Statistics and Its Interface Volume 10 (2017) 81–92

Testing trend stationarity of functional time series with application to yield and daily price curves*

Piotr Kokoszka† and Gabriel Young

CONTENTS

Econometric and financial data often take the form of a collection of curves observed consecutively over time. Examples include intraday price curves, term structure curves, and intraday volatility curves. Such curves can be viewed as functional time series. A fundamental issue that must be addressed, before an attempt is made to statistically model or predict such series, is whether they can be assumed to be stationary with a possible deterministic trend. This paper extends the KPSS test to the setting of functional time series. We propose two testing procedures: Monte Carlo and asymptotic. The limit distributions of the test statistics are specified, the procedures are algorithmically described and illustrated by application to yield curves and daily price curves.

Keywords and phrases: Functional data, Daily price curves, Integrated time series, Random walk, Trend stationarity.

1. INTRODUCTION

Many econometric and financial data sets take the form of a time series of curves, or functions. The best known and most extensively studied data of this form are yield curves. Even though they are observed at discrete maturities, in financial theory they are viewed as continuous functions, one function per month or per day. The yield curves can thus be viewed as time series of curves, functional time series. Other examples include intraday price, volatility or volume curves. Intraday price curves are smooth, volatility and volume curves are noisy and must be smoothed before they can be effectively treated as curves. As with scalar and vector valued time series, it is important to describe the random structure of a functional time series. A fundamental question, which has received a great deal of attention in econometric research, is whether the time series has a random walk, or unit root, component. The present paper addresses this issue in the context of functional time series by proposing extensions of the KPSS test of [15] and applying them to several data sets.

The work of [15] was motivated by the fact that unit root tests developed by [5, 6], and [24] indicated that most aggregate economic series had a unit root. In these tests, the null hypothesis is that the series has a unit root. Since such tests have low power in samples of sizes occurring in many applications, [15] proposed that trend stationarity should be considered as the null hypothesis, and the unit root should be the alternative. Rejection of the null could then be viewed as convincing evidence in favor of the unit root hypothesis. It was soon realized that the KPSS test of [15] has a much broader utility. For example, [16] and [8] used it to detect long memory, with short memory as the null hypothesis; [4] developed a robust version of the KPSS test. The work of [17] is crucial because he observed that under temporal dependence, to obtain parameter–free limit null distributions, statistics similar to the KPSS statistic must be normalized by the long run variance rather than by the sample variance.

In the functional setting, the null hypothesis of trend stationarity is stated as follows:

$$H_0: \quad X_n(t) = \mu(t) + n\xi(t) + \eta_n(t), \tag{1.1}$$

*We thank the editors and two referees for providing advice on how to refocus and restructure the paper. We thank Professor Hong Miao of the Department of Finance and Real Estate at Colorado State University for providing the data used in Section 5.
†Corresponding author.

where n is the serial number of the day in our applications, and t refers to "time" for each function. For example, for the intraday price curves, t is the time within a trading day, measured in minutes or at an even finer resolution. For the yield curves, t does not correspond to physical time but to time until expiration, the maturity horizon of a bond. The functions μ and ξ correspond, respectively, to the intercept and slope. The errors η_n are also functions which model random departures of the observed functions X_n from a deterministic model. Under the alternative, the model contains a random walk component:

$$H_A: \quad X_n(t) = \mu(t) + n\xi(t) + \sum_{i=1}^{n} u_i(t) + \eta_n(t), \tag{1.2}$$

where $u_1, u_2, \ldots$ are mean zero identically distributed random functions.

Our approach to testing exploits the ideas of functional data analysis (FDA), mostly those related to functional principal component expansions; several monographs, e.g. [23] and [9], explain them in detail. Application of FDA methodology in an econometric context is not new. Among others, [13] studied prediction of yield curves, [19] considered functional modeling of volatility, [14] used a regression type model to explain the shapes of price curves. A contribution most closely related to the present work is that of [11] who developed a test of level stationarity. Incorporating a possible trend leads to different limit distributions and more complex numerical implementations.

The remainder of the paper is organized as follows. After introducing the required concepts and notation in Section 2, we present in Section 3 the large sample results needed to construct the tests. The resulting testing procedures are described in Section 4. Section 5 presents their applications to data representing bond, equity, forex and commodity markets. In this last section, we also examine and discuss finite sample properties of the tests.

2. PRELIMINARIES

To understand the construction of the tests in the setting of functional time series, we must introduce some notation and definitions. This is the objective of the present section.

All random functions and deterministic functional parameters μ and ξ are assumed to be elements of the Hilbert space $L^2 = L^2([0,1])$ with the inner product $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$. This means that the domain of all functional observations, e.g. of the daily price or yield curves, has been normalized to be the unit interval. If the limits of integration are omitted, integration is over the interval $[0,1]$. All random functions are assumed to be square integrable, i.e., $E\,||\eta_n||^2 < \infty,\ E\,||u_n||^2 < \infty$, where the norm is generated by the inner product, i.e. $||f||^2 = \int f^2(t)dt$.

[15] assumed that the errors η_n are iid, but subsequent research extended their work to errors which form a stationary time series, see, e.g., [8] and the references therein. In the case of scalar observations, temporal dependence can be quantified in many ways, e.g., via structural, mixing or cumulant conditions, and a large number of asymptotic results established under such assumptions can be used. For functional time series, the corresponding results are much fewer and fall into two categories: 1) those derived assuming a linear, ARMA type, structure, see, e.g., [1]; 2) those assuming a nonlinear moving average representation (Bernoulli shifts) with the decay of dependence specified by a moment condition. We have established the asymptotic validity of our tests assuming very general conditions falling into the second category. Detailed formulations of these conditions are presented in [25]. In essence, the error functions η_n and u_i need not be iid, but merely must form stationary and weakly dependent sequences.

Next we define the long–run covariance function of the errors η_n and its estimator. The long–run covariance function is defined as

$$c(t,s) = E\eta_0(t)\eta_0(s) + \sum_{i=1}^{\infty} \left(E\eta_0(t)\eta_i(s) + E\eta_0(s)\eta_i(t)\right). \tag{2.1}$$

The series defining the function $c(t,s)$ converges in $L^2([0,1]\times[0,1])$, see [10]. The function $c(t,s)$ is positive definite. Therefore there exist eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$, and orthonormal eigenfunctions $\phi_i(t)$, $0 \leq t \leq 1$, satisfying

$$\lambda_i \phi_i(t) = \int c(t,s)\phi_i(s)ds, \quad 0 \leq i \leq \infty. \tag{2.2}$$

The eigenvalues λ_i play a crucial role in our tests. They are estimated by the sample, or empirical, eigenvalues defined by

$$\hat{\lambda}_i \hat{\phi}_i(t) = \int \hat{c}(t,s)\hat{\phi}_i(s)ds, \quad 0 \leq i \leq N, \tag{2.3}$$

where $\hat{c}(\cdot,\cdot)$ is an estimator of (2.1), and N is the sample size of the functional time series. We use a kernel estimator similar to that introduced by [10], but with suitably defined residuals in place of the centered observations X_n. To define model residuals, consider the least squares estimators of the functional parameters $\xi(t)$ and $\mu(t)$ in model (1.1):

$$\hat{\xi}(t) = \frac{1}{s_N} \sum_{n=1}^{N} \Big(n - \frac{N+1}{2}\Big) X_n(t) \tag{2.4}$$

with

$$s_N = \sum_{n=1}^{N} \Big(n - \frac{N+1}{2}\Big)^2 \tag{2.5}$$

and

$$\hat{\mu}(t) = \bar{X}(t) - \hat{\xi}(t)\Big(\frac{N+1}{2}\Big). \tag{2.6}$$

The functional residuals are therefore

$$e_n(t) = (X_n(t) - \bar{X}(t)) - \hat{\xi}(t)\Big(n - \frac{N+1}{2}\Big), \tag{2.7}$$

where $1 \le n \le N$. Defining their empirical autocovariances by

$$\hat{\gamma}_i(t,s) = \frac{1}{N}\sum_{j=i+1}^{N} e_j(t)e_{j-i}(s), \quad 0 \le i \le N-1, \tag{2.8}$$

leads to the kernel estimator

$$\hat{c}(t,s) = \hat{\gamma}_0(t,s) + \sum_{i=1}^{N-1} K\Big(\frac{i}{h}\Big)(\hat{\gamma}_i(t,s) + \hat{\gamma}_i(s,t)). \tag{2.9}$$

It can be shown that under the usual assumptions on the kernel function K and the bandwidth h ($h \to \infty,\ h/N \to 0$),

$$\iint [\hat{c}(t,s) - c(t,s)]^2 dtds \xrightarrow{P} 0, \quad \text{as } N \to \infty, \tag{2.10}$$

details are presented in [25].

