NEAR UNIT ROOTS AND THE PREDICTIVE POWER OF YIELD SPREADS FOR CHANGES IN LONG-TERM INTEREST RATES

Markku Lanne*

Abstract — The ability of yield spreads to predict changes in long-term interest rates implied by the expectations hypothesis is usually rejected. It is suggested that this rejection is often caused by high persistence in the spread when standard inference is employed. Instead, the asymptotically valid method of Cavanagh et al. (1995) is applied to monthly U.S. data from 1952:1–1991:2. The persistence of the spreads seems to have varied over time, and in subsample analysis, the expectations hypothesis cannot be rejected at the long end of the maturity spectrum.

I. Introduction

Empirical evidence on the expectations hypothesis of the term structure of interest rates is mixed. Most empirical studies reject the theory, but results seem to depend considerably on which methods are used, or to which sample period or maturities they are applied. In this paper we shall take a closer look at a common test based on measuring the predictive power of the yield spread. Under rational expectations, the theory implies that the current long-term interest rate should equal the expected average of current and future short-term rates. This suggests that the spread between the long-term and short-term rates should be able to predict future short-term changes in the yield on the longer-term bond. A common finding, especially with U.S. data, is that the expectations hypothesis is rejected for virtually all maturities in tests of this implication. (See, e.g., Shiller (1990) and references therein; more recent evidence is presented by Campbell and Shiller (1991), Hardouvelis (1994), and, for the short-term interest rates, Evans and Lewis (1994).)

It is suggested in this paper that, in some cases, the reason the expectations hypothesis is rejected in tests of the ability of the spread to predict changes in the longer-term rate is that standard inference is applied in a situation where a large autoregressive root is involved. It is a common finding that interest rate series have autoregressive roots close to unity. If the spreads between interest rates also have this property, it is possible that there is a problem of overrejection in the test of the expectations hypothesis mentioned above when critical values from the standard normal distribution are used. In all of the studies cited above, standard inference was used, and only Hardouvelis (1994) did pretests to find out the order of integration of the spread. However, as documented by Elliott (1994), Elliott and Stock (1994), and Cavanagh et al. (1995), size distortions in $t$ tests can be substantial when the explanatory variable has a large, possibly unit root, even if a pretest is used to decide between $I(1)$ and $I(0)$ and critical values from the subsequent distribution are employed. One solution, which controls size uniformly over the values of the largest autoregressive root, was recently presented by Cavanagh et al. (1995), whose approach is based on simultaneous confidence intervals. In this paper, one of their methods will be applied to testing for the expectations hypothesis with U.S. monthly data from the years 1952 to 1991.

The recent results of Hall et al. (1992), and Fuhrer (1996), inter alia, suggest that inference from the entire sample may not be appropriate because that period includes different monetary policy regimes. Our results also lend support to the conjecture that the persistence of yield spreads differs between the interest-rate targeting period 1952:1–1979:9 and the rest of our sample, 1979:10–1991:2. Therefore, the tests are conducted separately for the interest-rate targeting period and the latter period using an asymptotically valid procedure, and support is found for the expectations hypothesis.

The plan of the paper is as follows. The test of the expectations hypothesis is discussed in section II. The pitfalls of standard inference in the presence of a large autoregressive root of the spread are described and the method due to Cavanagh et al. is presented. Section III contains the empirical results with McCulloch and Kwon’s (1993) U.S. term structure data. Some Monte Carlo simulations are also presented to evaluate the small sample performance of the different testing procedures. Section IV concludes.

II. Testing the Expectations Hypothesis

According to the expectations hypothesis, the term structure of yields to maturity is explained by investors’ expectations about future interest rates. For yields on continuously compounded zero-coupon bonds, the expectations hypothesis can be expressed as (Campbell & Shiller, 1991)

$$ r^n_t = \frac{1}{n} \sum_{i=0}^{n-1} E_t r_{t+i}^1 + L_n, \quad (1) $$

where $r_{t+i}^1$ is the one period and $r^n_t$ the $n$-period interest rate.

