THE JOURNAL OF FINANCE • VOL. LIV, NO. 3 • JUNE 1999

# Ex Ante Bond Returns and the Liquidity Preference Hypothesis

JACOB BOUDOUKH, MATTHEW RICHARDSON, TOM SMITH, and ROBERT F. WHITELAW*

**ABSTRACT**

We provide a formal test of the liquidity preference hypothesis (LPH), that is, the monotonicity of ex ante term premiums, using nonparametric estimates that do not require a structural model for conditional expected returns. Although the point estimates of the term premiums are consistent with previous conclusions in the literature regarding violations of the LPH, the test statistics are generally insignificant, even when powerful conditioning information is used. These results illustrate the importance of correctly accounting for correlations across maturities and of formally testing the inequality restrictions implied by the LPH.

THERE IS A VAST LITERATURE on bond returns and how they vary across the maturity spectrum. One of the more prominent theories is the liquidity preference hypothesis (LPH) (see, e.g., Hicks (1946) and Kessel (1965)), which states that the ex ante return on government securities is a monotonically increasing function of time to maturity. That is, conditional on all available information, the expected monthly return on a T-bill with one year to maturity should exceed the expected monthly return on a six-month T-bill, which should be greater than the certain yield on a one-month T-bill, and so forth. The LPH implies this condition, regardless of the shape of the term structure or any other economic variables contained in the agent's information set. The underlying intuition behind the LPH is that longer term bonds are *riskier*; that is, they are more sensitive to interest rate changes than shorter term bonds. Individuals need to be compensated for the risk of holding these bonds, hence the higher expected return.

Although the LPH is not a necessary condition of bond market equilibrium, it is consistent with a variety of term structure models. For example, in an economic environment in which future production possibilities are independent of the current economic state, Benninga and Protopapadakis (1986)

* Boudoukh and Richardson are from the Stern School of Business, New York University and the NBER; Smith is from the Australian Graduate School of Management, University of New South Wales; and Whitelaw is from the Stern School of Business, New York University. We thank an anonymous referee, the editor, René Stulz, Dong-Hyun Ahn, Rob Stambaugh, and seminar participants at the University of Iowa and the Western Finance Association meetings for helpful comments and suggestions. Support from the Geewax-Terker program in financial instruments is much appreciated.

examine a general equilibrium economy for quite general specifications of utility and production functions. They find that in complete markets the LPH's main conclusions are valid. The intuition is that longer bonds are a poor hedge for the representative agent's shorter term consumption, thus requiring a premium to hold these bonds over shorter horizons. Alternatively, in the Cox, Ingersoll, and Ross (1985) one-factor model of interest rates, term premiums on instantaneous holding periods are monotonic in the maturity of the bond. These results do not necessarily carry through to more complex economic settings; nevertheless, the LPH is one of the older models in finance and has been the object of numerous empirical studies.

Using post-1963 data on T-bills, direct tests of the unconditional version of the liquidity preference hypothesis have been performed by Fama (1984), McCulloch (1987), and Richardson, Richardson, and Smith (1992). Though there is some disagreement concerning the reliability of the data in the 1964–1972 period, the evidence suggests that expected returns are monotonic in maturity (throughout, monotonic is taken to mean monotonically increasing). For example, Richardson et al. (1992) find that, when one correctly tests for monotonicity using inequality constraints, there is little evidence against the LPH.

This evidence, however, comprises only unconditional tests of the theory. These tests are expected to have low power because the econometrician is ignoring the information available to economic agents. The LPH relates conditional expected returns across maturities; thus, unconditional tests provide very weak tests of the underlying theory. In fact, asset pricing theory suggests that the current term structure contains important information for expected returns on bonds of different maturities. For example, suppose that expected returns are monotonically increasing in maturity when the term structure is upward sloping, yet decreasing in maturity when the term structure is downward sloping. Since upward sloping term structures occur more often, unconditional tests will not be able to reject the monotonicity of returns because the tests average over all term structure shapes.

A strand of the literature recognizes this problem, and documents time-varying expected returns on bonds (e.g., Fama (1986), Fama and Bliss (1987), Stambaugh (1988), Fama and French (1989), and Klemkosky and Pilotte (1992)). All of these papers suggest that the *fitted* values of ex ante bond returns are not always increasing with the maturity of the bond. In order to correctly interpret these results, however, it is necessary to consider the joint statistical properties of these estimates of ex ante bond returns across maturities. Consequently, it may not be surprising that no formal test of the LPH (using conditioning information) has been performed. The difficulty is that the LPH implies a set of inequality restrictions on the ex ante returns on bonds of different maturities. Since these ex ante returns are unobservable, and statistical methods for testing inequality restrictions have only recently been developed, only anecdotal evidence regarding the LPH appears in the finance literature.