We conclude this section by stating the definitions of Gaussian stochastic processes which are needed to construct the limit distributions of the test statistics. Recall that if $\{W(x), 0 \le x \le 1\}$ is a standard Brownian motion (Wiener process), then the Brownian bridge is defined by $B(x) = W(x) - xW(x),\ 0 \le x \le 1$. The second–level Brownian bridge is defined by

$$\begin{aligned} V(x) &= W(x) + \Big(2x - 3x^2\Big)W(1) \\ &\quad + \Big(-6x + 6x^2\Big)\int_0^1 W(y)dy, \quad 0 \le x \le 1. \end{aligned} \tag{2.11}$$

Both the Brownian bridge and the second–level Brownian bridge are special cases of the generalized Brownian bridge introduced by [18] who studied the asymptotic behavior of partial sums of polynomial regression residuals. Process (2.11) appears as the null limit of the KPSS statistic of [15]. We will see in Section 3 that for functional data the limit involves an infinite sequence of independent and identically distributed second-level Brownian bridges $V_1(x), V_2(x), \ldots$.

3. LARGE SAMPLE LIMITS

[11] developed tests of level–stationarity of a functional time series, i.e., of the null hypothesis $X_n(t) = \mu(t) + \eta_n(t)$, using the partial sum process

$$\begin{aligned} U_N(x,t) &= \frac{1}{\sqrt{N}}\sum_{n=1}^{\lfloor Nx \rfloor}\Big(X_n(t) - \bar{X}(t)\Big) \\ &= S_N(x,t) - \frac{\lfloor Nx \rfloor}{N}S_N(1,t), \end{aligned}$$

where $S_N(x,t)$ is the partial sum process of the curves $X_1(t), X_2(t), \ldots, X_N(t)$ is defined by

$$S_N(x,t) = \frac{1}{\sqrt{N}}\sum_{n=1}^{\lfloor Nx \rfloor} X_n(t), \quad 0 \le t, x \le 1. \tag{3.1}$$

The process $U_N(x,t)$ has the form of a functional Brownian bridge. Their main statistic

$$\begin{aligned} T_N &= \iint U_N^2(x,t)dtdx \\ &= \int \|U_N(x,\cdot)\|^2 dx, \quad 0 \le t, x \le 1, \end{aligned}$$

is asymptotically distributed, under the null, as $\sum_{i=1}^{\infty}\lambda_i \int B_i^2(x)dx$, where $\lambda_1, \lambda_2, \ldots$ are eigenvalues of the long–run covariance function of the observations X_n, and $B_1, B_2, \ldots$ are iid Brownian bridges. In the case of trend stationarity, a different distribution arises; the B_i must be replaced by second level Brownian bridges, and the λ_i are defined differently. The remainder of this section explains the details.

The test statistic for the trend-stationary case is based on the partial sum process of residuals (2.7), i.e., on the two–parameter process

$$Z_N(x,t) = \frac{1}{\sqrt{N}}\sum_{n=1}^{\lfloor Nx \rfloor} e_n(t). \tag{3.2}$$

A suitable test statistic is given by

$$\begin{aligned} R_N &= \iint Z_N^2(x,t)dtdx \\ &= \int \|Z_N(x,\cdot)\|^2 dx, \quad 0 \le t, x \le 1. \end{aligned} \tag{3.3}$$

It can be shown, [25], that under the null hypothesis,

$$R_N \xrightarrow{\mathcal{D}} \sum_{i=1}^{\infty}\lambda_i \int V_i^2(x)dx, \tag{3.4}$$

where $\lambda_1,\ \lambda_2, \ldots$ are the eigenvalues of the long–run covariance function (2.1), and $V_1, V_2, \ldots$ are iid second–level Brownian bridges.

We now explain the issues arising in the functional case by comparing our result to that obtained by [15]. If all curves are constant functions ($X_i(t) = X_i$ for $t \in [0,1]$), the statistic R_N given by (3.3) is the numerator of the KPSS test statistic of [15], which is given by

$$\mathrm{KPSS}_N = \frac{1}{N^2\hat{\sigma}_N^2}\sum_{n=1}^{N} S_n^2 = \frac{R_N}{\hat{\sigma}_N^2},$$

where $\hat{\sigma}_N^2$ is a consistent estimator of the long-run variance σ^2 of the residuals. In the scalar case, (3.4) reduces to $R_N \xrightarrow{\mathcal{D}} \sigma^2 \int_0^1 V^2(x)dx$, where $V(x)$ is a second–level Brownian bridge. If $\hat{\sigma}_N^2$ is a consistent estimator of σ^2, the result of [15] is recovered, i.e. $\text{KPSS}_N \xrightarrow{\mathcal{D}} \int_0^1 V^2(x)dx$. In the functional case, the eigenvalues λ_i can be viewed as long–run variances of the residual curves along the principal directions determined by the eigenfunctions of the kernel $c(\cdot,\cdot)$ defined by (2.1). To obtain a test analogous to the scalar KPSS test, with a parameter free limit null distribution, we must construct a statistic which involves a division by consistent estimators of the λ_i. We use only d largest eigenvalues in order not to increase the variability of the statistic caused by division by small empirical eigenvalues. A suitable statistic is

$$R_N^0 = \sum_{i=1}^{d} \frac{1}{\hat{\lambda}_i} \int_0^1 \langle Z_N(x,\cdot), \hat{\phi}_i\rangle^2 dx, \tag{3.5}$$

where the sample eigenvalues $\hat{\lambda}_i$ and eigenfunctions $\hat{\phi}_i$ are defined by (2.3). Statistic (3.5) extends the statistic KPSS_N. It can be shown that under suitable assumptions, [25],

$$R_N^0 \xrightarrow{\mathcal{D}} \sum_{i=1}^{d} \int_0^1 V_i^2(x)dx, \tag{3.6}$$

with the $V_i, 1 \le i \le d$, the same as in (3.4).

Section 4 describes how the tests based on relations (3.4) and (3.6) are implemented.

4. ALGORITHMIC DESCRIPTION OF THE TEST PROCEDURES

This section provides step–by–step descriptions of the test procedures based on limit relations (3.4) and (3.6).

ALGORITHM 4.1 [Monte Carlo test based on relation (3.4)]

1. Estimate the null model (1.1) and compute the residuals defined in equation (2.7).
2. Select kernel K and a bandwidth h in (2.9) and compute the eigenvalues $\hat{\lambda}_i\ \hat{\phi}_i,\ 1 \le i \le N$, defined by (2.3).
3. Simulate a large number, say $G = 10{,}000$, of vectors $[V_1, V_2, \ldots, V_N]$ consisting of independent second level Brownian bridge processes V_i defined in (2.11). Find the 95th percentile, R_{critical}, of the G replications of
$$R_N^{\star} = \sum_{i=1}^{N} \hat{\lambda}_i \int_0^1 V_i^2(x)dx.$$
4. Compute the test statistic R_N defined in (3.3). If $R_N \ge R_{\text{critical}}$, reject H_0 at the 5% significance level.

