact on the price of a longer maturity bond. Consideration of the $MAYE$ is also intuitively appealing from a practitioners point of view since bonds are often considered in terms of their yield as opposed to price. Finally, since any point between the bid-ask spread is a viable price another measure of performance is to compute the frequency of times that the reconstructed price is within the quoted bid-ask spread. This is measured by the hit rate $HR = \frac{1}{N} \sum_{i=1}^{N} I_{[BID^i \leqslant P^i \leqslant ASK^i]}$ where $I_{[BID^i \leqslant \widehat{P^i} \leqslant ASK^i]}$ is the indicator function that takes the value 1 when the reconstructed price of the i^{th} bond is between the observed bid price BID^i and ask price ASK^i and 0 otherwise.

A novel feature of the Bliss (1997) comparison methodology is that pricing error is defined in terms of the reconstructed bond price relative to the observed bid and ask prices. More specifically the pricing error of i^{th} observed bond is computed as $e_i^P = \widehat{P^i} - ASK^i$ when $\widehat{P^i} > ASK^i$, $e_i^P = \widehat{P^i} - BID^i$ when $\widehat{P^i} < BID^i$, and $e_i^P = 0$ when $BID^i \leqslant \widehat{P^i} \leqslant ASK^i$. This notion of pricing error is intuitive since all values between the bid and ask prices are viable. All performance measures considered here are based on "mispricing outside the bid-ask spread" including the mean absolute yield error, which is based on the yield-to-maturities implied by the fitted bond price, bid price and ask price; denoted $\widehat{Y^i}$, Y_{BID}^i and Y_{ASK}^i respectively. In particular $e_i^Y = \widehat{Y^i} - Y_{ASK}^i$ when $\widehat{Y^i} < Y_{ASK}^i$, $e_i^Y = \widehat{Y^i} - Y_{BID}^i$

[14] The three criteria considered in Bliss (1997) are: i) the mean absolute pricing error, ii) the inverse-duration weighted mean absolute pricing error, and iii) the hit-rate.

when $\widehat{Y^i} > Y^i_{BID}$, and $e^Y_i = 0$ when $Y^i_{BID} \geqslant \widehat{Y^i} \geqslant Y^i_{ASK}$. Note however that the Bliss (1997) notion of pricing error is only used for performance evaluation. When implementing the term structure extraction procedures of section 2 all errors are with respect to mid-point of bid and ask prices, that is all errors are of the form $\varepsilon_i = P^i - \widehat{P^i}$, where $P^i = \frac{1}{2}(BID^i + ASK^i)$.

3.3 Results

The above performance measures are computed for each of the three highly flexible term structure extraction procedures presented in Section 2 with results reported in Tables 1 and 2. The two implementations of the McCulloch cubic spline procedure are referred to as McCulloch-A and McCulloch-B with the first using the originally suggested $\sqrt{N}$ number of knot points and the second extending the number of knot points to $2\sqrt{N}$. To determine whether these performance measures vary across the maturity dimension we also compute these measures for different segments of the term structure. This is achieved by partitioning securities into time-to-maturity regions, namely $0 \leqslant \tau < 1$, $1 \leqslant \tau < 3$, $3 \leqslant \tau < 5$ and $5 \leqslant \tau < 10$, where τ represents the time-to-maturity of the security. Since a primary goal of term structure extraction procedures is to estimate the price of a bond that is not actually traded at the time of interest we wish to assess the relative interpolation performance across term structure extraction procedures. This is achieved by separating the sample of bonds on each day into an estimation subsample and a hold out subsample and computing both in-sample and out-of-sample performance measures. The estimation subsample is obtained by selecting every other bond from the sample where the longest maturing bond is always included. The remaining bonds constitute the hold out sample. Calculating performance measures using out-of-sample data can also be viewed as controlling for the problem of over-fitting associated with non-parsimonious methods. The in-sample and out-of-sample results appear in Tables 1 and 2 respectively.

TABLE 1 ABOUT HERE

By construction, the Fama-Bliss (1987) method provides a perfect fit of all in-sample bond prices. The only errors are from the bonds additionally filtered out using the Fama-Bliss filtering rules. Consequently in-sample performance comparisons are heavily biased in favor of the Fama-Bliss bootstrapping procedure and should be ignored. This accounts for the high Hit-Rate measure in Table 1 indicating that on average 95% of bonds priced using the Fama-Bliss extracted term structure are within quoted bid-ask spreads. This is notably higher than the average hit-rates generated using other procedures, which range from 63% to 73%. Further in-sample errors from the Fama-Bliss procedure yields an error reduction of approximately 50% relative to other term structure extraction methods. In particular for bonds with a time-to-maturity of less than one year, comprising of Treasury bills, the average absolute error is approximately 20 to 40 times smaller.

Focusing on the other term structure extraction methods we find that on average the in-sample absolute pricing error is 3.27 cents, 3.69 cents and 4.23 cents outside the bid-ask spread for the LMNT, McCulloch-B and McCulloch-A methods respectively. Note that these errors are listed in "performance order" with the scheme providing the smallest mean absolute pricing error first. However Table 1 shows that the mean absolute pricing error is consistently smaller at the short end of the term structure and that error magnitudes increase as the maturity of bonds increases. Roughly speaking, the mean absolute pricing error outside the bid-ask spread for all three fitting procedures grows from 1 cent for short maturity bonds to 7.5 cents for long maturity bonds. Given this heteroskedastic behavior in the error, Bliss (1997) suggests weighting errors by their respective bond's duration inverse to prevent the pricing errors of long term bonds dominating the comparison results. The resulting in-sample weighted mean absolute pricing errors outside the bid-ask spread are 1.21 cents, 1.91 cents and 2.23 cents for the LMNT, McCulloch-B and McCulloch-A methods respectively. Note that the performance order remains unchanged. Considering bonds in terms of their yield-to-maturity instead of their price shows that larger yield errors occur at the very short end of the term structure. This is particularly true for both implementations of the McCulloch procedure. Ordering the in-sample mean absolute yield-to-maturity errors results in an unchanged performance order with mean absolute yield errors outside of the bid-ask spread of 1.51 basis points, 2.47 basis points and 3.06 basis points respectively for the LMNT, McCulloch-B and McCulloch-A procedures. Consideration of hit-rates again leaves the performance order unchanged with respective overall hit-rates of 73%, 67% and 63%.