Received for publication February 26, 1997. Revision accepted for publication September 21, 1998.

* Department of Economics, University of Helsinki.

I am indebted to Tom Berglund, Graham Elliott, Seppo Honkapohja, Erkki Koskela, Anne Mikkola, Pentti Saikkonen, Allan Timmermann, John Y. Campbell (co-editor), and the two anonymous referees for useful comments. The usual disclaimer applies. This paper is a part of the research program of the Research Unit on Economic Structures and Growth (RUESG) in the Department of Economics of the University of Helsinki. Financial support from the Yrjö Jahnsson Foundation and Säästöpankkien tutkimussäätiö is gratefully acknowledged.

© 1999 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology
\( L_n \) is a term premium dependent only on \( n \), and \( E_t \) denotes mathematical expectation conditional on public information at time \( t \).

Equation (1) states that the longer-term interest rate is an arithmetic average of the present and expected future one-period interest rates over the life of the longer-term bond plus a constant premium term \( L_n \). By rearranging terms in equation (1), the spread between the \( n \)-period and one-period rates can be written as a forecast of the future change in the yield on the longer-term bond. (See, e.g., Campbell and Shiller (1991).) This suggests a test on the slope coefficient in the following model

\[
(n - 1)(r_{t-1}^n - r_{t-1}^1) = \alpha + \gamma (r_{t-1}^n - r_{t-1}) + \epsilon_{2t}.
\]

Under the expectations hypothesis, \( \gamma \) should equal unity. A common finding, especially with U.S. data, is that the expectations hypothesis is almost always rejected by a \( t \) test of \( \gamma = 1 \) in equation (2).

Typically inference on \( \gamma \) in equation (2) has relied on standard asymptotic distribution theory, i.e., Gaussian critical values for the \( t \)-statistic. According to the recent results of Elliott (1994), Elliott and Stock (1994), and Cavanagh et al. (1995), however, standard inference may not be asymptotically valid here if the spread \( r_t^n - r_t^1 \) has a large autoregressive root. Specifically, Elliott and Stock considered a model in which the regressor (here the spread) is generated from an autoregressive model

\[
r_t^n - r_t^1 = \mu + \rho (r_{t-1}^n - r_{t-1}) + \epsilon_t, \quad t = 1, \ldots, T
\]

where \( b(L) \epsilon_t = \epsilon_{1t}, b(L) \) is a \( 4 \)-th order lag polynomial with \( b_0 = 1 \), and \( \epsilon_t = (\epsilon_{1t}, \epsilon_{2t})' \) is a martingale difference sequence with \( \text{sup} \mathbb{E} \epsilon_t^4 < \infty, i = 1, 2 \), and a constant conditional covariance matrix \( \Sigma \), and \( \text{E} \epsilon_{10} \leq \infty \).

If \( \rho = 1 \), then the spread is an \( I(1) \) variable; if \( \rho \) is fixed and \( |\rho| < 1 \), then it is \( I(0) \). If the order of integration of the spread is unknown, asymptotic distribution theory corresponding to these two cases may work poorly, and it is reasonable to apply local-to-unity asymptotics, i.e., parametrize \( \rho \) as \( \rho = 1 + c/T \) where \( c \) is a constant. Using this approach, Elliott and Stock reached two main conclusions. First, the asymptotic null distribution of the \( t \)-statistic on \( \gamma \) in equation (2), \( t_\gamma \), is nonstandard and depends on both the local-to-unity parameter \( c \) and the correlation between \( \epsilon_{1t} \) and \( \epsilon_{2t} \) (i.e., on the simultaneity in the system). They showed that overrejection is worse the higher the simultaneity. As \( c \) gets smaller (or, equivalently, \( c \) gets larger), the limiting distribution gets closer to the standard normal distribution. (In the limit \( c \to -\infty \), \( t_\gamma \) is distributed as a standard normal variable; See Cavanagh et al. (1995).) Second, the often-applied procedure of pretesting for a unit root in the spread and subsequently employing the ensuing distribution theory also leads to overrejection. This is because asymptotically the rejection probability of the null hypothesis \( \rho = 1 \) is less than one, meaning that the pretest is not consistent in the local-to-unity case. Hence, the nonstandard critical values corresponding to the \( I(1) \) case are too often selected for \( t_\gamma \) and overrejection prevails.