In most applications, the $\hat{\lambda}_i$ decay very quickly to zero, so if N is large, it can be replaced in Algorithm 4.1 by a smaller number, e.g by $d = 20$, and the empirical distribution of the $R_N^{\star}$ can be replaced by that of the $R_d^{\star}$. In Algorithm 4.1 the critical value must be obtained via Monte Carlo simulations for each data set. In Algorithm 4.2, tabulated critical values

Table 1. Critical values of the distribution of the variable $R^0(d)$ given by (4.1)

	d	1	2	3	4	5
	10%	0.1201	0.2111	0.2965	0.3789	0.4576
Size	5%	0.1494	0.2454	0.3401	0.4186	0.5068
	1%	0.2138	0.3253	0.4257	0.5149	0.6131
	d	6	7	8	9	10
	10%	0.5347	0.6150	0.6892	0.7646	0.8416
Size	5%	0.5909	0.6687	0.7482	0.8252	0.9010
	1%	0.6960	0.7799	0.8574	0.9487	1.0326

can be used. These depend on the number d of the functional principal components used to construct statistic R_N^0. Typically d is a small, single digit, number. Table 1 lists selected critical values. They have been obtained by simulating $G = 10{,}000$ vectors $[V_1, V_2, \ldots, V_d]$ and finding the percentiles of the G replications of

$$R^0(d) = \sum_{i=1}^{d} \int_0^1 V_i^2(x)dx. \tag{4.1}$$

ALGORITHM 4.2 [Asymptotic test based on relation (3.6)]

1. Perform steps 1 and 2 of Algorithm 4.1.
2. Choose the smallest d such that $\sum_{i\le d} \hat{\lambda}_i / \sum_{i\le N} \hat{\lambda}_i > 0.85$.
3. Calculate the statistic R_N^0 given by (3.5) and reject H_0 if $R_N^0 > R_{\text{critical}}^0$, with the critical value given in Table 1.

The 85% rule in Step 2 is a rule of thumb; asking for 85% of the variance to be explained is based on good empirical results, leading to our choice above. In some applications, Step 2 may be replaced by a selection of d based on a visual fit of the truncated principal component expansion

$$X_n^{(d)}(t) = \hat{\mu}(t) + \sum_{j=1}^{d} \left\langle X_n, \hat{\phi}_j \right\rangle \hat{\phi}_j(t)$$

to the observed curves $X_n(t)$. In other applications, existing theory or experience may support certain choices of d. This is the case for the yield curves, which we use to illustrate the application of our tests (mean level plus $d = 2$ principal components). For financial data, d is generally small, with $d = 2, 3, 4$ being the typical values. However, for other types of data, e.g. for environmental data, d exceeding 10 may be needed. In such cases, caution is recommended in the application of Algorithm 4.2 as the resulting test may be numerically and statistically unstable due to the division by small $\hat{\lambda}_j$ which may exhibit large sampling variability.

An important step is the choice of h needed to estimate the long run covariance function. A great deal of research in this direction has been done for scalar and vector time

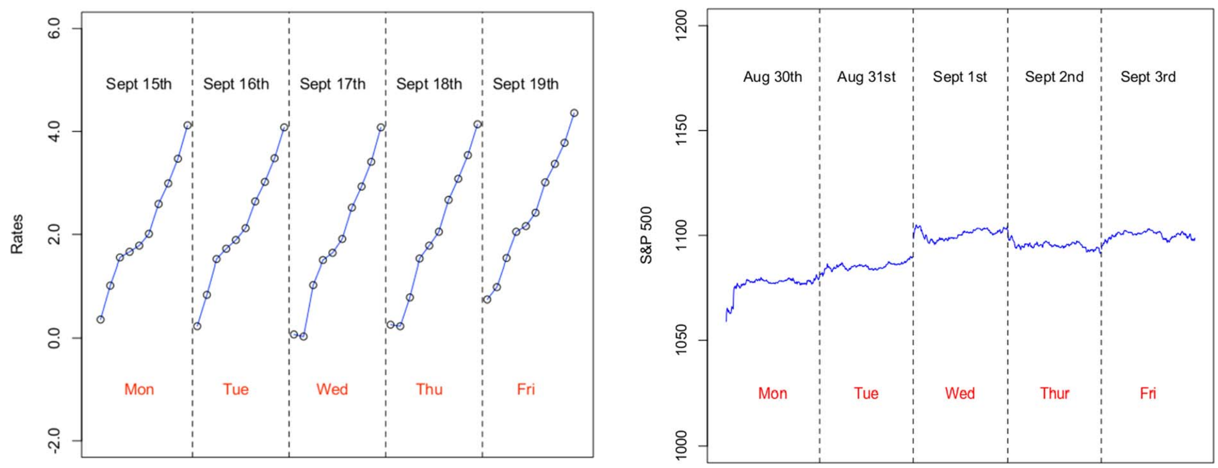

Figure 1. Left: five consecutive yield curves. Right: prices of the S&P 500 index over five consecutive days.

series. For functional time series, the method proposed by [12] often gives good results. It uses the flat top kernel

$$K(t) = \begin{cases} 1, & 0 \le t < 0.1 \\ 1.1 - |t|, & 0.1 \le t < 1.1 \\ 0, & |t| \ge 1.1 \end{cases} \tag{4.2}$$

advocated by Politis and Romano ([21], [22]) and [20], and a data–driven selection of h. This method performs well if the series length N is larger than several hundred, longer than the series we consider. In the simulations reported in [25] a deterministic bandwidth $h = N^{2/5}$ (combined with the flat top kernel) produced good size and power. The optimal selection of h is not a focus of this paper, this complex issue must be investigated in a separate work. As in the scalar case, it is however unlikely that a selection procedure that is uniformly optimal for all dependence structures can be found. In testing problems, it is useful to use several values of h and trust results which do not depend on h in a reasonable range.

5. APPLICATION TO YIELD AND DAILY PRICE CURVES

In this section, we apply the test procedures of Section 4 to several financial data sets which can be viewed as time series of functions. The most extensively studied series of this type is the series of yield curves. In the past, the series of monthly yield curves have been typically studied, but in recent years high quality data at the daily frequency have become available. On a given day, a yield curve shows the yield (interest) earned on a fixed income instrument as a function of maturity. In most economic studies, these are yields on bills and bonds issued by central banks. The shape and level of these curves reflect the expectations of investors on the future direction of a specific economic area, see e.g., Chapter 10 of [2] or [7]. Figures 1 (left panel) and 2 show, respectively, five consecutive yield curves and two sets of 250 yield curves. The question we want to answer is whether the time series of yield curves can be treated as stationary time series with trend, or if they contain a random walk component. Visual examination and economic interpretation of these data leads to the conclusion that a pure trend model will not hold over very long periods of time which include periods of growth and recession and changes in central bank policies. Over shorter periods of time, the trend model may however hold, and may be useful to investors in fixed income securities.

The second type of functional time series we study are daily price curves like those shown in the right panel of Figure 1. As noted in the introduction, whether a time series of closing prices on a specific asset contains a random walk (is a unit root process) has been one of the most extensively studied topics in finance. In contrast to these studies, we consider the series of price curves. Out of a large number of assets that are of interest, we selected the S&P 500 index, the US dollar index and light crude oil futures. These assets represent, respectively, the equity, currency and commodity markets. As for the bond market, trend stationarity will not hold over long periods of time, but our tests can identify periods for which it does hold.