This paper adds to the current literature by performing formal ex ante tests of the LPH. Using information contained in the yield curve, we estimate conditional mean returns of bonds of different maturities. These means are then compared cross-sectionally using recently developed techniques in the inequality constraints econometric literature (see, e.g., Wolak (1989)). Contrary to current thinking in this area, our results provide little evidence to support previous conclusions regarding violations of the LPH. For example, though the term structure of ex ante bond returns is estimated to be downward sloping in some environments, these estimates cannot be differentiated from term structures generated from the LPH. This statistical result shows the importance of correctly accounting for correlations across maturities and of formally testing the inequality restrictions implied by the LPH.

The paper is organized as follows. Section I presents the methodology for testing the monotonicity of term premiums in a conditional setting. Section II provides two applications of this methodology. The first application reinvestigates Fama's (1986) seminal work on time-varying term premiums and illustrates several benefits from the inequality testing procedure. The second application looks at the LPH directly using a cross section of short-term and long-term bond returns. Section III provides some concluding remarks.

## I. Conditional Tests of the Maturity Structure of Term Premiums

Over the years, a substantial empirical literature has developed on the topic of the term structure of term premiums (see Campbell and Shiller (1991), Engle and Ng (1993), Fama (1984, 1986), Fama and Bliss (1987), Klemkosky and Pilotte (1992), Shiller, Campbell and Schoenholtz (1983), and Stambaugh (1988), among many others). However, there have been few formal tests of term structure hypotheses due to the inequality restrictions that these hypotheses imply and the difficulty in incorporating conditioning information.

Below, we show how to perform formal tests of hypotheses related to the term structure of term premiums (e.g., the LPH). Of particular note, these tests incorporate conditioning information that does not require a structural model of ex ante bond returns. However, for the test to be powerful, the conditioning set must provide useful information about alternative theories—that is, states in which these hypotheses may not be valid. To motivate our choice of conditioning information in the empirical work that follows, recall that the slope of the term structure can be decomposed into two components: expected changes in future interest rates and risk premiums (e.g., Shiller et al. (1983) and Engle and Ng (1993)). If future expected short-term rates are equal to the current short-term rate, then the sign of the risk premium coincides with the slope of the term structure. More generally, as long as variations in future short-term rates do not completely explain variations in the yield spread, this spread will contain information about the risk premiums. Existing evidence suggests that the relation between short rates and yield spreads is positive, yet far from one-to-one (e.g., see Fama and Bliss (1987) and Campbell and Shiller (1991)). Thus, the shape of the yield curve may produce a conditioning set that will provide a high hurdle for tests of term structure patterns of premiums.

Define $r_{t,t+j}(\tau)$ as the log return on a $\tau$-period bond that is purchased at time $t$ and held for $j$ periods. Previous research has investigated (albeit informally) the relation between ex ante bond returns, $E_t[r_{t,t+j}(\tau)]$, of various holding periods $j$ and maturities $\tau$. That is, researchers have been interested in hypotheses of the following form:

$$E_t\begin{pmatrix} r_{t,t+j_2}(\tau_2) - r_{t,t+j_1}(\tau_1) \\ r_{t,t+j_3}(\tau_3) - r_{t,t+j_2}(\tau_2) \\ \cdots \\ r_{t,t+j_n}(\tau_n) - r_{t,t+j_{n-1}}(\tau_{n-1}) \end{pmatrix} \geq 0, \tag{1}$$

where $j_k$ and $\tau_k$ ($k = 1,\ldots,n$) refer to a set of (not necessarily different) holding periods and bond maturities, respectively.

There are several issues in developing formal tests of equation (1). First, the term premiums, $E_t[r_{t,t+j_k}(\tau_k) - r_{t,t+j_{k-1}}(\tau_{k-1})]$, are unobservable. In theory, the researcher could posit a model for these term premiums, but then the tests would become a joint hypothesis of equation (1) and the model. Since equation (1) most probably forms much weaker restrictions than the model, this approach is subject to substantial amounts of type I error. Second, even if the unobservability issue can be dealt with in a consistent fashion, equation (1) imposes a vector of inequality restrictions that need to be tested. The statistics required for inequality constraints are different from the more usual equality restrictions tested elsewhere. Fortunately, a new literature has emerged that simplifies the problem dramatically (see Wolak (1989)). In particular, Boudoukh, Richardson, and Smith (1993) show how to extend the inequality testing methodology to the use of conditioning information. Since the method is conceptually straightforward, we provide a brief outline for our particular application.

