# LONG-TERM EQUILIBRIA OF YIELDS ON TAXABLE AND TAX-EXEMPT BONDS

LAWRENCE KRYZANOWSKI and KUAN XU

## ABSTRACT

Engle and Granger (1987), Campbell and Shiller (1987), Bradley and Lumpkin (1992) and Shea (1992) test the cointegration relations among yields on taxable long- and short-term bonds. This paper extends this literature by examining the cointegration of yields on both taxable and tax-exempt bonds across the term structure. It finds that yields on taxable and tax-exempt bonds with the same and near proximity terms-to-maturity are cointegrated and have long-term equilibria. Although the yields from both extremes of the one-to-thirty-year term structure do not exhibit clear pairwise cointegration, the two types of yields at the two extremes of the term structure are linked by three cointegrating vectors based on the Johansen (1988) and Johansen and Juselius (1990) estimations and tests. These cointegrating vectors reflect the strengths of the long-term links. In addition, there is at least one stochastic factor in the vector autoregression representation of the yields at both extremes of the term structure which cannot be represented by the cointegrating factors. *JEL: E43, C32, G12*

## I. INTRODUCTION

Since Engle and Granger (1987) proposed cointegration analysis for characterizing long-term equilibrium, various authors (e.g., Campbell & Shiller, 1987; Bradley & Lumpkin, 1992; Shea, 1992) have tested for cointegration relations among yields on taxable long- and short-term US Treasury bonds.[1] This paper extends that literature by examining

*Direct all correspondence to:* Lawrence Kryzanowski, Department of Finance, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, P.Q. Canada H3G 1M8; Kuan Xu, Department of Economics, Dalhousie University, Halifax, N.S. Canada B3J 3H5.

**International Review of Economics and Finance, 6(2):** 119-143
ISSN: 1059-0560

the cointegration relations among yields on both taxable and tax-exempt bonds across a longer (but complete) term structure.

This research is important for at least four reasons. First, past research on the relationship between yields on taxable and tax-exempt bonds rests on the untested assumption of long-term equilibrium. It estimates the mean differences of yield spreads for various maturities (e.g., Mussa & Kormendi, 1979), explores the factors affecting yield spreads (Trzcinka, 1982; Skelton, 1983; Buser & Hess, 1986), and examines various explanations for the downward sloping term structure of yield spreads (Kochin & Parks, 1988; Green, 1993; Kryzanowski, Xu & Zhang, 1995). Second, knowing to what extent the taxable and tax-exempt market are related helps US municipal governments better position "munis" to be competitive with respect to US Federal Government Treasury securities, and helps investors in constructing diversified bond portfolios. Third, while past research examines the monthly yields on taxable bonds with terms-to-maturity of up to twenty-five years (Shea, 1992), this research covers bonds with one to thirty year terms-to-maturity.[2] Fourth, this research extends existing work by examining both taxable and tax-exempt bonds with more complete data to determine if longer bonds are cointegrated with shorter bonds. Thus, this research aims to fill the void in the literature and to provide new empirical evidence concerning the long-term equilibria of yields on taxable and tax-exempt bonds.

We find that yields on taxable and tax-exempt bonds with the same and near proximity terms-to-maturity are cointegrated and have long-term equilibria. When the analysis is extended to the one to thirty year segment, the yields from both extremes of the term structure do not exhibit clear pairwise cointegration based on the single equation cointegration analysis. While very long-term bonds behave differently due to different time- and term-varying risk premia for both types of bonds, the two types of yields from both extremes of the term structure are cointegrated in a more comprehensive system. The Johansen-type estimation and tests provide further evidence on the long-term equilibria of yields on taxable and tax-exempt bonds.

The remainder of this paper is organized as follows. In section two, the theoretical framework and cointegration relations are discussed. In section three, the data and unit root test results for yields on taxable and tax-exempt bonds are presented. In section four, the results of conducting cointegration analyses based on Engle and Granger (1987), Johansen (1988) and Johansen and Juselius (1990) are reported and analyzed. Some concluding remarks follow in the last section.

## II. THEORETICAL FRAMEWORK AND COINTEGRATION RELATIONS

If long-term equilibria exist among various yields, then cointegration equations can be used to characterize long-term equilibria and random deviations can be attributed to short-term shocks. Such cointegration relations for conventionally-reported yield data for coupon bonds can be illustrated using the Shiller, Campbell and Schoenholtz (1983) [SCS] model.[3] Under the expectations hypothesis, the yield on an $i$-period coupon bond is a weighted average of expected future one-period yields where more distant yields are given lower weight. Given that Campbell and Shiller (1984) find that the risk premium is both *time- and term-varying*, the SCS model can be modified to:

$$R_t^{(i)} = \sum_{k=0}^{i-1} W(k)E_t R_{t+k}^{(1)} + V_t^{(i)}, \tag{1}$$

where $R_t^i$ is the yield to maturity on an $i$-period bond at time $t$; $W(k) = g^k(1-g)/(1-g^i)$ ($k = 1, 2, \ldots, i-1$) are weights; $g = 1/(1+\bar{R})$ is a constant discount factor; $E_t$ is the mathematical expectation conditional on the information set available at time $t$; and $V_t^{(i)}$ is the risk premium for the yield of an $i$-period bond at time $t$. Equation (1) reflects Macaulay's duration, $D_i$, (Macaulay, 1938) by defining $W(k) = g^k/D_i$, where $D_i = (1-g^i)/(1-g)$ when the coupon rate, $c_i$, equals the average yield, $\bar{R}$.

Let $R_t^{\tau(i)}$ and $R_t^{e(i)}$ represent the yields on taxable and tax-exempt $i$-period bonds at time $t$, respectively. Their returns can be expressed as:

$$R_t^{\tau(i)} = \sum_{k=0}^{i-1} W(k)^{\tau} E_t R_{t+k}^{\tau(1)} + V_t^{\tau(i)}, \tag{2}$$

and

$$R_t^{e(i)} = \sum_{k=0}^{i-1} W(k)^{e} E_t R_{t+k}^{e(1)} + V_t^{e(i)} \tag{3}$$

where $V_t^{\tau(i)}$ and $V_t^{e(i)}$ denote the time- and term-varying risk premia for $i$-period taxable and tax-exempt bonds, respectively, at time $t$. $W(k)^{\tau}$ in equation (2) differs from $W(k)^{e}$ in equation (3) because $\bar{R}$ is equal to the taxable and tax-exempt coupon rates, respectively, in the definition of $g$.

The taxable and tax-exempt yields, $R_t^{\tau(i)}$ and $R_t^{e(i)}$, are related as follows:

$$R_t^{\tau(i)} = \varphi_0 + \varphi R_t^{e(i)} + \varepsilon_t^{(i)}, \tag{4}$$

where $\varphi_0 = \sum_{k=0}^{i-1} W(k)^{\tau} E_t R_{t+k}^{\tau(1)} - \varphi \sum_{k=0}^{i-1} W(k)^{e} E_t R_{t+k}^{e(1)}, \varepsilon_t^{(i)} = V_t^{\tau(i)} - \varphi V_t^{e(i)}$, and $\varphi$ is a free parameter. Only if $\varepsilon_t^{(i)}$ is a stationary process are $R_t^{\tau(i)}$ and $R_t^{e(i)}$ cointegrated in the long-term given each is integrated of order one.[4] Our hypothesis is that yields on taxable and tax-exempt bonds with the same terms-to-maturity, which only differ by a stationary tax rate, should be cointegrated. As is shown later in the paper, this conjecture is supported by the empirical evidence.

If the long- and short-term yields, $R_t^{l(j)}$ and $R_t^{s(i)}$ are related regardless of tax status, then it follows for $i < j$,

$$R_t^{s(j)} = \psi_0 + \psi R_t^{l(i)} + \varepsilon_t^{(j-i)}, \tag{5}$$

where $\psi_0 = \sum_{k=0}^{j-1} W(k)^{s} E_t R_{t+k}^{s(1)} - \psi \sum_{k=0}^{i-1} W(k)^{l} E_t R_{t+k}^{l(1)}, \varepsilon_t^{(j-i)} = V_t^{s(j)} - \psi V_t^{l(i)}$, and $\psi$ is a free parameter and $s, l = \tau$ and/or $e$. Only if $\varepsilon_t^{(j-i)}$ follows a stationary process are the long- and short-term yields cointegrated given each is integrated of order one. Our hypothesis is that equation (5) represents a cointegration relation, which can be viewed as an equilibrium condition across bonds of different tax status and terms-to-maturity, respectively. As is shown later in the paper, the yields with similar and near proximity terms-to-maturity have similar time-varying risk-premia, and are cointegrated. The yields from both extremes of the term structure have different time- and term-varying risk-premia, and are not strongly cointegrated.

The above specifications of cointegration relations for pairwise comparisons of two yields also can be used in multivariate cointegration analyses. For example, $X_t' = [R_t^{\tau(1)}, \ldots, R_t^{\tau(30)}, R_t^{e(1)}, \ldots, R_t^{e(30)}]$ can be used in the following standard vector autoregression (VAR) model:

$$X_t = \mu + \Pi_1 X_{t-1} + \ldots + \Pi_k X_{t-k} + \varepsilon_t, \tag{6}$$

where $X_t$ is a $(p \times 1)$ vector of yields; $\varepsilon_t$ is a sequence of *i.i.d.* $p$-dimensional Gaussian random vectors following the distribution $N(0, \Sigma)$; $\Pi_i$ are $(p \times p)$ matrices of parameters; and $\mu$ is a $(p \times 1)$ vector of constant terms. This system allows for a test of the number of cointegrating vectors in the multivariate cointegration model to determine if the model is able to encompass long-term equilibria. As is shown later, if the shortest and longest taxable and tax-exempt yields are considered in the framework of a VAR model, there are at least three cointegrating vectors which can represent the long-term equilibria, and one stochastic factor exists which cannot be represented by the cointegrating vectors.

