The Econometrics Journal

Econometrics Journal (2013), volume **16**, pp. 285–308.
doi: 10.1111/ectj.12006

Predictability of shapes of intraday price curves

PIOTR KOKOSZKA[†] AND MATTHEW REIMHERR[‡]

[†]*Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA.*
E-mail: Piotr.Kokoszka@colostate.edu

[‡]*Department of Statistics, University of Chicago, Chicago, IL 60637, USA.*
E-mail: mreimherr@uchicago.edu

First version received: July 2012; final version accepted: February 2013

Summary We develop a statistical framework, based on functional data analysis, for testing the hypothesis of the predictability of shapes of intraday price curves. We derive test statistics based on signs of the scores of the functional principal components. We establish its asymptotic properties under the null and alternative hypotheses, and demonstrate via simulations that it has excellent finite sample properties. A small empirical study shows that the shapes of the intraday price curves of large US corporations are not predictable.

Keywords: *Functional data analysis, Intraday returns, Predictability.*

1. INTRODUCTION

The problem of the predictability of prices of risky assets and its relationship to the research on market efficiency have occupied a central place in financial economics for over a century. The contributions of Bachelier, Working, Cowles, Granger, Fama, Samuelson and Mandelbrot are discussed in many textbooks. Up to the present day, there is however still a keen interest in specifying empirical or theoretical conditions under which some form of predictability can be claimed. Empirical evidence based on statistical tests points out that predictability defined as some form of departure from the assumption of independent returns holds for many asset classes, but it may be due to structural changes and market frictions rather than inefficiency of the markets. It is relatively broadly believed that once such factors are appropriately accounted for, the *direction* of returns on broadly traded securities cannot be predicted, even though some forms of predictability can exist. The research literature is very extensive, and we do not attempt to discuss the issues involved. Some of them are discussed in Fama (1991), Pesaran and Timmermann (1995), Campbell and Yogo (2006) and Pettenuzzo and Timmermann (2011), to name only a handful of hundreds of contributions. This paper develops a framework similar to the classical martingale theory, but for testing the predictability of the *shapes* of intraday price curves.

1.1. Objectives of the study

Intraday price curves have been the subject of interest in the investment community, see, e.g. Folger and Leibfarth (2007) and Coles and Hawkins (2011), with the focus on technical analysis

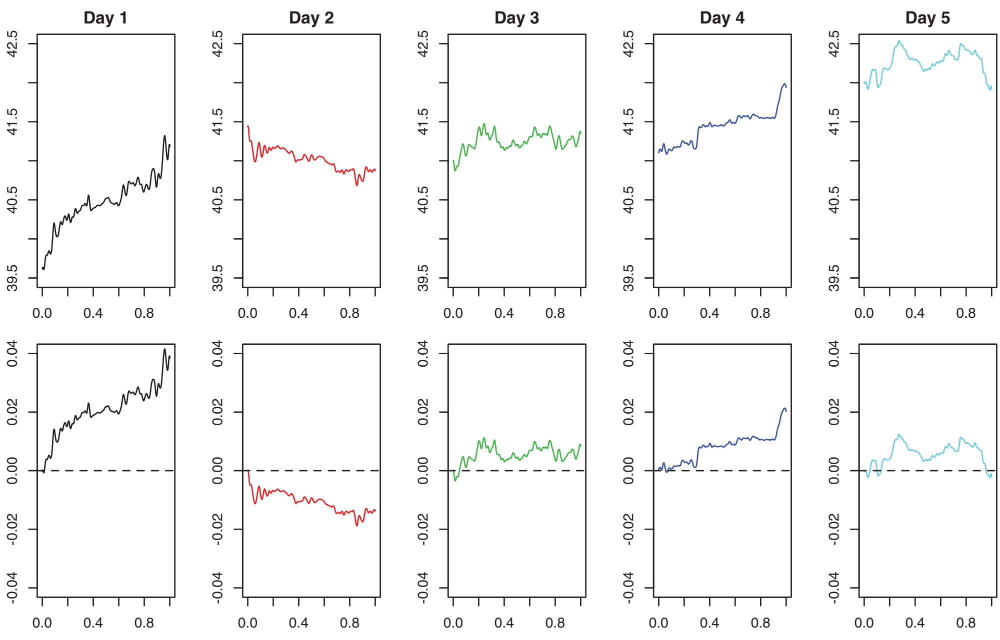

Figure 1. Price curves and cumulative intraday returns.