In summary irrespective of the performance measure the in-sample results suggest that the LMNT method is preferred to the McCulloch-B scheme, which in turn is preferred to the McCulloch-A scheme. The arguments for this ordering are based on overall performance measures however the results provided in Table 1 shows that this ordering is preserved within each bond maturity region as well.

TABLE 2 ABOUT HERE

As expected the out-of-sample performance measures provided in Table 2 are not as good as their in-sample counterparts. This is particularly true for the Fama-Bliss bootstrapping procedure. It is interesting to observe that based on overall out-of-sample mean absolute pricing error the Fama-Bliss procedure is now the worst performer with an average absolute pricing error of 6.01 cents outside the bid-ask spread. The best performer is the LMNT procedure with 4.96 cents then the McCulloch-B and McCulloch-A procedures with 5.34 cents and 5.49 cents respectively. Once again notable heteroskedastic behavior in the maturity dimension is observed suggesting that ordering the performance of curve fitting methods based on mean absolute pricing error may be misleading. In particular mean absolute pricing errors at the long end are in the order of 15 cents whereas they are approximately 1 cent at the short end. Focusing on both the overall weighted mean absolute pricing error and the overall mean absolute yield error, which attempt to correct for this heteroskedastic behavior, improves the relative performance of the Fama-Bliss scheme to the point where it now performs better than both McCulloch schemes. In particular the Fama-Bliss, McCulloch-B and McCulloch-A methods result in an out-of-sample overall weighted mean absolute pricing error of 1.61, 2.25 and 2.49 cents respectively with corresponding overall mean absolute yield errors of 2.05, 2.94 and 3.44 basis points. Note however that the Fama-Bliss procedure does not perform notably better or worse than the LMNT procedure. The LMNT method has an overall weighted mean absolute pricing error of 1.56 cents and a overall mean absolute yield error of 2.03 basis points both of which are negligibly better than the corresponding measures provided by the Fama-Bliss method. The comparable performance between the Fama-Bliss and LMNT methods is reiterated with a small 0.7% difference in the out-of-sample overall hit-rate in favor of the Fama-Bliss procedure. The reason for the improvement in the Fama-Bliss method can be obtained by observing what happens across bond-maturity regions. Notice that the only region where the Fama-Bliss bootstrapping procedure performs better than other schemes is in the Treasury bill region - bonds with a time-to-maturity of less than one year. Within this region the Fama-Bliss procedure yields a mean absolute pricing error of 0.36 cents and a mean absolute yield error of 1.17 basis points and these small errors have a significant impact on the $WMAPE$ and the $MAYE$ performance measures. To further emphasize the heavy influence of this notably high performance at the short end observe that the Fama-Bliss procedure is consistently the worst performer when bond maturity is greater than one year.

How to choose from the term structure extraction schemes based on out-of-sample performance is not clear cut. This is because the performance order does vary within maturity regions, in particular the striking change in performance order of the Fama-Bliss method. However, based on overall performance and controlling for the heteroskedastic behavior along the maturity dimension [by using the weighted mean absolute pricing error and the mean absolute yield error metrics] suggests the performance ordering: LMNT, Fama-Bliss, McCulloch-B and then McCulloch-A. This is consistent with the in-sample ranking given the exclusion of the Fama-Bliss scheme because of its in-sample construction. The only deviation from this ordering is a reversal in order of the Fama-Bliss and LMNT procedures when using the overall hit-rate metric. We choose to ignore this reversal given: i) that this reversal is based on a mere 0.7% difference in hit-rate, ii) our previous observation that the LMNT scheme is consistently superior to the Fama-Bliss scheme in all bond maturity regions beyond one year, and iii) for bonds with a maturity of less than one year the difference in mean absolute pricing error is 0.26 cents and the difference in mean absolute yield error is 1.45 basis points both of which are economically small.

A feature that has been ignored in the above discussion is how these performance measures behave over time. For ease of exposition only the time series of the overall out-of-sample LMNT performance measures are provided in Figure 1 as all other schemes result in similar behavior both in-sample and out-of-sample.

FIGURE 1 ABOUT HERE

It is interesting to observe that the decade prior to 1980 is a time period when term structure extraction procedures work relatively well. However the early to mid-1980s, encapsulating the period of what is referred to as the "fed-experiment", marked a time of relatively poor performance with absolute errors more than doubling. After the mid to late 1980 absolute error magnitudes steadily fell nearing levels observed in the 1970s even though the number of bonds out of sample more than doubled; from January 1989 to December 1998 there were a total of 8333 out-of-sample bonds, whereas from January 1970 to December 1979 the out-of-sample bonds numbered 3791.

When comparing term structure extraction procedures it is important to control for the systematic variation in performance measure magnitudes over time. This is achieved by using each metric to determine which term structure extraction method is preferred on each day when performing pair-wise comparisons. We then compute the fraction of time that one scheme is strictly preferred to another noting that the times when both schemes perform equally well are excluded. We refer to this comparison metric as "percentage preference":

$$Percentage\ Preference_{metric} = \frac{\sum_{t=1}^{T} I_{[X_t > Y_t]}}{\sum_{t=1}^{T} I_{[X_t > Y_t]} + \sum_{t=1}^{T} I_{[Y_t > X_t]}}$$

where $T = 348$, which is the number of days in the sample and $I_{[X_t > Y_t]}$ is the indicator function taking on the value 1 when term structure extraction procedure X indicates a better fit than term structure extraction procedure Y at time t under a given *metric* (namely one of $MAPE$, $WMAPE$, $MAYE$ or HR). By aggregating preference orders over time in this manner we are ensuring that the high error magnitudes observed in the 1980s do not dominate the conclusions drawn. In-sample and out-of-sample percentage preferences for each comparison metric appear in Tables 3 and 4 respectively. For completeness the in-sample percentage preference of the Fama-Bliss procedure is included in Table 3 even though it should be ignored since, by construction, its percentage preference will be close to 100%.

TABLE 3 ABOUT HERE

From Table 3 it is apparent that in-sample the LMNT procedure consistently out performs the original McCulloch scheme (McCulloch-A); within all maturity regions the LMNT procedure is consistently preferred 70%-85% of the time under all performance measures. The extended McCulloch scheme, that is McCulloch-B, should perform better in-sample than the original McCulloch-A scheme.