## III. DATA AND UNIT ROOT TESTS

The data are drawn from *Analytical Record of Yields and Yield Spreads*, which is published by Salomon Brothers Inc. The monthly government bond yields with terms-to-maturity of 1, 2, 5, 10, and 20 years are available from the monthly yield curves over the period from January 1950 to December 1990. The monthly government bond yields with a term-to-maturity of 30 years, are available from May 1953 to December 1990.[5] The prime grade municipal bond yields with the same maturities are estimates on the first day of each month, and are available from January 1950 to December 1988. Thus, the common time period of these series, May 1953 to December 1988, is studied herein.[6]

In the subsequent analyses, yields on taxable (tax-exempt) bonds refer to the yields on taxable U.S. government (prime grade municipal) bonds with terms-to-maturity of 1, 2, 5, 10, 20 and 30 years, and are denoted as g1, g2, g5, g10, g20 and g30 (m1, m2, m5, m10, m20 and m30), respectively. The first-differences of these yields are denoted as dg1, dg2, dg5, dg10, dg20, dg30, dm1, dm2, dm5, dm10, dm20, and dm30 for the series.[7]

A visual inspection of Figures 1 and 2 suggests nonstationarity in each of the series.[8] As a test of this observation, the sample autocorrelation functions are computed for each series up to lag 40. Based on the estimates for lags from one to six (reported in Table 1), all the sample autocorrelation functions are positive, fail to damp, and have very smooth and persistent movements. Based on the partial autocorrelation functions computed for each series up to lag 40, each series has a very large and highly significant value at lag 1, and insignificant values thereafter. This suggests that the yield series are first-order nonstationary