The main objective of the empirical analysis presented in this section is to uncover commonalities and differences between the various classes of assets with respect to the trend behavior of specific daily functions. The analysis will also illustrate the statistical properties of the tests we propose.

5.1 Data description

As an example of the time series of yield curves we use the daily United States Federal Reserve yield curves defined for maturities of 1, 3, 6, 12, 24, 36, 60, 84, 120 and 360 months. The available data covers all business days from January 2001 to December 2013.

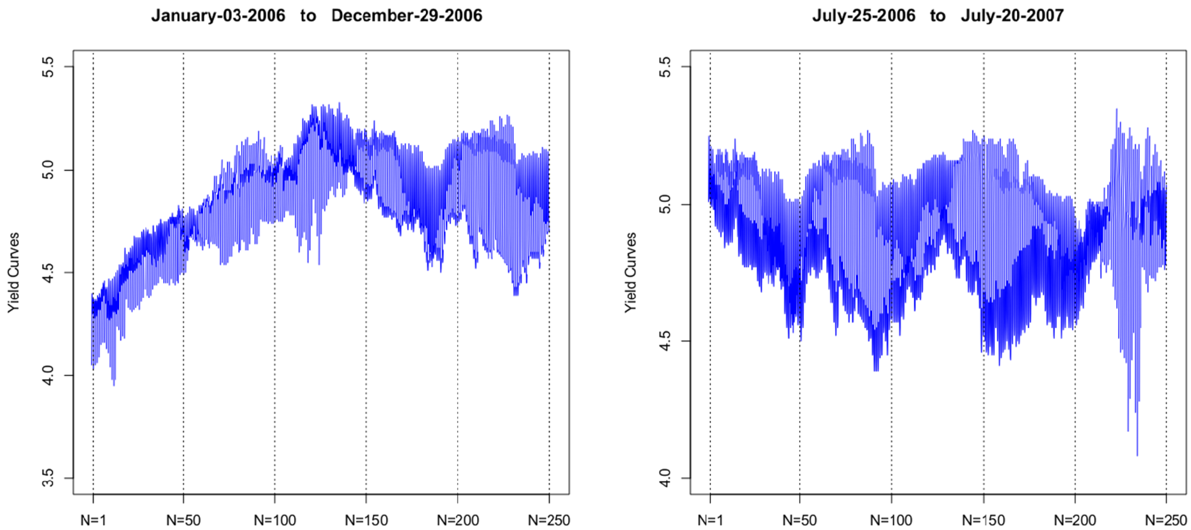

Figure 2. Consecutive yield curves over two time periods. Vertical lines show the location of sample sizes $N = 50, 100, 150, 200, 250$.

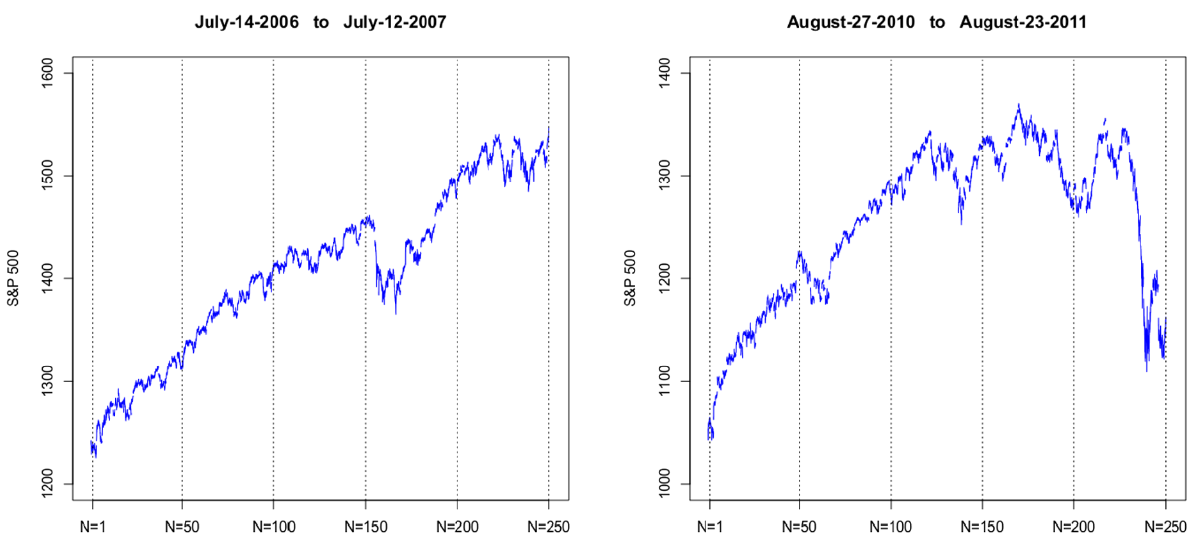

Figure 3. S&P 500 index for two different time periods.

The second data set is the Standard & Poor's 500 financial index (S&P 500) in one minute resolution. The index is a weighted average of stock values of the largest 500 U.S. companies. At each trading day, we consider a price curve. The last value on day $n-1$ is not the same as the first value on day n. An overnight jump of over half a percent is not unusual. Figure 3 shows the S&P 500 index over two periods. The available data cover the period of 23 years from January 1989 to December 2012.

The third data set is the U.S. dollar index. As a weighted average of exchange rates against several major currencies, it measures the value of the U.S. dollar relative to a collection of other foreign currencies. A higher index indicates that the U.S. dollar is stronger compared to foreign currencies. The index is traded and used for the construction of derivative instruments. Similar to the S&P 500 index, we use values in one minute resolution and consider one day as a single functional observation. However, instead of considering business days, we only exclude Saturdays for this data set. Figure 4 shows the U.S. dollar index over two different sampling periods. The available data cover the period of 23 years from January 1989 to December 2012.

The fourth data set consists of light crude oil futures. Light sweet crude oil futures and options are one of the worlds most highly traded energy products. Similar to the S&P 500 and the U.S. dollar index, we use minute–by–minute prices, and consider one day as a single functional observation. We exclude only Saturdays for this data set. Figure 5 shows the light crude oil futures over two different sampling periods. The available data cover

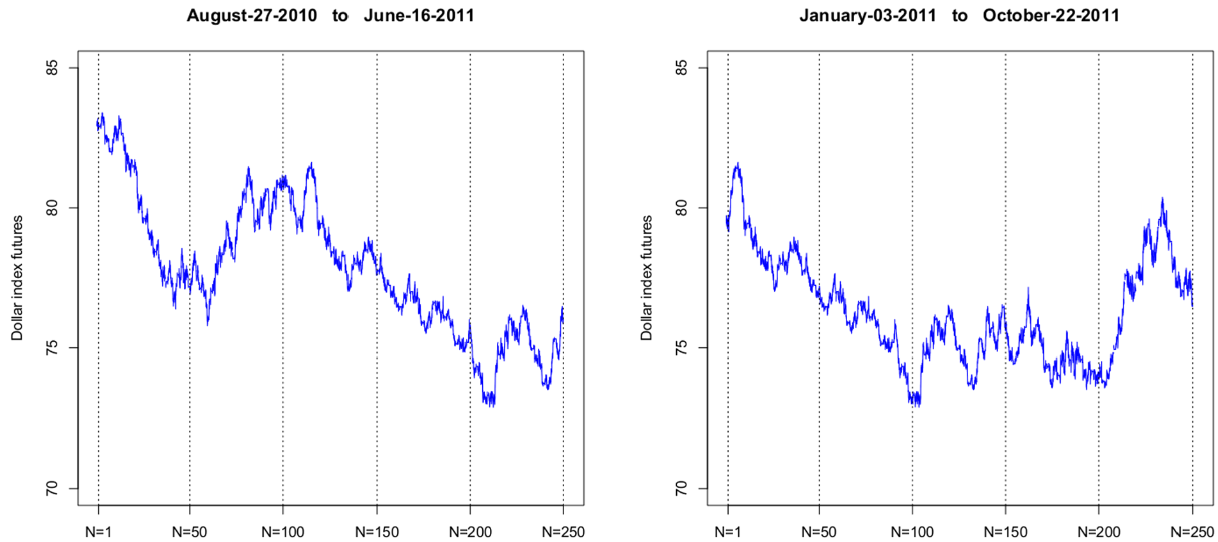

Figure 4. U.S. dollar index for two different time periods.

