

**Finance and Economics Discussion Series  
Divisions of Research & Statistics and Monetary Affairs  
Federal Reserve Board, Washington, D.C.**

**Linear Cointegration of Nonlinear Time Series  
with an Application to Interest Rate Dynamics**

Barry E. Jones and Travis D. Nesmith

**2007-03**

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

# Linear Cointegration of Nonlinear Time Series with an Application to Interest Rate Dynamics

Barry E. Jones

Binghamton University

Travis D. Nesmith\*

Board of Governors of the Federal Reserve System

November 29, 2006

## Abstract

We derive a definition of linear cointegration for nonlinear stochastic processes using a martingale representation theorem. The result shows that stationary linear cointegrations can exhibit nonlinear dynamics, in contrast with the normal assumption of linearity. We propose a sequential nonparametric method to test first for cointegration and second for nonlinear dynamics in the cointegrated system. We apply this method to weekly US interest rates constructed using a multirate filter rather than averaging. The Treasury Bill, Commercial Paper and Federal Funds rates are cointegrated, with two cointegrating vectors. Both cointegrations behave nonlinearly. Consequently, linear models will not fully replicate the dynamics of monetary policy transmission.

*JEL Classification:* C14; C32; C51; C82; E4

*Keywords:* cointegration; nonlinearity; interest rates; nonparametric estimation

---

\*Corresponding author: 20th and C Sts., NW, Mail Stop 188, Washington, DC 20551

E-mail: [travis.d.nesmith@frb.gov](mailto:travis.d.nesmith@frb.gov)

Melvin Hinich provided technical advice on his bispectrum computer program. Hermann Bierens provided technical advice on implementing his nonparametric cointegration tests. We thank William Barnett, Florenz Plassmann, Eric Verhoogen and seminar participants at George Washington University, the IMF Institute, the 2006 Midwest Macroeconomics Conference, and the 2006 Udine Workshop for helpful comments and suggestions. We also thank Samia Husain for research assistance. The views presented are solely those of the authors and do not necessarily represent those of the Federal Reserve Board or its staff. Any remaining errors are the responsibility of the authors.

# 1 Introduction

Cointegration is the primary econometric model of system dynamics for nonstationary time series. Cointegration is normally defined as the existence of a stationary linear combination of nonstationary time series. The fact that the combination is linear does not necessarily imply linear dynamics for the resulting stationary stochastic process. Cointegration is nevertheless strongly associated with linear dynamics, because cointegration was initially developed within the linear Box-Jenkins framework. In particular, the standard model of cointegration—the vector error correction model (VECM)—does assume linear dynamics. Linearity makes econometric models tractable, but linear models can only reproduce a restricted class of dynamic behavior. Most economic models are nonlinear: producing richer dynamics.

A broader definition of cointegration is necessary in order to incorporate nonlinear dynamics. Our motivation for broadening the class of dynamics is based on the simple observation that nonlinearity is a dominant property in the sense that a linear combination of nonlinear processes is itself generally nonlinear. Nonstationarity is similarly dominant. Cointegration is a special case where adding two or more nonstationary processes together results in a stationary process. But if any of the cointegrated series are nonlinear, the linear combination generally produces a nonlinear stationary process. For example, let  $x_t = x_{t-1} + \varepsilon_t$  so that  $x_t$  is a random walk and let  $y_t = x_t + z_t$  where  $z_t$  is a stationary and nonlinear stochastic process. Then  $[1 \quad -1]^T$  is a linear cointegrating vector for  $[y_t \quad x_t]^T$  as  $y_t - x_t = z_t$  and  $z_t$  is stationary. Since  $z_t$  is nonlinear, the cointegrating relation  $y_t - x_t$  is a nonlinear stochastic process.

We developed this motivation in Barnett, Jones, and Nesmith (2000), mainly as a critique of Johansen’s maximum likelihood cointegration estimator which assumes linearity.<sup>1</sup> Although we start from the same observation—that linear combinations of nonlinear series are generally nonlinear—this paper is more constructive. We derive a definition of linear cointegration from Hall and Heyde’s (1980) martingale representation theorem for stationary stochastic processes. This extended definition, suggested by Bierens (1997), does not *a priori* restrict the dynamic behavior to be linear as did previous definitions of cointegration. As a result, the stationary dynamics of the cointegrated system may exhibit nonlinear dynamics. We also develop an asymptotically valid procedure to test linear cointegrations for the nonlinear stationary dynamics.

Our definition of cointegration, and the associated concept of nonlinear dynamics, differs from nonlinear cointegration introduced by Granger (1991). Intuitively, non-

---

<sup>1</sup>Johansen’s estimator assumes linearity because it is based on a VECM model driven by Gaussian innovations. Calza and Sousa (2006) cite our critique in their rejection of Johansen’s estimator.

linear cointegration occurs when a nontrivial nonlinear combination of nonstationary time series is stationary. In contrast, our extended definition still uses linear combinations to produce stationarity. The definition of nonlinear cointegration actually says nothing about the dynamics of the resulting stationary process, which is our focus. In practice, nonlinear cointegration has been defined by a VECM model with a nonlinear error correction term. Consequently, nonlinear cointegration has been predicated on the assumption that the stationary dynamics are linear. The martingale based definition could be further extended to allow for nonlinear cointegration with potentially nonlinear stationary dynamics, but testing for such complicated dynamics would be challenging.

To test for the presence of nonlinear dynamics in a linearly cointegrated system, we implement a sequential procedure. In the first stage, cointegrating vectors are estimated using Bierens's (1997, 2005) nonparametric test for cointegration. Although Bierens (1997) assumed linearity for clearer exposition, the test is still valid when applied to nonlinear processes. If a cointegration is found, it defines a new stationary process representing the long-run economic equilibrium. In the second stage, the stationary cointegration is tested for nonlinearity. At this stage we are testing a system of economic variables, or an equilibrium economic relation, for nonlinear dynamics even though existing tests for nonlinearity are univariate.

The nonlinearity test used in the second stage should be conservative. A conservative test reduces the chances of finding nonlinearity due to inappropriately accepting the hypothesis of stationarity in the first stage. The possibility that a nonstationary linear time series could be identified as nonlinear is not a new problem. Tests for nonlinearity require stationarity as a maintained assumption. Given the strong evidence that many economic series are nonstationary, this requirement implies that tests for nonlinearity are almost always conditional on the correct removal of nonstationarity, for example through correct detrending or differencing. Failure to remove any nonstationarity can lead to spurious acceptance of nonlinearity (Lee, Kim, and Newbold, 2005/6).

Based on Monte Carlo comparisons of various tests for nonlinearity, we use Hinich's (1982) nonparametric bispectral test, which was found to be conservative.<sup>2</sup> We do not, however, implement the surrogate data and bootstrap methods introduced by Hinich, Mendes, and Stone (2005) to improve the power of Hinich's test, as the theoretical validity of the sequential testing rests on an asymptotic argument. Furthermore, bootstrapping only the second-stage estimator would be inappropriate.<sup>3</sup>

---

<sup>2</sup>See Barnett, Gallant, Hinich, Jungeilges, Kaplan, and Jensen (1995), Barnett, Gallant, Hinich, Jensen, and Jungeilges (1996) and Barnett, Gallant, Hinich, Jungeilges, Kaplan, and Jensen (1997),

<sup>3</sup>There are also potential problems with applying surrogate methods to testing for nonlinearity. The

To demonstrate the two-stage nonparametric testing method, we test a system of short-term U.S. interest rates: specifically the rates for short-term Commercial Paper, short-term Treasury Bills, and Federal Funds. Short-term interest rates on Federal Funds and on unsecured corporate and government debts are frequently included in studies of the business cycle, money demand, and the monetary transmission mechanism. Since short-term interest rates are likely to respond more quickly to monetary policy than other economic variables, the dynamic interaction between Federal Funds and other short-term rates is critical to understanding how changes in monetary policy are transmitted through the economy.

Besides their economic relevance, interest rates are available on a daily basis for a long period of time and the nonparametric tests perform better with more data. Using the business day data does, however cause difficulties with missing values due to holidays. To avoid this problem, we sample the daily data at a weekly frequency, but only after appropriately filtering the data to prevent aliasing. This multirate filter, produced by applying the anti-aliasing filter and resampling, appears to be a new approach in econometrics and improves the performance of both Bierens' and Hinich's tests.

Correcting for aliasing may also be the reason we unequivocally find that U.S. interest rates contain a unit root. Whether or not interest rates contain unit roots has been heavily debated. Although many authors have found that U.S. interest rates are integrated (Nelson and Plosser, 1982, Psaradakis, Sola, and Spagnolo, 2006, Rapach and Weber, 2004, Rose, 1988), other research has suggested that interest rates are better described as long-memory or fractionally integrated series (Backus and Zin, 1993, Gil-Alana, 2004, Tsay, 2000). The empirical case for long-memory is usually based on conflicting results from various tests of the unit root and stationarity hypotheses. In contrast, we find uniform agreement among a variety of univariate tests that the levels of the interest rates are nonstationary and the differences are stationary. Other authors, such as Pfann, Schotman, and Tschernig (1996) and Maki (2003), have suggested that interest rates exhibit nonlinear dynamics which affect the power of stationarity tests. Our results do not seem to suffer from low power, despite finding evidence of nonlinearity.

Since the interest rates are both integrated and nonlinear, we apply our two-stage method. Bierens' nonparametric test shows that the weekly interest rates are linearly cointegrated. For comparison, we also perform Johansen (1988)'s standard parametric tests. The nonparametric and parametric results are very similar: identifying the same two cointegrating vectors. In addition, the cointegration estimates are robust to remov-

---

method developed by Hinich et al. (2005) solves these problems for a large subset of univariate linear processes. But their method needs to be extended before it can be applied to systems of cointegrated variables.

ing the period starting near the third quarter of 1979 through the first quarter of 1984 when the Federal Reserve shifted its monetary policy instrument away from interest rates (Rudebusch, 1995). Estimates for the prior and subsequent sub-sample find the same cointegrating vectors as the estimates for the full sample.