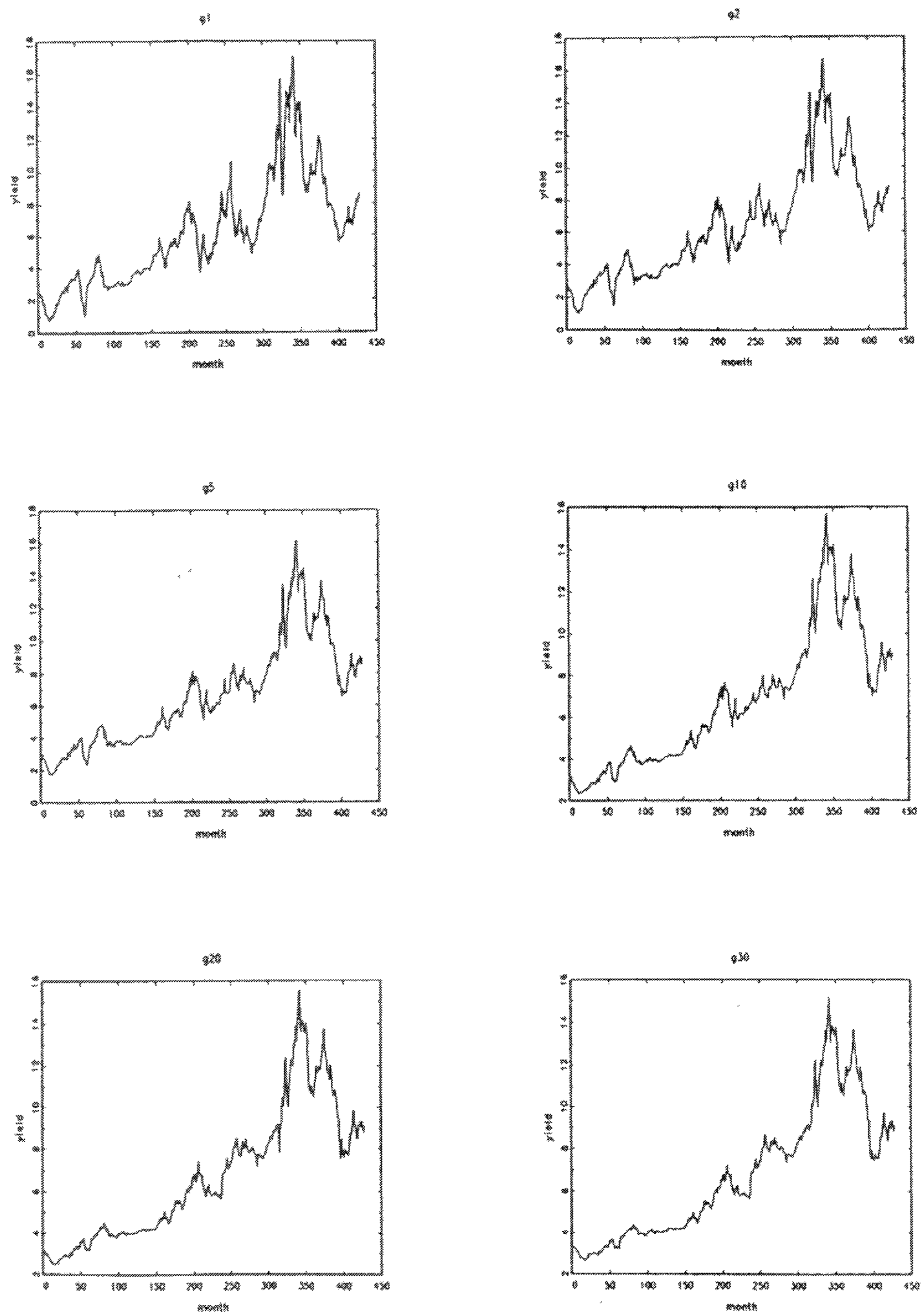


*Figure 1.* Treasury Bond Yields

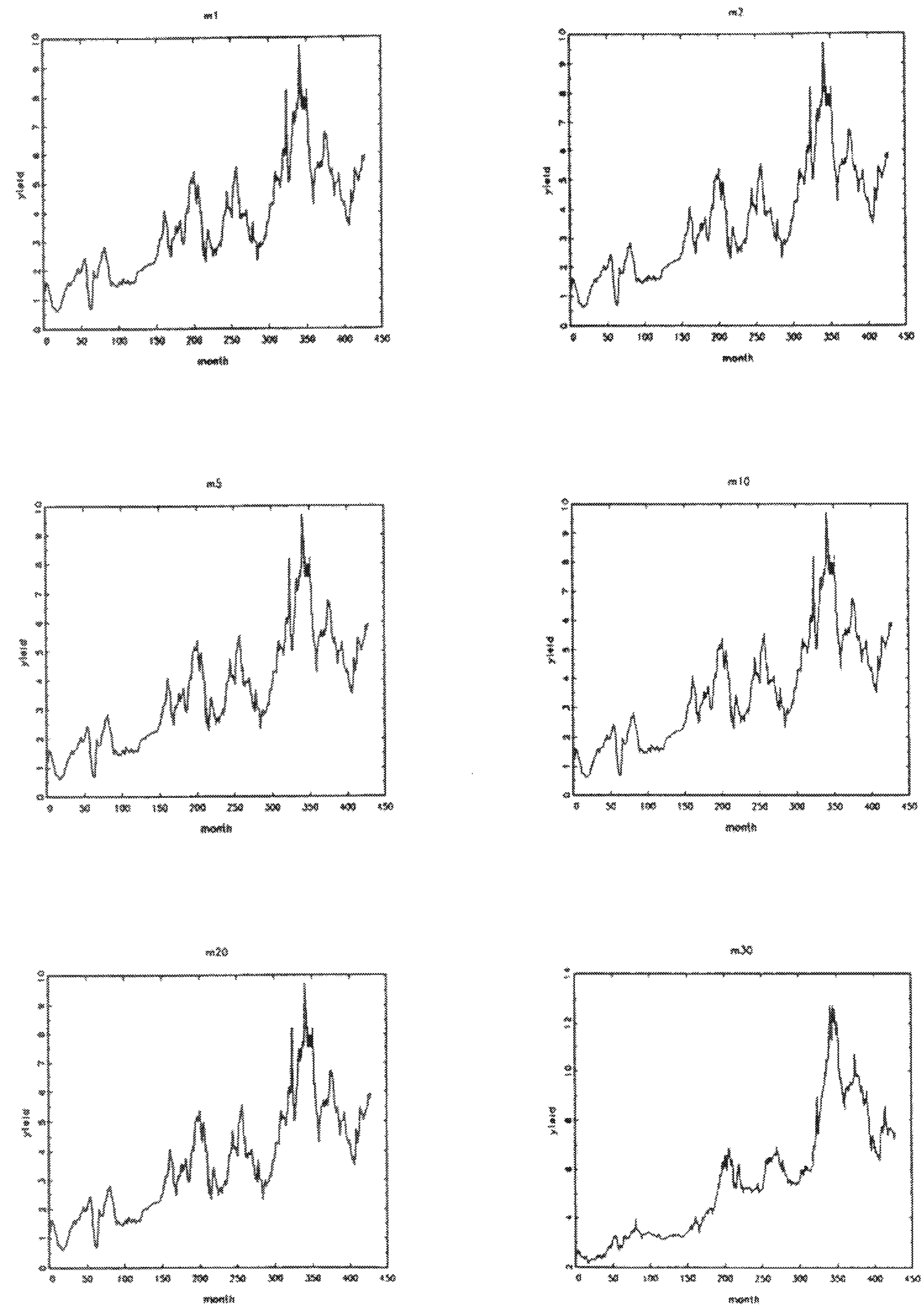


*Figure 2.* Municipal Bond Yields

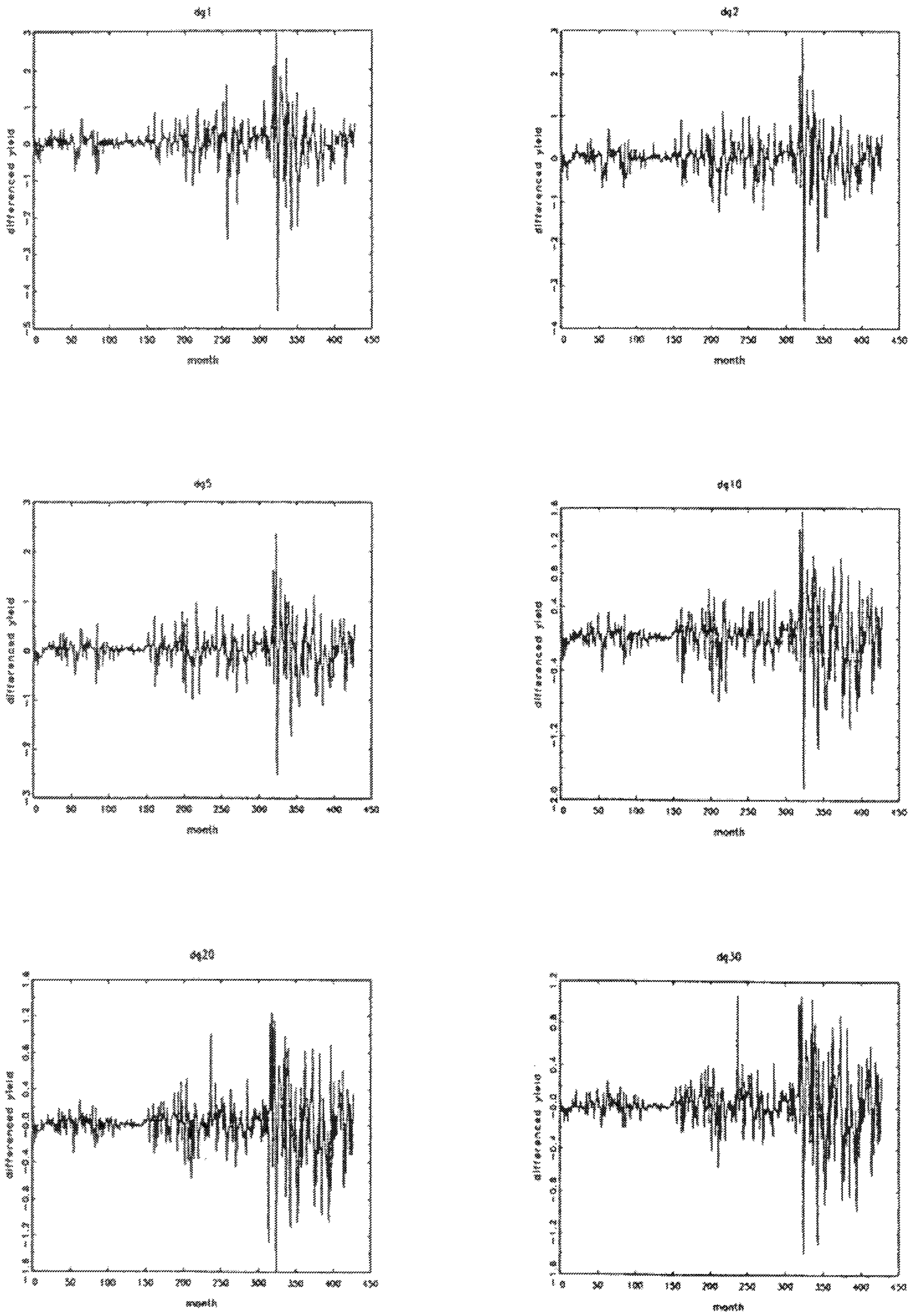


*Figure 3.* Differenced Treasury Bond Yields

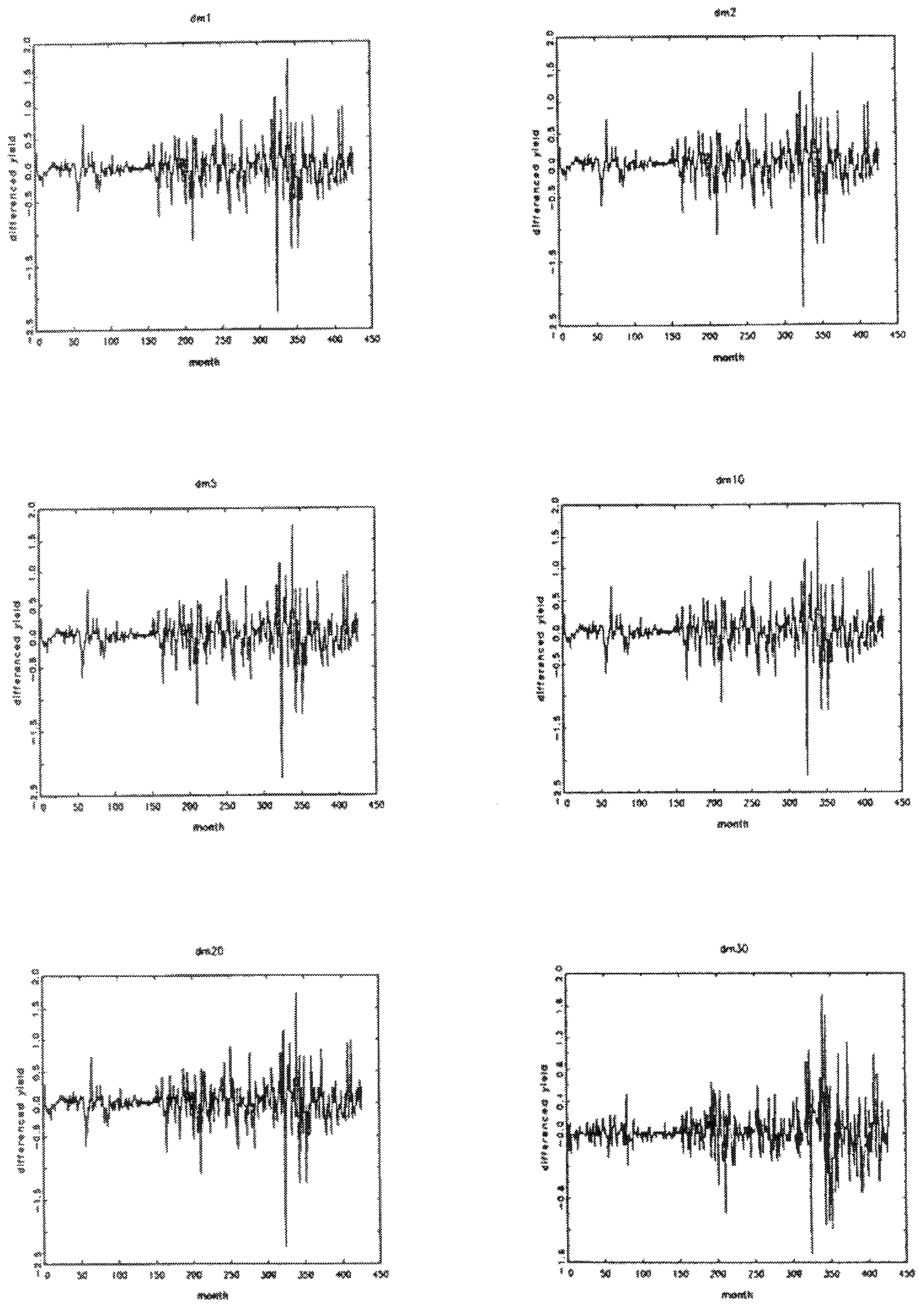


*Figure 4.* Differenced Municipal Bond Yields

processes. When the first differences are taken for each of the yield series, each series is transformed into a mean-reverting stationary process which exhibits persistent periods of high and low volatility (see Figures 3 and 4). This "volatility clustering" is a fundamental characteristic of rates of return for many financial assets (see Engle, Lilien & Robins, 1987, for example). The sample autocorrelation functions reported in Table 2 collaborate these findings, since only the autocorrelation functions at lags 1, 3 and 5 are outside the confidence interval ± 0.0967.

To further examine if the yield series are integrated of order one, the augmented Dickey-Fuller (DF) test is used.[9] Under the null hypothesis of $\rho$ equal to one, the limiting distribution of $n(\hat{\rho} - 1)$ has a special form (see Fuller, 1976; Dickey & Fuller, 1979). The results for such a test are summarized in Table 3, where the lower tail area of the DF test

*Table 1.* Autocorrelation Functions of the Monthly Bond Yield Series

| *Series*[a] | $\hat{r}_1^b$ | $\hat{r}_2$ | $\hat{r}_3$ | $\hat{r}_4$ | $\hat{r}_5$ | $\hat{r}_6$ |
|---|---|---|---|---|---|---|
| g1 | .982 | .960 | .939 | .922 | .908 | .893 |
| g2 | .985 | .967 | .951 | .937 | .926 | .913 |
| g5 | .989 | .976 | .965 | .955 | .945 | .935 |
| g10 | .992 | .983 | .974 | .966 | .957 | .949 |
| g20 | .993 | .984 | .976 | .969 | .962 | .955 |
| g30 | .993 | .986 | .978 | .972 | .965 | .957 |
| m1 | .980 | .957 | .938 | .919 | .901 | .881 |
| m2 | .981 | .961 | .945 | .929 | .914 | .987 |
| m5 | .984 | .970 | .957 | .945 | .932 | .917 |
| m10 | .988 | .977 | .966 | .957 | .947 | .934 |
| m20 | .990 | .981 | .973 | .966 | .957 | .946 |
| m30 | .990 | .982 | .975 | .968 | .959 | .949 |

*Notes:* [a] *gi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*; and *mi* refers to the yield series of the municipal (tax-exempt) bond with term-to-maturity *i*.
[b] $\hat{r}_j$ refers to the sample autocorrelation functions with lag *j*.

*Table 2.* Autocorrelation Functions of the Differenced Monthly Bond Yield Series

| *Series*[a] | $\hat{r}_1^b$ | $\hat{r}_2$ | $\hat{r}_3$ | $\hat{r}_4$ | $\hat{r}_5$ | $\hat{r}_6$ |
|---|---|---|---|---|---|---|
| g1 | .128 | −.051 | −.121 | −.081 | .024 | −.074 |
| g2 | .151 | −.086 | −.112 | −.066 | .049 | −.042 |
| g5 | .117 | −.082 | −.101 | −.014 | .030 | .047 |
| g10 | .108 | −.035 | −.093 | .017 | .039 | .040 |
| g20 | .084 | −.023 | −.122 | −.014 | .080 | .078 |
| g30 | .139 | −.026 | −.118 | .030 | .050 | .065 |
| m1 | .065 | −.065 | −.043 | .022 | .028 | −.067 |
| m2 | .024 | −.076 | −.055 | .014 | .064 | −.029 |
| m5 | −.063 | −.007 | −.056 | .021 | .063 | −.032 |
| m10 | −.062 | −.012 | −.093 | .035 | .143 | .009 |
| m20 | −.109 | −.014 | −.081 | .092 | .125 | .052 |
| m30 | −.117 | −.032 | −.063 | .111 | .136 | .025 |

*Notes:* [a] *gi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*; and *mi* refers to the yield series of the municipal (tax-exempt) bond with term-to-maturity *i*.
[b] $\hat{r}_j$ refers to the sample autocorrelation functions with lag *j*.

*Table 3.* Tests for Autoregressive Unit Roots

| *Series*[a] | $k$[b] | $\hat{\mu}$ | $t(\hat{\mu})$[c] | $\hat{\gamma}$ | $t(\hat{\gamma})$[c] | $\hat{\rho} - 1$ | $t(\hat{\rho} - 1)$[c] | *Lower Tail Area*[d] |
|---|---|---|---|---|---|---|---|---|
| g1 | 1 | .062 | .919 | .0010 | 2.521 | −.048 | −3.399 | .053 |
| g2 | 2 | .057 | .978 | .0008 | 2.288 | −.040 | −2.944 | .154 |
| g5 | 1 | .061 | 1.255 | .0007 | 2.220 | −.035 | −2.813 | .202 |
| g10 | 1 | .051 | 1.279 | .0005 | 1.887 | −.026 | −2.363 | .449 |
| g20 | 0 | .049 | 1.281 | .0004 | 1.590 | −.022 | −2.009 | .671 |
| g30 | 1 | .044 | 1.290 | .0005 | 1.791 | −.022 | −2.196 | .557 |
| m1 | 0 | .031 | .822 | .0007 | 2.753 | −.052 | −3.336 | .062 |
| m2 | 0 | .039 | 1.029 | .0007 | 2.802 | −.054 | −3.389 | .054 |
| m5 | 0 | .044 | 1.189 | .0007 | 2.742 | −.050 | −3.250 | .076 |
| m10 | 0 | .042 | 1.164 | .0006 | 2.334 | −.039 | −2.817 | .202 |
| m20 | 0 | .042 | 1.160 | .0005 | 2.097 | −.032 | −2.543 | .340 |
| m30 | 0 | .048 | 1.307 | .0005 | 2.020 | −.031 | −2.467 | .385 |

*Notes:* [a] *gi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*; and *mi* refers to the yield series of the municipal (tax-exempt) bond with term-to-maturity *i*.
[b] $k$ refers to the number of lags used in estimating $\Delta x_t = \mu + \gamma t + (\rho - 1)x_{t-1} + \sum_{i=1}^{k} \phi_i \Delta x_{t-i} + \varepsilon_t, t = 1, \ldots, 421$.
[c] $t(.)$ equals the ratio of the estimated coefficient to its standard error.
[d] The lower tail area is the one side $p$-value that should be compared with the significance level chosen.

statistic is computed based on MacKinnon's method (1991). At the 0.05 level, the null hypothesis of the existence of a unit root cannot be rejected. This implies that each series is integrated of order one.