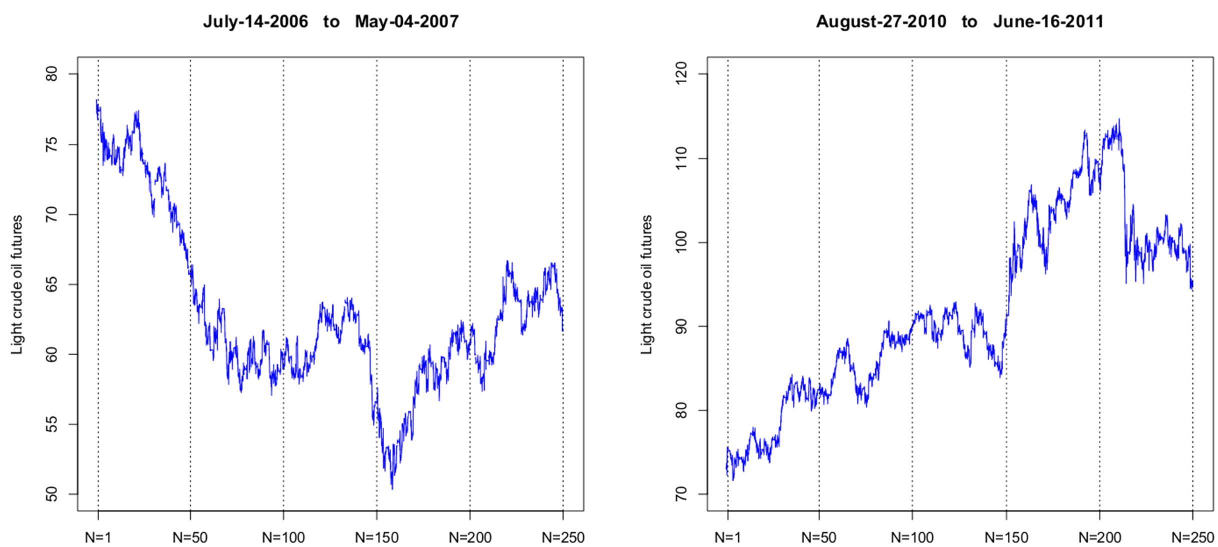

Figure 5. Light crude oil futures over two different time periods.

the period of 22 years from January 1989 to December 2011.

In Section 5.2, we show how the P–values broadly depend on the length of the series, N, and on the time period. We complement this analysis by focusing in Section 5.3 on selected time periods, those shown in Figures 2, 3, 4 and 5. This additional, more detailed analysis allows us to gain insights about the properties of the tests.

5.2 Long term trend characteristics of the curves

In this section, we display the P–values of the Monte Carlo test described in Algorithm 4.1 applied to the data described in Section 5.1. We also computed the P–values for the test based on Algorithm 4.2. While the P–values for the two tests are different, their general patterns are very similar, so to conserve space we focus on Algorithm 4.1 with the bandwith $h = N^{2/5}$. We take a closer look at the differences between the two algorithms and the effect of the bandwidth in Section 5.3.

The main finding of our analysis is that for time periods of length $N = 100$ days, what corresponds roughly to the number of business days in four months, it is not uncommon that the null hypothesis of trend stationarity is not rejected. For periods covering the whole year, the null hypothesis is generally rejected. However, the proportion and temporal pattern of rejections are different for different assets. For example, for the yield curves there is hardly any period when H_0 can

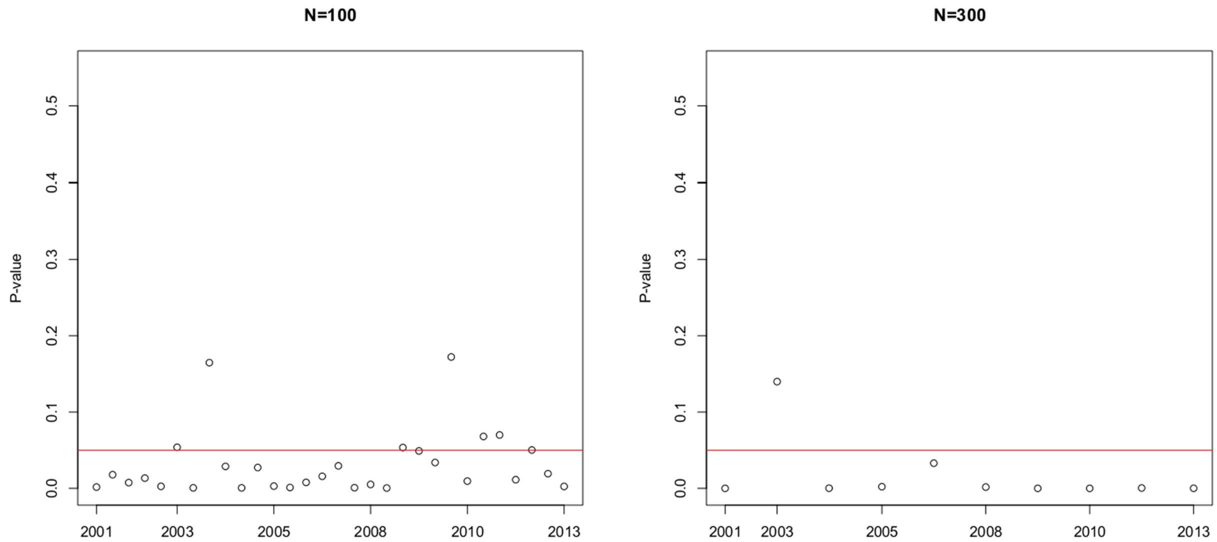

Figure 6. P–values of the test based on Algorithm 4.1 applied to the **Treasury Yield Curves***. The plot on the left shows thirty 100 day periods and the plot on the right shows ten 300 day periods.*

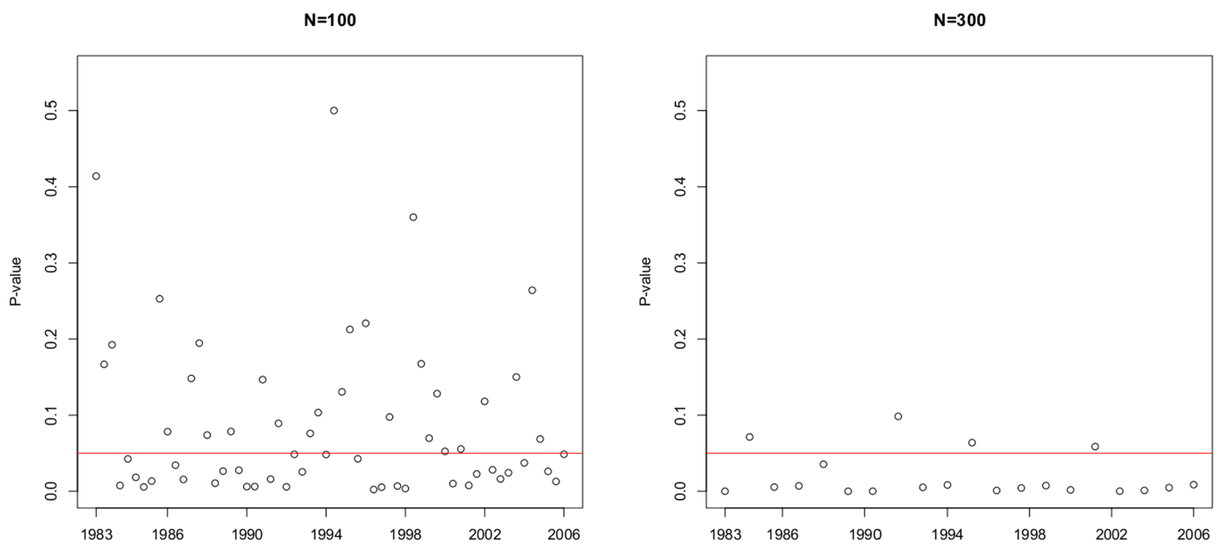

Figure 7. P–values of the test described in Algorithm 4.1 applied to the **S&P 500 index***. The plot on the left shows sixty 100 day periods and the plot on the right shows twenty 300 day periods.*