After identifying two cointegrations in the first stage, we subsequently test each cointegration for nonlinearity. Linearity is rejected for both using the full data set. This result is not completely robust as linearity can be accepted for the first sub-sample. This result may stem from the reduction in the power of Hinich's test that stems from the relatively short span of data. However linearity can be rejected for the second sub-sample, which suggests that the nonlinearity is not produced only by switching regimes.<sup>4</sup>

We conclude that stationary interest rate dynamics are nonlinear. A simple explanation is that the adjustment mechanism that corrects deviations from the long-run interest rate equilibrium is nonlinear. Since we find two cointegrations, it is possible that nonlinearity also describes movements within the cointegration space. Regardless of whether nonlinearity can be isolated as a disequilibrium phenomena or not, the equilibrium dynamics are not simply characterized by the individual dynamics as would be expected from a linear system. This complexity points to the need for further work modeling the interest rate dynamics.

The paper is organized as follows. Section 2 clarifies the difference between linearity and nonlinearity for stationary processes. We also discuss the bispectrum to provide intuition for Hinich's test. Section 3 contains the theoretical contribution. Using a martingale representation for integrated processes, we derive a definition of cointegration that is applicable to nonlinear stochastic processes. This extended definition of linear cointegration is compared to the standard VECM model of both linear and nonlinear cointegration. Section 4 contains the empirical results and Section 5 concludes. The appendix reviews the aliasing problem and our multirate anti-aliasing filter design.

## 2 Nonlinear Processes

Whether a process is linear or nonlinear is determined by its serial dependence structure. For stationary processes, the difference between linear and nonlinear dynamics can be clarified by looking at the restrictions implied by linearity for both the Wold decomposition and the Volterra representation. The discussion in this section assumes

---

<sup>4</sup>Several authors have suggested regime shifts as a source on nonlinearity in interest rates. See Pfann et al. (1996) for a discussion.

stationarity. We defer formally defining stationarity until the discussion of integration and cointegration in the next section.

Under mild regularity conditions, a stationary stochastic process  $X_t$  has a representation of the form:

$$X_t = \sum_{u=-\infty}^{\infty} g_u \epsilon_{t-u}, \quad (2.1)$$

where  $g_u$  is a sequence of coefficients, and  $\epsilon_t$  is a serially uncorrelated white noise input sequence. This is a consequence of the Wold decomposition theorem. The Wold decomposition therefore shows that a stationary process, such as that produced by cointegration, can be represented as the output of a moving average filter applied to uncorrelated white noise input.

At first glance, this representation seems to suggest that every stationary process can be represented as an infinite-order moving average process. This impression is misleading. The process may be nonlinear because the input process is uncorrelated but is not necessarily stochastically independent.  $X_t$  can be represented as the output of a time-invariant linear filter applied to white noise input, but  $X_t$  is a *linear process* only if  $\epsilon_t$  is stochastically independent.<sup>5</sup> In general, whiteness is not sufficient for stochastic independence unless the white noise sequence is Gaussian.

For linear models, the coefficients of the moving average representation completely characterize the effect of a shock. The response of a linear sequence to a shock is completely characterized by the *transfer function* of the filter:

$$G(f) = \sum_{u=-\infty}^{\infty} g_u e^{-i(2\pi f)u}. \quad (2.2)$$

If the input to a linear sequence is a sine wave of frequency  $f$ , the output will also be a sine wave with frequency  $f$ . The amplitude will be scaled by  $|G(f)|$ , and the phase will be shifted by  $\tan^{-1}(\text{Im } G(f)/\text{Re } G(f))$  where the operator  $|\cdot|$  denotes the complex modulus.

A general model for a stationary stochastic process is

$$X_t = h(\dots, \epsilon_{t-2}, \epsilon_{t-1}, \epsilon_t, \epsilon_{t+1}, \epsilon_{t+2}, \dots) \quad (2.3)$$

where, unlike the Wold representation  $\epsilon_t$  is stochastically independent. If  $X_t$  is causal, it does not depend on the future values of  $\epsilon_t$  (making this common assumption would not substantively affect our discussion). If  $h$  is a well-behaved function it can be rep-

<sup>5</sup>See Hinich (1982), Hinich and Patterson (1989) and Priestley (1988, pp. 13-16).

resented as a Volterra series:<sup>6</sup>

$$X_t = \sum_{u=-\infty}^{\infty} g_u \varepsilon_{t-u} + \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} g_{u,v} \varepsilon_{t-u} \varepsilon_{t-v} + \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} g_{u,v,w} \varepsilon_{t-u} \varepsilon_{t-v} \varepsilon_{t-w} + \dots \quad (2.4a)$$

If  $X_t$  is linear then only the first term in the Volterra representation exists; for linear processes the Wold and Volterra representations are identical implying that the impulse process in the Wold decomposition must be independent in this case.

The existence of higher-order terms in the Volterra expansion implies that  $X_t$  is a *nonlinear process*. Unlike a linear process, the response of the nonlinear sequence to a shock will depend on *generalized transfer functions* of the form:

$$G(f) = \sum_{u=-\infty}^{\infty} g_u e^{-i(2\pi f)u}, \quad G(f, g) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} g_{u,v} e^{-i2\pi(fu+gv)}, \quad \dots \quad (2.5)$$

If the input to a nonlinear sequence contains components with frequencies  $f$  and  $g$ , then the output will contain components with frequencies  $f, g, (f+g), 2f, 2g, 2(f+g), 3f, 3g, 3(f+g), \dots$ , and the amplitudes and phases of these components will depend on the generalized transfer functions.

Tests for linearity and Gaussianity can be based on the properties of these generalized transfer functions as reflected in a process' higher-order polyspectra. In general, the  $k^{th}$ -order polyspectrum is the Fourier transform of the  $k^{th}$ -order cumulant function. The first three cumulants are defined by  $c_X(t) = E[X_t]$ ,  $c_{XX}(t_1, t_2) = E[X_{t_1} X_{t_2}]$ , and  $c_{XXX}(t_1, t_2, t_3) = E[X_{t_1} X_{t_2} X_{t_3}]$ .<sup>7</sup> Strict stationarity (or even third-order stationarity) implies  $c_X(t) = 0$  for all  $t$ ,  $c_{XX}(t_1, t_2)$  is a function only of  $\tau = (t_1 - t_2)$ , and  $c_{XXX}(t_1, t_2, t_3)$  is a function only of  $\tau_1 = (t_1 - t_2)$  and  $\tau_2 = (t_2 - t_3)$ . The second and third-order cumulant functions for stationary processes can be denoted by  $c_{XX}(\tau)$  and  $c_{XXX}(\tau_1, \tau_2)$  respectively. These functions are assumed to be absolutely summable.

The *power spectrum* is then defined as the Fourier transform of  $c_{XX}(\tau)$ :

$$P_X(f) = \sum_{\tau=-\infty}^{\infty} c_{XX}(\tau) e^{-i2\pi f\tau}, \quad |f| < \frac{1}{2}. \quad (2.6)$$

where  $f$  denotes the frequency measured in units of inverse time.<sup>8</sup> The *bispectrum* is

<sup>6</sup>For details on Volterra representations, see Schetzen (1980, 1981) and Rugh (1981).

<sup>7</sup>Cumulants and moments are equivalent up to the third-order. This is not true for higher orders.

<sup>8</sup>Multiplying these frequencies by  $2\pi$  converts them to radians.



defined as the second-order Fourier transform of  $c_{XXX}(\tau_1, \tau_2)$ :

$$B_X(f, g) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_{XXX}(\tau_1, \tau_2) e^{-i2\pi(f\tau_1 + g\tau_2)}, \quad (2.7)$$

$(f, g) \in D = \{(f, g) \mid 0 < f < (1/2), g < f, 2f + g < 1\}$  which is called the principal domain (Hinich and Messer, 1995). If the second and third-order cumulant functions are absolutely summable, then the power spectrum and the bispectrum exist and are well defined. The integral of the power spectrum is equal to the variance of the sequence,  $c_{XX}(0)$ , and the power spectrum can be interpreted as a decomposition of the variance by frequency. Similarly, the bispectrum decomposes the skewness of the sequence,  $c_{XXX}(0, 0)$ , by pairs of frequencies.

Define the *skewness function*,  $\Gamma_X(f, g)$ , as the normalized square modulus of the bispectrum:

$$\Gamma_X(f, g) = \frac{|B_X(f, g)|^2}{P_X(f)P_X(g)P_X(f+g)}. \quad (2.8)$$

Let  $\varepsilon_t$  be a stochastically independent sequence, then  $P_\varepsilon(f) = c_{\varepsilon\varepsilon}(0)$  and  $B_\varepsilon(f, g) = c_{\varepsilon\varepsilon\varepsilon}(0, 0)$  for all  $(f, g) \in D$ . This implies that a linear process has a constant skewness function equal to  $\Gamma_X(f, g) = c_{\varepsilon\varepsilon\varepsilon}(0, 0)^2 c_{\varepsilon\varepsilon\varepsilon}(0)^{-3}$ , because  $P_X(f) = |G(f)|^2 P_\varepsilon(f)$  and  $B_X(f, g) = G(f)G(g)G^*(f+g)B_\varepsilon(f, g)$ . If the stochastically independent input sequence is also Gaussian then  $c_{\varepsilon\varepsilon\varepsilon}(0, 0) = 0$  and  $\Gamma_X(f, g)$  will be identically zero.