As mentioned above, existing theory *suggests* that expected bond returns may move with the shape of the term structure. As an illustration, let us condition on monotonic and nonmonotonic yield curves. To generate testable restrictions implied by equation (1) using information in the term structure, first define

$$I_t = \begin{cases} 1 & \text{if the term structure is inverted or humped} \\ 0 & \text{if the term structure is monotonically upward sloping.} \end{cases} \tag{2}$$

For normalization purposes, we define the instrument $z_t$ as $z_t \equiv I_t/E[I_t]$, so that $E[z_t] = 1$. Because $z_t$ is a nonnegative random variable, equation (1) can be rewritten as

$$E_t \begin{pmatrix} r_{t,t+j_2}(\tau_2) - r_{t,t+j_1}(\tau_1) \\ r_{t,t+j_3}(\tau_3) - r_{t,t+j_2}(\tau_2) \\ \cdots \\ r_{t,t+j_n}(\tau_n) - r_{t,t+j_{n-1}}(\tau_{n-1}) \end{pmatrix} \times z_t \geq 0. \tag{3}$$

Rearranging equation (3) and applying the law of iterated expectations,

$$E \begin{pmatrix} [r_{t,t+j_2}(\tau_2) - r_{t,t+j_1}(\tau_1)]z_t - \theta_1 \\ [r_{t,t+j_3}(\tau_3) - r_{t,t+j_2}(\tau_2)]z_t - \theta_2 \\ \cdots \\ [r_{t,t+j_n}(\tau_n) - r_{t,t+j_{n-1}}(\tau_{n-1})]z_t - \theta_{n-1} \end{pmatrix} = 0, \tag{4}$$

where under the null hypothesis the vector $\theta \equiv (\theta_1, \theta_2, \ldots, \theta_{n-1}) \geq 0$.

To test the hypothesis $\theta \geq 0$, we first estimate $\theta$ as the sample mean of the term premiums, conditional on $z_t$:

$$\hat{\theta} = \frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} r_{t,t+j_2}(\tau_2) - r_{t,t+j_1}(\tau_1) \\ r_{t,t+j_3}(\tau_3) - r_{t,t+j_2}(\tau_2) \\ \cdots \\ r_{t,t+j_n}(\tau_n) - r_{t,t+j_{n-1}}(\tau_{n-1}) \end{pmatrix} \times z_t. \tag{5}$$

Equation (5) provides a set of moment conditions that identify the vector $\theta$ in terms of observables—the ex post returns on bonds and the shape of the term structure. These moments have a particular interpretation due to the normalization of $z_t$. Specifically, the vector $\hat{\theta}$ equals the average term premiums, *conditional on nonmonotonic term structures.*

The next step is to estimate the same mean, but now under the restriction that it must be nonnegative. Denote this restricted estimator $\hat{\theta}^R$. A natural test statistic of the restriction, $\theta \geq 0$, is to compare the vector of unrestricted conditional means to the vector of restricted conditional means. One way to do this is to apply a multivariate one-sided *Wald* statistic; that is,

$$W \equiv T\,(\hat{\theta}^R - \hat{\theta})'\hat{\Omega}^{-1}(\hat{\theta}^R - \hat{\theta}), \tag{6}$$

where $\hat{\Omega}^{-1}$ is the sample covariance matrix of the conditional term premiums.[1] The statistic can then be evaluated at some appropriate level of significance using its asymptotic distribution,

$$\sum_{k=0}^{N} Pr[\chi_k^2 \geq c] w\left(N, N-k, \frac{\hat{\Omega}}{T}\right), \tag{7}$$

where $c \in R^+$ is the critical value for a given size, $N$ is the number of restrictions, and the weight $w(N, N-k, (\hat{\Omega}/T))$ is the probability that $\hat{\theta}^R$ has exactly $N-k$ positive elements.

The description so far conditions on whether a state occurs or does not occur, and may ignore other relevant information. For example, it may be the case that term premiums are negative only in periods of sharply inverted term structures. Thus, it may be important to put more weight on these periods in the empirical analysis. As an illustration, suppose we want to condition not only on downward sloping term structures, but also on the magnitude of the slope. In this case, we choose $I_t^*$ such that, if the term structure is downward sloping, it equals the difference between the short-term and long-term rates. Using these "informative" instruments, equation (5) still provides a set of moment conditions that identify the vector $\theta$ in terms of observables—the ex post return on bonds, $r_{t,t+j}(\tau)$, and now both the shape and *magnitude* of the slope of the term structure, $z_t^*$. The vector of parameters $\theta$ has a new economic interpretation; it now equals the weighted average term premiums, where the weights correspond to the steepness of the yield curve (adjusted by the probability of such events).