This is indeed reflected in Table 3 however it is interesting to observe that the majority of the benefit appears at the long end of the term structure; all performance measures suggest that the McCulloch-B scheme is preferred only 55% of the time in the Treasury bill region however this preference for the McCulloch-B method grows to 70% and higher in the five to ten year maturity region. This variation in percentage preference across maturity explains the 76% overall preferential given by the $MAPE$ metric, which over-weights the importance of the long end, whereas both the $WMAPE$ and $MAYE$ metrics, which standardize for error heteroskedasticity, provide an overall percentage preference of 57%. Comparing the LMNT and McCulloch-B procedures we again see variation in percentage preference across maturity. Within the one year or less bond maturity region the LMNT procedure is preferred to the McCulloch-B procedure more than 75% of the time but this percentage preference slowly decreases to approximately 50% in the five to ten year bond maturity region. It is clear however that even though there is no clear preference between the LMNT and McCulloch-B schemes at the long end of the term structure there is a distinct preference for the LMNT procedure for shorter maturity bonds. This is reiterated by the overall percentage preference figures ranging from 76% to 84% across all performance metrics in favor of the LMNT procedure relative to the McCulloch-B procedure. In summary the in-sample results suggest a clear performance ordering, First the LMNT procedure followed by the McCulloch-B scheme and finally the McCulloch-A method.

TABLE 4 ABOUT HERE

Turning to the out-of-sample results reveals some striking differences to their in-sample counterparts. First comparing the two McCulloch implementations we see that the heavy in-sample preference of the McCulloch-B method over the McCulloch-A scheme at the longer end of the term structure has eroded and in fact in the three to five year bond maturity bin the McCulloch-A scheme is slightly preferred. This is suggestive that the superior in-sample performance of the McCulloch-B method is due to over-fitting. The overall percentage performance figures still suggest a slight benefit of the McCulloch-B procedure however this is primarily due to the marginally better performance in maturities less than three years. The hit-rate performance measure is the strongest supporter of the McCulloch-B scheme over the McCulloch-A method suggesting an overall percentage preference of 62%. However we must realize that the hit-rate measure does not distinguish between error magnitudes outside the bid-ask spread and hence does not penalize those methods that result in larger errors. The remaining metrics ($MAPE$, $WMAPE$ and $MAYE$) do not suffer from this deficiency and provide overall percentage preference figures in the range of 52% to 54%, which suggests the benefit of adding parameters to the original McCulloch cubic spline procedure is negligible.

Comparing the LMNT procedure to both McCulloch implementations indicates a 75% percentage preference for the LMNT method when using the overall maturity spectrum across all metrics.

Consistently the LMNT procedure is preferred to both McCulloch schemes approximately 75% of the time in the Treasury bill region, 65-69% of the time in the one to three year bond maturity region, 52-60% of the time in the three to five year maturity region, and approximately 62% of the time in the five to ten year maturity region. Occasional exceptions to these figures arise when considering the hit-rate metric where a lower preference for the LMNT method is sometimes observed. Given the hit-rate measure's lack of penalty toward larger errors outside the bid-ask spread and the stronger preference for the LMNT method when considering other performance measures, we choose not to place an emphasis on hit-rate percentage preference numbers.

Finally considering the Fama-Bliss scheme reveals several interesting observations. In the Treasury bill region the Fama-Bliss method is strongly preferred out-of-sample to both McCulloch implementations and the LMNT procedure. Depending on the performance measure considered the percentage preference is an extraordinary 81-88% over the LMNT procedure and a massive 90-97% over both McCulloch schemes. However this strong preference rapidly dwindles almost to the other extreme when considering bonds with a maturity of greater than one year. In the one to three, three to five, and five to ten year bond maturity bins the approximate out-of-sample percentage preference over the Fama-Bliss method by i) both McCulloch schemes is 65%, 74% and 56%, and ii) the LMNT procedure is a massive 71%, 77% and 64%.[15] These observations provide strong evidence to suggest that beyond the one-year maturity region the Fama-Bliss procedure is the last method of choice, however in the Treasury bill maturity region it is the preferred method. Aggregating such a diverse ranking across maturity is virtually an impossible task since it depends on the users subjective weightings regarding which maturity regions are more important. If all bonds are treated as equally important, irrespective of their maturity, then the overall percentage preference figures should be used. Adopting either the overall $WMAPE$ or the $MAYE$ based percentage preference measures to rank term structure estimation procedures, because they attempt to standardize error magnitudes across time and maturity, the Fama-Bliss procedure is preferred out-of-sample to both McCulloch schemes at least 70% of the time and the LMNT scheme is preferred marginally to the Fama-Bliss scheme (55% of the time out-of-sample).

To summarize, in-sample results suggest a clear ranking of the term structure estimation procedures studied; first the LMNT procedure followed by the McCulloch-B scheme and finally the McCulloch-A method. The Fama-Bliss scheme is excluded from this ranking since by its construc-

[15] These numbers do not include the percentage preference figures based on the hit-rate metric however this metric does in part support the above arguement. The only exception is in the five to ten year bond maturity spectrum where the hit-rate percentage preference suggests a slight preference for the Fama-Bliss scheme over other methods. Once again we explain this result by resorting to the deficiancy of the hit-rate metric, namely it fails to penalize larger errors.

tion it will almost always be ranked first. Ignoring the Fama-Bliss scheme the out-of-sample results reaffirm the above ranking. Placing the Fama-Bliss scheme within this ranking unfortunately is not as clear-cut. If one is predominantly interested in the short end of the term structure (less than one year time-to-maturity) then the Fama-Bliss scheme is the preferred choice above the LMNT procedure. If one is interested in maturities greater than one year then the Fama-Bliss scheme ranks last, below the LMNT, McCulloch-B, and McCulloch-A schemes.