These properties form the basis of Hinich's (1982) tests of Gaussianity and linearity. The intuition is that the skewness function will be flat for linear processes and identically zero for Gaussian processes. If the skewness function is significantly rough then linearity is rejected.

Hinich's test is conservative, not only in practice but also in theory. The test is conservative in theory because the null hypothesis that  $\Gamma_X(f, g)$  is constant for all frequency pairs is a necessary, but not sufficient condition for linearity. Nonlinearity could be detected in higher order polyspectra, even if the normalized bispectrum is flat.<sup>9</sup> Nevertheless, Ashley, Patterson, and Hinich (1986) found that Hinich's bispectral test had substantial power against many common nonlinear time series models including bilinear models, nonlinear moving-average and autoregressive models, and linear and nonlinear threshold autoregressive models.

A key aspect of Hinich's tests is that (at least third-order) stationarity is assumed. However, economic time series often appear to be subject to permanent shocks, and it

<sup>9</sup>Tests of higher-order polyspectra are generally not applicable in econometrics, because most economic time series are not long enough for consistent estimation of even the fourth-order polyspectrum.

has become a standard practice to model these time series as non-stationary integrated processes. As is the norm in testing for nonlinearity, if the process is nonstationary Hinich's test can falsely reject linearity. Consequently, individual economic series are usually differenced or detrended before being tested for nonlinearity. Cointegration can provide a richer model of nonstationarity and an alternate method to recover stationary dynamics for a system of economic variables.

### 3 Integration and Cointegration

Cointegration as it is normally defined is incompatible with nonlinear dynamics. Cointegration was developed within the framework of vector error-correction models. Linearity of the stationary dynamics was explicitly assumed, because the VECM model is linear and the innovation process was assumed to be independent or Gaussian. However, there is no compelling reason for this restriction.

Using Hall and Heyde's martingale representation, we show that the innovation process of a integrated series is not in general a linear stochastic process. It is then straightforward to define cointegration for a vector of integrated processes using the martingale representation.

For clarity, our martingale-based definition is contrasted with the standard VECM definition of cointegration including the extension to nonlinear cointegration. The representation theorem shows that nonlinearity is more general than just nonlinear cointegration, as nonlinear dynamics can be present even when the cointegrating relationship is linear.

Initially, we establish some definitions and notational conventions. The definitions are standard and can be found in a number of references. For all time periods, let  $S_t$  denote a  $q$ -dimensional vector random sequence,  $S_t = (S_{1t}, \dots, S_{qt})^T$  on a probability space  $(\Omega, \mathcal{F}, \mu)$ .

**Martingale Definition.** A vector martingale is an adapted sequence  $(S_t, \mathcal{F}_t)$  where  $\mathcal{F}_t$  is an increasing sequence of  $\sigma$ -algebras contained in  $\mathcal{F}$  such that  $S_t$  is integrable and satisfies

$$E(S_t | \mathcal{F}_{t-1}) = S_{t-1} \text{ a.s.}$$

for every  $t$ . The first difference of a martingale,  $\Delta S_t = S_t - S_{t-1} = Y_t$  is referred to as a martingale difference sequence; it is integrable and satisfies  $E(Y_t | \mathcal{F}_{t-1}) = 0 \text{ a.s.}$

Let  $T : \Omega \rightarrow \Omega$  denote a one to one ergodic measure-preserving shift transformation. If  $X_0(\omega)$  is a random variable on the probability space, then  $X_t(\omega) = X_0(T^t \omega)$

defines a *strictly stationary ergodic sequence*. A stochastic sequence is said to be *integrated of order one*,  $I(1)$ , if the first difference of the sequence is strictly stationary.<sup>10</sup> A martingale difference sequence is strictly stationary by definition, so martingales are  $I(1)$ . The concept of integration can be extended to higher orders.

Hall and Heyde (1980, pp. 136) prove the following representation theorem:

**Hall and Heyde Representation Theorem.** *Any stationary ergodic sequence  $X_t$  can be represented in the form:*

$$X_t = Y_t + Z_t - Z_{t+1}, \quad (3.1)$$

where  $Y_t(\omega) = Y_0(T^t\omega)$  is a stationary martingale difference sequence, and  $Z_t(\omega) = Z_0(T^t\omega)$  such that  $Z_0(\omega)$  is in  $L^1$ .<sup>11</sup> Explicit formulas for the representation are given by:

$$Y_0 = \sum_{k=-\infty}^{\infty} (E[X_k | F_0] - E[X_k | F_{-1}]); \quad (3.2)$$

and,

$$Z_0 = \sum_{k=0}^{\infty} (E[X_k | F_{-1}]) - \sum_{k=-\infty}^{-1} (X_k - E[X_k | F_{-1}]), \quad (3.3)$$

where  $\{F_s\}$  is the filtration generated by the shift transform.

From the Hall and Heyde representation, we derive a representation for an  $I(1)$  sequence:

**$I(1)$  Representation Corollary.** *If the stationary first-difference of an  $I(1)$  sequence is ergodic, then the nonstationary level of the integrated sequence is represented by*

$$S_t = \sum_{s=0}^t Y_s - Z_{t+1} + Z_1 + S_0. \quad (3.4)$$

*Proof.* If the stationary first-difference of an  $I(1)$  sequence is ergodic, from the representation theorem it has the following representation:

$$S_t = Y_t + Z_t - Z_{t+1},$$

where  $Y_t$  is a stationary vector martingale difference sequence and  $Z_t$  is a stationary

<sup>10</sup>Engle and Granger (1987) add the condition that the stationary moving average representation of the first difference be invertible.

<sup>11</sup>Hall and Heyde also require  $E|X_0| < \infty$ . Alternatively, the theorem can be proved under mixingale assumptions, see Davidson (1994, pp. 247-252).

vector sequence. Equation (3.4) is derived by solving this representation of the first difference for  $S_{t-1}$ , advancing the index one period and recursively substituting for  $S_{t+1}$ .  $\square$

*Remark 1.* The level of the  $I(1)$  sequence is dominated by the accumulated martingale difference sequence which gives rise to the permanent shocks.

*Remark 2.* The components of  $\Delta S_t$ , the first difference of  $S_t$ , have the form:

$$\Delta S_{jt} = Y_{jt} + Z_{jt} - Z_{j,t+1}. \quad (3.5)$$

From (3.2) and (3.3), both  $Y_{jt}$  and  $Z_{jt}$  exhibit dependence, although  $Y_{jt}$  is a martingale difference and is non-forecastable in the mean square metric, see Hinich and Patterson (1987).

A system of integrated time series is cointegrated if some linear combinations of the time series are stationary. Cointegration can be defined as a reduced rank condition involving the covariance matrix of the vector martingale difference. We need the following lemma for the form of the covariance matrix for a vector martingale difference sequence.

**Lemma 1.** *The covariance matrix of a martingale difference sequence has the form:*

$$E[Y_t Y_s^T] = \begin{cases} \mathbf{C}\mathbf{C}^T & \text{if } s = t \\ \mathbf{0} & \text{if } s \neq t \end{cases} \quad (3.6)$$

*Proof.* Vector martingale differences are serially uncorrelated and have positive semi-definite covariance matrices.  $\square$

Cointegrating vectors for an  $I(1)$  sequence that are based on the martingale representation are defined by:

**Theorem 1** (Martingale Cointegration). *If  $\mathbf{C}$  in (3.6) has reduced rank,  $(q - r)$ , then there will exist  $r$  non-trivial vectors  $\beta_1, \dots, \beta_r$ , called cointegration vectors where, the linear combinations  $\beta_j^T S_t$ , called cointegration relations, are stationary for all  $j = 1, \dots, r$ .*

*Proof.* Choose  $\beta$ , a  $q$  by  $r$  matrix  $[\beta_1 \beta_2 \dots \beta_r]$  that spans the null space of  $\mathbf{C}$ . Then by definition  $\beta_j^T \mathbf{C} = 0^T$ , for all  $j = 1, \dots, r$ . These vectors define stationary processes

because

$$\begin{aligned} E_0 \left[ \left( \sum_{s=0}^t \beta_j^T Y_s \right) \left( \sum_{s=0}^t \beta_j^T Y_s \right)^T \right] &= \beta_j^T E_0 \left[ \left( \sum_{s=0}^t Y_s \right) \left( \sum_{s=0}^t Y_s \right)^T \right] \beta_j \\ &= \beta_j^T \mathbf{C} \mathbf{C}^T \beta_j = \mathbf{0}^T \mathbf{0} = 0. \end{aligned} \quad (3.7)$$

□

The proof also supports the following corollary:

**Corollary 2.** *Denote the  $q$  by  $(q - r)$  orthogonal complement matrix of  $\beta$  by  $\beta_\perp$ , so that  $\beta_\perp$  has the property  $\beta^T \beta_\perp = 0$ . The common stochastic trend  $\beta_\perp^T S_t$ , which has dimension  $(q - r)$ , is integrated but not cointegrated. The  $q$ -dimensional sequences  $\Delta S_t$  and  $\begin{bmatrix} \beta^T & \beta_\perp^T \Delta \end{bmatrix} S_t$  are both stationary. In the absence of cointegration, the two transformations are equivalent. If  $r = 0$ , then  $\beta_\perp$  is full rank and can be taken as the identity matrix.*

In contrast to extant definitions, the martingale based definition of cointegration does not require independence, Gaussianity, or linearity of the stationary components of the process. Previous definitions of linear cointegration are a special case, much like independence is a special case of the martingale property. The difference can be made clearer by looking at the expectation of the cointegration relations,

$$E_0 \left[ \beta_j^T S_t \right] = E_0 \left[ \beta_j^T (Z_1 - Z_{t+1}) \right] + \beta_j^T S_0. \quad (3.8)$$

These expectations have been purged of the effects of the permanent shocks generated by the martingale difference and are stationary. When viewed as a new stochastic process, there are no restrictions on the dependence structure of  $\beta_j^T S_t$ , aside from stationarity and ergodicity.