The above test results may be affected by the changes in monetary regimes during the late 1970's and early 1982 (see Huizinga & Mishkin, 1984; Clarida & Friedman, 1984). During the period of 1979:10-1982:9, the Fed increased its emphasis on control of bank reserves and money supply (M1) and reduced its emphasis on short-term interest rates. Within the same period, the inflation rate was reduced significantly (i.e., from 11–12% to 4–5%), and short-term interest rates were greater than ever before. Further, the taxable and tax-exempt yields are also affected by changes in the tax regime such as the Economic Recovery Tax Act of 1981 and the Tax Reform Act of 1986 (see Poterba, 1989). Furthermore, the yields may be affected by the Fed's policy in 1985 when the Fed attempted to bring down the dollar by driving the growth rates for M1 and M2, and interest rates simultaneously decreased rapidly (Bernanke & Mishkin, 1992, p. 194).

The Zivot and Andrews (1992) (ZA) test for a unit root, which allows for a flexible break point, is used for this purpose.[10] The test results summarized in Tables 4 and 5 provide stronger evidence for the existence of a unit root even when breaks are allowed. The plausible breaks for the series occur mainly during 1984 and 1985 (Table 4), and the null hypothesis of a unit root is not rejected for each series even if a break is allowed (Table 5). According to the asymptotic distribution of the ZA test statistic, the one-side critical values at 1% and 5% significance levels are −5.57 and −5.08, respectively. Since the $t(\hat{\rho} - 1)$ values for all series given in Table 5 are greater than these critical values, the null hypotheses of a unit root cannot be rejected.

Although conventional practice is to use the test for a unit root directly for a time series without questioning if the series is better represented by a higher-order autoregressive process, Dickey and Pantula (1987) propose a test procedure (the DP test) for the

*Table 4.* Minimum *t* Statistics

| *Series*[a] | *t-statistic*[b] | $T^c$ | $T_B{}^d$ | *Month*[e] |
|---|---|---|---|---|
| g1 | −4.793 | 421 | 370 | 84:11 |
| g2 | −3.067 | 421 | 370 | 84:11 |
| g5 | −4.402 | 421 | 376 | 85:05 |
| g10 | −3.810 | 421 | 368 | 84:09 |
| g20 | −3.960 | 421 | 376 | 85:05 |
| g30 | −4.018 | 421 | 376 | 85:05 |
| m1 | −4.230 | 421 | 370 | 84:12 |
| m2 | −4.007 | 421 | 370 | 84:12 |
| m5 | −3.905 | 421 | 383 | 85:12 |
| m10 | −3.552 | 421 | 383 | 85:12 |
| m20 | −3.155 | 421 | 383 | 85:12 |
| m30 | −3.096 | 421 | 383 | 85:12 |

*Notes:* [a] *gi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*; and *mi* refers to the yield series of the municipal (tax-exempt) bond with term-to-maturity *i*.
[b] *t*-statistic refers to the lowest *t* statistics for $(e - 1)$ given $\xi \varepsilon \Xi = [.001, .999]$ in

$$\Delta x_t = \mu + \gamma t + \zeta DU_t(\xi) + (\rho - 1)x_{t-1} + \sum_{i=1}^{k} \phi_i \Delta x_{t-i} + \varepsilon_t,$$

where $t = 1, \ldots, T$. This test is applied to all the series.
[c] $T$ refers to the number of observations used in estimation.
[d] $T_B$ refers to the break point.
[e] The month corresponds to the break point $T_B$.

*Table 5.* Tests for Autoregressive Unit Roots Allowing a Break Point

| *Series*[a] | $k^b$ | $\hat{\mu}$ | $t(\hat{\mu})^c$ | $\hat{\gamma}$ | $t(\hat{\gamma})^c$ | $\hat{\xi}$ | $t(\hat{\xi} - 1)^c$ | $\hat{\rho} - 1$ | $t(\hat{\rho} - 1)^c$ |
|---|---|---|---|---|---|---|---|---|---|
| g1 | 1 | .075 | 1.218 | .002 | 4.266 | −.438 | −3.461 | −.081 | −4.793 |
| g2 | 2 | .129 | 1.975 | .001 | 1.347 | .087 | .928 | −.042 | −3.067 |
| g5 | 1 | .098 | 2.090 | .002 | 4.114 | −.338 | −3.539 | −.067 | −4.402 |
| g10 | 0 | .062 | 1.599 | .001 | 3.983 | −.282 | −4.178 | −.046 | −3.810 |
| g20 | 1 | .079 | 2.089 | .001 | 3.886 | −.261 | −3.682 | −.051 | −3.960 |
| g30 | 1 | .075 | 2.193 | .001 | 3.967 | −.247 | −3.810 | −.048 | −4.018 |
| m1 | 1 | .050 | 1.408 | .001 | 3.775 | −.146 | −2.286 | −.072 | −4.230 |
| m2 | 0 | .058 | 1.586 | .001 | 3.620 | −.144 | −2.273 | −.068 | −4.007 |
| m5 | 0 | .072 | 2.016 | .001 | 3.540 | −.151 | −2.232 | −.066 | −3.905 |
| m10 | 0 | .064 | 1.830 | .001 | 3.293 | −.156 | −2.390 | −.053 | −3.552 |
| m20 | 1 | .061 | 1.698 | .001 | 3.010 | −.173 | −2.596 | −.044 | −3.155 |
| m30 | 1 | .068 | 1.866 | .001 | 2.950 | −.172 | −2.605 | −.043 | −3.096 |

*Notes:* [a] *gi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*; and *mi* refers to the yield series of the municipal (tax-exempt) bond with term-to-maturity *i*.
[b] $k$ refers to the number of lags used in estimating $\Delta x_t = \mu + \gamma t + \zeta DU_t(\xi) + (\rho - 1)x_{t-1} + \sum_{i=1}^{k} \phi_i \Delta x_{t-i} + \varepsilon_t$, sample size $T = 421$.
[c] $t(.)$ equals the ratio of the estimated coefficient to its standard error.

number of unit roots in a stochastic autoregressive process. This test is used herein to verify the conclusions resulting from the use of the less general augmented DF test for a unit root. Because the DP test is a three-step sequential test and is often not used, a detailed description of the test is given in Appendix A. The test results for the series of

*Table 6.* The Dickey-Pantula Sequential Tests for Taxable Yields

| *Series*[a] | *Step* | $\hat{\pi}_1$ | $t^*_{1,T}(3)$[b] | $\hat{\pi}_2$ | $t^*_{2,T}(3)$[b] | $\hat{\pi}_3$ | $t^*_{3,T}(3)$[b] | $\hat{\tau}^c_{500,.05}$ | $\hat{\tau}^c_{500,.05}$ |
|---|---|---|---|---|---|---|---|---|---|
| g1 | 1 | | | | | −1.397 | −31.165 | −1.95 | 1.28 |
| | 2 | | | −.930 | −14.453 | −.932 | −19.116 | −1.95 | 1.28 |
| | 3 | −.0022 | −.535 | −.927 | −14.339 | −.934 | −19.095 | −1.95 | 1.28 |
| g2 | 1 | | | | | −1.360 | −29.803 | −1.95 | 1.28 |
| | 2 | | | −.941 | −14.874 | −.889 | −18.278 | −1.95 | 1.28 |
| | 3 | −.0010 | −.297 | −.940 | −14.784 | −.890 | −18.242 | −1.95 | 1.28 |
| g5 | 1 | | | | | −1.386 | −30.726 | −1.95 | 1.28 |
| | 2 | | | −.968 | −14.969 | −.903 | −18.529 | −1.95 | 1.28 |
| | 3 | −.0001 | −.024 | −.967 | −14.905 | −.903 | −18.480 | −1.95 | 1.28 |
| g10 | 1 | | | | | −1.420 | −31.971 | −1.95 | 1.28 |
| | 2 | | | −.931 | −14.286 | −.954 | −19.511 | −1.95 | 1.28 |
| | 3 | .0004 | .193 | −.932 | −14.241 | −.953 | −19.447 | −1.95 | 1.28 |
| g20 | 1 | | | | | −1.441 | −32.827 | −1.95 | 1.28 |
| | 2 | | | −.941 | −14.028 | −.971 | −19.838 | −1.95 | 1.28 |
| | 3 | .0005 | .260 | −.942 | −14.196 | −.970 | −19.770 | −1.95 | 1.28 |
| g30 | 1 | | | | | −1.404 | −31.367 | −1.95 | 1.28 |
| | 2 | | | −.900 | −14.028 | −.955 | −19.523 | −1.95 | 1.28 |
| | 3 | .0006 | .315 | −.900 | −13.993 | −.954 | −19.453 | −1.95 | 1.28 |

*Notes:* [a] *gi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*.
[b] $t^*_{i,T}$ (3) equals the ratio of ith estimated coefficient to its standard error.
[c] The critical values for $\alpha = 0.05, 0.95$ when $T = 500$.