The size distortions can be substantial. The computations of Cavanagh et al. suggest that the size can be as high as 40% for a 5% nominal level test. These size distortions could be avoided if the local-to-unity parameter \( c \) could be consistently estimated, because then the asymptotic null distribution could be considered known in each case and the test would have correct size asymptotically. Unfortunately, this is not the case.3

It has been suggested that bootstrapping could be used to overcome the overrejection problem. (See, e.g., Bekar et al. (1997).) Recently, however, Stock (1997) pointed out that bootstrapping fails to provide an asymptotically valid method of inference in models like equation (2) when the order of integration of the regressor is unknown. This follows from the fact that bootstrapping depends on a consistent estimate of the local-to-unity parameter \( c \) that is not available. The asymptotic invalidity of the bootstrap has been proved in at least two closely related cases. Stock and Watson (1996) showed that the bootstrap does not control size in inference on the parameters of the cointegrating vector when the regressors do not necessarily have a unit root. Basawa et al. (1991), on the other hand, demonstrated that in the special case of the AR(1) model with a unit root, the bootstrap distribution of the slope coefficient converges to a random distribution, even if the error distribution is known to be normal.

In this paper, we employ the method based on Scheffe-type simultaneous confidence intervals presented by Cavanagh et al. to test the expectations hypothesis. The general idea of their method is to find all the values of \( \gamma \) that are consistent with any reasonable value of \( c \), then eliminating this nuisance parameter. A Scheffe-type confidence interval having at least \( 100(1 - \eta)/\% \) confidence level for \( \gamma \) is obtained by inverting a joint \( \eta \)-level Wald test for the null hypothesis \( \gamma = \gamma_0 \) and \( c = c_0 \), and projecting the resulting confidence region onto the \( \gamma \) axis. In other words, the

3 One might conjecture that a consistent estimator of \( c \) is provided by \( T(\hat{\rho} - 1) \), where \( \hat{\rho} \) is the OLS estimator of \( \rho \). To see that this is false, note that (see, e.g., Phillips (1987, Theorem 1))

\[
T(\hat{\rho} - 1) = T(\hat{\rho} - \rho) + T(\rho - 1) = T(\hat{\rho} - \rho) + c + \int_0^1 J_s(s) dW(s) - \int_0^1 J_s(s)^2 ds + dW(s) = T(0),
\]

where \( \Rightarrow \) denotes weak convergence, \( W(s) \) is a standard Brownian motion, \( J_s(s) \) is an Ornstein-Uhlenbeck diffusion process defined by the stochastic differential equation \( dJ_s(s) = cJ_s(s)ds + dW(s) \) with \( J(0) = 0 \). Thus, \( T(\hat{\rho} - 1) \) converges to a distribution that depends on the true value of \( c \) which is, of course, unknown. Therefore, the value that \( T(\hat{\rho} - 1) \) yields is just a random draw from the distribution above.
Scheffé-type interval consists of all the values of \( \gamma_0 \) for which there is some \( c_0 \) such that the joint null hypothesis is not rejected.