be accepted. This implies that this functional time series is not stationary even if a deterministic trend is allowed. This finding has implications for the prediction of yield curves; many methods assume a stationary model, some form of autoregression for factor coefficients. However, [3] obtained better prediction by assuming that the yield curves form only a locally stationary functional time series, i.e. stationary only on short subintervals. Our inferential procedures confirm the validity of such an approach. In the remainder of this section, we systematically present and discuss the results for all four data sets. For each asset, we consider all available consecutive, nonoverlapping periods of $N = 100$ and $N = 300$ days.

Figure 6 exhibits the pattern of P–values for the Daily United States Federal Reserve yield curves. Focusing first on periods of length $N = 100$, we see that 23 out of the 30 periods show P–values below the significance level of 0.05. As the sample size increases to $N = 300$, we see that 9 out of the 10 longer periods have P–values below the significance level of 0.05. As noted above, an overriding conclusion is that the yield curves do not follow a stationary model even with a trend, and a presence of a random walk component or some other changes in the the stochastic structure must be taken into account.

Figure 7 shows the P–values for the S&P 500 curves. For $N = 100$, 32 out of the 60 periods have P–values smaller

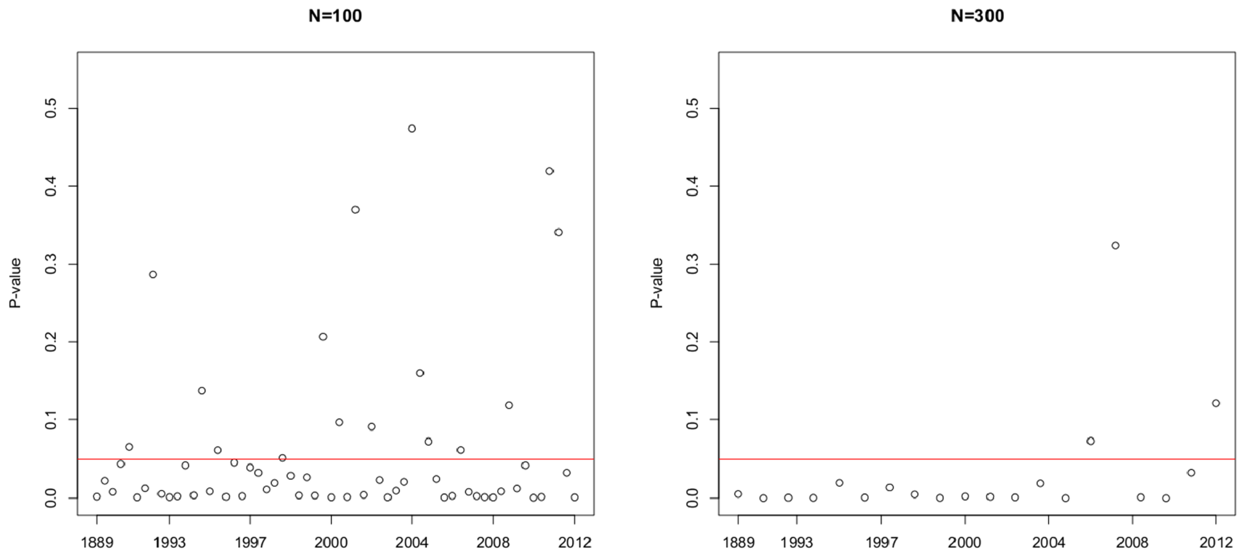

Figure 8. P–values of the test described Algorithm 4.1 applied to the **U.S. dollar index**. *The plot on the left shows sixty 100 day periods and the plot on the right shows twenty 300 day periods.*

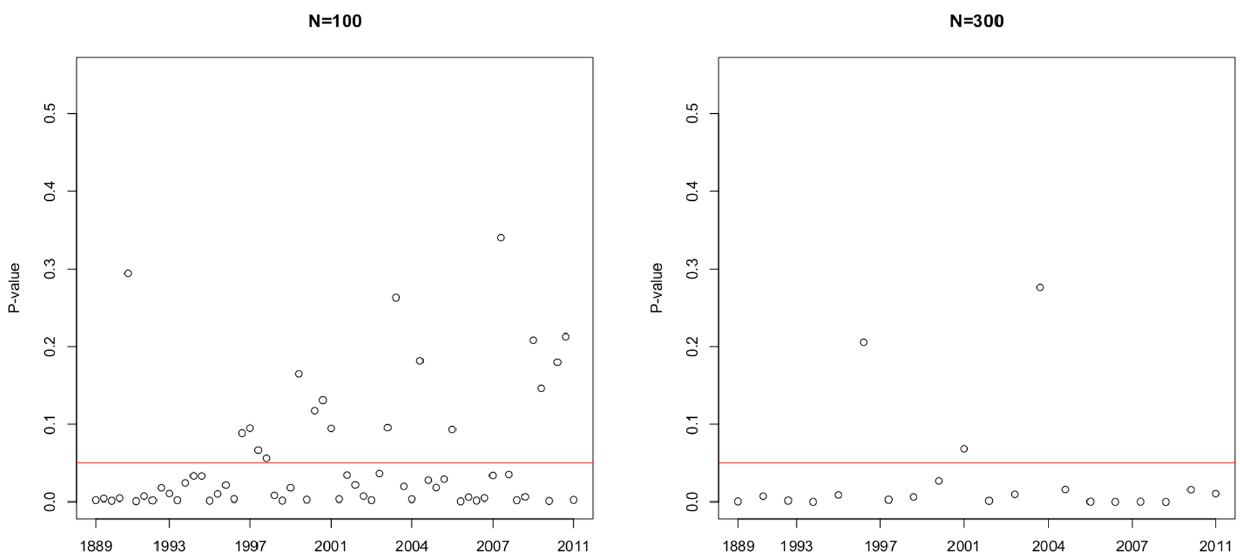

Figure 9. P–values of the test described Algorithm 4.1 applied to the **Oil Futures**. *The plot on the left shows sixty 100 day periods and the plot on the right shows twenty 300 day periods.*

than 0.05. In contrast to the yield curves, this shows that a stationary model with a trend can be suitable for many periods extending over several months; in most cases this corresponds to a persistent bull market, cf. Figure 3. However, as the sample size increases to $N = 300$, we see that 16 out of the 20 longer periods have P–values below 0.05; a bull market cannot last forever.

Figure 8 shows the P–values for the U.S. dollar index. For $N = 100$, 44 out of the 60 periods have P–values below 0.05. As the sample size increases to $N = 300$, we see that 17 out of the 20 longer periods have P–values smaller than 0.05. In terms of the trend behavior, the currency index is somewhere between the yield curves and the equity index. There are periods of trend stationarity but they are less frequent than for equities.