*Table 7.* The Dickey-Pantula Sequential Tests for Tax-Exempt Yields

| *Series*[a] | *Step* | $\hat{\pi}_1$ | $t^*_{1,T}(3)$[b] | $\hat{\pi}_2$ | $t^*_{2,T}(3)$[b] | $\hat{\pi}_3$ | $t^*_{3,T}(3)$[b] | $\hat{\tau}^c_{500,.05}$ | $\hat{\tau}^c_{500,.05}$ |
|---|---|---|---|---|---|---|---|---|---|
| m1 | 1 | | | | | −1.431 | −32.454 | −1.95 | 1.28 |
| | 2 | | | −.999 | −14.991 | −.931 | −19.100 | −1.95 | 1.28 |
| | 3 | −.0009 | −.224 | −.997 | −14.879 | −.932 | −19.046 | −1.95 | 1.28 |
| m2 | 1 | | | | | −1.448 | −33.144 | −1.95 | 1.28 |
| | 2 | | | −1.050 | −15.437 | −.923 | −18.940 | −1.95 | 1.28 |
| | 3 | −.0005 | −.137 | −1.094 | −15.340 | −.923 | −18.885 | −1.95 | 1.28 |
| m5 | 1 | | | | | −1.525 | −36.729 | −1.95 | 1.28 |
| | 2 | | | −1.073 | −15.060 | −.989 | −20.231 | −1.95 | 1.28 |
| | 3 | .0001 | .030 | −1.073 | −14.988 | −.989 | −20.168 | −1.95 | 1.28 |
| m10 | 1 | | | | | −1.522 | −36.552 | −1.95 | 1.28 |
| | 2 | | | −1.075 | −15.133 | −.984 | −20.155 | −1.95 | 1.28 |
| | 3 | .0004 | .139 | −1.076 | −15.072 | −.984 | −20.086 | −1.95 | 1.28 |
| m20 | 1 | | | | | −1.542 | −37.554 | −1.95 | 1.28 |
| | 2 | | | −1.133 | −15.603 | −.975 | −19.956 | −1.95 | 1.28 |
| | 3 | .0006 | .262 | −1.135 | −15.554 | −.974 | −19.882 | −1.95 | 1.28 |
| m30 | 1 | | | | | −1.537 | −37.264 | −1.95 | 1.28 |
| | 2 | | | −1.164 | −15.980 | −.955 | −19.547 | −1.95 | 1.28 |
| | 3 | .0008 | .328 | −1.166 | −15.938 | −.954 | −19.475 | −1.95 | 1.28 |

*Notes:* [a] *mi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*.
[b] $t^*_{i,T}$ (3) equals the ratio of ith estimated coefficient to its standard error.
[c] The critical values for $\alpha = 0.05, 0.95$ when $T = 500$.

tax-exempt and taxable yields based on the sequential $t$-test procedure are summarized in Tables 6 and 7, respectively. The results are illustrated by concentrating on g1 in Table 6. During step one, $H_3$ (that three stationary roots exist) is rejected because $t^*_{3,T}(3) = -31.165 < \hat{\tau}_{500,.05} = -1.95$.[11] During step two, $H_2$ (that two unit roots and one stationary root exist) is rejected because $t^*_{3,T}(3) = -19.116 < \hat{\tau}_{500,.05} = -1.95$ and $t^*_{2,T}(3) = -14.453 < \hat{\tau}_{500,.05} = -1.95$. During step three, $H_1$ (that exactly one unit root and two stationary roots exist) is not rejected because $t^*_{1,T}(3) = -.535 > \hat{\tau}_{500,.05} = -1.95$ although both $t^*_{1,T}$ and $t^*_{2,T}$ are less than $\hat{\tau}_{500,.05}$. Based on the results in both tables, all the yield series (tax-exempt and taxable) have only one unit root, and, thus, are integrated of order one.

## IV. COINTEGRATION ANALYSIS

To provide additional empirical evidence about the long run equilibrium relations among taxable and tax-exempt yields, the single equation cointegration analysis due to Engle and Granger (1987) is used initially on every pair of bond yields across types of bonds and terms-to-maturity. To examine the possible cointegrating vectors that exist in the two types of yields from the two extremes of the term structures, Johansen (1988) and Johansen and Juselius (1990) estimation methods and tests are used subsequently.

In the bond markets, yields on various types of bonds with the same term-to-maturity are expected to be cointegrated because they are affected by the same set of economic factors such as the expected rate of inflation and the real growth rate of the economy. Differences in yields on different types of bonds may reflect issue-specific factors such as tax status and risk level, which cannot be eliminated completely by arbitrage.

The tests for non-cointegration are implemented first for each pair of yields on taxable and tax-exempt bonds for the same term-to-maturity:

$$R_t^{\tau(i)} = \varphi_0 + \varphi_1 t + \varphi_2 R_t^{e(i)} + \varepsilon_t^{(i)}, \tag{7}$$

where $i = 1, 2, 5, 10, 20$, and $30$, represents term-to-maturity, and $t = 1, ..., T$ denotes time. Equation (7) is equation (4) with an additional time trend term, $t$, to capture possible time effects.[12] The test of a unit root is conducted on the residuals of each equation. The lower tail area of the DF Test statistic is computed using MacKinnon's method (1991). Based on the results summarized in Table 8, the paired yields are cointegrated for all terms-to-maturity at the 0.05 level.

The cointegrating vector $(\hat{\varphi}_0, \hat{\varphi}_1, \hat{\varphi}_2)$ for each model in Table 8 demonstrates the long-term equilibria conditions for the yields of comparable term-to-maturity. $\hat{\varphi}_0$ represents the fixed markup or markdown. It is a relatively small markdown for one- and two-year yields, and becomes a greater markup for longer-term yields. $\hat{\varphi}_1$ is a parameter for a tend. Since it is trivial, it is not a important factor for all yields. $\hat{\varphi}_2$ denotes the proportional markup on top of the tax-exempt yield such that the two types of yields could have a long-term cointegration relation. As we can see in Table 8, the proportional markup for one-year tax-exempt yields is about 83%. This markup is reduced significantly as the term-to-maturity gets longer. It reaches 6% at the long-term end of the term structure. The fixed and proportional markups are determined by the marginal tax rate as well as by other

*Table 8.* Test for Non-Cointegration in Each Term-to-Maturity of Bond Market

| $i$[a] | $\phi_0$[b] | $\phi_1$[b] | $\phi_2$[b] | $R^2$[b] | *D.F. Test*[c] | *Lower Tail Area*[d] |
|---|---|---|---|---|---|---|
| 1 | −.0520 | −.0011 | 1.8313 | .9427 | −6.0661 | .0001 |
| 2 | −.0446 | −.0003 | 1.7257 | .9556 | −6.5130 | .0000 |
| 5 | .1213 | .0003 | 1.5288 | .9580 | −5.8569 | .0001 |
| 10 | .2569 | .0025 | 1.2436 | .9516 | −4.4362 | .0081 |
| 20 | .1935 | .0034 | 1.0849 | .9603 | −4.4246 | .0084 |
| 30 | .1042 | .0034 | 1.0599 | .9578 | −4.0287 | .0264 |

*Notes:* [a] $i$ refers to $i$ term-to-maturity.
[b] The estimates of the coefficients are for the model

$$R_t^{\tau(i)} = \varphi_0 + \varphi_1 t + \varphi_2 R_t^{e(i)} + \varepsilon_t^{(i)}$$

[c] D.F. Test refers to the Dickey-Fuller test. It is applied to the residuals of the linear equation for each pair of taxable and tax-exempt bonds.
[d] The lower tail area is the one side $p$-value that should be compared with the significance level chosen.

factors such as expectations, risk premia, and so on. The observed cointegrating vectors show that the tax-exempt yields with the shorter terms-to-maturity need greater proportional markups and smaller fixed markdowns, while the tax-exempt yields with the longer terms-to-maturity tend to have greater fixed markups and smaller proportional markups. These markups represent the long-term equilibria, and hence provide some quantitative explanations to the issue of why researchers observe decreasing "marginal tax rates" over the terms-to-maturity. Although the cointegration relations become somewhat weaker at the long-term end of the term structure (i.e., 30 years), the overall results confirm the hypothesis that yields on taxable and tax-exempt bonds with the same terms-to-maturity are cointegrated. Whether some factor exerts an influence on yield spreads at the long-term end of the term structure is dealt with below.

The cointegration relations are evaluated next within each term structure (taxable and tax-exempt), and across the two term structures as follows: (i) The non-cointegration test is conducted on each possible pair of yields on *taxable bonds* with *different terms-to-maturity*; namely:

$$R_t^{\tau(j)} = \psi_0 + \psi_1 t + \psi_2 R_t^{\tau(i)} + \varepsilon_t^{(j-i)}, \tag{8}$$

where $i, j = 1, 2, 5, 10, 20$ and $30$, $i \neq j$ and $t = 1, \ldots, T$. Based on the results presented in Table 9, g1 is not cointegrated with g10, g20 and g30; and g2 is not cointegrated with g30 at the 0.05 level, while all the other pairs of yields are cointegrated. (ii) The non-cointegration test is applied to each possible pair of yields of *tax-exempt bonds* with *different terms-to-maturity*; namely:

$$R_t^{e(j)} = \psi_0 + \psi_1 t + \psi_2 R_t^{e(i)} + \varepsilon_t^{(j-i)}, \tag{9}$$

where $i, j = 1, 2, 5, 10, 20$ and $30$, $i \neq j$ and $t = 1, \ldots, T$. Based on the results summarized in Table 10, m1 and m2 are not cointegrated with m10, m20 and m30 at the 0.05 level and all the other pairs of yields are cointegrated. (iii) The non-cointegration test is applied to each possible pair of yields *across types of bonds* and *terms-to-maturity*; namely:

*Table 9.* Test for Non-Cointegration in the Taxable Bond Market

| | *g1*[a] | *g2* | *g5* | *g10* | *g20* | *g30* |
|---|---|---|---|---|---|---|
| g1 | — | | | | | |
| | — | | | | | |
| g2 | −5.4429 | — | | | | |
| | (.0004) | — | | | | |
| g5 | −4.0625 | −4.6688 | — | | | |
| | (.0239) | (.0041) | — | | | |
| g10 | −3.4776 | −4.0463 | −5.7585 | — | | |
| | (.1085) | (.0251) | (.0002) | — | | |
| g20 | −3.5900 | −4.0442 | −5.3915 | −5.5125 | — | |
| | (.0835) | (.0252) | (.0005) | (.0003) | — | |
| g30 | −3.2677 | −3.6946 | −4.6228 | −4.1652 | −9.4533 | — |
| | (.1727) | (.0651) | (.0047) | (.0178) | (.0000) | — |

*Notes:* [a] *gi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*. Each cell contains a non-cointegration test based on the regression between each possible pair of yields of taxable bonds with different terms-to-maturity; namely:

$$R_t^{\tau(j)} = \psi_0 + \psi_1 t + \psi_2 R_t^{\tau(i)} + \varepsilon_t^{(j-i)}$$

The number in each cell without the parentheses is the D.F. test statistic. The number within the parenthesis is the lower tail area.