## II. Empirical Tests

### *A. Do Term Premiums Change with Maturity? A Reinvestigation of Fama (1986)*

A substantive literature has developed that looks at the monotonicity of ex ante bond returns in a conditional setting. For example, in a seminal piece, Fama (1986) documents time-varying movements in term premiums that depend on the business cycle. He states that

> term premiums are generally interpreted as rewards for risk. In this view, the changes from upward sloping term structures of expected returns during good times to humped and inverted term structures of ex-

[1] The sample covariance matrix can be constructed to take account of serial correlation in the data, as we have here, due to month-to-month correlation in term structure shapes. We employ the procedure suggested by Andrews (1991) with a quadratic spectral kernel and a bandwidth determined by fitting an AR(1) model for each element of $\hat{\theta}$.

pected returns during recessions imply that the ordering of risks and rewards across maturities changes with the business cycle and is not always monotonic. (p. 176)

Fama (1986) finds that the term structure of short-term premiums can be humped, inverted, or decreasing in maturity depending on the stage of the business cycle. Moreover, these stages tend to coincide with humped, inverted, or downward sloping forward-rate term structures. Though this evidence is not a direct contradiction of the standard form of the LPH due to Fama's (1986) choice of maturities and holding periods, it is worthwhile to provide a formal test of these relations as an illustration both of the inequality testing methodology and of the importance of conditioning information.

We use data from the Fama bond files for 1- to 11-month bills (in yearly percentage) over the period November 1971 to July 1984.[2] Both the period and the maturities are chosen to coincide with Fama (1986). We perform two tests: (i) a formal test of Fama's (1986) analysis given in his Table 1 that documents average term premiums in different term structure environments, and (ii) a test of the analysis given in his Table 2 that relates term premiums over particular maturities to forward rates at these same maturities.

In both cases (i) and (ii), Fama (1986) looks at four annualized holding period log returns on bills: $r_{t,t+1}(1)$, $r_{t,t+2}(3)$, $r_{t,t+3}(6)$, and $r_{t,t+6}(12)$.[3] Under the usual form of the LPH, the conditional expectation of the annualized return ($E_t[r_{t,t+j}(\tau)]$) should be monotonic in maturity ($\tau$), fixing the holding period ($j$). However, Fama allows both the holding period and the maturity to vary, and our analysis replicates this feature.

In terms of the methodology described in Section I, the corresponding conditions on these term premiums can be written as

$$E_t\begin{pmatrix} r_{t,t+2}(3) - r_{t,t+1}(1) \\ r_{t,t+3}(6) - r_{t,t+2}(3) \\ r_{t,t+5}(11) - r_{t,t+3}(6) \end{pmatrix} \geq 0. \tag{8}$$

To coincide with Fama (1986), we choose two different sets of instruments, both based on forward rates corresponding to the holding period returns. Specifically, we use $f_t(0,1)$, $f_t(1,3)$, $f_t(3,6)$, and $f_t(6,11)$, where $f_t(\tau_1,\tau_2)$ is the annualized forward rate at time $t$ between maturities $t + \tau_1$ and $t + \tau_2$. Fama

[2] We do not include the bill with 12 months to maturity for two reasons. First, there are a substantial number of missing observations for this maturity. Second, the 12-month bill is actually defined in the data to be bills of at least 11 months and 10 days. As such, we consider its definition too unreliable for our analysis.

[3] Recall that our estimation uses data from the Fama files for 1- to 11-month bills (in yearly percentage), so we replace $r_{t,t+6}(12)$ with $r_{t,t+5}(11)$.

**Table I**

**Replication of Fama (1986)**

Panels A and B provide tests of whether particular expected returns are increasing cross-sectionally, motivated by results in Fama (1986, Tables 1 and 2). The data are collected from the Fama files for 1- to 11-month bills (percentage annualized) over the period 1971–1984. Each panel provides the average term premiums ($\hat{\theta}$) and corresponding standard errors (s.e.), conditional on either nonmonotonic forward rates or corresponding declining forward rates over a particular maturity (i.e., case (i) and case (ii), respectively, as described in Section II). $r_{t,t+j}(\tau)$ is the return on a $\tau$-period bond that is purchased at time $t$ and held for $j$ periods. The zero/one and magnitude-based columns denote the type of conditioning information used. The table also provides one-sided multivariate Wald test statistics (W) of the hypothesis that the term premiums are positive, with the corresponding $p$-values. The tests for each instrument are performed jointly across the maturities. The statistic's $p$-value is calculated using a Monte Carlo simulation.

| | Zero/One | | Magnitude-Based | |
|---|---|---|---|---|
| Term Premium | $\hat{\theta}$ | (s.e.) | $\hat{\theta}$ | (s.e.) |
| Panel A: Case (i), Nonmonotonicity of Forward Rates | | | | |
| $E_t[r_{t,t+2}(3) - r_{t,t+1}(1)]$ | 0.568 | (0.205) | 0.530 | (0.360) |
| $E_t[r_{t,t+3}(6) - r_{t,t+2}(3)]$ | 0.170 | (0.277) | −0.181 | (0.356) |
| $E_t[r_{t,t+6}(12) - r_{t,t+3}(6)]$ | −0.614 | (0.595) | −1.558 | (1.345) |
| W ($p$-value) | 1.066 | (0.382) | 2.170 | (0.301) |
| Panel B: Case (ii), Declining Forward Premium (per Maturity) | | | | |
| $E_t[r_{t,t+2}(3) - r_{t,t+1}(1)]$ | −0.345 | (0.206) | −0.675 | (0.518) |
| $E_t[r_{t,t+3}(6) - r_{t,t+2}(3)]$ | −0.659 | (0.290) | −1.260 | (0.602) |
| $E_t[r_{t,t+6}(12) - r_{t,t+3}(6)]$ | −0.761 | (0.693) | −2.895 | (2.102) |
| W ($p$-value) | 8.388 | (0.011) | 8.426 | (0.012) |