4 A Proxy for the Short Rate

Several recent articles have raised the issue of which interest rate to use as a proxy for the short term interest rate. Duffee (1996) argues that the three month Treasury bill rate is preferable to the one month rate and a recent paper by Chapman-Long-Pearson (1999) shows that severe biases can arise in the short rate's drift and diffusion estimates when using these Treasury bill rates as short rate proxies. An additional issue that is not addressed in either study is the effect of observation error in bond prices. If observation errors are present a reasonable conjecture is that using an estimated short rate from a curve fitting procedure should be superior to both the one month and three month Treasury bill rate proxies. This conjecture is based on the idea that a curve fitting procedure averages observations at the short end of the term structure and extrapolates back to the origin providing a short rate estimate with higher precision. Evidence to support this is provided below via a simulation where it is demonstrated that: i) the estimated short rate from the LMNT curve fitting procedure results in a smaller mean error and a smaller error variance, and ii) estimated drift and diffusion coefficients for the short rate process are more accurate when using the LMNT based short rate estimate as a proxy for the short rate. The LMNT term structure extraction method is used here since it was concluded in Section 3 that the LMNT procedure performs better than both McCulloch cubic spline implementations. The Fama-Bliss scheme is not considered since it is not a curve fitting procedure in the sense that it is not regression based and hence does not have the ability to average observation errors. Note that even though it was concluded in section 3 that the Fama-Bliss method for estimating the term structure is preferred in the Treasury bill region, it is also true that Treasury bills with very short maturities [less than one month] were excluded from the sample and under ideal circumstances these Treasury bills should be used to determine a proxy the short rate. As will be seen in the following simulation, in less then ideal circumstances it is in these very short term securities that large errors occur when estimating the short rate.

4.1 The Simulation

When estimating a term structure model it is typical that the data consists of Treasury bill, note and bond prices. To simulate the prices of these securities a time-series of zero-coupon bond yield curves is simulated. This is achieved using the Cox-Ingersoll-Ross (1985) term structure model, which is chosen purely for the convenience of having a closed for solution. In particular the evolution of the short rate is characterized by the mean-reverting process

$$dr(t) = \kappa \left(\theta - r(t)\right) dt + \sigma \sqrt{r(t)} dW(t),$$

where $r(t)$ is the short rate at time t and $W(t)$ is the Brownian motion characterizing uncertainty in the bond market. Further the market price of interest rate risk takes the form $\frac{\lambda}{\sigma}\sqrt{r(t)}$ and the resulting functional form for the yield curve at time t with time-to-maturity τ is

$$y(\tau) \;=\; -\frac{\phi_3}{\tau} \log\left(\frac{\phi_1 e^{\phi_2 \tau}}{\phi_2\left(e^{\phi_1 \tau} - 1\right) + \phi_1}\right) + \frac{1}{\tau}\left(\frac{e^{\phi_1 \tau} - 1}{\phi_2\left(e^{\phi_1 \tau} - 1\right) + \phi_1}\right) r(t),$$

where $\phi_1 = \sqrt{(\kappa + \lambda)^2 + 2\sigma^2}$, $\phi_2 = \left(\kappa + \lambda + \phi_1\right)/2$ and $\phi_3 = 2\kappa\theta/\sigma^2$. For the simulation the short rate process is simulated daily over a ten year time horizon using parameters from Chen and Scott (1993); $\kappa = 0.4697$, $\theta = 0.06182$, $\sigma = 0.08248$, $\lambda = -0.04544$. Given this time-series realization of yield curves the prices of Treasury securities are constructed each with a face value of \$100. To ensure that the maturity and coupon structure of these securities are as realistic as possible we use the structure observed in the CRSP bond data set from January 1989 to December 1998 on a daily basis, which provides 2501 days of simulated prices.[16] We choose the structure implied by government securities over the most recent ten year period purely because more securities are observed in later years. Finally these simulated security prices are contaminated with Normally distributed i.i.d. random numbers that are mean zero and have a standard deviation of 5 cents, 15 cents, 25 cents and 35 cents for securities with a time-to-maturity τ in years of $0 < \tau \leqslant 1$, $1 < \tau \leqslant 3$, $3 < \tau \leqslant 5$ and $5 < \tau \leqslant 10$ respectively.

Using the above contaminated prices three proxies for the short rate are computed; the yields implied by the one and three month Treasury bills and the estimated short rate from the LMNT procedure. Table 5 presents the mean and standard deviation of the error associated with each short rate proxy. From this table the LMNT based short rate estimate has a lower bias and a higher precision (its mean error is 6.8 / 3.6 times smaller and its error standard deviation is approximately 4.6 / 1.7 times smaller than the one / three month Treasury bill rate based proxy).

TABLE 5 ABOUT HERE

[16] The only regular exception is the addition of both a one-month and a three-month Treasury bill.

The benefit of using the LMNT short rate estimate as opposed to the above two Treasury bill proxies is further highlighted when estimating the short rate process itself. In particular suppose that we do not know the functional form of the process describing the evolution of the short rate and we are interested in learning about it through the data. This can be achieved using nonparametric techniques such as those in Aït-Sahalia (1996a), Stanton (1997), Jiang and Knight (1997) and Bandi and Phillips (1998). Typically these methods assume that the short rate process follows a time-homogeneous univariate Markov diffusion and the objective is to estimate the drift and diffusion coefficients nonparametrically. Here we use Stanton's method where the drift and diffusion estimators are based on standard kernel smoothing methods with the normal kernel $K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$. The drift and diffusion estimators are respectively

$$\widehat{\mu}(r) = \frac{\sum_{i=1}^{n-1} \left(r(t_{i+1}) - r(t_i)\right) K\left(\frac{r(t_i)-r}{h}\right)}{\Delta \sum_{i=1}^{n-1} \frac{1}{h} K\left(\frac{r(t_i)-r}{h}\right)}$$

$$\widehat{\sigma}(r) = \sqrt{\frac{\sum_{i=1}^{n-1} \left(r(t_{i+1}) - r(t_i)\right)^2 K\left(\frac{r(t_i)-r}{h}\right)}{\Delta \sum_{i=1}^{n-1} \frac{1}{h} K\left(\frac{r(t_i)-r}{h}\right)}}$$

where n is the sample size and $\Delta = t_{i+1} - t_i$ for all i. The bandwidth parameter h chosen by Stanton is $4\sqrt{V}n^{-1/5}$ where V is the sample variance of changes on the short rate.[17] The results from implementing these estimators using the three short rate proxies is provided in Figure 2 along with the "true" drift and volatility functions from the original Cox-Ingersoll-Ross model that formed the foundation of the simulated data.