Our method contrasts with the standard approach to cointegration. Stationary linear combinations of integrated variables are usually specified to follow a linear ARMA process or are included in linear structural models. The standard linear VECM has the form:

$$\Delta S_t = \alpha \beta^T S_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta S_{t-j} + \varepsilon_t. \quad (3.9)$$

If the model is cointegrated then the  $q$  by  $r$  parameter matrices,  $\alpha$  and  $\beta$ , have rank  $r$ . The cointegration relations enter the model linearly, through  $\alpha$ . The error-correction model is estimated under the assumption that  $\varepsilon_t$  is stochastically independent, which implies that the cointegration relations are linear stochastic processes. Our discussion

shows that cointegration does not generally imply linearity, therefore, there is no reason to expect  $\varepsilon_t$  to be either Gaussian or independent.

Granger (1991) proposes three nonlinear generalizations of cointegration. The first generalization is that nonlinear functions of the time series may be cointegrated in the sense that  $g_1(x_{1t})$  and  $g_2(x_{2t})$  have a dominant property that the linear combination of nonlinearly transformed variables  $z_t = g_1(x_{1t}) - Ag_2(x_{2t})$  does not exhibit. A second generalization is to allow time-varying cointegration vectors. A third generalization is nonlinear error correction, in which the cointegration relations would enter the error-correction model through a nonlinear function  $f$ , i.e.

$$\Delta S_t = f(\beta^T S_{t-1}) + \sum_{j=1}^{p-1} \Gamma_j \Delta S_{t-j} + \varepsilon_t. \quad (3.10)$$

Granger (1991) gives conditions under which  $f(z)$  is stationary.

A natural nonlinear error-correction specification is to allow mean reversion only for large deviations, so that  $f$  has the form:

$$f(z) = \begin{cases} -z & \text{if } |z| > k \\ 0 & \text{if } |z| \leq k \end{cases}. \quad (3.11)$$

In this case,  $z_t = \beta^T S_t$  behaves like a unit root in a neighborhood of its mean, but exhibits mean reversion when it is outside the neighborhood. This model is a straightforward generalization of the standard error-correction model that exhibits nonlinear dynamics, but the linear combination  $\beta^T S_t$  is not stationary.<sup>12</sup>

Although, the extended definition of cointegration could be further extended to allow for nonlinear cointegration, we limit ourselves to the case where the linear combination is stationary. Such stationary linear combinations can exhibit nonlinear dynamics. Differentiating between nonlinear error correction and stationary nonlinear dynamics is likely difficult in practice.

Our proposed method for testing for whether a cointegration is nonlinear is sequential. This sequential method allows us to test the stationary components of the system for nonlinear dynamics. We first estimate cointegrating vectors using Bierens' (1997) non-parametric test. Bierens' test is asymptotically valid for a nonlinear data generating processes due to Hall and Heyde's representation theorem. We then test the estimated cointegrating relations for Gaussianity and linearity using Hinich's (1982) tests.

---

<sup>12</sup>For example, the process defined by (3.11) is nonstationary and behaves like a unit root when near its mean.

Asymptotically, Hinich's test is also valid as the cointegrating vectors are stationary. In practice, as is the norm, the results of Hinich's tests are conditional on whether the first stage estimates do eliminate any nonstationarity.

## 4 Empirical Results

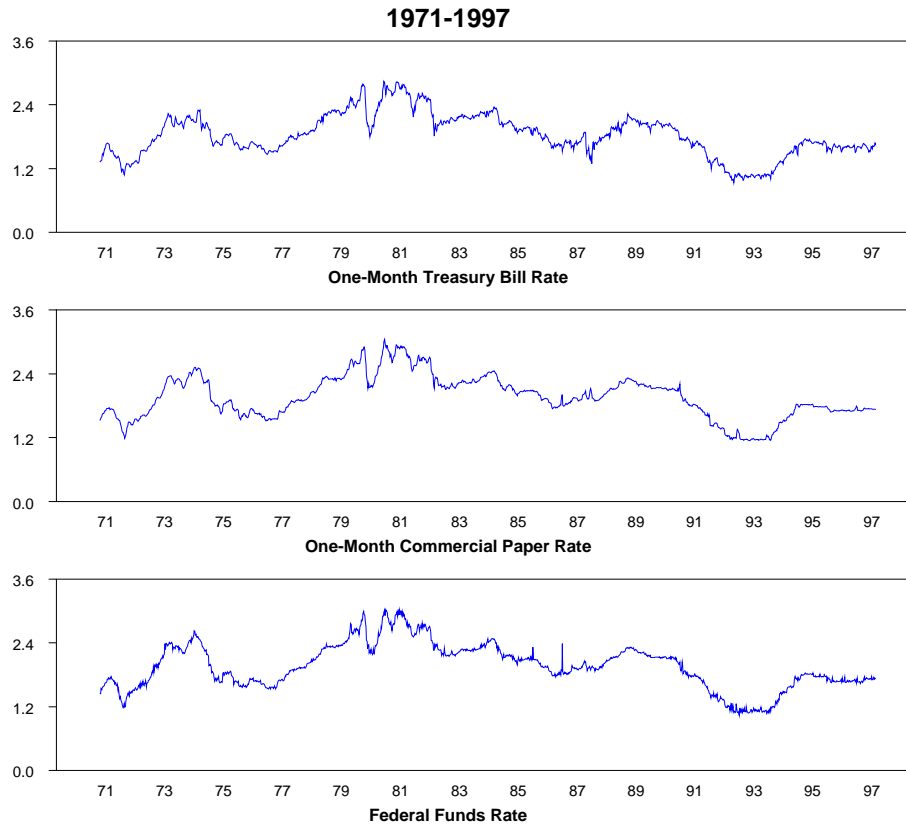
We apply our sequential procedure to a system of short-term U.S. interest rates. Short-term interest rates are available at a high frequency over an extended time period: constituting a larger sample size than many other business cycle variables, such as real output and inflation. In addition, interest rates directly capture the dynamics caused by monetary policy changes.

We use business daily data for the interest rates on one-month Commercial Paper (*CP*), the secondary market rate on one-month Treasury Bills (*TB*), and the Federal Funds (*FF*) from 4/08/1971 to 8/29/1997. The commercial paper and Federal Funds rates are available from the Federal Reserve Board's website. The commercial paper rate series was discontinued in August 1997. The Federal Reserve Bank of St. Louis provided us with the secondary market rate on one-month Treasury Bills. These interest rates are converted to one-month holding period yields on a bond interest basis, and are passed through an anti-aliasing filter. The anti-aliasing filter is designed to remove the high-frequency power in the daily rate series to minimize the bias caused by converting the daily time series to weekly time series either by direct sampling or weekly averaging. The daily rates are converted to weekly rates by sampling the filtered daily rates once per week. Figure 1 on the following page displays the natural logarithms of the filtered interest rates.

Correcting for aliasing does not impact the asymptotics of the cointegration estimator, because cointegration is related to the long-run dynamics while aliasing distorts higher frequency dynamics. Nevertheless, correcting for aliasing might improve the power of the cointegration estimators in a finite sample. In addition, Hinich and Patterson (1985, 1989) showed that aliasing does bias tests for nonlinearity towards accepting linearity. Aliasing is discussed in the appendix.

After applying the multirate filter, we test this data with our two-stage method: first testing for cointegration and then testing for nonlinearity. Two cointegrating vectors are found for the system of three interest rates over the period 1971 – 1997. We then run several tests on each cointegrating relation. We first test the cointegrations for an alternative form of nonstationarity considered by Hinich and Wild (2001). This alternative type of nonstationarity is rejected, so we test for Gaussianity. Gaussianity

Figure 1: Logarithm of Interest Rates



of both the real and imaginary parts of the bispectrum is strongly rejected. Finally, we test for nonlinearity. We find strong evidence that the cointegrations exhibit nonlinear dynamics.

#### 4.1 Univariate Tests

Before estimating cointegration relations, we run a battery of univariate tests. We first test the unit root and stationarity hypotheses on  $\ln(CP)$ ,  $\ln(TB)$ ,  $\ln(FF)$ , and their first differences ( $\Delta$ ) using several tests with different nulls. These tests include: an augmented Dickey-Fuller (ADF) test and a Phillips-Perron (PP) test of the unit root hypotheses against the alternative of stationarity; the KPSS test of the null of stationarity against the alternative of nonstationarity; and the Bierens (1997) non-parametric test for the existence of cointegration run as a univariate test of the unit root with drift hy-



Table 1: Univariate Stationarity Tests

<b>Variable</b>	<b>ADF</b>	<b>PP</b>	<b>KPSS1</b>	<b>Bierens</b>
$\ln(CP)$	-1.9132	-8.6587	0.9272	1.1441
$\ln(TB)$	-1.8289	-9.4498	1.0186	0.7741
$\ln(FF)$	-1.8174	-8.8747	1.0027	1.0203
$\Delta \ln(CP)$	-7.4386	-740.9803	0.1120	0.0000
$\Delta \ln(TB)$	-8.0123	-877.6077	0.1210	0.0000
$\Delta \ln(FF)$	-7.3217	-2118.753	0.1249	0.0000
$H0 :$	UR	UR	S	URD
$H1 :$	S	S	NS	TS
5% c.v.	< -3.86	< -14.0	> .436	< .025
10% c.v.	< -2.57	< -11.2	> .347	< .006

pothesis against trend stationarity on each variable.<sup>13</sup> For the ADF test, the lag length,  $p$ , is chosen by the formula  $p = 5(n)^{.25}$ . For the PP and KPSS tests, the truncation lag for the Newey-West estimator is also set with this formula. The test results are shown in Table 1 along with a mnemonic for the tests' hypotheses and the 5 and 10 percent critical values.