chooses these instruments in order to capture business cycle effects as described by the term structure of interest rates. For case (i), we consider all periods in which the forward-rate term structure is nonmonotonic; that is, periods in which the condition $f_t(0,1) \leq f_t(1,3) \leq f_t(3,6) \leq f_t(6,11)$ is violated. We also perform tests using more "informative" instruments by conditioning on the largest spread between the forward rates as long as it is negative. For case (ii), we choose a different instrument for each moment condition. This second case investigates term premiums over a particular maturity and holding period that exactly coincide with the nonmonotonicity of the forward-rate term structure at the corresponding maturity. For example, we look at the premium $r_{t,t+3}(6) - r_{t,t+2}(3)$ only in periods in which $f_t(3,6) < f_t(1,3)$. Again, tests are also performed for instruments that condition on the magnitude of the difference in forward rates.

Table I, Panel A, reports results for a conditional test of the monotonicity of Fama's (1986) term premiums based on case (i). With four holding period returns and one instrument, the system imposes three inequality restrictions. The term premiums are negative only at the longer end of the yield curve—that is, −0.614 percent (annualized) for the difference between the returns on 11-month and 6-month bills. When we condition on the magnitude of the nonmonotonicity present, this difference further declines to −1.558 percent, but the standard error also increases proportionately. The one-sided joint test statistics for conditional monotonicity of Fama's (1986) holding period returns equal 1.066 (with a $p$-value of 0.382) and 2.170 (with a $p$-value of 0.301), respectively.[4] These tests illustrate that, even though some premiums are individually negative, it is important to perform joint tests in a cross-sectional analysis across maturities.

Based on his Table 1 (and other evidence), Fama (1986) concludes that the forward-rate term structure coincides with the term structure of term premiums. The results above provide a sharp contrast. Although the individual point estimates suggest Fama's result, the statistical tests provide much weaker evidence. The reasons are threefold. First, the test takes into consideration the joint nature of the hypothesis and, in particular, the high cross-correlation patterns across the premiums. Second, autocorrelation in the data induced by serial correlation in the forward-rate term structures through time is explicitly accounted for. Third, the test is formal and therefore adjusts for the special distribution of the statistic under the null.

Of course, the sample size of Fama's (1986) study is small, and this may explain why monotonicity is not rejected. However, it is still inappropriate to consider the term premium estimates individually given their joint correlation properties across maturities. At the very least, these results show the different types of conclusions that can be reached by using tests for inequality restrictions. In particular, the apparent nonmonotonicities in the data are consistent with sampling error.

The low significance values suggest a potential lack of power. One way to address this issue is to partition the information into finer elements, as in our description of case (ii) above. Each term premium is associated with its corresponding forward-rate spread, so that we condition on states in which the forward rate is declining only at the particular maturity. Table I, Panel B, provides tests of the restrictions in case (ii). Using the maturity-specific instruments, the annualized differences in expected returns are all negative with values −0.345 percent, −0.659 percent, and −0.761 percent.

The appropriate multivariate one-sided test statistics for the average premiums are 8.388 (for the 0/1-based instruments) and 8.426 (for the informative instruments), which represent $p$-values of 0.011 and 0.012, respectively. Similar to Fama (1986, Table 2), and in contrast to our initial tests above, there is strong evidence that the expected returns time-vary and that they are nonmonotonic. These states are related to periods in which the term

[4] The $p$-value has a slightly different interpretation than under tests of equality constraints. Here, we calculate the distribution of the one-sided Wald test statistic for the least favorable value of the null hypothesis and thus of any size test. This can, but does not necessarily, lead to complications in determining the least favorable value of the null if the variance-covariance matrix of the estimators depends on the term premiums themselves (see Boudoukh et al. (1993) and Wolak (1989)). In any event, the test can always be interpreted locally.

structure of forward rates is nonmonotonic over maturities corresponding to the maturities of the bills used to calculate the term premiums. The contrast between these results and those in case (i) above further emphasizes the relation between conditioning information and the power of tests.

### *B. Tests of the LPH*

Two important motivations for this paper are (i) to explore how conditioning information impacts tests of the LPH, and (ii) to interpret these LPH tests cross-sectionally over various maturities. In particular, though some previous research has analyzed expected bond returns over both short- and long-maturity U.S. Treasury bills and notes (see Fama (1984), Fama and Bliss (1987), Fama and French (1989), and Klemkosky and Pilotte (1992)), the majority of work has focused on shorter term maturities. To the extent that both theory and empirical work comment on real rates of interest, longer term maturities are especially important because they correspond to the time horizons associated with business cycle fluctuations. In this section, we document and test properties of short-horizon holding period returns across the entire maturity spectrum in the context of the LPH.