For the presentation of the test, it is convenient to write the autoregressive model (3) in the standard augmented Dickey-Fuller (ADF) form:

\[
\Delta(r^n_t - r^1_t) = \tilde{\mu}_t + \beta(r^n_{t-1} - r^1_{t-1}) + p(L)\Delta(r^n_{t-1} - r^1_{t-1}) + \varepsilon_{tn},
\]

where \( \tilde{\mu}_t = b(1)\mu_t, \beta = (\rho - 1)b(1), \) and \( p_t = \Sigma_{i=j+1}^{k} \tilde{p}_i \) with \( \tilde{p}(L) = L - (1 - (1 - \rho)L)b(L). \) The Wald test statistic for the joint hypothesis \( (\gamma, c) = (\gamma_0, c_0) \) considered by Cavanagh et al. is

\[
W = \frac{1}{2} \phi_T \left[ \hat{\Sigma}^{-1} T^{-2} \sum_{i=2}^{T} (r^n_{t-1} - r^1_{t-1})^2 \right] \phi_T,
\]

where

\[
\phi_T = \left[ \frac{\hat{f}(\hat{\gamma} - \gamma_0)}{T(\hat{\gamma} - \gamma_0)} \right],
\]

\( \hat{\Sigma} = 1/(T-1) \Sigma_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}_t' \), \( \hat{\epsilon}_t = (\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t})' \) are the OLS residuals from equations (4) and (2), \( (r^n_{t-1} - r^1_{t-1})^2 = (r^n_{t-1} - r^1_{t-1}) - 1/(T-1) \Sigma_{t=2}^{T} (r^n_{t-1} - r^1_{t-1}), \) and \( \hat{\beta}, \hat{\gamma}, \) and \( \hat{b}(1) = 1 - \Sigma_{i=1}^{k} \tilde{p}_i \) are the OLS estimators. The asymptotic null distribution of \( W \) can be obtained by simulation; some critical values are tabulated by Cavanagh et al. (table 2, panel A). A 100(1 - \( \eta \)% Scheffé-type confidence interval for \( \gamma \) can be formed by first estimating the ADF regression (4) for the spread and the actual regression model (2), and then finding all the values of \( \gamma_0 \) such that the null \( (\gamma, c) = (\gamma_0, c_0) \) is not rejected for some \( c_0 \) using the test statistic \( W \). The critical values depend on \( c_0, \) so that the interval cannot be expressed in closed form but must be solved numerically for different values of \( c_0. \)

The asymptotic null rejection rates computed by Cavanagh et al. indicate that the Scheffé procedure is rather conservative with rejection rates varying between approximately 3% and 7% for a test with nominal asymptotic size at most 10%. Small-sample Monte Carlo simulation results with data generated from a model corresponding to the actual interest rate data will be presented in section III.

### III. Empirical Results

In this section, model (2) will be studied with U.S. data.

More specifically, the data consists of the zero-coupon yield curves constructed by McCulloch (1990) and updated by McCulloch and Kwon (1993). For the most part, this is the same data used by Campbell and Shiller (1991). However, we consider the updated monthly data from period 1952:1–1991:2 while they used the data from McCulloch (1990), ending in 1987:2.

In each case, the short-term rate is taken to be the one-month rate and the long-term rates include maturities of 2, 3, 4, 6, 9, 12, 24, 36, 48, 60, and 120 months. In addition to the entire sample, two subsample periods are considered: 1952:1–1979:9 and 1979:10–1991:2. The total sample thus comprises 470 observations, while the subsamples have 333 and 137 observations, respectively. This division into subperiods reflects changes in the interest rate targeting behavior of the Federal Reserve System. In the first period, interest rates were targeted, whereas the latter period is not as homogeneous: from October, 1979, until September, 1982, the Fed ceased to target interest rates, and, from 1982 onwards, these so-called “new operating procedures” were gradually abandoned (Meulendyke, 1989). Because the period of the new operating procedures was very short (36 observations), further division of the latter subsample does not seem to be feasible. This subsample division has been adopted in several recent studies of the U.S. term structure; also, Campbell and Shiller (1991) examined separately the 1952–1978 period, but the results were comparable to those for the entire sample. Also, the recent results of Hall et al. (1992) and Fuhrer (1996), inter alia, suggest that inference from the full sample may not be appropriate.