The logged interest rates are clearly  $I(1)$  processes: every test rejects stationarity of the levels at well above the 95% confidence level and fails to reject stationarity of the first differences even at the 80% confidence level. The consistency of the test results is important, because differences in these tests can be interpreted as evidence of long-memory rather than integration. For example, Karanasos, Sekioua, and Zeng (2006) interpret their simultaneous rejection of both the unit root hypothesis and stationarity as evidence for fractional integration and long-memory in real U.S. interest rates. Our results are not open to such interpretation.

Given the results of the stationarity tests, we test the stationary first difference of each interest rate for nonlinearity. We pre-whiten each of the components using an  $AR(6)$  filter to eliminate bias in the spectral estimation prior to testing and to decrease the likelihood of falsely rejecting the null of linearity. The tests (available on request) provide overwhelming evidence of nonlinear dynamics for the first differences of these short-term interest rates over the full sample.

We also tested the first differences for nonlinearity over two sub-periods: Sept. 13, 1974 through Sept. 19, 1979 and March 1, 1984 through Dec. 31, 1996. These are periods over which a target for the federal funds rate can be constructed, see Rudebusch

<sup>13</sup>Stationarity can be viewed as a special case of trend stationarity with the trend restricted to be zero. Consequently, running versions of the ADF, PP, and KPSS tests that test for trend stationarity produces results consistent with Bierens' test.

(1995). Effectively, we are dropping the period when the Federal Reserve shifted its intermediate target away from interest rates. This period is also when many interest rates were deregulated.

For these sub-samples, we can accept the null of linearity in the first sub-sample, but reject linearity in the second. The number of data points for the first sub-sample is 258 versus 669 for the second sub-sample and 1,363 for the full sample. The evidence reported in Ashley et al. (1986) would indicate that the power of these tests is substantially higher over both the second sub-sample and over the full sample. This provides one explanation for the inability to reject linearity in the first sub-sample. Another possible explanation for finding nonlinearity only in the second sub-sample could be that deregulation of interest rates transformed the dynamics going forward.

Regardless for the reasons for accepting linearity in the first sub-sample, finding evidence of nonlinearity in the second sub-sample is crucial. If linearity was rejected for both sub-samples, it would appear that the nonlinearity found over the full sample was driven solely by a regime shift. Rejecting linearity in the second sub-sample does not rule out a break in the dynamics due to the policy, but it does rule out the shift being the only source of nonlinearity. Consequently, we continue analyzing the full sample, although we also check the results for the two sub-samples.

## 4.2 Cointegration

The cointegration analysis used the system

$$S_t = [\ln(CP1M), \ln(TB1M), \ln(FF)]^T .$$

The cointegration analysis is conducted in two steps: rank identification and estimation. The rank identification, which determines the number of cointegration relations, is based on the non-parametric test procedure developed by Bierens (1997, 2005). The number of cointegration relations is determined by a set of hypothesis tests, called  $\lambda$ -min tests, that are essentially non-parametric versions of the well-known Johansen (1988) parametric  $\lambda$ -max tests. The  $\lambda$ -min tests are non-parametric because the matrices involved are constructed from the data independently of the data-generating process. The number of cointegration relations can also be estimated using a function of the eigenvalues  $\hat{g}_m(r)$ . The value of  $r$  that minimizes  $\hat{g}_m(r)$  is a consistent estimate of the true number of cointegration relations.

The number of cointegrations determined by both the  $\lambda$ -min test and estimating  $\hat{g}_m(r)$  is 2. The  $\lambda$ -min tests are reported in Table 2 on the following page.  $M$  is the

Table 2: Nonparametric Cointegration Tests

Hypothesis	Test Stat	Critical Region	M	Conclusion
$H0 : r = 0$	0.00000	20% (0, .006)	3	Reject
$H1 : r = 1$		10% (0, .017)	4	Reject
		5% (0, .008)	4	Reject
$H0 : r = 1$	0.00054	20% (0, .077)	3	Reject
$H1 : r = 2$		10% (0, .034)	3	Reject
		5% (0, .017)	3	Reject
$H0 : r = 2$	0.76618	20% (0, .341)	3	Accept
$H1 : r = 3$		10% (0, .187)	3	Accept
		5% (0, .111)	3	Accept

M is the smoothing parameter for the nonparametric estimator

smoothing parameter; the value is set optimally for the different confidence levels following Bierens (1997). The tests are run in sequence, starting with the null hypothesis that the number of cointegrating vectors is zero, followed by a test of the null hypothesis that there is one cointegrating vector, and so on until the null cannot be rejected. We find that  $r = 0$  (no cointegration) is decisively rejected, as is the hypothesis that  $r = 1$  (one cointegrating vector), but we cannot reject the hypothesis that  $r = 2$  (two cointegrating vectors).

For comparison, we also estimate the parametric maximum likelihood  $\lambda$ -max and trace tests of Johansen (1988) using the *CATS* package (Hansen and Juselius, 2006). The  $I(1)$  maximum likelihood method estimates a finite-order VECM, as in (3.9), where the coefficient matrices  $\Pi, \Gamma_1, \dots, \Gamma_{p-1}$  are  $3 \times 3$ . If the system is cointegrated then the matrix  $\Pi$  has reduced rank  $r < 3$ , and can be decomposed into  $\Pi = \alpha\beta^T$ . The matrices  $\alpha$  and  $\beta$  are full rank  $3$  by  $r$  matrices, and the columns of  $\beta$  are the cointegration vectors.

Pantula (1989) and Johansen (1992) suggested a procedure to jointly identify the deterministic components and the rank of  $\Pi$ . The idea is to test the models sequentially, beginning with the most restrictive model considered. Each hypothesis can be tested using either the trace or  $\lambda$ -max test statistics. We conducted these tests for a set of lag lengths  $p = 4, 5, \dots, 20$ . These tests uniformly find that there are two cointegration vectors and that the correct deterministic component is a constant that is restricted to the cointegration space. This specification is therefore extremely robust to the lag length and agrees with the rank determination of the non-parametric test. Table 3 on the next page reports these tests for a lag length of  $p = 6$ .<sup>14</sup>

<sup>14</sup>We computed various information criteria for the VECM. The Schwartz criteria indicated a lag length of 4 and the Akaike criterion indicated a length of 20. We estimated the model over this range of lags. The

Table 3: Parametric Cointegration Tests

Hypothesis		$\lambda - \max$	90% c.v.	Trace	90% c.v.	95% c.v.
$H_0 : r = 0$	rest. const	103.48	14.09	158.63	31.88	34.78
$H_0 : r = 0$	const.	103.47	13.39	158.60	26.70	29.38
$H_0 : r = 0$	const., trend	114.26	16.13	172.02	39.08	42.20
$H_0 : r = 1$	rest. const	50.59	10.29	55.15	17.79	19.99
$H_0 : r = 1$	const.	50.59	10.60	55.13	13.31	15.34
$H_0 : r = 1$	const., trend	51.41	12.39	57.77	22.95	25.47
$H_0 : r = 2$	rest. const	4.55	7.50	4.55	7.50	9.13

The results from the nonparametric and parametric estimators are very similar. The non-parametric estimate of the cointegration vectors is  $\beta_{NP} = [\beta_{1,NP} \ \beta_{2,NP}]$ , where  $\beta_{1,NP} = (1, -1.075, 0)^T$  and  $\beta_{2,NP} = (0, 1, -0.863)^T$ . The parametric estimate of the cointegration vectors is  $\beta_P = [\beta_{1P} \ \beta_{2P}]$ , where  $\beta_{1P} = (1, -1.031, 0)^T$  and  $\beta_{2P} = (0, 1, -0.913)^T$ .<sup>15</sup> The parametric estimate is statistically equivalent to the nonparametric estimate. For both estimators, the first basis vector  $\beta_1$  reflects the near stationarity of the spread between the logarithms of the Commercial Paper and Treasury Bill rates. Similarly, the second basis vector  $\beta_2$  reflects the near stationarity of the spread between the Treasury Bill rate and Federal Funds rates.<sup>16</sup> The non-parametric estimates of the two cointegration relations are shown in Figure 2 on the following page. The differences between the nonparametric and parametric estimates, also included in the figure, are an order of magnitude smaller.

This consistency of the nonparametric and parametric contrasts with the results of Coakley and Fuertes (2001) and Calza and Sousa (2006) where the parametric and non-parametric results differ. In these papers, the authors argue for accepting the nonparametric results because Bierens estimator is valid for a broader range of data generating processes. In particular, Coakley and Fuertes (2001) argue that the maximum likelihood estimates are distorted due to nonlinear mean reversion in exchange rates which would imply nonlinear cointegration. The consistency between our nonparametric and parametric estimates reveals no evidence of nonlinear cointegration between interest rates.