*Table 10.* Test for Non-Cointegration in the Tax-Exempt Bond Market

| | *m1*[a] | *m2* | *m5* | *m10* | *m20* | *m30* |
|---|---|---|---|---|---|---|
| m1 | — | | | | | |
| | — | | | | | |
| m2 | −7.0562 | — | | | | |
| | (.0000) | — | | | | |
| m3 | −4.1196 | −5.4774 | — | | | |
| | (.0203) | (.0004) | — | | | |
| m10 | −3.0162 | −3.5433 | −5.9239 | — | | |
| | (.2852) | (.0932) | (.0001) | — | | |
| m20 | −2.7848 | −3.1244 | −4.2845 | −5.8827 | — | |
| | (.4199) | (.2319) | (.0126) | (.0001) | — | |
| m30 | −2.7124 | −3.0442 | −4.2286 | −5.2585 | −4.2286 | — |
| | (.4659) | (.2707) | (.0148) | (.0007) | (.0148) | — |

*Notes:* [a] *mi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*. Each cell contains a non-cointegration test based on the regression between each possible pair of yields of taxable bonds with different terms-to-maturity; namely:

$$R_t^{e(j)} = \psi_0 + \psi_1 t + \psi_2 R_t^{e(i)} + \varepsilon_t^{(j-i)}.$$

The number in each cell without the parentheses is the D.F. test statistic. The number within the parenthesis is the lower tail area.

$$R_t^{\tau(j)} = \psi_0 + \psi_1 t + \psi_2 R_t^{e(i)} + \varepsilon_t^{(j-i)}, \tag{10}$$

where $i, j = 1, 2, 5, 10, 20$ and $30$, $i \neq j$ and $t = 1, \ldots, T$. Based on Table 11, m1 is not cointegrated with g10, g20 and g30; and m2 is not cointegrated with g30 at the 0.05 level, while all the other pairs of yields are cointegrated.

*Table 11.* Test for Non-Cointegration Across the Taxable and Tax-Exempt Bond Markets

| | *m1*[a] | *m2* | *m5* | *m10* | *m20* | *m30* |
|---|---|---|---|---|---|---|
| g1[a] | −6.0661 (.0001) | | | | | |
| g2 | −5.7949 (.0002) | −6.5130 (.0000) | | | | |
| g5 | −4.6115 (.0048) | −5.5383 (.0003) | −5.8569 (.0001) | | | |
| g10 | −3.6824 (.0670) | −4.3784 (.0096) | −5.0034 (.0015) | −4.4362 (.0081) | | |
| g20 | −3.1280 (.2303) | −4.1541 (.0184) | −4.8107 (.0027) | −4.4921 (.0069) | −4.4246 (.0084) | |
| g30 | −3.1280 (.2304) | −3.6986 (.0644) | −4.2276 (.0148) | −3.9323 (.0347) | −3.8471 (.0442) | −4.0287 (.0264) |

*Notes:* [a] *gi* refers to the yield series of the government (taxable) bond with term-to-maturity *i*. Each cell contains a non-cointegration test based on the regression between each possible pair of yields of taxable bonds with different terms-to-maturity; namely:

$$R_t^{\tau(j)} = \psi_0 + \psi_1 t + \psi_2 R_t^{e(i)} + \varepsilon_t^{(j-i)}.$$

The number in each cell without the parentheses is the D.F. test statistic. The number within the parenthesis is the lower tail area.

Based on the above findings, long-term equilibria exist in all the pairs of taxable and tax-exempt yields on bonds with identical terms-to-maturity, although this relationship is somewhat weaker at the long-term end of the term structure. For both types of bonds (taxable and tax-exempt), yields from both extremes of the term structure do not exhibit a clear pairwise cointegration relationship.[13] It is not clear whether these results for the longest and shortest yields are reliable because some information has been ignored in the single equation cointegration analysis. To use more information from both extremes of the term structure, the following multivariate analysis is used.

From a statistical point of view, it is appealing for $X_t$ in equation (6) to include all the yields on both taxable and tax-exempt bonds. However, such a system exhibits high multicollinearity since yields of adjacent terms-to-maturity have similar comovements. More importantly, our focus is on the two extremes of the taxable and tax-exempt term structures to evaluate jointly the equilibrium relations between the shortest and longest yields. In this context, a more restricted version of VAR is desirable. The choice of $X_t' = [\mathrm{g1}, \mathrm{g30}, \mathrm{m1}, \mathrm{m30}]$ is useful for identifying the cointegrating vectors and possible stochastic factor(s) among the selected yields from the two extremes of the term structure.

The order of the VAR model is determined using the estimations of various lag lengths, inspection of the autocorrelation functions of the residuals, and comparisons of the Akaike (1974) and Schwarz (1978) information criteria. Doing so leads to the selection of the following VAR(1) representation:

$$X_t = \mu + \Pi_1 X_{t-1} + \varepsilon_t, \tag{11}$$

where $t = 1, 2, \ldots, T$. Subtracting $X_{t-1}$ from both sides of equation (11) yields:

*Table 12.* Autocorrelation Functions of the Residuals of VAR(1)

| Series[a] | $\hat{r}_1$[b] | $\hat{r}_2$ | $\hat{r}_3$ | $\hat{r}_4$ | $\hat{r}_5$ | $\hat{r}_6$ | $\hat{r}_7$ | $\hat{r}_8$ |
|---|---|---|---|---|---|---|---|---|
| g1 | .0495 | −.0136 | −.0415 | −.0280 | .0106 | −.0237 | −.0485 | .0227 |
| g30 | .0087 | −.0051 | −.0119 | .0007 | .0029 | .0048 | −.0044 | −.0025 |
| m1 | .0108 | −.0030 | −.0014 | .0074 | .0074 | −.0015 | −.0104 | .0002 |
| m30 | −.0093 | −.0037 | −.0070 | .0099 | .0119 | .0030 | −.0043 | −.0017 |

*Notes:* [a] *gi* refers to the yield of the government (taxable) bond with term-to-maturity *i*; and mi refers to the yield series of the municipal (tax-exempt) bond with term-to-maturity *i*.
[b] $\hat{r}_j$ refers to the sample autocorrelation functions with lag *j*.

$$\Delta X_t = \mu + \Gamma_{t-1}\Delta X_{t-1} - \Pi X_{t-2} + \varepsilon_t, \tag{12}$$

where $t = 1, 2, \ldots, T$; $\Delta X_t = X_t - X_{t-1}$; $\Gamma_{t-1} = \Pi_{t-1} - I$; and $\Pi = I - \Pi_1$. To determine if the model is well fitted, the autocorrelation functions of the residuals are computed for each equation in the system. Based on Table 12, all of the autocorrelation functions are within the confidence interval of ± .0974. Thus, the residual series are close to white noise.

The number of cointegration relations in the bond yields specified by the VAR(1) system is tested using the procedures proposed by Johansen (1988) and Johansen and Juselius (1990). The rank of the parameter matrix Π in equation (12) has the following interpretation. A full and zero rank of Π implies that the yield vector $X_t$ is stationary and an integrated process, respectively. Π with a rank *r*, which is less than full and greater than zero rank, implies that r cointegrating relations exist. In this case, Π can be expressed as a product of two matrices, $\Pi = \omega\beta'$. The first is a $(p \times r)$ matrix, ω, and the second is a $(p \times r)$ matrix of cointegrating vectors, β. The number of stochastic trends in the VAR(1) system is represented by $p - r$, which is the number of factors that cannot be represented by cointegrating vectors. The parameters ω and β cannot be estimated since they form an overparametrization of the model.[14] However, we can rely on the data to determine the cointegrating space as is explained later in the paper.

If $\mu \neq 0$ and μ is unrestricted in equation (12), then a constant trend exists in the process. If $\mu \neq 0$ and μ is restricted (i.e., $\mu = \omega\beta_0'$), then no constant trend exists in the process. To test the number of cointegration relations jointly, the likelihood-ratio test is used. The empirical percentile tables of the test statistic for both restricted and unrestricted μ are given in Johansen and Juselius (1990). A detailed description of the test procedure is given in Appendix B.

The test results for the models with unrestricted and restricted μ are reported in Tables 13 and 14, and in Tables 15 and 16, respectively. The computed eigenvalues and their normalized eigenvectors for the unrestricted (restricted) models are denoted as $\hat{\lambda}(\hat{\lambda}_i^*)$ and $\hat{e}_i(\hat{e}_i^*)$, respectively. The results of the likelihood ratio test summarized in Table 14 show that the null hypotheses that at most two or three cointegrating vectors exist cannot be rejected at the 0.05 level. Based on Table 16, the null hypotheses that at most one, two, or three cointegrating vectors exist cannot be rejected at the 0.05 level.

The hypothesis that no constant trend exists in the process is tested using another test statistic proposed by Johansen and Juselius (1990). When three cointegrating vectors are assumed to exist, the test statistic is $-T\ln[(1 - \hat{\lambda}_4^*)/(1 - \hat{\lambda}_4) = .25$, which has a $\chi^2(1)$ distribution. When two cointegrating vectors are assumed, the test statistic is

*Table 13.* Eigenvalues $\hat{\lambda}_i$ and Eigenvectors of $\hat{e}_i$ $|\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}| = 0$, μ Unrestricted

| Coef. | $\hat{\lambda}_i$ .1999 $\hat{e}_i$ | .0920 | .0310 | .0021 |
|---|---|---|---|---|
| g1 | 1.0000 | −0.2680 | 0.1121 | −0.0162 |
| g30 | 0.1620 | 1.0000 | 0.5054 | 0.2441 |
| m1 | −0.1907 | −0.4566 | 1.0000 | −0.1784 |
| m30 | 0.0574 | −0.3299 | 0.0569 | 1.0000 |

*Note:* Eigenvectors are normalized here.

*Table 14.* Test Results for $H_0$: There are at Most $r$ Cointegration Vectors; μ Unrestricted

| Test | | Percentiles | | |
|---|---|---|---|---|
| $H_0$ | $-2\ln(Q)$ | .900 | .950 | .975 |
| $r \leq 3$ | .86 | 6.69 | 8.08 | 9.65 |
| $r \leq 2$ | 14.10 | 15.58 | 17.84 | 19.61 |
| $r \leq 1$ | 54.64 | 28.43 | 31.25 | 34.06 |
| $r = 0$ | 148.30 | 45.24 | 48.41 | 51.80 |

*Table 15.* Eigenvalues $\hat{\lambda}_i^*$ and Eigenvectors $\hat{e}_i^*$ of $|\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}| = 0$, μ Unrestricted

| Coef. | $\hat{\lambda}_i^*$ .1999 $\hat{e}_i^*$ | .0921 | .0328 | .0027 | .0000 |
|---|---|---|---|---|---|
| g1 | 1.0000 | −0.2680 | 0.1053 | −0.0429 | −0.0358 |
| g30 | 0.1619 | 1.0000 | 0.5219 | 0.1959 | −0.2143 |
| m1 | −0.1901 | −0.4615 | 1.0000 | −0.3176 | −0.1268 |
| m30 | −0.0572 | −0.3313 | 0.0634 | 1.0000 | −0.5627 |
| 1 | 0.0142 | −0.0403 | 0.2781 | 0.5628 | 1.0000 |

*Note:* Eigenvectors are normalized here.