which uses various charts. Intraday price plots are now readily accessible even to small investors at numerous websites similar to Google Finance, with technical tools to analyse them available for a small fee. Figure 1 shows price curves (top panel) and cumulative intraday returns (bottom panel) on five consecutive days for XOM. The question we attempt to address in this paper is whether there is any non-random or stochastic pattern in the *shapes* of these curves. In econometric research, testing for an existence of a stochastic pattern is termed as testing of predictability. It thus may be said that we seek to investigate the existence of predictability in the shapes of intraday price curves and related curves of cumulative intraday returns (CIDRs), which are defined later.

Tests of predictability of scalar returns are sensitive to long-term stochastic and deterministic trends and regime changes. To mitigate this issue, we shift the curves so that they start at zero on every day. This also emphasises our focus on the shapes rather than absolute prices. Over a period of several years the range of prices within a day may increase. Even though this effect is small for the period we study, we consider CIDRs defined as follows: Suppose $P_n(t)$ is the price of an asset on day n at time t within that day (trading hours). We define the CIDR on day n as the curve

$$r_n(t) = 100(p_n(t) - p_n(0)), \quad p_n(t) = \log P_n(t). \tag{1.1}$$

The usual daily log return is defined as $p_n(1) - p_{n-1}(1)$. The CIDRs are not directly comparable to daily returns. They do not include the overnight price change $p_n(0) - p_{n-1}(1)$. Since $r_n(t) \approx 100(P_n(t) - P_n(0))/P_n(0)$, with $P_n(0)$ being a constant for a given day n, the curves $r_n(t)$ and $P_n(t)$ have similar shapes but different scales and origins. In Section 1.2, we propose a statistical framework for testing the predictability of the shapes of the CIDRs akin the paradigm of the sequence/reversal and runs tests.

1.2. Functional framework for testing the predictability of shapes

Every day, the curves $r_n(\cdot)$ exhibit a certain pattern, typically with some upward or downward momentum or with a broadly parabolic shape. To quantify such behaviour, we use the framework of functional data analysis (FDA), which has grown over the last 20 years into an important and expanding branch of statistics. At its core is the idea that curves should be treated as complete statistical objects, rather than as collections of individual observations. This agrees with our objectives, as we are not interested in price changes from one trade to another, but in the dynamics of the shapes. In the following sections, we describe the tools of FDA that we need, but we cannot offer a broader introduction; Horváth and Kokoszka (2012) and Hörmann and Kokoszka (2012) contain the most relevant background. We also recommend the monograph of Ramsay and Silverman (2005), a more theoretical exposition of Bosq (2000), and a recent collection of Ferraty and Romain (2011).

We work with mildly smoothed CIDRs on stocks of broadly traded US corporations. The smoothing is a result of expanding the data with respect to a finite number of basis functions in the `R` package `fda`; here we use 80 b-splines basis functions. We use this step merely to perform the calculations in a time efficient manner. Figure 2 which displays smoothed CIDRs (left hand panel) and raw CIDRs (right hand panel) shows that the important shape characteristics are preserved. For each stock, we view the sequence $r_n(\cdot)$ as a time series of functions. Each function is viewed as an element of the Hilbert space L^2 of square integrable functions with the usual inner product. (In the theoretical work, we normalise the trading day to be the unit time interval [0, 1], so the inner product is $\int_0^1 f(t)g(t)dt$.) We assume that the $r_n(\cdot)$ form a strictly stationary L^2 valued sequence. In particular, for every stock, the $r_n(\cdot)$ are assumed to have the same distribution in L^2 (the population distribution).