Bierens (1997) argued that hypothesis tests in the parametric model have higher

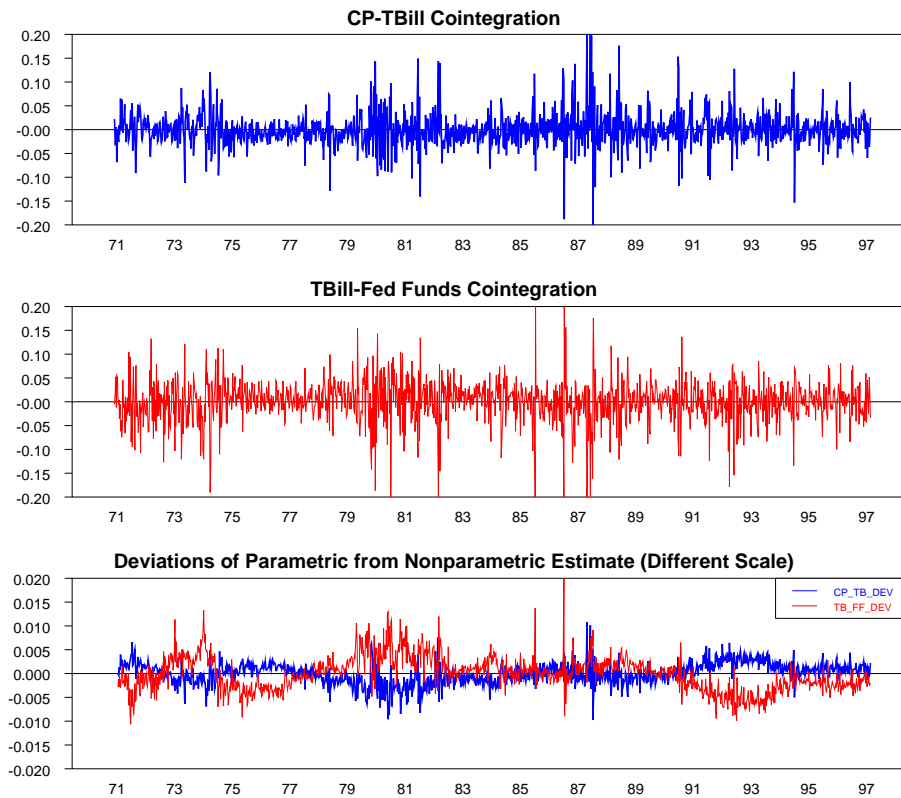
---

results were not greatly affected by the choice of lag length within this range. The model with  $p = 6$  is fairly parsimonious and passed tests for absence of first and fourth order auto-correlation.

<sup>15</sup>The basis for the cointegration space has been transformed into a basis with one zero in each vector and the estimated restricted constant is subtracted from the cointegration relations. This transformation does not change any results.

<sup>16</sup>Chi-squared tests in the VECM(6) accept the hypothesis that  $\beta_1 = (1, -1, 0)^T$  but reject the hypothesis that  $\beta_2 = (0, 1, -1)^T$ . The values of the test statistics are 1.31 and 10.87 respectively. These tests are  $\chi^2(1)$ .

Figure 2: Nonparametric cointegrations and the difference from the parametric estimates



power than comparable tests in the non-parametric model. This argument does not necessarily hold because the argument and the hypothesis tests are predicated on linearity. Despite the parametric estimator's consistency with the nonparametric estimates, the parametric estimator is likely misspecified since the individual interest rates are nonlinear. Since the nonparametric and parametric cointegrations are indistinguishable, we can safely sidestep the issue of misspecification by focusing solely on the nonparametric results.

#### 4.2.1 Robustness

As already discussed, our results are robust to the type of estimator and lag length. Before moving to the second stage of our approach and testing for nonlinearity, we also tested the results for robustness to the Federal Reserve's choice of policy instrument,

Table 4: Nonparametric Cointegration Tests for Sub-Samples

Hypothesis	Test Stat: 1974 – 1979	Test Stat: 1984 – 1996	Critical Region	M	Conclusion
$H_0 : r = 0$	0.00000	0.00000	20% (0, .006)	3	Reject
$H_1 : r = 1$			10% (0, .017)	4	Reject
			5% (0, .008)	4	Reject
$H_0 : r = 1$	0.00523	0.00008	20% (0, .077)	3	Reject
$H_1 : r = 2$			10% (0, .034)	3	Reject
			5% (0, .017)	3	Reject
$H_0 : r = 2$	1.33438	2.33982	20% (0, .341)	3	Accept
$H_1 : r = 3$			10% (0, .187)	3	Accept
			5% (0, .111)	3	Accept

M is the smoothing parameter for the nonparametric estimator

by examining the integration and cointegration properties of the data over two sub-periods: Sept. 13, 1974 through Sept. 19, 1979 and March 1, 1984 through Dec. 31, 1996.

The results of the non-parametric cointegration tests for the two sub-samples are reported in Table 4. The results show that the rank identifications are consistent with those from the full sample. The parametric estimators also identified two cointegrating vectors for each sub-period. Further, the estimated cointegration vectors are consistent with the estimated vectors from the full sample; we cannot reject the joint hypothesis,  $H_0 : \beta_{1P} = (1, -1.031, 0)^T$  and  $\beta_{2P} = (0, 1, -0.913)^T$ , for either sub-sample. These tests are  $\chi^2(1)$ . For the 1974-1979 sub-sample, the test statistic is 3.26 (p-value of .2), and for the 1984-1996 sub-sample, the test statistic is .38 (p-value of .83).

### 4.3 Tests for Nonlinearity of the Cointegration Relations

The stationary components of the system consist of the two cointegration relations and the first difference of the common stochastic trend. We test the estimated cointegration relations for nonlinear serial dependence using the bispectrum tests. The cointegration vectors,  $\beta_1$  and  $\beta_2$  are basis vectors for the cointegration space, so that any linear combination of  $\beta_1$  and  $\beta_2$  are also stationary. Thus, evidence of nonlinearity in one of the cointegration relations is actually evidence that the stationary components of the system are nonlinear. Prior to testing for nonlinearity, each of the cointegration relations is pre-whitened by an  $AR(6)$  filter to eliminate bias in the spectral estimation prior and to decrease the likelihood of falsely rejecting the null of linearity.

For robustness, we test these relations for stationarity using the frequency domain

Table 5: Stationarity, Gaussianity and Time Reversability Tests (1971 – 1997)

Cointegration	HW	Gauss1	Gauss2
Nonparametric			
$\beta_{1,NP}^T S_t$	25.2276 (0.8620)	638.6265 (0.0000)	707.1962 (0.0000)
$\beta_{2,NP}^T S_t$	34.2187 (0.4572)	387.4622 (0.0000)	465.5005 (0.0000)
HW test statistic is $\chi^2(34)$ under $H_0$ : stationarity			
Gauss1 test statistic is $\chi^2(34)$ under $H_0$ : $\text{Re}(B(f, g)) = 0 \forall (f, g) \in D$			
Gauss2 test statistic is $\chi^2(34)$ under $H_0$ : $\text{Im}(B(f, g)) = 0 \forall (f, g) \in D$			

test derived by Hinich and Wild (2001). The Hinich and Wild (HW) test checks for residual non-stationarity due to the existence of a waveform with random phase and amplitude. This test has a very different alternative hypothesis than the cointegration test, and should detect nonstationarity at seasonal frequencies. The test is  $\chi^2(34)$  under the null of stationarity. The HW-stationarity tests, reported in Table 5, confirm that the cointegration relations are stationary.

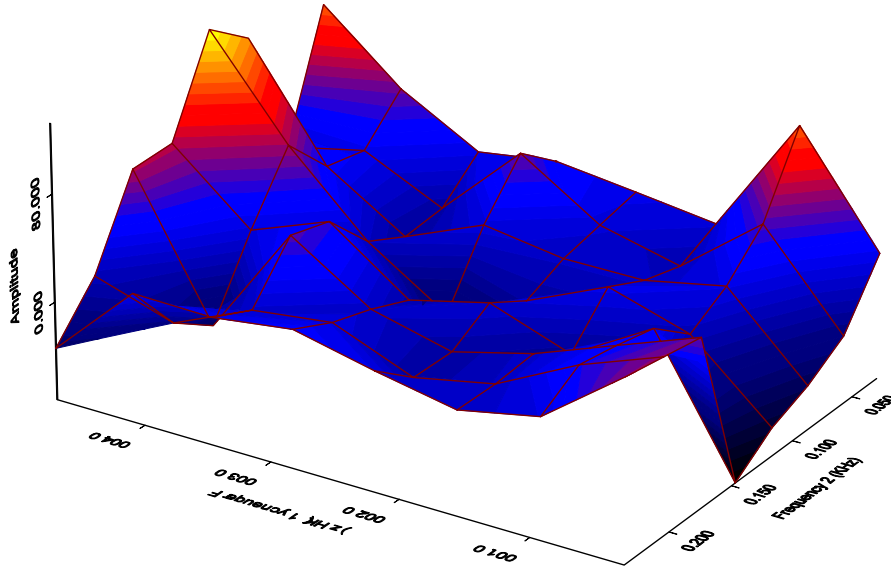
If the time series are Gaussian, then the real and imaginary components of the bispectrum are zero. The test statistics for these two hypotheses, called Gauss1 and Gauss2 respectively, are also reported in Table 5. If either the real or imaginary components of the bispectrum are non-zero then Gaussianity is rejected. If the imaginary component is non-zero then the sequence is not time-reversible. The results indicate that the stationary components of the system are highly non-Gaussian and are not time-reversible.

Rejecting Gaussianity is necessary but not sufficient to reject linearity. Table 6 gives the results of Hinich’s test for nonlinearity over the full sample. The  $Z_\rho$  test statistics are independent and normally distributed under the null of linearity, and we treat these tests as two-tailed, as Ashley et al. (1986) found that one-tailed tests may fail to detect certain types of nonlinearity. The tests are computed for the non-parametric estimates of the cointegrations  $\beta_{1,NP}^T S_t$  and  $\beta_{2,NP}^T S_t$ .