FIGURE 2 ABOUT HERE

The results in Figure 2 indicate that the benefits from using a curve fitting based estimate of the short rate is not inconsequential. Both the drift and volatility estimates are notably closer to the true drift and volatility when using the LMNT based short rate proxy. All drift estimates demonstrate a notable error in their slope, corresponding to the speed of mean reversion, and the volatility estimates are upward biased. The cause of the bias in the volatility estimate is clear. It stems from the additional noise in the short rate proxy from observation noise. It is also interesting to observe the dramatic improvement of the three month Treasury bill rate proxy over the one month proxy even though the former is theoretically closer to the short rate if observation noise is ignored. The improvement by using the three month Treasury bill arises from the fact that for a given error

[17] This bandwidth choice by Stanton is reported in Chapman and Pearson (1999) footnote 4.

in price the corresponding error in yield will be greater for shorter maturity securities. The presence of errors in short rate proxies can also explain the error in the slope of the drift function's estimate. To see this consider the case where the short rate is indeed reverting toward some mean level θ. Any proxy for the short rate that has noise will have higher volatility than the true short rate. Consequently the proxy will cross the value θ more frequently than the true short rate. This makes the estimation of the rate of mean reversion more difficult and the higher the proxy error variance the more difficult it will be.

In short, given observation noise in prices, the LMNT based short rate proxy provides a better estimate to the short rate than the three month Treasury bill rate, which in turn is a better estimate than that provided by the one month Treasury bill rate. The importance of using a better estimate is highlighted when using it to estimate the time series properties of the short rate. In particular the variance of the noise in the proxy determines the bias in short rate volatility and strongly influences the ability to estimate the slope of the drift function.

5 Summary and Conclusion

Using month end bond price data from January (1970) to December (1998) three highly flexible term structure extraction methods have been compared: McCulloch's (1975) cubic spline, the Fama-Bliss (1987) bootstrapping procedure and the LMNT (2000) nonparametric estimation method. Two versions of the McCulloch cubic spline procedure were considered, namely that originally proposed by McCulloch (1975) and a version with twice as many knot points. Despite this additional flexibility of the latter version, out-of-sample results suggest that there is almost no benefit of increasing the number of knot points beyond that proposed by McCulloch. Ranking the performance of the three methods, based on analyzing both in-sample and out-of-sample pricing and yield errors, we conclude that the LMNT (2000) method is the preferred method of choice when estimating the term structure beyond one year. This is followed by McCulloch's cubic spline and then the Fama-Bliss bootstrapping method. It is interesting to observe that when considering the maturity region of less than one year the Fama-Bliss procedure is now the preferred method followed by the LMNT procedure and then McCulloch's cubic spline. Given that the Fama-Bliss procedure is the least parsimonious method, since all in-sample bonds are priced exactly, the last result suggests that additional flexibility in the Treasury bill region needs to be incorporated into the other two term structure estimation methods.

Use of LMNT curve fitting method is shown via simulation to aid in the estimation of the unobservable short rate. Even though it was concluded that the Fama-Bliss method is the preferred term structure extraction method in the Treasury bill region it is also true that this conclusion was based on reconstructing Treasury bill prices and yields with maturities ranging from one to twelve months. When focusing on estimating the short rate the simulation shows that if observation noise is present in prices then using the yield from a very short maturity Treasury bill [maturity of one month] as a proxy for the short rate has a lower precision then a yield from a longer term Treasury bill [maturity of three months]. This is because a small error in price has a larger impact on the error in yield as the maturity of the Treasury bill decreases. The benefit of using the LMNT procedure arises from the fact that errors from a number a short dated Treasury bills are averaged providing a short rate estimate with higher precision. The simulation further demonstrates that using the LMNT based estimate is not inconsequential when estimating the time series properties of the short rate. The higher precision of the short rate estimate results in a lower bias in the short rate volatility estimate and smaller error in the slope estimate of the short rate's drift function.

References

[1] Aït-Sahalia, Y. (1996a) Nonparametric Pricing of Interest Rate Derivative Securities. Econometrica 64, 3, 527-560.

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

[3] Aït-Sahalia, Y. (1998). Maximum Likelihood Estimation of Discretely-Sampled Data: A Closed Form Approach. Working paper.

[4] Amihud, Y. and Mendelsohn, H. (1991) Liquidity, Maturity and the Yields on U.S. Treasury Securities, Journal of Finance, 46, 1411-1425.

[5] Bandi, F. and P. Phillips (1998). Econometric Estimation of Diffusion Models. Unpublished paper, Yale University.

[6] Bliss, R. (1997). Testing Term Structure Estimation Methods. Advances in Futures and Options Research, 9, 197-231.

[7] Chambers, D. R., Carleton, W. T. and Waldman, D. W. (1984). A New Approach to Estimation of the Term Structure of Interest Rates, Journal of Financial and Quantitative Analysis, 19, 233-269.

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

[9] Chapman, D. A., J. B. Long and N.D. Pearson (1999). Using Proxies For the Short Rate: When Are Three Months Like an Instant? Reviews of Financial Studies (Forthcoming).

[10] Chapman, D. A. and N. D. Pearson (1999). Is the Short Rate Drift Actually Nonlinear? Journal of Finance (Forthcoming).

[11] Chen R., and L. Scott. (1993). Maximum Likelihood Estimation For a Multifactor Equilibrium Model of the Term Structure of Interest Rates. Journal of Fixed Income, December, 14-31.

[12] Conley, T., P. L. Hansen, E. Luttmer, and J. Scheinkman (1997) Short-Term Interest Rate as Subordinated Diffusison. Review of Financial Studies, 10, 10(3), 525-77.

[13] Cox, J., J. Ingersol and S. Ross (1985). A Theory of the Term Structure of Interest Rates. Econometrica, 53, 385-406.

[14] Duffee, D. (1996). Idioyncratic Variation of Treasury Bill Yields. Journal of Finance, 51, 527-551.

[15] Fama, E. and R. Bliss (1987). The Information in Long-Maturity Forward Rates. The American Economic Review, 77, 680-692.

[16] Fisher, M., D. Nychka, and D. Zervos. (1995). Fitting the Term Structure of Interest Rates with Smoothing Splines. Working paper 95-1, Finance and Economics Discussion Series, Federal Reserve Board.

[17] Heath,D., R. Jarrow, A. Morton (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. Econometrica 60, 77-105.