Finally, we turn to the light crude oil futures. Figure 9 shows the P–values. For $N = 100$, 42 out of the 60 periods have P–values smaller than 0.05. In this case, an interesting temporal pattern of these P–values can be seen. Starting from 1997, there are periods with increasing P–values, indicting that a trend model might often be suitable. This agrees with a persistent, almost linear, decline in prices from December 1996 to December 1998 followed by a long rise from January 1999 up to the summer of 2008, just before

Table 2. P–values of the tests of Section 4 applied to **Treasury Yield Curves**. *The data are shown in Figure 2*

Time period	Sample size	Bandwidth	R_N	R_N^0
01/03/2006 – 03/15/2006	$N = 50$	$N^{1/3}$	0.0200	0.0206
		$N^{2/5}$	0.0432	0.0561
		$N^{1/2}$	0.0878	0.1361
01/03/2006 – 05/25/2006	$N = 100$	$N^{1/3}$	0.0022	0.0024
		$N^{2/5}$	0.0108	0.0116
		$N^{1/2}$	0.0410	0.0661
01/03/2006 – 08/07/2006	$N = 150$	$N^{1/3}$	0.0003	0.0005
		$N^{2/5}$	0.0013	0.0007
		$N^{1/2}$	0.0117	0.0086
01/03/2006 – 10/18/2006	$N = 200$	$N^{1/3}$	0.0000	0.0000
		$N^{2/5}$	0.0005	0.0001
		$N^{1/2}$	0.0015	0.0050
01/03/2006 – 12/29/2006	$N = 250$	$N^{1/3}$	0.0000	0.0000
		$N^{2/5}$	0.0000	0.0001
		$N^{1/2}$	0.0011	0.0051
Time period	Sample size	Bandwidth	R_N	R_N^0
07/25/2006 – 10/03/2006	$N = 50$	$N^{1/3}$	0.0065	0.0109
		$N^{2/5}$	0.0164	0.0290
		$N^{1/2}$	0.0404	0.0704
07/25/2006 – 12/14/2006	$N = 100$	$N^{1/3}$	0.0075	0.0169
		$N^{2/5}$	0.0272	0.0590
		$N^{1/2}$	0.0967	0.1885
07/25/2006 – 02/28/2007	$N = 150$	$N^{1/3}$	0.0002	0.0005
		$N^{2/5}$	0.0027	0.0044
		$N^{1/2}$	0.0230	0.0447
07/25/2006 – 05/09/2007	$N = 200$	$N^{1/3}$	0.0030	0.0081
		$N^{2/5}$	0.0166	0.0451
		$N^{1/2}$	0.1035	0.1965
07/25/2006 – 07/20/2007	$N = 250$	$N^{1/3}$	0.0000	0.0000
		$N^{2/5}$	0.0002	0.0010
		$N^{1/2}$	0.0090	0.0240

Table 3. P–values of the tests of Section 4 applied to **S&P 500 index**. *The data are shown in Figure 3*

Time period	Sample size	Bandwidth	R_N	R_N^0
07/14/2006 – 09/22/2006	$N = 50$	$N^{1/3}$	0.1364	0.2407
		$N^{2/5}$	0.1580	0.2584
		$N^{1/2}$	0.1443	0.1960
07/14/2006 – 12/04/2006	$N = 100$	$N^{1/3}$	0.2242	0.2620
		$N^{2/5}$	0.2669	0.2320
		$N^{1/2}$	0.2915	0.1707
07/14/2006 – 02/16/2007	$N = 150$	$N^{1/3}$	0.0001	0.0003
		$N^{2/5}$	0.0001	0.0013
		$N^{1/2}$	0.0057	0.0076
07/14/2006 – 05/01/2007	$N = 200$	$N^{1/3}$	0.0001	0.0001
		$N^{2/5}$	0.0001	0.0011
		$N^{1/2}$	0.0075	0.0120
07/14/2006 – 07/12/2007	$N = 250$	$N^{1/3}$	0.0013	0.0063
		$N^{2/5}$	0.0090	0.0254
		$N^{1/2}$	0.0519	0.1240
Time period	Sample size	Bandwidth	R_N	R_N^0
08/27/2010 – 11/05/2010	$N = 50$	$N^{1/3}$	0.0186	0.0457
		$N^{2/5}$	0.0324	0.0682
		$N^{1/2}$	0.0607	0.0967
08/27/2010 – 01/19/2011	$N = 100$	$N^{1/3}$	0.0329	0.0465
		$N^{2/5}$	0.0671	0.0916
		$N^{1/2}$	0.1365	0.1207
08/27/2010 – 03/31/2011	$N = 150$	$N^{1/3}$	0.0032	0.0088
		$N^{2/5}$	0.0112	0.0266
		$N^{1/2}$	0.0439	0.0762
08/27/2010 – 06/13/2011	$N = 200$	$N^{1/3}$	0.0000	0.0001
		$N^{2/5}$	0.0001	0.0009
		$N^{1/2}$	0.0026	0.0086
08/27/2010 – 08/23/2011	$N = 250$	$N^{1/3}$	0.0000	0.0000
		$N^{2/5}$	0.0001	0.0004
		$N^{1/2}$	0.0012	0.0055

the financial crisis. These long periods were punctuated by short periods of reversals, so only in 3 out of 20 longer periods a trend model is accepted.

5.3 Properties of the tests

In this section, we elaborate on the findings of Section 5.2 in two ways: 1) we zoom in on specific time periods, those displayed in Figures 2, 3, 4 and 5, to establish a more direct connection between the data and the P–values, 2) we apply to these fewer periods both algorithms of Section 4 and use a selection of bandwiths h. The results are shown in Tables 2, 3, 4 and 5.

We begin by analyzing Table 2 which pertains to the yield curves shown in Figure 2. As for most other periods, the null hypothesis is rejected, except for a few cases corresponding to the bandwidth $h = N^{1/2}$. Simulations reported in [25] show that for artificial data which resemble the yield curves, this bandwidth is too large. It makes the statistic too small and so the tests are too conservative. We also see that while the test based on the Monte Carlo distribution, statistic R_N, and the pivotal test based on R_N^0 generally give different P–values, the differences are small, and generally do not affect significance statements. Turning to the S&P 500 index, for the two periods shown in Figure 3, Table 3 shows rejections, except for the first 100 days in the left panel of Figure 3. In some cases these rejections are weak if $h = N^{1/2}$; both tests again give the same conclusions in almost all cases. The conclusions for the U.S. Dollar index and Oil Futures are qualitatively the same as for the S&P 500 data.

The conclusion is that bandwidths $h = N^{1/3}$ or $h = N^{2/5}$ can be used for sample sizes in the range from 50 to 300. Both algorithms presented in Section 4 give practically the same results. We note that for the data we studied d was small, typically 2 or 3. If d is large, Algorithm 4.2 must be used with caution, as explained in Section 4.