*Table 16.* Test Results for $H_0$: There are at Most $r$ Cointegration Vectors; μ Restricted

| Test | | Percentiles | | |
|---|---|---|---|---|
| $H_0$ | $-2\ln(Q)$ | .900 | .950 | .975 |
| $r \leq 3$ | .00 | 7.56 | 9.09 | 10.70 |
| $r \leq 2$ | 1.13 | 17.95 | 20.16 | 22.20 |
| $r \leq 1$ | 15.13 | 32.09 | 35.06 | 37.60 |
| $r = 0$ | 55.71 | 49.92 | 53.34 | 56.44 |

$(-T\sum_{i=3}^{4}\ln[(1-\hat{\lambda}_i^*)/(1-\hat{\lambda}_i)]) = 1.03$, which has a $\chi^2(2)$ distribution. Given the test statistics for either the two or three cointegrating vector cases, the null hypothesis that no constant trend exists in the process cannot be rejected. Thus, the process with no constant trend is assumed, and the number of cointegrating vectors is at most three. Thus, at least one stochastic factor cannot be represented by the cointegrating vectors.

Although β can not be estimated in the usual way, the cointegrating space represented by β can be represented by the following choice (see Johansen & Juselius, 1990, p. 177):

$$\hat{\beta} = (\hat{e}_1, \ldots, \hat{e}_r),$$

for a model with unrestricted μ, and

$$\hat{\beta}^* = (\hat{e}_1^*, \ldots, \hat{e}_r^*),$$

for a model with restricted μ, where $\hat{e}_1, \ldots, \hat{e}_r$ and $\hat{e}_1^*, \ldots, \hat{e}_r^*$ are first r normalized eigenvectors from two different models. Since the model selected for VAR(1) is the one with μ restricted, $\hat{\beta}^*$ is the suitable representation of the cointegrating space.

In Table 15, the first three columns of the normalized eigenvectors can be interpreted as the representation of the cointegrating vectors $\hat{\beta}$ for g1, g30, m1, m30 and the constant term 1. These cointegrating vectors reflect not only the wide links among four yields from both extremes of the term structure but also the measures of the relative strength of the links. For example, the first cointegrating vector is [1.000, 0.1619, –0.1901, –0.052, 0.0142]′, which shows that g1 is closely linked to m1 and g30. This may reflect the fact that g1 and m1 are both one-year yields, and g1 and g30 coexist in the same taxable bond market. The second cointegrating vector, [–0.2680, 1.0000, –0.4615, –0.3313, –0.0403]′, reflects that g30 is closely connected with m1, m30, and g1. This may reflect the fact that g30 and m30 are both thirty-year yields, and g1 and g30 coexist in the same taxable bond market. The third cointegrating vector, [0.1053, 0.5219, 1.0000, 0.0634, 0.2781]′, demonstrates that m1 has stronger links to g30 and g1 but a weaker one to m30. One important observation is about the cointegration between the yields from both extremes of the term structure. While the single equation analysis normally treats the weaker relation as a sign of no cointegration, the multivariate analysis utilizes more information and couches individual weaker relations in a broader cointegrating space. Although the multivariate analysis shows that the strength of links among variables in the VAR model does vary considerably, it provides further evidence that long-term equilibria indeed exist among the yields from both extremes of the term structure, and that there are at least three meaningful cointegrating vectors.

## V. CONCLUDING REMARKS

This paper extends the existing literature by examining the cointegration relations among taxable (tax-exempt) yields across a complete term structure over a longer period of time; taxable and tax-exempt yields with same (different) terms-to-maturity across the whole term structure over an extended period of time; and taxable and tax-exempt yields from both extremes of the term structure in a vector autoregression (VAR) framework. We find that yields of bonds with the same and near proximity terms-to-maturity are cointegrated and have long-term equilibria regardless of the type of bond (i.e., taxable or tax-exempt). This relationship becomes somewhat weaker at the long-term end of the term structure. In addition, the single equation analysis indicates that the very long-term (thirty-year term-to-maturity) yields are not closely cointegrated with the very short-term (one-year term-to-maturity) counterparts for both types of bonds. To fully employ the information available, we adopt the VAR model to evaluate the cointegration relations among yields from the two extremes of the term structure. The Johansen-type estimation and tests indicate that there are at most three cointegrating vectors for yields g1, g30, m1 and m30. These cointegrating vectors suggest that g1, g30, m1 and m30 are cointegrated and have long-term equilibria. However, there is at least one stochastic factor which cannot be represented by cointegrating vectors in the VAR representation of the long- and short-term yields across the types of bonds.

## ACKNOWLEDGMENTS

Research support from Fonds pour la formation de chercheurs et l'aide á la researche (FCAR), Social Sciences and Humanities Research Council of Canada (SSHRC), Concordia Professional Development Fund, and Dalhousie University Research Fund are gratefully acknowledged. We would like to thank an anonymous referee for very helpful comments on an earlier version of this paper. We would also thank Monique Comeau and Heather Lennox for their excellent secretarial assistance. We are of course responsible for any remaining errors.

## APPENDIX A: THE DICKEY-PANTULA TEST PROCEDURE

Let $x_t$ be a time series that satisfies:

$$x_t = \sum_{i=1}^{p} \alpha_i x_{t-i} + \varepsilon_t,$$

where $t = 1, \ldots, T$; and $p$ is equal to 3 for our purposes. This equation is reparameterized as:

$$y_t = \eta_1 x_{t-1} + \eta_2 z_{t-1} + \eta_3 w_{t-1} + \varepsilon_t,$$

where $t = 1, \ldots, T$; $z_{t-1} = x_t - x_{t-1}$, $w_{t-1} = z_t - z_{t-1}$, and $y_t = w_t - w_{t-1}$. That is, $z_t$, $w_t$, and $y_t$ are first, second, and third differences of the series $x_t$, respectively. It can be shown that $\alpha_1 = 3 + \eta_1 + \eta_2 + \eta_3$, $\alpha_2 = -(3 + \eta_2 + 2\eta_3)$, $\alpha_3 = 1 + \eta_3$, $\eta_1 = -(1 - r_1)(1 - r_2)(1 - r_3)$, $\eta_2 = 2\eta_1 - (1 - r_1)(1 - r_2) - (1 - r_2)(1 - r_3) - (1 - r_3)(1 - r_3)$, and $\eta_3 = r_1 r_2 r_3 - 1$, where $r_1$, $r_2$ and $r_3$ are the roots of the characteristic equation $r^3 - \alpha_1 r^2 - \alpha_2 r - \alpha_3 = 0$.

The tests can be conducted on various hypotheses. The hypothesis of three stationary roots can be expressed as $H_3$: $\eta_1 = \eta_2 = \eta_3 = 0$. The hypothesis of two unit roots and one stationary root can be expressed as $H_2$: $\eta_1 = \eta_2 = 0$, $\eta_3 < 0$. The hypothesis of exactly one unit root and two stationary roots can be expressed as $H_1$: $\eta_1 = 0$, $\eta_2 < 0$ and some restrictions on $\eta_2$ and $\eta_3$ (viz. $-2 < \eta_3 < 0$, $0 < 4 + \eta_2 + 2\eta_3$). The hypothesis of three stationary roots is expressed as $H_0$: $\eta_1 < 0$ and some restrictions on $\eta_2$ and $\eta_3$(viz. $-12 < \eta_2 + 2\eta_1 < 0$, $-2 < \eta_3 < 0$, etc..). See Dickey and Pantula (1987) for details.

Dickey and Pantula note that using the t statistics, $t_{i,T}(3)(i = 1, 2, 3)$, from the regression of $y_t$ on $x_{t-1}$, $z_{t-1}$, and $w_{t-1}$ is not appropriate because the asymptotic distribution of

$t_{1,T}(3)$ for testing one unit root depends on the number of unit roots present in the third-order autoregressive representation of the stochastic process. They also note that the sequential $F$ test procedure proposed by Pantula (1986) has less power than the sequential pseudo $t$ test procedure.

The pseudo $t^*_{i,T}(p)$ is computed and compared with $\hat{\tau}_{T,\alpha}$ given by Fuller (1976). $t^*_{i,T}(p)$ is the $t$ statistic for the coefficient of $(1-L)^{i-1}x_{t-1}$ in the regression of $(1-L)^p x_{t-1}$ on $(1-L)^{i-1}x_{t-1}$, $(1-L)^i x_{t-1}$, ..., and $(1-L)^{p-1}x_{t-1}$, where $L$ is the lag-operator. Under the null hypothesis $H_i$ and the alternative hypothesis $H_{i-1}$, a common set of restrictions is implicit when $i > 1$, i.e., $\eta_1 = ... = \eta_{i-1} = 0$. The pseudo $t^*_{i,T}(p)$ for $\eta_i = 0$ is computed for $i = 3, 2, 1$. The specific test procedure is as follows:

In step one, test the null hypothesis $H_3$: $\eta_1 = \eta_2 = \eta_3 = 0$ against the alternative hypothesis $H_2$: $\eta_1 = \eta_2 = 0$, $\eta_3 < 0$. The common set of restrictions is $\eta_1 = \eta_2 = 0$. $t^*_{3,T}(3)$ is computed in the regression of $y_t$ on $\omega_{t-1}$. If $t^*_{3,T}(3) \leq \hat{\tau}_{T,\alpha}$, $H_3$ of three unit roots is rejected and go to step two. Otherwise, conclude that $H_3$ is true.

In step two, test the null hypothesis $H_2$: $\eta_1 = \eta_2 = 0$, $\eta_3 < 0$ against the alternative hypothesis $H_1$: $\eta_1 = 0$, $\eta_2 < 0$. The common restriction is $\eta_1 = 0$. $t^*_{2,T}(3)$ is computed in the regression of $y_t$ on $z_{t-1}$ and $w_{t-1}$. If $t^*_{3,T}(3) \leq \hat{\tau}_{T,\alpha}$, and $t^*_{2,T}(3) \leq \hat{\tau}_{T,\alpha}$, reject $H_2$ of exactly two unit roots and go to step three. Otherwise, conclude that $H_2$ is true.

In step three, test the null hypothesis $H_1$: $\eta_1 = 0$, $\eta_2 < 0$ against the alternative hypothesis $H_0$: $\eta_1 < 0$ and some restrictions on $\eta_2$ and $\eta_3$. Based on the regression of $y_t$ on $x_{t-1}$, $z_{t-1}$ and $w_{t-1}$, compute $t^*_{1,T}(3)$. If $t^*_{1,T}(3) \leq \hat{\tau}_{T,\alpha}$, $(i = 1, 2, 3)$, reject $H_1$ of exactly one unit root in favor of $H_0$ of no unit roots. Otherwise, conclude that $H_1$ is true.

For more details, see Dickey and Pantula (1987).