[18] Hull, J. and A. White (1990). Pricing Interest Rate Derivative Securities. Review of Financial Studies 3 (4), 573-592.

[19] Jeffrey, A. (1995). Single Factor Heath-Jarrow-Morton Term Structure Models Based on Markov Spot Interest Rate Dynamics. Journal of Financial and Quantitative Analysis, 619-642.

[20] Jeffrey, A. (1998). An Empirical Examination of a Path-Dependent Term Structure Model. Working paper, Yale School of Managment.

[21] Jiang, G. and J. Knight (1997). A Nonparametric Approach to the Estimation of Diffusion Processes, With an Application to a Short Term Interest Rate Model. Econometric Theory, 13, 616-645.

[22] Linton, O., E. Mammen, J. Nielsen and C. Tanggaard (2000). Estimating Yield Curves by Kernel Smoothing Methods. Journal of Econometrics (Forthcoming).

[23] McCulloch, J. H. (1971). Measuring the Term Structure of Interest Rates. The Journal of Business, 44, 19-31.

[24] McCulloch, J. H. (1975). The Tax-Adjusted Yield Curve. Journal of Finance, 30, 811-830.

[25] Mishkin, F. R. (1988). The Information in the Term Structure: Some Further Results. Journal of Applied Econometrics, 3, 307-314.

[26] Nelson, C. R. and A. F. Siegel (1987). Parsimonious Modeling of Yield Curves. Journal of Business, 60, 473-489.

[27] Sarig O. and A. Warga (1989). Bond Price Data and Bond Market Liquidity. Journal of Financial and Quantitative Analysis, 24, 367-378.

[28] Stanton, R. (1997). A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk. Journal of Finance, 52, 1973-2002.

[29] Vasicek, O. A. and Fong, H. G. (1982). Term Structure Modelling Using Exponential Splines. Journal of Finance, 37, 339-348.

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

[31] Waggoner, D. (1997). Spline Method for Extracting Interest Rate Curves From Coupon Bond Prices. Working paper, Federal Reserve Bank of Atlanta.

TABLE 1
In-Sample Performance: January 1970 to December 1998

The four performance measures are presented for each term structure extraction procedure; LMNT, McCulloch and Fama-Bliss. Two versions of the McCulloch scheme are implemented, the originally proposed scheme denoted McCulloch-A and one where twice the number of knot points is used which is denoted McCulloch-B. The sample consists of $N = 18233$ bonds observed at month end over the January 1970 to December 1998 period with a time-to-maturity T of less than or equal to 10 years. Variation in performance across time-to-maturity is obtained by calculating the performance measures for bonds grouped by their time-to-maturity. In particular the groupings consist of $0 < T < 1$ where $N = 4613$, $1 < T < 3$ where $N = 6666$, $3 < T < 5$ where $N = 3809$, and $5 < T < 10$ where $N = 3145$.

	$0 < T < 1$	$1 < T < 3$	$3 < T < 5$	$5 < T < 10$	Overall $0 < T < 10$
MEAN ABSOLUTE PRICING ERROR					
LMNT	0.005697	0.019393	0.053003	0.070863	0.032697
McCulloch-A	0.016603	0.024800	0.066291	0.088555	0.042304
McCulloch-B	0.013147	0.021399	0.062155	0.076589	0.036877
Fama-Bliss	0.000715	0.014230	0.030867	0.029903	0.016648
WEIGHTED MEAN ABSOLUTE PRICING ERROR					
LMNT	0.005938	0.017413	0.050523	0.070635	0.012078
McCulloch-A	0.017536	0.022757	0.063674	0.090501	0.022333
McCulloch-B	0.014901	0.019140	0.059508	0.078987	0.019053
Fama-Bliss	0.000368	0.012991	0.029426	0.030967	0.004711
MEAN ABSOLUTE YIELD-TO-MATURITY ERROR					
LMNT	0.000241	0.000108	0.000157	0.000136	0.000151
McCulloch-A	0.000716	0.000142	0.000197	0.000174	0.000306
McCulloch-B	0.000605	0.000119	0.000184	0.000153	0.000247
Fama-Bliss	0.000015	0.000081	0.000095	0.000063	0.000063
HIT RATE					
LMNT	0.666278	0.822913	0.681897	0.631494	0.729445
McCulloch-A	0.468617	0.771555	0.618559	0.564353	0.631289
McCulloch-B	0.511434	0.796227	0.639704	0.610843	0.671944
Fama-Bliss	0.993225	0.931474	0.946780	0.962618	0.956376

TABLE 2
Out-of-Sample Performance: January 1970 to December 1998

The four performance measures are presented for each term structure extraction procedure; LMNT, McCulloch and Fama-Bliss. Two versions of the McCulloch scheme are implemented, the originally proposed scheme denoted McCulloch-A and one where twice the number of knot points is used which is denoted McCulloch-B. The sample consists of $N = 18057$ bonds observed at month end over the January 1970 to December 1998 period with a time-to-maturity T of less than or equal to 10 years. Variation in performance across time-to-maturity is obtained by calculating the performance measures for bonds grouped by their time-to-maturity. In particular the groupings consist of $0 < T < 1$ where $N = 4616$, $1 < T < 3$ where $N = 6655$, $3 < T < 5$ where $N = 3814$, and $5 < T < 10$ where $N = 2972$.

	$0 < T < 1$	$1 < T < 3$	$3 < T < 5$	$5 < T < 10$	Overall $0 < T < 10$
MEAN ABSOLUTE PRICING ERROR					
LMNT	0.006262	0.026827	0.078006	0.144572	0.049601
McCulloch-A	0.017002	0.029838	0.075482	0.152052	0.054858
McCulloch-B	0.013376	0.028281	0.079640	0.152175	0.053446
Fama-Bliss	0.003636	0.032672	0.104921	0.160128	0.060094
WEIGHTED MEAN ABSOLUTE PRICING ERROR					
LMNT	0.006387	0.023688	0.074736	0.143162	0.015632
McCulloch-A	0.017886	0.027118	0.071977	0.150687	0.024884
McCulloch-B	0.015220	0.024957	0.076113	0.151089	0.022468
Fama-Bliss	0.002846	0.029423	0.100946	0.158993	0.016064
MEAN ABSOLUTE YIELD-TO-MATURITY ERROR					
LMNT	0.000260	0.000148	0.000231	0.000282	0.000203
McCulloch-A	0.000731	0.000169	0.000222	0.000301	0.000344
McCulloch-B	0.000616	0.000156	0.000235	0.000297	0.000294
Fama-Bliss	0.000117	0.000183	0.000310	0.000312	0.000205
HIT RATE					
LMNT	0.663389	0.779331	0.605307	0.457735	0.681398
McCulloch-A	0.464781	0.753508	0.591462	0.436348	0.604877
McCulloch-B	0.502854	0.762980	0.586943	0.454743	0.629343
Fama-Bliss	0.787265	0.750037	0.546485	0.474790	0.688797