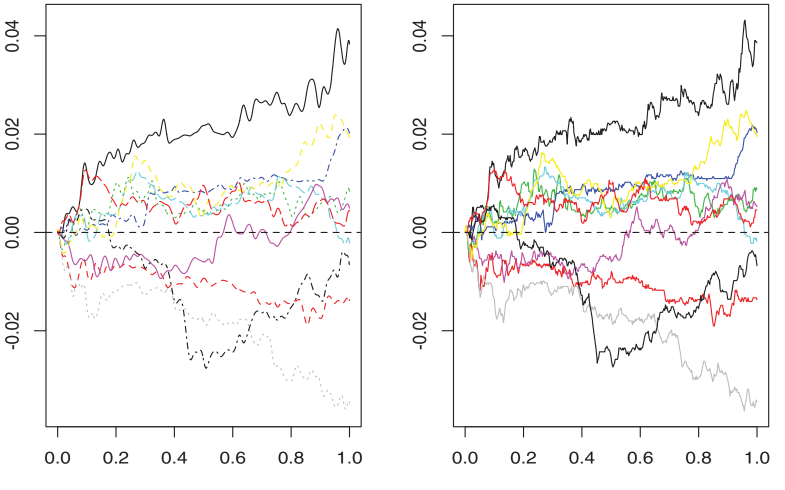

Figure 2. Ten CIDRs on Exon Mobil (XOM).

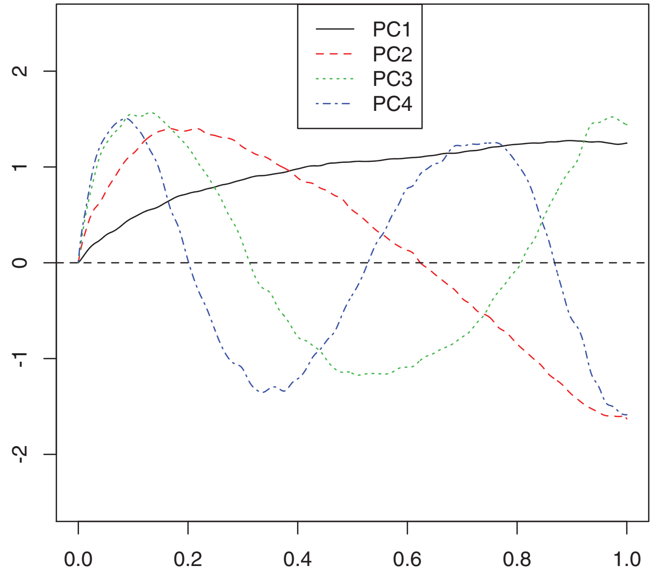

Figure 3. First four EFPCs of XOM.

The shape of a curve $r(\cdot)$ randomly selected from a population representing a given stock is most conveniently summarised by the functional principal components (FPCs) v_k, see, e.g. Chapters 3 and 16 of Horváth and Kokoszka (2012). In our context, it is most useful to view them as the orthonormal functions in the Karhunen–Loéve expansion

$$r(t) = \mu(t) + \sum_{k=1}^{\infty} \xi_k v_k(t), \tag{1.2}$$

where $\mu(t) = Er(t)$ and $\xi_k = \int (r(t) - \mu(t)) v_k(t) dt$. The v_k are population parameters (defined up to a sign) which are estimated by the empirical FPCs (EVPCs) $\hat{v}_k$. The expansion in (1.2) exists for any stochastic process satisfying the necessary moment conditions. Therefore, most well-behaved traditional financial models for $r(t)$ will have such an expansion. For example, for an Itō diffusion process based on Brownian motion, one can use the Itō isometry to see that (1.2) will be satisfied if the volatility is square integrable.

The shape of the curve r_n observed on day n is quantified by the vector of the scores corresponding to the first p EFPCs

$$[\hat{\xi}_{1n}, \hat{\xi}_{2n}, \ldots, \hat{\xi}_{pn}]^T, \quad \hat{\xi}_{kn} = \int \{r_n(t) - \hat{\mu}(t)\} \hat{v}_k(t) dt$$

and by the sample mean function $\hat{\mu}(t)$. For our data, the mean function is not significantly different from zero, see Section 3, so only the scores are needed to describe the evolution of shapes. Figure 3 shows the first four EFPC of Exon Mobil (XOM), the graphs for other stocks are very similar. The sequence of the scores $\hat{\xi}_{1n}$ shows 'how much' component $\hat{v}_1$ is present on day n. The first component $\hat{v}_1$ has a clear interpretation as the principal shape describing an increasing or a decreasing tendency of the CIDRs: if the sign of $\hat{\xi}_{1n}$ is positive, the r_n is mostly increasing. The remaining $\hat{\xi}_{kn}$ describe departures from this main tendency and typically contribute to the overall shape with smaller weights. Their shapes are to a large degree determined by the condition of mutual orthogonality. We focus on testing the predictability of the signs of the scores $\hat{\xi}_{kn}$. For example, if a test finds no predictability in the signs of the $\hat{\xi}_{1n}$, then we conclude that the past pattern of broadly increasing or decreasing shapes does not allow us to predict the increasing or decreasing tendency on the next trading day.