Table 6: Linearity Tests (1971 – 1997)

Cointegration	$Z_{.1}$	$Z_{.2}$	$Z_{.4}$	$Z_{.6}$	$Z_{.8}$	$Z_{.9}$
Nonparametric						
$\beta_{1,NP}^T S_t$	-2.81	-4.09	-5.39	-2.65	2.78	2.19
$\beta_{2,NP}^T S_t$	-2.26	-2.66	-0.93	-0.38	1.29	1.94
$Z_\rho$ is $N(0, 1)$ under $H_0$ : $B(f, g)$ is constant $\forall (f, g) \in D$ .						
Linearity is rejected if $ Z_\rho $ exceeds the critical value.						
c.v. 1.65 (90%) 1.96 (95%) 2.58 (99%)						

Figure 3: Skewness function of CP-TBill cointegration



The strongest evidence of nonlinearity is found in the first cointegration relation. Linearity is rejected for  $\beta_{1,NP}^T S_t$  by  $Z_{.1}$ ,  $Z_{.2}$ ,  $Z_{.4}$ ,  $Z_{.6}$ , and  $Z_{.8}$  using the 99% critical values and by  $Z_{.9}$  using the 95% critical values. Flatness of the skewness function is a necessary condition for linearity. Figure 3 shows that the skewness function for the first cointegration is clearly far from flat, which is what should be expected from the statistical tests.

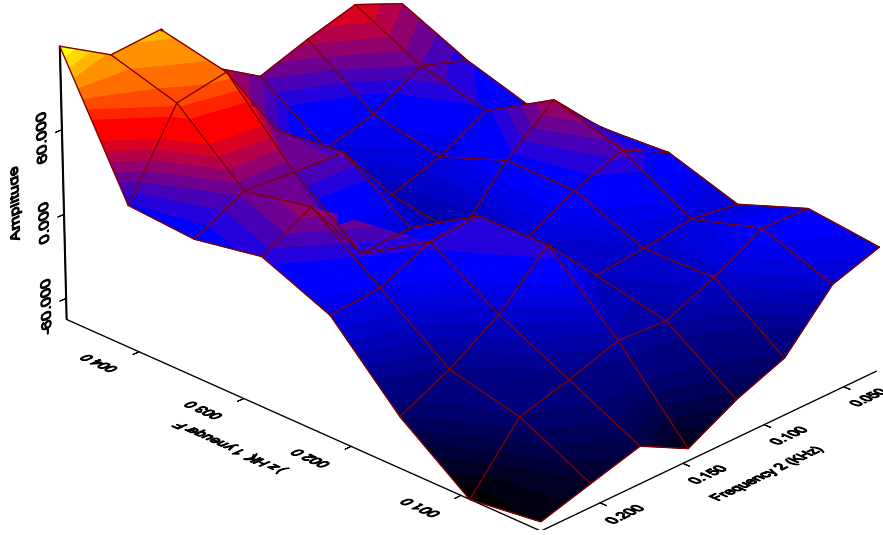
Evidence of nonlinearity is also found in the second cointegration, although this evidence is somewhat weaker. Linearity is rejected for  $\beta_{2,NP}^T S_t$  by  $Z_{.1}$  and by  $Z_{.9}$  using the 95% critical values, and by  $Z_{.2}$  using the 99% critical values. Figure 4 on the following page shows the skewness function for the second cointegration. Again, the skewness function is not flat, but it is flatter than the skewness function of the first cointegration, reflecting weaker evidence of nonlinearity in the second cointegration. However, nonlinearity in either cointegration implies that the cointegrated system exhibits nonlinear dynamics.

#### 4.3.1 Robustness

Structural shifts over the long period being analyzed could be mistaken for nonlinear dynamics. As previously discussed, to address this issue we delete the period 1979 – 1983 and consider the two sub-samples. Table 7 on the next page presents



Figure 4: Skewness function for TBill-Fed Funds cointegration



the test for nonlinearity over these two sub-samples. Similar to the univariate results, linearity cannot be rejected for the first subsample but can for the second. As previously mentioned, Hinich's test has relatively low power for the first sub-sample. Rejecting linearity for the second sub-sample implies that the shift in policy regime does not cause the nonlinearity *per se*. The results could alternatively be interpreted as the result of interest rate deregulation rather than low power. Post-deregulation, the interest rate dynamics seem to become more complex, even though the long-run equilibrium

Table 7: Linearity Tests for Sub-Samples

Cointegration	$Z_{.1}$	$Z_{.2}$	$Z_{.4}$	$Z_{.6}$	$Z_{.8}$	$Z_{.9}$
<b>Sub-Sample #1: 1974 – 1979</b>						
$\beta_{1,NP}^T S_t$	-1.07	-1.39	-1.99	-0.36	0.98	1.12
$\beta_{2,NP}^T S_t$	-1.15	-1.72	0.50	0.79	0.15	0.67
<b>Sub-Sample #2: 1984 – 1996</b>						
$\beta_{1,NP}^T S_t$	-2.00	-2.95	-4.43	-3.88	1.95	1.96
$\beta_{2,NP}^T S_t$	-1.99	-2.82	-2.14	0.05	2.40	1.91
$Z_\rho$ is $N(0, 1)$ under $H_0 : B(f, g)$ is constant $\forall (f, g) \in D$ .						
Linearity is rejected if $ Z_\rho $ exceeds the critical value.						
c.v. 1.65(90%)1.96(95%)2.58(99%)						

relations were unchanged.

There is strong evidence of nonlinearity in the stationary components of the system. Although the evidence is not completely robust to the sample of data tested, the nonlinearity does not stem solely from a structural break caused by the change in targeting approach in the early 1980s or the deregulation of interest rates, as the second sub-sample shows strong evidence of nonlinearity.

## 5 Conclusion

We have shown that cointegration relations in  $I(1)$  systems generally produce nonlinear dynamics. Our approach follows advancements in probability theory where many results that required independence, and therefore implied linearity, have been extended using martingales to allow for more general dependence or nonlinearity. Because the cointegration relations derived from  $I(1)$  systems are stationary, they can be tested for nonlinear serial dependence using standard polyspectral techniques. A feature of our two-stage method is that it tests a system of economic variables, or an equilibrium economic relation, for nonlinearity even though existing tests for nonlinearity, including the bispectrum test, are univariate.

Tests for the existence of nonlinear dynamics require large sample sizes and may be adversely affected by aliasing and other problems associated with time aggregation. Interest rates are measured with high frequency and aliasing can be controlled by adequate attention to filter design. For these reasons, the conditions are more favorable to testing interest rate data for nonlinear dynamics than for most other variables that are important to the business cycle, money demand, and the monetary transmission mechanism. We found that short-term US interest rates are cointegrated and that the stationary components of the system are nonlinear. The Hinich nonlinearity test is conservative, which strengthens our finding of nonlinear interest rate dynamics.

These results suggest that the untested assumption of linearity may be incorrect. The failure to find robust evidence of nonlinearity in lower frequency macroeconomic time series may be due to the small sample sizes that can be obtained for those time series, in addition to problems associated with sampling and time aggregation. Our particular example shows that the spreads between the Commercial Paper, Treasury Bill, and Federal Funds rates exhibit nonlinear dynamics. Our results are consistent with work that suggests there are asymmetric effects of monetary policy on interest rates, such as Choi (1999). Our results suggest that better forecasts of these spreads might be obtained with nonlinear models, such as bilinear models.

## Appendix: Aliasing and Constructing Anti-Aliasing Filters

Let  $X_t$  be a stationary continuous-time series that is sampled at regular intervals of time,  $0, \delta T, 2\delta T, \dots, (N-1)\delta T$ .  $\delta T$  is called the sampling interval, and  $1/\delta T$  is the sampling rate. The sampled sequence is denoted  $X_{k\delta T}, k = 0, \dots, N-1$ .

The power spectrum of the continuous-time series is

$$g(f) = \int_{-\infty}^{\infty} c_{XX}(\tau) e^{-i(2\pi f)\tau} d\tau.$$

The power spectrum of the discrete-time sampled sequence,  $g_{\delta T}(f)$ , is given by the following:

$$g_{\delta T}(f) = \sum_{j=-\infty}^{\infty} g\left(f + \frac{j}{\delta T}\right) \quad (5.1)$$

for  $|f| \leq (1/2\delta T)$  (see Koopmans, 1975, pp. 66-73). The frequency  $f_N = (1/2\delta T)$  is called the Nyquist folding frequency. If  $g(f) = 0$  for all  $f \geq |f_N|$  then the power spectrum of the continuous-time series and the discrete-time sampled sequence are equal. If the continuous-time series does not have this property then the power spectrum at frequency,  $f$ , of the sampled sequence is equal to the sum of the values of the power spectrum of the continuous-time series at all frequencies of the form  $f + (j/\delta T)$  for  $j = 0, \pm 1, \pm 2, \dots$ . Thus, the low frequency harmonics are made indistinguishable from the combined power of higher frequency harmonics because of sampling. This phenomenon is called aliasing.

It is very important to eliminate any power in a time series at frequencies that exceed the Nyquist folding frequency prior to sampling, because failure to do so will lead to biased estimation due to aliasing. Aliasing has traditionally been described in the frequency domain, but Hinich and Rothman (1998) showed that aliasing corrupts the impulse response functions in the time domain and therefore leads to serious identification problems.