TABLE 3
In-Sample Percentage Preference

Using in-sample data the percentage preference for a each pair of term structure extraction procedures is computed based on each of the four performance measures *MAPE*, *WMAPE*, *MAYE* and *HIT_RATE*. The notation $X > Y$ refers to the consideration of the case where term structure extraction procedure X is preferred to procedure Y with the corresponding number indicating the fraction of time that X is indeed preferred to Y under the given metric within the given maturity bin. Maturity bins include $0 < T < 1$, $1 < T < 3$, $3 < T < 5$, $5 < T < 10$ and the overall spectrum $0 < T < 10$ where T denotes the security's time-to-maturity. Percentage preference of method X relative to Y is computed via

$$\text{percentage preference} = \frac{Nunber\,(X > Y)}{Number\,(X > Y) + Number\,(Y > X)}$$

where $Number\,(X > Y)$ is the number of times that term structure extraction procedure X is strictly preferred to method Y. Note that the number of times that procedures X and Y perform equally well is excluded from this metric and hence the percentage preference of method Y relative to X is one minus the percentage preference of X relative to Y.

	$0 < T < 1$	$1 < T < 3$	$3 < T < 5$	$5 < T < 10$	Overall $0 < T < 10$
LMNT > McCulloch-B					
MAPE	78.70%	69.02%	68.54%	46.86%	75.79%
WMAPE	77.78%	71.04%	67.88%	49.45%	83.57%
MAYE	77.78%	71.33%	69.00%	49.08%	83.86%
HIT-RATE	75.52%	70.74%	64.95%	53.85%	79.73%
LMNT > McCulloch-A					
MAPE	82.39%	83.39%	77.20%	70.61%	89.05%
WMAPE	77.01%	85.99%	78.50%	74.55%	86.74%
MAYE	77.01%	85.95%	79.08%	74.19%	86.71%
HIT-RATE	83.75%	82.79%	68.78%	69.76%	86.20%
McCulloch-B > McCulloch-A					
MAPE	57.52%	68.59%	61.02%	74.73%	76.15%
WMAPE	52.51%	69.55%	61.98%	75.09%	57.18%
MAYE	52.80%	69.87%	61.66%	75.45%	57.47%
HIT-RATE	57.38%	64.68%	57.75%	69.75%	67.20%
Fama-Bliss > LMNT					
MAPE	97.59%	79.42%	91.34%	91.51%	93.01%
WMAPE	98.28%	80.87%	92.06%	90.73%	95.44%
MAYE	98.28%	79.71%	89.89%	91.51%	95.74%
HIT-RATE	99.65%	92.05%	98.15%	98.02%	99.07%
Fama-Bliss > McCulloch-B					
MAPE	98.41%	85.42%	93.67%	91.82%	95.66%
WMAPE	99.04%	85.76%	94.67%	91.82%	97.98%
MAYE	99.04%	85.08%	93.33%	92.19%	97.69%
HIT-RATE	99.68%	97.13%	98.98%	96.64%	99.71%
Fama-Bliss > McCulloch-A					
MAPE	98.78%	87.58%	94.14%	94.58%	98.27%
WMAPE	99.39%	88.24%	94.46%	94.58%	99.13%
MAYE	99.39%	87.58%	93.49%	94.58%	98.84%
HIT-RATE	100.00%	96.94%	99.33%	97.83%	99.71%

TABLE 4
Out-of-Sample Percentage Preference

Using out-of-sample data the percentage preference for a each pair of term structure extraction procedures is computed based on each of the four performance measures *MAPE*, *WMAPE*, *MAYE* and *HIT_RATE*. The notation $X > Y$ refers to the consideration of the case where term structure extraction procedure X is preferred to procedure Y with the corresponding number indicating the fraction of time that X is indeed preferred to Y under the given metric within the given maturity bin. Maturity bins include $0 < T < 1$, $1 < T < 3$, $3 < T < 5$, $5 < T < 10$ and the overall spectrum $0 < T < 10$ where T denotes the security's time-to-maturity. Percentage preference of method X relative to Y is computed via

$$\text{percentage preference} = \frac{Nunber\,(X > Y)}{Number\,(X > Y) + Number\,(Y > X)}$$

where $Number\,(X > Y)$ is the number of times that term structure extraction procedure X is strictly preferred to method Y. Note that the number of times that procedures X and Y perform equally well is excluded from this metric and hence the percentage preference of method Y relative to X is one minus the percentage preference of X relative to Y.

	$0 < T < 1$	$1 < T < 3$	$3 < T < 5$	$5 < T < 10$	Overall $0 < T < 10$
LMNT > Fama-Bliss					
MAPE	18.59%	71.52%	77.12%	64.87%	75.29%
WMAPE	18.27%	70.90%	77.12%	63.92%	55.17%
MAYE	18.01%	70.90%	77.74%	63.61%	54.60%
HIT-RATE	12.20%	68.40%	71.73%	43.69%	44.95%
LMNT > McCulloch-B					
MAPE	76.15%	64.78%	58.93%	61.39%	75.79%
WMAPE	75.23%	64.78%	59.56%	61.71%	76.08%
MAYE	75.23%	65.19%	60.19%	60.76%	75.79%
HIT-RATE	78.24%	61.86%	62.73%	46.43%	76.09%
LMNT > McCulloch-A					
MAPE	78.76%	65.74%	52.02%	65.94%	77.01%
WMAPE	75.52%	68.83%	52.34%	62.50%	75.86%
MAYE	75.52%	69.04%	52.50%	63.64%	76.15%
HIT-RATE	79.86%	66.98%	57.78%	53.33%	80.06%
Fama-Bliss > McCulloch-B					
MAPE	89.75%	37.04%	24.84%	42.72%	38.79%
WMAPE	93.17%	37.35%	25.16%	42.41%	71.55%
MAYE	92.86%	37.46%	24.30%	42.41%	70.69%
HIT-RATE	95.89%	43.03%	32.00%	54.84%	72.93%
Fama-Bliss > McCulloch-A					
MAPE	92.22%	38.23%	26.32%	45.45%	42.82%
WMAPE	94.01%	38.53%	26.63%	45.77%	73.28%
MAYE	94.01%	38.77%	26.32%	46.08%	72.70%
HIT-RATE	97.07%	44.66%	39.00%	61.93%	79.22%
McCulloch-B > McCulloch-A					
MAPE	57.31%	53.89%	41.77%	49.36%	52.30%
WMAPE	52.05%	55.14%	42.41%	50.32%	54.31%
MAYE	52.34%	55.49%	42.86%	49.84%	54.18%
HIT-RATE	56.31%	56.42%	49.72%	57.05%	61.56%