1.3. Current approaches and difficulties

There is a fundamental difference between scalar (daily through annual) returns and the scores $\hat{\xi}_{kn}$: the scalar returns are directly observable but the $\hat{\xi}_{kn}$ are only estimators of unobservable quantities. Even if the $\xi_{kn} = \int r_n(t)v_k(t)dt$ form an i.i.d. sequence, the $\hat{\xi}_{kn} = \int r_n(t)\hat{v}_k(t)dt$ will not, because the $\hat{v}_k$ depend on the whole sample $r_1, r_2, \ldots, r_N$, and so are obviously dependent. This is a very delicate point when it comes to testing for independence. Fortunately, the transition from quantities defined in terms of the v_k to those defined in terms of $\hat{v}_k$ is facilitated by the bound

$$E\int (\hat{v}_k(t) - v_k(t))^2\, dt = O(N^{-1}), \tag{1.3}$$

which was established for i.i.d. curves by Dauxois et al. (1982), and then extended and refined by Bosq (2000) and Hörmann and Kokoszka (2010) (cf. Chapter 2 of Horváth and Kokoszka, 2012). This bound is useful for establishing asymptotic properties of statistics which can be expressed as L^2 continuous functionals of the observations. It is an important theoretical tool in the derivation of a portmanteau test of independence developed by Gabrys and Kokoszka (2007), which we discuss below. In this paper, we work with the signs of the $\hat{\xi}_{kn}$. A small change caused by replacing the $\hat{v}_k$ by v_k, can cause a change in the sign from 1 to –1 because $s_r(v) = \operatorname{sign}\int r(t)v(t)dt$ is not a continuous function of $v(\cdot)$. We therefore cannot use bound (1.3), and so we develop a different asymptotic analysis.

To provide a more specific illustration of this point, consider the Cowles–Jones statistic

$$\mathrm{CJ}_N^{(k)} = \frac{S^{(k)}}{N - S^{(k)}},$$

where $S^{(k)}$ is the count of sequences ($\xi_{kn} > 0, \xi_{k,n+1} > 0$ or $\xi_{kn} < 0, \xi_{k,n+1} < 0$) for the variables $\xi_{kn},\ 1 \le n \le N$. The statistic $\mathrm{CJ}_N^{(k)}$ is based on unobservable scores, which are independent if the r_n are so. In that case, its asymptotic distribution is known, see Chapter 2 of Campbell et al. (1997). However to implement the test, we must use the $\hat{\xi}_{kn}$, which are computed using the $\hat{v}_k$. Will this replacement have an effect on the asymptotic distribution of $\mathrm{CJ}_N^{(k)}$? Can we construct a test statistic that takes the projections on the first p EFPCs $\hat{v}_k,\ k = 1, 2, \ldots, p$ simultaneously into account? Before we pursue these issues in the following sections, we consider the application of the portmanteau test of Gabrys and Kokoszka (2007) which can be viewed as an extension to the functional setting of the usual portmanteau tests of the i.i.d. assumption based on correlations (Box and Pierce, 1970, Ljung and Box, 1978). The null hypothesis for this test is that the functions are i.i.d. in L^2, the alternative is that they are correlated in a suitably defined way. To apply this test, the number p of the EFPCs $\hat{v}_k$ to be used must be selected. The P-values are given in Table 1. (The stocks and time periods are described in Section 3.) If only the first EFPC $\hat{v}_1$ is used, we generally cannot reject the i.i.d. assumption, but as p increases, this conclusion remains

Table 1. Portmanteau test P-values as a function of the number of FPCs.

p	1	2	3	4
BOA	0.167	0.000	0.000	0.000
CITI	0.082	0.000	0.000	0.000
COCA	0.021	0.013	0.008	0.000
CVX	0.488	0.093	0.025	0.022
DIS	0.099	0.206	0.140	0.003
IBM	0.511	0.000	0.000	0.000
MCD	0.164	0.110	0.368	0.055
MSFT	0.227	0.145	0.005	0.000
WMT	0.032	0.008	0.000	0.000
XOM	0.054	0.091	0.367	0.000

Notes: Based on lag one correlations.

valid only for MCD. The conclusion that the main shape of increase or decrease throughout a trading day, cf. Figure 3, is not predictable, but that the less important components introduce predictability does not seem satisfactory and motivates our quest for a different approach and a more profound understanding of the behaviour of the shapes of intraday prices.