First, the results using critical values from the normal distribution are presented. They, of course, accord with the previous literature. Unity is never included in the 95% confidence intervals in table 1 as far as the entire sample and the targeting period are concerned. In the latter subperiod, the expectations hypothesis cannot, in general, be rejected (with the exception of the cases of the two-, six-, and nine-month rates).

In order to see in which cases standard inference is likely to fail, the largest autoregressive roots of the spreads, \( \rho \), are estimated. Model (4) for each spread was estimated by OLS.

### Table 1. 95% Confidence Intervals for the Slope Coefficient in the Regression of \( \mu_t \) on \( r^n_t - r^1_t \)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-0.401, 0.405</td>
<td>-0.520, 0.059</td>
<td>-0.193, 0.897</td>
</tr>
<tr>
<td>3</td>
<td>-0.845, 0.555</td>
<td>-0.818, -0.083</td>
<td>-0.937, 1.346</td>
</tr>
<tr>
<td>4</td>
<td>-1.462, 0.771</td>
<td>-0.993, -0.031</td>
<td>-1.763, 1.537</td>
</tr>
<tr>
<td>6</td>
<td>-1.921, 0.252</td>
<td>-1.143, 0.075</td>
<td>-2.585, 0.857</td>
</tr>
<tr>
<td>9</td>
<td>-2.175, 0.341</td>
<td>-1.175, 0.302</td>
<td>-2.890, 0.829</td>
</tr>
<tr>
<td>12</td>
<td>-2.591, 0.528</td>
<td>-1.596, 0.280</td>
<td>-3.042, 1.021</td>
</tr>
<tr>
<td>24</td>
<td>-3.647, 0.754</td>
<td>-2.443, 0.528</td>
<td>-4.425, 1.922</td>
</tr>
<tr>
<td>36</td>
<td>-4.403, 0.630</td>
<td>-2.880, 0.643</td>
<td>-5.517, 2.078</td>
</tr>
<tr>
<td>48</td>
<td>-5.042, 0.527</td>
<td>-3.088, 0.732</td>
<td>-6.439, 2.012</td>
</tr>
<tr>
<td>60</td>
<td>-5.613, 0.390</td>
<td>-3.353, 0.803</td>
<td>-7.206, 1.837</td>
</tr>
<tr>
<td>120</td>
<td>-8.282, -0.142</td>
<td>-4.984, 0.720</td>
<td>-11.115, 1.833</td>
</tr>
</tbody>
</table>

Note: For \( n \approx 9, r^{-1} \) is approximated by \( r_t^1 \).
and the estimates of the largest autoregressive root were obtained as \( \hat{\rho} = 1 + \hat{\beta}(1 - \hat{\rho}(1)) \), where \( \hat{\beta} \) and \( \hat{\rho}(1) \) are the OLS estimators from equation (4). The estimates of \( \rho \) and the implied values of \( c \) (computed as \( T(\hat{\rho} - 1) \)) are presented in table 2. Although these implied values are not consistent estimates of \( c \), they may still give some idea of whether a large autoregressive root is a problem. The estimates indicate clear differences both between the sample periods and across maturities. First, the estimates of the largest autoregressive roots of spreads involving maturities less than two years are small, and the implied values of the local-to-unity parameters large in absolute value, so that there does not seem to be any problem with applying standard asymptotic theory for the shortest-term rates. Second, especially for the spreads involving longer-term rates, the largest roots are estimated to be clearly larger in the targeting period than in the latter subsample. Consequently, the implied values of \( c \) are of comparable magnitude in these periods and, in general, close enough to zero to invalidate standard inference. In the entire sample, there does not seem to be a near unit root problem, so that standard inference is asymptotically valid assuming that the value of \( \rho \) really did not change. However, if the switch in monetary policy indeed induced a change in \( \rho \), then tests based on regression (2) using data from the entire sample are not sensible, because the distributions of the \( t \)-statistics are dependent on \( \rho \) via \( c \), and \( c \) is assumed to be constant in the distribution theory discussed in section II. In summary, we would expect overrejection to be a problem in tests using standard inference concerning the 36–120-month rates in the two subsamples, where the values of \( c \) implied by the estimates of \( \rho \) are relatively close to zero. Hence, the use of the methods presented in the previous section is likely to overturn the rejections in the targeting period in table 1 in these cases.