Table 4. P–values of the tests of Section 4 applied to **U.S. dollar index**. The data are shown in Figure 4

Time period	Sample size	Bandwidth	R_N	R_N^0
08/27/2010 – 10/24/2010	$N = 50$	$N^{1/3}$	0.0538	0.0350
		$N^{2/5}$	0.0980	0.0410
		$N^{1/2}$	0.1742	0.0562
08/27/2010 – 12/21/2010	$N = 100$	$N^{1/3}$	0.0001	0.0003
		$N^{2/5}$	0.0006	0.0014
		$N^{1/2}$	0.0079	0.0165
08/27/2010 – 02/18/2011	$N = 150$	$N^{1/3}$	0.0016	0.0012
		$N^{2/5}$	0.0093	0.0059
		$N^{1/2}$	0.0533	0.0243
08/27/2010 – 04/18/2011	$N = 200$	$N^{1/3}$	0.0022	0.0018
		$N^{2/5}$	0.0126	0.0087
		$N^{1/2}$	0.0711	0.0371
08/27/2010 – 06/16/2011	$N = 250$	$N^{1/3}$	0.0045	0.0089
		$N^{2/5}$	0.0219	0.0347
		$N^{1/2}$	0.1041	0.1146
Time period	Sample size	Bandwidth	R_N	R_N^0
01/03/2011 – 03/01/2011	$N = 50$	$N^{1/3}$	0.0039	0.0199
		$N^{2/5}$	0.0129	0.0455
		$N^{1/2}$	0.0350	0.1033
01/03/2011 – 04/30/2011	$N = 100$	$N^{1/3}$	0.0001	0.0006
		$N^{2/5}$	0.0003	0.0031
		$N^{1/2}$	0.0033	0.0178
01/03/2011 – 06/27/2011	$N = 150$	$N^{1/3}$	0.0001	0.0003
		$N^{2/5}$	0.0003	0.0021
		$N^{1/2}$	0.0029	0.0158
01/03/2011 – 08/24/2011	$N = 200$	$N^{1/3}$	0.0000	0.0001
		$N^{2/5}$	0.0001	0.0012
		$N^{1/2}$	0.0019	0.0107
01/03/2011 – 10/22/2011	$N = 250$	$N^{1/3}$	0.0038	0.0184
		$N^{2/5}$	0.0179	0.0609
		$N^{1/2}$	0.0691	0.1746

Table 5. P–values of the tests of Section 4 applied to **Oil Futures**. The data are shown in Figure 5

Time period	Sample size	Bandwidth	R_N	R_N^0
07/14/2006 – 09/10/2006	$N = 50$	$N^{1/3}$	0.0161	0.0185
		$N^{2/5}$	0.0333	0.0340
		$N^{1/2}$	0.0691	0.0781
07/14/2006 – 11/07/2006	$N = 100$	$N^{1/3}$	0.0115	0.0243
		$N^{2/5}$	0.0318	0.0477
		$N^{1/2}$	0.0963	0.1116
07/14/2006 – 01/07/2007	$N = 150$	$N^{1/3}$	0.0000	0.0001
		$N^{2/5}$	0.0004	0.0007
		$N^{1/2}$	0.0055	0.0084
07/14/2006 – 03/06/2007	$N = 200$	$N^{1/3}$	0.0001	0.0004
		$N^{2/5}$	0.0001	0.0037
		$N^{1/2}$	0.0128	0.0342
07/14/2006 – 05/04/2007	$N = 250$	$N^{1/3}$	0.0000	0.0000
		$N^{2/5}$	0.0001	0.0006
		$N^{1/2}$	0.0028	0.0110
Time period	Sample size	Bandwidth	R_N	R_N^0
08/27/2010 – 10/24/2010	$N = 50$	$N^{1/3}$	0.1117	0.1159
		$N^{2/5}$	0.1771	0.1931
		$N^{1/2}$	0.2752	0.2694
08/27/2010 – 12/21/2010	$N = 100$	$N^{1/3}$	0.0882	0.2514
		$N^{2/5}$	0.1567	0.3883
		$N^{1/2}$	0.2343	0.4744
08/27/2010 – 02/19/2011	$N = 150$	$N^{1/3}$	0.0007	0.0079
		$N^{2/5}$	0.0044	0.0267
		$N^{1/2}$	0.0176	0.0620
08/27/2010 – 04/18/2011	$N = 200$	$N^{1/3}$	0.0030	0.0174
		$N^{2/5}$	0.0144	0.0589
		$N^{1/2}$	0.0551	0.1189
08/27/2010 – 06/15/2011	$N = 250$	$N^{1/3}$	0.0164	0.0704
		$N^{2/5}$	0.0530	0.1682
		$N^{1/2}$	0.1707	0.2828

Received 17 March 2015

REFERENCES

[1] Bosq, D. (2000). *Linear Processes in Function Spaces*. Springer, New York. MR1783138

[2] Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press, New Jersey.

[3] Chen, Y. and Niu, L. (2014). Adaptive dynamic Nelson–Siegel term structure model with applications. *Journal of Econometrics* **180** 98–115. MR3188914

[4] de Jong, R. M., Amsler, C. and Schmidt, P. (1997). A robust version of the KPSS test based on indicators. *Journal of Econometrics* **137** 311–333. MR2354947

[5] Dickey, D. A. and Fuller, W. A. (1979). Distributions of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* **74** 427–431. MR0548036

[6] Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with unit root. *Econometrica* **49** 1057–1074. MR0625773

[7] Diebold, F. and Rudebusch, G. (2013). *Yield Curve Modeling and Forecasting: The Dynamic Nelson–Siegel Approach*. Princeton University Press. MR3235466

[8] Giraitis, L., Kokoszka, P. S., Leipus, R. and Teyssière, G. (2003). Rescaled variance and related tests for long memory in volatility and levels. *Journal of Econometrics* **112** 265–294. MR1951145

[9] Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer.

[10] Horváth, L., Kokoszka, P. and Reeder, R. (2013). Estimation of the mean of functional time series and a two sample problem. *Journal of the Royal Statistical Society (B)* **75** 103–122. MR3008273

[11] Horváth, L., Kokoszka, P. and Rice, G. (2014a). Testing stationarity of functional time series. *Journal of Econometrics* **179** 66–82. MR3153649

[12] Horváth, L., Rice, G. and Whipple, S. (2014b). Adaptive bandwith selection in the long run covariance estimator of functional time series. *Computational Statistics and Data Analysis* **00** 000–000. Forthcoming.

[13] Kargin, V. and Onatski, A. (2008). Curve forecasting by functional autoregression. *Journal of Multivariate Analysis* **99** 2508–2526. MR2463404

[14] Kokoszka, P., Miao, H. and Zhang, X. (2015). Functional dynamic factor model for intraday price curves. *Journal of Financial Econometrics* **13** 456–477.

[15] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *Journal of Econometrics* **54** 159–178.

[16] Lee, D. and Schmidt, P. (1996). On the power of the KPSS test of stationarity against fractionally integrated alternatives. *Journal of Econometrics* **73** 285–302. MR1410008

[17] Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica* **59** 1279–1313.

[18] MacNeill, I. B. (1978). Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *The Annals of Statistics* **6** 422–433. MR0474645

[19] Müller, H.-G., Sen, R. and Stadtmüller, U. (2011). Functional data analysis for volatility. *Journal of Econometrics* **165** 233–245. MR2846647

[20] Politis, D. N. (2011). Higher-order accurate, positive semidefinite estimation of large sample covariance and spectral density matrices. *Econometric Theory* **27** 1469–4360. MR2822363

[21] Politis, D. N. and Romano, J. P. (1996). On flat–top spectral density estimators for homogeneous random fields. *Journal of Statistical Planning and Inference* **51** 41–53. MR1394143

[22] Politis, D. N. and Romano, J. P. (1999). Multivariate density estimation with general fat-top kernels of infinite order. *Journal of Multivariate Analysis* **68** 1–25. MR1668848

[23] Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer. MR2168993

[24] Said, S. E. and Dickey, D. A. (1984). Testing for unit roots in autoregressive–moving average models of unknown order. *Biometrika* **71** 599–608. MR0775407

[25] Young, G. (2016). Inference for functional time series with applications to yield curves and intraday cumulative returns. PhD thesis, Colorado State University, Fort Collins, CO, USA.

Piotr Kokoszka
Department of Statistics
Colorado State University
Fort Collins, CO 80522
USA
E-mail address: Piotr.Kokoszka@colostate.edu

Gabriel Young
Department of Statistics
Colorado State University
Fort Collins, CO 80522
USA
E-mail address: gabrieljaggeryoung@gmail.com