Define the term premium for a bond with maturity $\tau$ as the one-period return in excess of the yield on a one-period bond; that is,

$$P_{\tau,t+1} \equiv r_{t,t+1}(\tau) - r_{t,t+1}(1). \tag{9}$$

The liquidity preference hypothesis implies that expected term premiums increase with the maturity of the bond ($\tau$); that is,

$$E_t[P_{\tau_k,t+1}] \geq E_t[P_{\tau_{k-1},t+1}] \geq \cdots \geq E_t[P_{\tau_1,t+1}],\ \tau_i > \tau_{i-1}. \tag{10}$$

The model in equation (10) implies that, conditional on all information available to the market at time $t$, expected returns are larger for longer maturity bonds. Of particular interest, the available information contains the entire term structure and thus the market's expectations about future rates. Using the methodology described in Section I, we can generate a vector of moment conditions with testable restrictions:

$$E[(P_{\tau_i,t+1} - P_{\tau_{i-1},t+1}) \times z_t - \theta] = 0, \tag{11}$$

where under the null model of the LPH, $\theta \geq 0$.

From the Fama bond files, we collect data on one-month T-bills and monthly holding period returns on 2–120+ month T-bills, T-notes, and T-bonds. We then form six equally weighted portfolios of 2–6 month bills, 7–12 month bills, 12–36 month notes, 36–60 month notes, 60–120 month notes, and

120+ month notes. The maturities and the monthly horizon cover those looked at by Klemkosky and Pilotte (1992).[5] We investigate the sample period January 1972 to December 1994.

To be consistent with the discussion in Section I, we choose a set of instruments to enhance the power of the inequality testing methodology by focusing on states of the economy which are likely to be the least supportive of the LPH. In particular, data on one- to six-month bills and one- to five-year spot rates are available from the Fama bond files over the sample period. We define a nonmonotonic yield curve as one in which the yield on one of the six maturities used (1–6, 12, 24, 36, 48, and 60 months) is lower than the yield on the shorter nearby maturity bond (where the shortest yield, 1–6 months, is the average yield on one- through six-month Treasuries). In addition to conditioning on the shape of the yield curve, we provide more information by conditioning on its magnitude. That is, if the yield curve is nonmonotonic in maturity, we let $I_t^*$ equal the maximum of the difference between any of the above six maturities (as long as they are negative). The idea is that extreme declines in the yield curve are associated with large drops in forward rates, and thus corresponding falls in term premiums (Fama (1986)).

One of the problems with conditioning on nonmonotonicity over the entire yield curve is that potentially important information is thrown away. For example, it may be that term premiums on shorter maturity instruments are declining only for downward sloping term structures at the short-end of the yield curve. To examine this possibility, we condition on periods in which the short-end lies above the long-end of the yield curve.

In order to produce a clearer picture of the conditioning information, Table II provides summary statistics on the instruments. In particular, we provide the in-sample probability of each state, and the state's serial dependence properties (its transition probabilities and corresponding autocorrelation). Over the sample period, a substantial fraction of the periods are captured by nonmonotonic states. The unconditional probability of the yield curve being nonmonotonic is 0.366, and more than 50 percent of these states (i.e., 0.185) involve a negative spread between the short-end and long-end of the yield curve. Note, though, that once a state occurs, such as monotonicity or nonmonotonicity, the probability of remaining in that state from month to month is very high—for example, 0.908 for nonmonotonicity. Thus, if term premiums time-vary depending on the particular state, these premiums will be autocorrelated. Hence, from a statistical viewpoint, it may be important to adjust the distribution of conditional bond returns for serial dependence.

[5] We choose a priori to form portfolios of these bonds over the different maturities to avoid reducing the power of the inequality testing methodology. For example, if we break the portfolio of 36–60 month notes into three bonds (as in Klemkosky and Pilotte (1992)) of 36, 48, and 60 months and the returns on these bonds are highly cross-correlated, then it is likely that the benefit of the added information does not offset the increase in the degrees of freedom of the weighted $\chi^2$ test.