The 95% Scheffe-type confidence intervals of \( \gamma \) in equation (2) for the two subsample periods are presented in table 3 for maturities of at least two years. For the shorter-term rates, the computation of the confidence intervals is likely to be sensible, because the distributions of the OLS estimators from equation (4). The estimates of \( r_t \) obtained as

### Table 2.—Estimates of the Largest Autoregressive Root of the Spread \( \frac{r_t - r_{t-1}}{c} \)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\rho} )</td>
<td>( T(\hat{\rho} - 1) )</td>
<td>( T(\hat{\rho} - 1) )</td>
<td>( T(\hat{\rho} - 1) )</td>
<td>( T(\hat{\rho} - 1) )</td>
<td>( T(\hat{\rho} - 1) )</td>
</tr>
<tr>
<td>2</td>
<td>0.117</td>
<td>−415</td>
<td>0.270</td>
<td>−243</td>
<td>0.298</td>
</tr>
<tr>
<td>3</td>
<td>0.353</td>
<td>−304</td>
<td>0.410</td>
<td>−196</td>
<td>0.387</td>
</tr>
<tr>
<td>4</td>
<td>0.451</td>
<td>−258</td>
<td>0.476</td>
<td>−174</td>
<td>0.455</td>
</tr>
<tr>
<td>6</td>
<td>0.572</td>
<td>−201</td>
<td>0.623</td>
<td>−125</td>
<td>0.556</td>
</tr>
<tr>
<td>9</td>
<td>0.639</td>
<td>−170</td>
<td>0.669</td>
<td>−110</td>
<td>0.618</td>
</tr>
<tr>
<td>12</td>
<td>0.706</td>
<td>−138</td>
<td>0.702</td>
<td>−99</td>
<td>0.678</td>
</tr>
<tr>
<td>24</td>
<td>0.802</td>
<td>−93</td>
<td>0.823</td>
<td>−59</td>
<td>0.760</td>
</tr>
<tr>
<td>36</td>
<td>0.837</td>
<td>−77</td>
<td>0.867</td>
<td>−44</td>
<td>0.794</td>
</tr>
<tr>
<td>48</td>
<td>0.855</td>
<td>−68</td>
<td>0.892</td>
<td>−36</td>
<td>0.760</td>
</tr>
<tr>
<td>60</td>
<td>0.867</td>
<td>−63</td>
<td>0.905</td>
<td>−32</td>
<td>0.766</td>
</tr>
<tr>
<td>120</td>
<td>0.887</td>
<td>−53</td>
<td>0.921</td>
<td>−26</td>
<td>0.795</td>
</tr>
</tbody>
</table>

Notes: The estimates of \( \rho \) are computed as \( 1 + \beta(1 - \hat{\rho}(1)) \) from the ADF regression (4). \( T \) is the sample size. The columns marked by \( l \) give the lag length of the AIC. The lag length \( k \) in equation (4) was selected by the AIC.

### Table 3.—95% Scheffe-type Confidence Intervals for the Slope Coefficient in the Regression of \( (n - 1)(c^2 - c_{t-1}) \) on \( c_{t-1} - c_{t-1} \)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>( 24 )</td>
<td>( 36 )</td>
<td>( 48 )</td>
</tr>
<tr>
<td>( T(\gamma) )</td>
<td>( -2.829, 0.947 )</td>
<td>( -3.334, 1.157 )</td>
<td>( -3.654, 1.354 )</td>
</tr>
<tr>
<td>( T(\gamma) )</td>
<td>( -5.764, 1.483 )</td>
<td>( -7.553, 1.523 )</td>
<td>( -11.895, 2.060 )</td>
</tr>
</tbody>
</table>

Notes: The confidence intervals were solved on a grid with the values of \( c \) ranging from \(-40.0 \) to \( 9.5 \). The lag length \( k \) in equation (4) was selected by the AIC.