The same problem results if a discrete-time sequence is sampled at a lower frequency, such as sampling a daily interest rate at weekly intervals. In this case, the sampling interval is  $\delta T = 7$  and the Nyquist folding frequency is  $(1/2\delta T) = (1/14)$ . If the daily interest rates have power at frequencies exceeding  $(1/14)$  then aliasing will occur. The solution to this problem is to filter the daily interest rates in such a way that the power spectrum of the filtered rates will be zero at frequencies exceeding the Nyquist. If  $\{g_j\}$  are the filter weights then the power spectrum of the filtered sequence

equals the power spectrum of the underlying sequence multiplied by the gain of the filter  $|G(f)|^2$  where  $G(f) = \sum_{j=-\infty}^{\infty} g_j e^{-i(2\pi f)j}$ . The solution to the aliasing problem would be to design a filter with gain:

$$|G(f)|^2 = \begin{cases} 1 & \text{if } |f| \leq f_N \\ 0 & \text{if } |f| > f_N \end{cases}. \quad (5.2)$$

This gain function corresponds to the ideal symmetric low-pass filter with weights

$$g_j = \begin{cases} \sin(2\pi f_N / \pi k) & k = \pm 1, \pm 2, \dots \\ 2f_N = 1/\delta T & k = 0 \end{cases} \quad (5.3)$$

which cannot be realized with a finite data sample. In fact, the rate of decrease of the filter weights is too slow to simply truncate the filter at some finite number of leads and lags. The usual solution is to taper the weights of the ideal filter. We taper the ideal weights using a Hanning cosine taper. This filter is referred to as an anti-aliasing filter in the text.

Applying the anti-aliasing filter produces a business day series that should not contain power at frequencies higher than every two weeks. This series still suffers from problems created by missing values caused by holidays. We subsequently resample our series on every Wednesday to avoid the missing value problem. Since the Nyquist frequency is then every two weeks, the resulting weekly series should avoid aliasing. The combination of applying the anti-aliasing filter and then decimating to the weekly sample produces a multi-rate filter.

The common approach in economics is to report unweighted weekly averages of daily interest rates. Weekly averages are also effectively produced by a multi-rate filter: combining a filter with decimation. The filter is an unweighted averaging filter that has a wider main lobe and much larger side lobes than the anti-aliasing filter we use. Weekly averages therefore potentially are strongly aliased. Monthly and quarterly averages that are often used in studies of the real interest rate are similarly aliased.

## References

- Ashley, R., D. M. Patterson, and M. J. Hinich (1986): "A diagnostic test for non-linear serial dependence in time series fitting errors," *Journal of Time Series Analysis*, 7, 165–178.
- Backus, D. K. and S. E. Zin (1993): "Long-memory inflation uncertainty: evidence from the term structure of interest rates," *Journal of Money, Credit, and Banking*, 25, 681–700.
- Barnett, W. A., A. R. Gallant, M. J. Hinich, J. A. Jungeilges, D. T. Kaplan, and M. J. Jensen (1995): "Robustness of nonlinearity and chaos tests to measurement error, inference method, and sample size," *Journal of Economic Behavior and Organization*, 27, 301–320.
- Barnett, W. A., A. R. Gallant, M. J. Hinich, J. A. Jungeilges, D. T. Kaplan, and M. J. Jensen (1997): "A single-blind controlled competition among tests for nonlinearity and chaos," *Journal of Econometrics*, 82, 157–192.
- Barnett, W. A., R. Gallant, M. J. Hinich, M. Jensen, and J. Jungeilges (1996): "Comparisons of the available tests for nonlinearity and chaos," in W. Barnett, G. Gandolfo, and C. Hillinger, eds., *Dynamic Disequilibrium Modeling: Theory and Applications*, Cambridge: Cambridge University Press, 313–346.
- Barnett, W. A., B. E. Jones, and T. D. Nesmith (2000): "Time series cointegration tests and non-linearity," in W. A. Barnett, D. F. Hendry, S. Hylleberg, T. Terasvirta, D. Tjostheim, and A. Wurtz, eds., *Nonlinear Econometric Modeling in Time Series*, Cambridge: Cambridge University Press, 9–30.
- Bierens, H. J. (1997): "Nonparametric cointegration analysis," *Journal of Econometrics*, 77, 379–404.
- Bierens, H. J. (2005): *EasyReg International*, Department of Economics, Pennsylvania State University, University Park, PA, URL <http://econ.la.psu.edu/~hbierens/EASYREG.HTM>.
- Calza, A. and J. Sousa (2006): "Output and inflation responses to credit shocks: are there threshold effects in the euro area?" *Studies in Nonlinear Dynamics and Econometrics*, 10, Article 3.
- Choi, W. G. (1999): "Asymmetric monetary effects on interest rates across monetary policy stances," *Journal of Money, Credit, and Banking*, 31, 386–416.
- Coakley, J. and A.-M. Fuertes (2001): "Nonparametric cointegration analysis of real exchange rates," *Applied Financial Economics*, 11, 1–8.
- Davidson, J. (1994): *Stochastic Limit Theory*, New York: Oxford University Press.
- Engle, R. F. and C. W. J. Granger (1987): "Co-integration and error correction: representation, estimation, and testing," *Econometrica*, 55, 251–276.

- Gil-Alana, L. A. (2004): "Long memory in the U.S. interest rate," *International Review of Financial Analysis*, 13, 265–276.
- Granger, C. W. J. (1991): "Some recent generalizations of cointegration and the analysis of long-run relationships," in R. F. Engle and C. W. J. Granger, eds., *Long-run Economic Relationships: Readings in Cointegration*, New York: Oxford University Press, 65–83.
- Hall, P. and C. C. Heyde (1980): *Martingale Limit Theory and Its Application*, New York: Academic Press.
- Hansen, H. and K. Juselius (2006): *CATS in RATS: Cointegration analysis of time series*, Illinois: Estima, Version 2.
- Hinich, M. J. (1982): "Testing for gaussianity and linearity of a stationary time series," *Journal of Time Series Analysis*, 3, 169–176.
- Hinich, M. J., E. M. Mendes, and L. Stone (2005): "Detecting nonlinearity in time series: surrogate and bootstrap approaches," *Studies in Nonlinear Dynamics and Econometrics*, 9, 13, article 3.
- Hinich, M. J. and G. R. Messer (1995): "On the principle domain of the discrete bispectrum of a stationary signal," *IEEE Transactions on Signal Processing*, 43, 2130–2134.
- Hinich, M. J. and D. M. Patterson (1985): "Evidence of nonlinearity in daily stock returns," *Journal of Business and Economic Statistics*, 3, 69–77.
- Hinich, M. J. and D. M. Patterson (1987): "A new diagnostic test of model inadequacy which uses the martingale difference criterion," *Journal of Time Series Analysis*, 13, 233–252.
- Hinich, M. J. and D. M. Patterson (1989): "Evidence of non-linearity in the trade by trade stock market return generating process," in W. Barnett, J. Geweke, and K. Shell, eds., *Economic Complexity: Chaos, Sunspots, Bubbles and Non-linearity*, Cambridge: Cambridge University Press, 383–409.
- Hinich, M. J. and P. Rothman (1998): "Frequency-domain test of time reversibility," *Macroeconomic Dynamics*, 2, 72–88.
- Hinich, M. J. and P. Wild (2001): "Testing time-series stationarity against an alternative whose mean is periodic," *Macroeconomic Dynamics*, 5, 380–412.
- Johansen, S. (1988): "Statistical analysis of cointegrating vectors," *Journal of Economic Dynamics and Control*, 12, 231–254.
- Johansen, S. (1992): "Determination of cointegration rank in the presence of a linear trend," *Oxford Bulletin of Economics and Statistics*, 54, 383–97.

- Karanasos, M., S. H. Sekioua, and N. Zeng (2006): "On the order of integration of monthly US ex-ante and ex-post real interest rates: new evidence from over a century of data," *Economics Letters*, 90, 163–169.
- Koopmans, L. H. (1975): *The Spectral Analysis of Time Series*, New York: Academic Press.
- Lee, Y.-S., T.-H. Kim, and P. Newbold (2005/6): "Spurious nonlinear regressions in econometrics," *Economics Letters*, 87, 301–306.
- Maki, D. (2003): "Nonparametric cointegration analysis of the nominal interest rate and expected inflation rate," *Economics Letters*, 81, 349–354.
- Nelson, C. R. and C. R. Plosser (1982): "Trends and random walks in macroeconomic time series: some evidence and implications," *Journal of Monetary Economics*, 10, 139–162.
- Pantula, S. G. (1989): "Testing for unit roots in time series data," *Econometric Theory*, 5, 256–271.
- Pfann, G. A., P. C. Schotman, and R. Tschernig (1996): "Nonlinear interest rate dynamics and implications for the term structure," *Journal of Econometrics*, 74, 149–176.
- Priestley, M. B. (1988): *Non-linear and non-stationary time series analysis*, London; New York: Academic Press.
- Psaradakis, Z., M. Sola, and F. Spagnolo (2006): "Instrumental-variables estimation in markov switching models with endogeneous explanatory variables: an application to the term structure of interest rates," *Studies in Nonlinear Dynamics and Econometrics*, 10, Article 1.
- Rapach, D. E. and C. E. Weber (2004): "Are real interest rates really nonstationary? New evidence from tests with good size and power," *Journal of Macroeconomics*, 26, 409–430.
- Rose, A. K. (1988): "Is the real interest rate stable?" *Journal of Finance*, 43, 1095–1112.
- Rudebusch, G. D. (1995): "Federal Reserve interest rate targeting, rational expectations, and the term structure," *Journal of Monetary Economics*, 35, 245–274.
- Rugh, W. J. (1981): *Nonlinear System Theory: The Volterra-Wiener Approach*, Baltimore: John Hopkins University Press.
- Schetzen, M. (1980): *The Volterra and Wiener Theories of Nonlinear Systems*, New York: Wiley.
- Schetzen, M. (1981): "Nonlinear system modeling based on the Wiener theory," *Proceedings of the IEEE*, 69, 1557–1573.
- Tsay, W.-J. (2000): "Long memory story of the real interest rate," *Economics Letters*, 67, 325–330.