**Table II**

**Statistical Properties of the Instruments**

Summary statistics are provided for different states of the world over the sample period January 1972 to December 1994. These instruments are defined as follows: $z_{1t} = 1$ when the yield curve is monotonic, $z_{2t} = 1$ when the yield curve is nonmonotonic, and $z_{3t} = 1$ when the yield curve is downward sloping. The column $\text{acr}(z_{it})$ is the autocorrelation of the instrument, $q$ is the probability that $z_{it+1} = 1$ given that $z_{it} = 1$, and $p$ is the probability that $z_{it+1} = 0$ given that $z_{it} = 0$.

| Instrument | $Pr(z_{it} = 1)$ | $\text{acr}(z_{it})$ | $q$ | $p$ |
|---|---|---|---|---|
| Monotonicity ($i = 1$) | 0.634 | 0.750 | 0.842 | 0.908 |
| Nonmonotonicity ($i = 2$) | 0.366 | 0.750 | 0.908 | 0.842 |
| Downward sloping ($i = 3$) | 0.185 | 0.783 | 0.960 | 0.824 |

In practice, although the serial-correlation-adjusted standard errors are slightly larger than those which do not adjust for serial dependence, the effect on the results is qualitatively insignificant.

Table III, Panel A, provides the differences in average term premiums across the maturities, conditional on nonmonotonic yield curves. Also given are the corresponding standard errors of these differences, adjusted for serial dependence and heteroskedasticity, and the corresponding one-sided Wald statistics. Although most of the individual differences in premiums for the nonmonotonic instrument are negative (ranging from 0.187 percent to $-0.315$ percent), their standard errors are fairly large. Nevertheless, it is still inappropriate to interpret these estimates individually and conclude that there is no evidence against the LPH. That is, the correlation structure across the term premium estimators can still matter. As it turns out, the one-sided test conditional on nonmonotonic yield curves has a Wald statistic of only 0.105. This represents a $p$-value of 0.513. In this particular case, the statistical result is consistent with the type of conclusion one might draw from the individual point estimates and standard errors.

We also break up nonmonotonic periods into states in which the short end of the yield curve lies above its long end. Table III, Panel B, provides the differences in average term premiums of the bond portfolios, the corresponding standard errors, and the one-sided Wald statistics for this case. The differences in premiums are always negative, with an average 1.63 percent difference between nearby bond portfolios and an average standard error of 1.34. Though some of the individual point estimates hover around 5 percent $p$-values for one-sided tests, the evidence is weaker in the appropriate joint test setting. The Wald statistic is 3.081 with a corresponding $p$-value of 0.092. In spite of the fact that the point estimates are all negative, the test statistic provides little evidence against the LPH at conventional levels.

The statistical evidence above is weaker than that implied by the current literature, and it is probably insufficient to overturn strong prior beliefs that the LPH is valid. The estimated premiums appear inconsistent with the theory, but the $p$-values indicate that estimation error might generate such

**Table III**

**Test of the Liquidity Preference Hypothesis across All Maturities: 1/72–11/94**

Panels A and B use data from the Fama bond files. One-month excess holding period returns (term premiums) are calculated for six equally weighted portfolios of 2–6 month bills, 7–12 month bills, 12–36 month notes, 36–60 month notes, 60–120 month notes, and 120+ month notes. The two panels report the average differences in the term premiums ($\hat{\theta}$) and the corresponding standard errors (s.e.). Panel A provides the differences in the average term premiums conditional on nonmonotonicity, and Panel B provides these differences conditional on whether the short end of the yield curve lies above the long end. The zero/one and magnitude-based columns denote the type of conditioning information used. Nonmonotonicity is defined as a state in which the yield curve is somewhere downward sloping. In Panel A, the corresponding magnitude-based instrument is defined as the maximum of the difference between any of the yields (as long as they are negative); in Panel B, the magnitude-based instrument is just the magnitude of the spread between the short rate and long rate. The table also provides one-sided multivariate Wald test statistics (W) of the hypothesis that the differences in these term premiums are positive, with the corresponding $p$-values. The tests for each instrument are performed jointly across the maturities. The statistic's $p$-value is calculated using a Monte Carlo simulation.

| | Zero/One | | Magnitude-Based | |
|---|---|---|---|---|
| Difference in Premiums | $\hat{\theta}$ | (s.e.) | $\hat{\theta}$ | (s.e.) |
| | Panel A: Nonmonotonicity | | | |
| $P_{7-12} - P_{2-6}$ | −0.100 | (0.649) | −0.094 | (0.738) |
| $P_{12-36} - P_{7-12}$ | 0.187 | (1.020) | 0.197 | (1.467) |
| $P_{36-60} - P_{12-36}$ | −0.115 | (0.924) | −1.228 | (1.246) |
| $P_{60-120} - P_{36-60}$ | −0.312 | (0.883) | −2.237 | (1.507) |
| $P_{120+} - P_{60-120}$ | −0.315 | (1.342) | −2.622 | (1.777) |
| W($p$-value) | 0.105 | (0.513) | 2.473 | (0.146) |
| | Panel B: Downward Sloping | | | |
| $P_{7-12} - P_{2-6}$ | −0.972 | (0.627) | −0.248 | (0.916) |
| $P_{12-36} - P_{7-12}$ | −1.307 | (0.949) | −0.189 | (1.340) |
| $P_{36-60} - P_{12-36}$ | −1.626 | (1.110) | −2.117 | (1.651) |
| $P_{60-120} - P_{36-60}$ | −1.545 | (1.086) | −2.744 | (1.776) |
| $P_{120+} - P_{60-120}$ | −2.774 | (1.647) | −3.201 | (2.166) |
| W($p$-value) | 3.081 | (0.092) | 2.555 | (0.133) |