## APPENDIX B: THE JOHANSEN TEST PROCEDURE

The following computation procedure is used to compute the likelihood-ratio test statistic for testing the number of cointegrating vectors in the VAR model:

First, compute two matrices of residuals, $R_{0t}$ from the regression of $\Delta X_t$ on $\Delta X_{t-1}$, and $R_{kt}$ from the regression of $X_{t-k}$ on $\Delta X_{t-1}$. For the case where $\mu \neq 0$ and $\mu$ is unrestricted, a constant one is added as an additional regressor to both regression equations. For the case where $\mu \neq 0$ and $\mu$ is restricted (i.e., $\mu = \omega\beta'_0$), a constant 1 is stacked onto $X_{t-k}$ (i.e., $X^*_{t-k} = [X'_{t-k}, 1]'$), and $X_{t-k}$ is replaced with $X^*_{t-k}$.

Second, the moment matrices $S_{00}$, $S_{kk}$ and $S_{k0}$, where $S_{ij} = T^{-1}\sum_{t=1}^{T} R_{it}R'_{jt}$ with $i,j \in \{0, k\}$, are obtained. Then, the equation $|\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}| = 0$ is solved for $p$ eigenvalues $\lambda_i$ and $p$ eigenvectors.

Third, the following likelihood-ratio test statistic is formed to test $H_0$: there are at most $r$ cointegrating vectors:

$$-2\ln(Q) = -T \sum_{i=r+1}^{p} \ln(1-\lambda_i),$$

where $\lambda_{r+1}, \ldots, \lambda_p$ are the $p - r$ smallest eigenvalues.

For more details, see Johansen and Juselius (1990).

## NOTES

1. Hall, Anderson and Granger (1992) conduct a cointegration analysis of Treasury Bill yields.

2. Although Bradley and Lumpkin (1992) examine cointegration relations for Treasury securities with terms-to-maturity of three months to thirty years, they exclude twenty-year bonds.

3. While other models may be used for this purpose, they often impose more restrictions to test the expectation hypothesis rather than to find long-term equilibria. For example, Shea (1992) uses the linearized term-structure equation of Campbell and Shiller (1987), and finds that the model is too restrictive for the data. Similarly, Hall, Anderson and Granger (1992) propose a similar model for testing cointegration among treasury bills, which is restricted to an exact one-to-one correspondence (i.e., the coefficients of the cointegration vectors are either positive one or negative one).

4. Engle and Granger (1987) provide the definition of cointegration.

5. Given a choice among coupons, the yields of higher coupon issues in the longer maturities are chosen. For greater details, see *Analytical Record of Yields and Yield Spreads*.

6. Engle and Granger (1987) use the period from February 1952 to December 1982, and Campbell and Shiller (1987) use the period from January 1959 to October 1983.

7. Whenever a methodological issue with respect to these series is discussed, the generic identifiers $x$ and $X$ are used to denote a bond yield series and a vector of such series, respectively. The differenced $x_t$ is given by $\Delta x_t = x_t - x_{t-1}$.

8. A stationary series $x$ is said to be integrated of order zero, and denoted as $x \sim I(0)$. If a series, $x$, needs to be differenced $d$ times to be $I(0)$, it is said to be integrated of order $d$, and denoted as $x \sim I(d)$.

9. To employ the test, each series, denoted as $x_t$, is first differenced, and then the differenced series $\Delta x_t$ is regressed on a constant term, $\mu$; the time trend, $t$; the lagged original series, $x_{t-1}$; and the lagged differenced series, $\Delta x_{t-i}$, $i = 1,\ldots,k$. Specifically:

$$\Delta x_t = \mu + \gamma t + (\rho - 1)x_{t-1} + \sum_{i=1}^{k} \phi_i \Delta x_{t-i} + \varepsilon_t,$$

where $t = 1, \ldots, T$. The value of $k$ is determined using model selection procedures such as the Akaike (1974) and Schwarz (1978) information criteria, and the Durbin and Watson or Durbin h statistics.

10. The ZA test procedure for a yield series is given below. The model for the series is:

$$\Delta x_t = \mu + \gamma t + \varsigma DU_t(\xi) + (\rho - 1)x_{t-1} + \sum_{i=1}^{k} \phi_i \Delta x_{t-i} + \varepsilon_t,$$

where $t = 1, ..., T$; and all the other variables and parameters are as defined before. $\varsigma$ is a coefficient; and $DU_t(\xi)$ is a dummy variable. $DU_t(\xi) = 1$ if $t > T\xi$, and $DU_t(\xi) = 0$ otherwise. In other words, the model allows for a flexible point of change in levels. $\xi = T_B/T$, where $T_B$ is a break point, and $T$ is the number of observations. $\xi$ ranges from $2/T$ to $(T-1)/T$. $\hat{\xi}$ is obtained by selecting the estimated regression which provides the lowest $t$ statistic, $t(\hat{\rho} - 1)$ given $\xi \in \Xi = [.001, .999]$. This model (A) is chosen because the preliminary tests of Models B and C in Zivot and Andrews indicate that they are not appropriate representations for yield series. Zivot and Andrews also find that Model A is a useful representation for interest rates. The $k$ is determined by eliminating all of the lagged changes associated with a coefficient less than 1.6 as in Zivot and Andrews (1992). Since the limiting distribution is unknown for this model, Zivot and Andrews (1992) provide the empirical percentile tables for the test statistics.

11. The sample size, $T$, is 421. However, $\hat{\tau}_{T,\alpha}$ is available only for $T = 250$ and 500 in Fuller (1976). For example, $\hat{\tau}_{250,.05} = -1.95$, $\hat{\tau}_{500,.05} = -1.95$, $\hat{\tau}_{250,.95} = 1.29$, and $\hat{\tau}_{500,.95} = 1.28$. $\hat{\tau}_{500,.05}$ is used, as $T$ is closer to 500.

12. The least squares estimator of a cointegration regression is "superconsistent" (Stock, 1987).

13. Bradley and Lumpkin (1992) present a cointegration equation with the yield on seven-year treasury notes as the dependent variable and the yields of other short- and long-term treasury bills and notes as the independent variables. However, as shown in their Appendix Table A1, the $t$-statistics for the coefficient estimates associated with the yields at both the short- and long-term ends of the term structure are very low and statistically insignificant, and the $t$-statistics for the five- and seven-year yields are very high and statistically significant. Their results implicitly confirm the results reported herein.

14. See Johansen (1988), and Johansen and Juselius (1990) for a detailed discussion.

## REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC-19, 716–723.

Bernanke, B., & Mishkin, F. (1992). Central bank behavior and the strategy of monetary policy: Observations from six industrialized countries. *NBER Macroeconomics Annual 1992*, The MIT Press.

Bradley, M.G., & Lumpkin, S.A. (1992). The treasury yield curve as a cointegrated system. *Journal of Financial and Quantitative Analysis*, *27*, 449–463.

Buser, S.A., & Hess, P.J. (1986). Empirical determinants of the relative yields on taxable and tax-exempt securities. *Journal of Financial Economics*, *17*, 335–355.

Campbell, J.Y., & Shiller, R.J. (1984). A simple account of the behavior of long-term interest tates. *American Economic Review*, *74*, 44–48.

Campbell, J.Y., & Shiller, R.J. (1987). Cointegration and tests of present value models. *Journal of Political Economy*, *95*, 1062–1088.

Clarida, R.H., & Friedman, B.M. (1984). The behavior of U.S. short-term interest rates since October 1979. *Journal of Finance*, *39*, 671–682.

Dickey, D.A., & Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427–431.

Dickey, D.A., & Pantula, S.G. (1987). Determining the order of differencing in autoregressive processes. *Journal of Business and Economic Statistics*, *5*, 455–461.

Engle, R.F., & Granger, C.W.J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*, 251–276.

Engle, R.F., Lilien, D.M., & Robins, R.P. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica*, *55*, 391–407.

Fuller, W.A. (1976). *Introduction to statistical time series.* John Wiley & Sons, Chapter 8, Table 8.5.1 (p.371) and Table 8.5.2 (p.373).

Green, R.C. (1993). A simple model of the taxable and tax-exempt yield curves. *The Review of Financial Studies*, *2*, 233–264.

Hall, A.D., Anderson, H.M., & Granger, C.W.J. (1992). A cointegration analysis of treasury bill yields. *Review of Economics and Statistics*, *74*, 116–126.

Huizinga, J., & Mishkin, F.S. (1984). Inflation and real interest rates on assets with different risk characteristics. *Journal of Finance*, *39*, 699–712.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration-with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, *52*, 169–209.

Kochin, L.A., & Parks, R.W. (1988). Was the tax-exempt bond market inefficient or were future expected tax rates negative? *Journal of Finance*, *38*, 913–931.

Kryzanowski, L., Xu, K., & Zhang, H. (1995). Determinants of the decreasing term structure of relative yield spreads for taxable and tax-exempt bonds. *Applied Economics*, *27*, 583–590.

Macaulay, F.R. (1938). Some theoretical problems suggested by the movements of interest rates, bond yields, and stock prices in the United States since 1856. *National Bureau of Economic Research.*

MacKinnon, J. (1991). Critical values for cointegration tests. In Engle, R.F., & Granger, C.W.J. (Ed.), *Long-run economic relationships.* Oxford University Press, 267–276.

Mussa, M.L., & Kormendi, R.C. (1979). *The taxation of municipal bonds.* American Enterprise Institute for Public Policy Research, Washington, D.C.

Pantula, S.G. (1986). Estimation for autoregressive processes with several unit roots. Unpublished manuscript, North Carolina State University, Department of Statistics.

Poterba, J. (1989). Tax reform and the market for tax-exempt debt. *Regional Science and Urban Economics*, 19, 537–562.

Shea, G.S. (1992). Benchmarking the expectations hypothesis of the interest-rate term structure: An analysis of cointegration vectors. *Journal of Business and Economic Statistics*, *10*, 347–366.

Shiller, R.J. (1990). The term structure of interest rates. In Friedman, B.M., & Hahn, F.H *(Eds.), Handbook of monetary economics*, Vol. 1. 425–722.

Shiller, R.J., Campbell, J.Y., & Schoenholtz, K.L. (1983). Forward rates and future policy: Interpreting the term structure of interest rates. *Brookings Papers on Economic Activity*, *1*, 173–217.

Skelton, J.L. (1983). Banks, firms and the relative pricing of tax-exempt and taxable bonds. *Journal of Financial Economics*, *12*, 343–355.

Stock, J. (1987). Asymptotic properties of least squares estimators of cointegrating vectors. *Econometrica*, *55*, 1035–1056.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.

Trzcinka, C. (1982). The pricing of tax-exempt bonds and the Miller hypothesis. *Journal of Finance*, *37*, 907–923.

Zivot, E., & D.W. Andrews (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, *10*, 251–270.