TABLE 5
Means and Standard Deviations of Proxy Errors

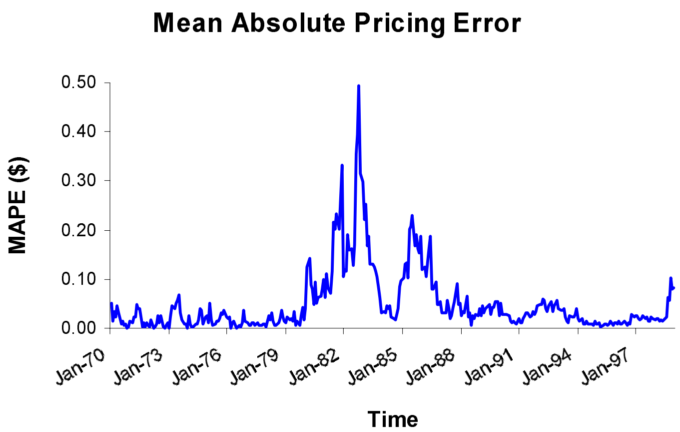
Three proxies for the short rate are the extrapolated short rate from the LMNT procedure, the one month Treasury Bill rate and the three month Treasury Bill rate. The LMNT estimate controls for the possibility of observation errors present in bond prices whereas the latter two do not. Based on a simulated Cox-Ingersoll-Ross (1985) model for the term structure's evolution over a ten year time-horizon and mean zero observation errors with calibrated standard deviations, the mean and standard deviation of the proxy errors are provided: the proxy error at time t is computed via

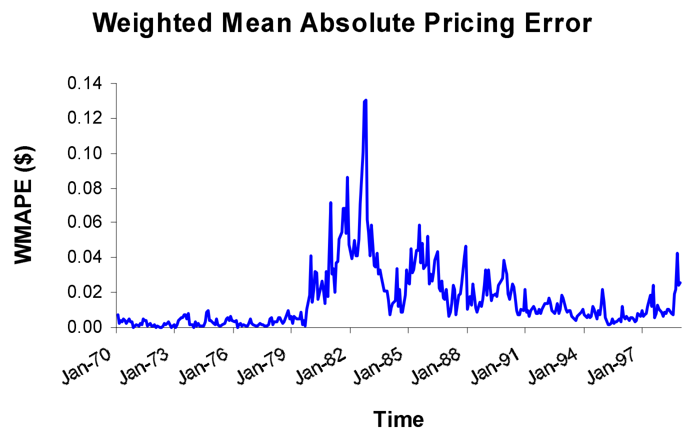
$$proxy\ error(t) = proxy(t) - r(t)$$

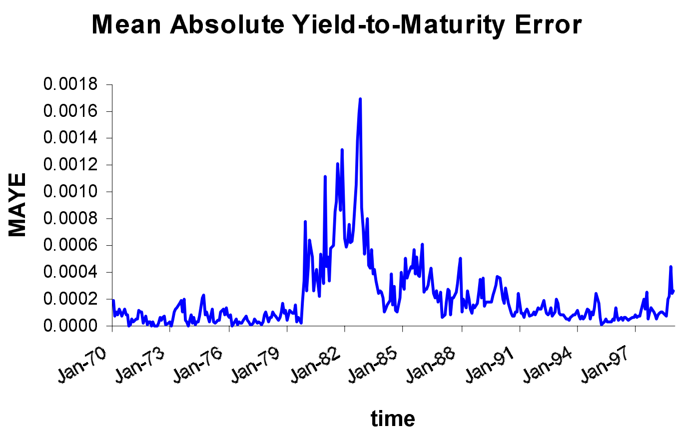
where $proxy(t)$ and $r(t)$ are the values of the proxy of the short rate and the short rate from the Cox-Ingersoll-Ross model respectively.

Proxy	Proxy Error Mean	Proxy Error Standard Deviation
LMNT Based Estimate	0.000243	0.006120
One Month T-Bill	0.000465	0.002178
Three Month T-bill	0.000069	0.001320

FIGURE 1
Time Series of Overall LMNT Out-of-Sample Performance Measures







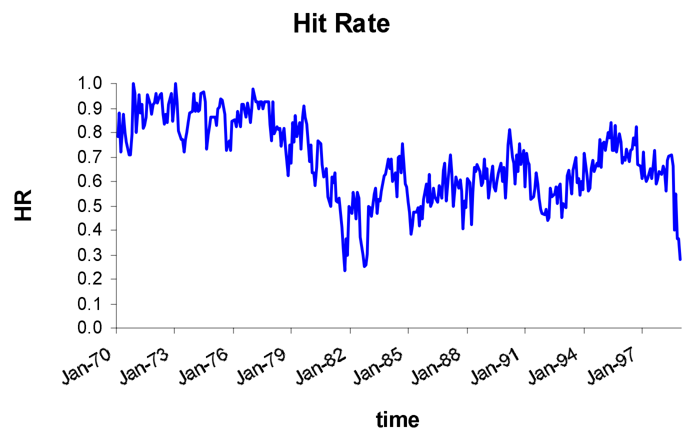


FIGURE 2
Nonparametric Drift and Diffusion Estimates of the Short rate Process Using Various Short Rate Proxies