results even if the LPH is correct. It is convenient to criticize these tests as having low power, but this criticism is not well founded. Inequality restrictions are generally weaker than the more standard equality restrictions; that is, by their nature, inequality constraints provide a higher hurdle. However, because the LPH implies a set of inequality restrictions with highly correlated variables, the real question is which inequality-constraints-based test has the most power. It is well known, for example, that Bonferroni-type procedures are generally weaker than the statistical techniques advocated here. Consequently, the appropriate interpretation of the results is not that the inequality tests have low power, but rather that the LPH has weak implications and that these implications are not inconsistent with the data.

As we discuss in Section II.A, a related power issue is that the above instruments ignore information about the magnitude of the nonmonotonicity of the yield curve. Table III provides extensions of the results to include instruments conditioned on the magnitude of the event. From Table III, Panel A, the joint one-sided test for the nonmonotonic case provides stronger (though not statistically significant) evidence against the null—the $W$-statistic equals 2.473, with a $p$-value 0.146. However, the differences in premiums increase substantially, with proportionate increases in the standard errors. Table III, Panel B, shows a similar pattern in terms of the magnitude of the premium differences; however, the increase in the estimation error associated with these weighted premiums actually leads to a reduction in the $p$-value (from 0.092 to 0.133). Although it is difficult to compare $p$-values across different test statistics, these results do point to the difficulty in measuring the means of short-horizon, long-term bond returns.

## III. Concluding Remarks

This paper provides a methodology for evaluating the pattern of ex ante bond returns across the maturity of those bonds. The novel aspect of this method is that it allows us to increase the power of the tests by looking at conditional data without a priori specifying a model for bond returns. Though the current application focuses primarily on the liquidity preference hypothesis, other applications can include investigations of the preferred habitat theory or even the term structure of volatilities. Of course, this will involve the researcher specifying relevant conditioning information up front. For example, to the extent that the preferred habitat theory is usually described in the context of price pressure at certain maturities, formal tests would require us to be able to specify the relevant periods of price pressure.

With respect to the empirical results documented in this paper, we show that, in a reinvestigation of Fama's (1986) work, it is very important to consider both the nature of the joint inequality constraints implied by the model and the type of conditioning information used in estimation. Motivated by these findings, we provide a formal test of the liquidity preference hypothesis across both short- and long-term maturities. In contrast to the negative findings documented elsewhere in the literature, our formal test results are generally consistent with the LPH.

## REFERENCES

Andrews, Donald W. K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* 59, 817–858.

Benninga, Simon, and Aris Protopapadakis, 1986, General equilibrium properties of the term structure of interest rates, *Journal of Financial Economics* 16, 389–410.

Boudoukh, Jacob, Matthew Richardson, and Tom Smith, 1993, Is the ex ante risk premium always positive? A new approach to testing conditional asset pricing models, *Journal of Financial Economics* 34, 387–408.

Campbell, John Y., and Robert J. Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495–514.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.

Engle, Robert F., and Victor Ng, 1993, Time-varying volatility and the dynamic behavior of the term structure, *Journal of Money, Credit and Banking* 25, 336–349.

Fama, Eugene F., 1984, Term premiums in bond returns, *Journal of Financial Economics* 13, 529–546.

Fama, Eugene F., 1986, Term premiums and default premiums in money markets, *Journal of Financial Economics* 17, 175–196.

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

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Hicks, John R., 1946, *Value and Capital* (Oxford University Press, London).

Kessel, Reuben A., 1965, The cyclical behavior of the term structure of interest rates, NBER occasional working paper no. 91.

Klemkosky, Robert C., and Eugene A. Pilotte, 1992, Time-varying term premiums on U.S. Treasury bills and bonds, *Journal of Monetary Economics* 30, 87–106.

McCulloch, J. Huston, 1987, The monotonicity of the term premium: A closer look, *Journal of Financial Economics* 18, 185–192.

Richardson, Matthew, Paul Richardson, and Tom Smith, 1992, The monotonicity of the term premium: Another look, *Journal of Financial Economics* 31, 97–106.

Shiller, Robert J., John Y. Campbell, and Kermit L. Schoenholtz, 1983, Forward rates and future policy: Interpreting the term structure of interest rates, *Brookings Papers on Economic Activity* 1, 173–217.

Stambaugh, Robert F., 1988, The information in forward rates: Implications for models of the term structure, *Journal of Financial Economics* 21, 41–70.

Wolak, Frank A., 1989, Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models, *Econometric Theory* 5, 1–35.