To assess the small-sample performance of the Scheffe procedure, some Monte Carlo simulation experiments were conducted. To generate data that obey the restrictions of the expectations hypothesis, the procedure due to Campbell and Shiller (1987, 1991) was used. First, a bivariate VAR model for a vector consisting of the difference of the one-month rate and the spread, \( \Delta r_t, r_t - r_{t-1} \), was estimated for each \( n \). Then, the estimated model was simulated by drawing random normal errors to generate the series of \( \Delta r_t \) and the spread, and using formula (6) in Campbell and Shiller (1991), the series of theoretical spreads (i.e., the spreads that would obtain if the expectations hypothesis were true) was computed. For each realization, the test of \( H_2 \) in equation (2) was conducted using both the normal approximation and the Scheffe-type procedure. Thus, the empirical size of each test is given by the rejection rate obtained with the theoreti-
cal spread series, whereas the rejection rate with the unrestricted simulated spread series gives the empirical power. The results are presented in table 4. The lag length of the VAR is selected by the AIC.

Overrejection clearly prevails in both subsample periods when the normal approximation is used. In the targeting period, the Scheffe-type procedure controls size well for all but the 24-month rate. This, of course, comes with the cost of lower (nominal) power. In the second subsample period, the conservative confidence intervals do not seem to work equally well, although the decrease in overrejection compared to the normal approximation is substantial. This is probably due to the small sample size.

Because the procedure seems to be quite conservative, “size-adjusted” confidence intervals for the interest-rate targeting period were also estimated, and the DGP of the Monte Carlo simulation experiments was used to compute the required size adjustment. The idea is to find the nominal level of significance at which the true level is (approximately) 5% according to the simulation. For the 36- and 48-month rates, that level turned out to be 7%, for the 60-month rate, 12%; and for the 120-month rate, 7.5%. All the size-adjusted confidence intervals also included unity, indicating that the nonrejection of the expectations hypothesis in the targeting period does not just follow from using a conservative test.

IV. Conclusion

This paper examines the predictive power of the yield spread. It provides one potential explanation for the common finding that—contrary to the expectations hypothesis—the yield spread does not seem to correctly predict changes in the future long-term rate. It is argued that this might be due to an omitted near unit root problem in the yield spread, which leads to overrejection of the true null hypothesis in case there is considerable simultaneity in the system consisting of the spread and the change in the long term rate. The Scheffe-type method proposed by Cavanagh et al. (1995) is applied to the problem of testing the expectations hypothesis with monthly U.S. data from 1952:1–1991:2.

Assuming that there has been no structural change, the expectations hypothesis is rejected in the entire sample. However, when testing separately for the interest-rate targeting period 1952:1–1979:9, we fail to reject it at the long end of the maturity spectrum, contrary to several previous studies that have relied on standard inference. Small-sample simulation results are provided to compare the validity of the normal approximation and the Scheffe-type procedure. These experiments show that overrejection prevails with standard inference, but the latter method controls size quite well in samples of fair size and much better than the normal approximation in relatively small samples.

Recently Bekaert et al. (1997) also addressed the problem of persistent yield spreads in regression tests of the expectations hypothesis. While our conclusions concerning the entire sample are similar to theirs, the methods are different. Bekaert et al. suggested a bootstrap procedure to find the finite sample distributions of the test statistics. As was mentioned in section II, the bootstrap does not yield asymptotically valid inference in regression models with a large autoregressive root, and so it cannot, in general, be applied to the test considered in this paper. Bekaert et al., however, only considered the entire period 1952–1995, and our results suggest that (assuming the absence of a change in the largest autoregressive root of the spreads) there the high persistence of the spreads is not a problem, whereas in the interest-rate targeting period it may be, so that the bootstrap cannot be applied. Bekaert et al. did not consider this subsample separately.

REFERENCES