2. DESCRIPTION OF THE TESTS AND ASYMPTOTIC RESULTS

We begin by formulating the assumptions. They are stated in terms of an arbitrary sequence $\{X_n\}$ of functions. In Section 3, we apply our results to the X_n equal to the CIDRs r_n constructed in Section 1. Throughout the paper, $\langle \cdot, \cdot \rangle$ denotes the inner product in the Hilbert space $L^2 = L^2([0, 1])$.

ASSUMPTION 2.1. *The sequence $\{X_n\}$ of random functions in L^2 is strictly stationary, ergodic, with $EX_n = 0$ and $E\|X_1\|^2 < \infty$. Let $\{(\lambda_j, v_j)\}$ be the eigenvalue/eigenfunction pairs of the covariance kernel$c(s, t) = E[X_1(s)X_1(t)]$ ordered by the magnitude of λ_j. Assume that $\lambda_1 > \ldots > \lambda_p > \lambda_{p+1}$ and that$P(\langle X_1, v_j \rangle = 0) = 0$ for $j = 1, \ldots, p$.*

Assumption 2.1 specifies the structure of the process under both the null and alternative hypotheses. Under the alternative, we might expect dependence between curves on consecutive days. The assumption of ergodicity ensures that time averages converge to appropriate expectations for all functionals $g(X_n)$, see, e.g. Chapter 9 of Kallenberg (1997), in particular his Lemma 9.5 and Theorem 9.6. A very general class of ergodic sequences is obtained by assuming that $X_n = f(\varepsilon_n, \varepsilon_{n-1}, \ldots)$, where the $\{\varepsilon_n\}$ are i.i.d. random elements taking values in some measurable space $\mathcal{S}$, and f is a measurable function $f : \mathcal{S}^\infty \to L^2$. Sequences of this type are called Bernoulli shifts. They have been used extensively in recent theoretical work, as all stationary time series models in practical use, in particular time series taking values in function spaces, admit representation as Bernoulli shifts, see Shao and Wu (2007), Aue et al. (2009) and Hörmann and Kokoszka (2010), among many other contributions. The assumption that the first p eigenvalues are different and positive implies that the first p eigenspaces are one-dimensional. Finally, the assumption that $\langle X_1, v_j \rangle$ cannot be zero with positive probability is needed to ensure that under the null hypothesis formulated below the conditional probability of a positive (negative) score is 1/2.

As eigenfunctions, the v_i are not uniquely defined, only the one-dimensional eigenspaces are uniquely defined. In the following, we fix any orthonormal vectors v_i belonging to these eigenspaces, and so we treat the v_i as population parameters. We formalise the above construction in Assumption 2.2. Our procedures will be invariant with respect to the signs of the FPCs and their estimators.

ASSUMPTION 2.2. *The FPCs $v_1, \ldots, v_p$ form a fixed orthonormal system.*

Assumption 2.1 specifies the stationarity and moment conditions valid under both the null and alternative hypotheses. Our last assumption specifies the null hypothesis.

ASSUMPTION 2.3. *(Null Hypothesis) The sequence $\{X_n\}$ is conditionally symmetric in the sense that*

$$\mathcal{L}(X_n | X_{n-1}, X_{n-2}, \ldots) = \mathcal{L}(-X_n | X_{n-1}, X_{n-2}, \ldots) \quad a.s.$$

Assumption 2.3 holds if the curves X_n are independent and have a symmetric distribution in L^2. It is however a weaker, martingale type, assumption satisfied by functional conditionally heteroscedastic models proposed in Section 3. It implies that the signs of the scores cannot be predicted based on the past curves.

We are now ready to state the theorems which lead to our tests. All proofs are presented in the Appendix.

2.1. *Tests based on a single component*

We first construct an analogue of the Cowles–Jones statistics for the sequence $\hat{\xi}_{kn} = \langle X_n, \hat{v}_k \rangle$. Define the triangular array

$$I_{N,n}^{(k)} = \text{sign}\{\langle X_n, \hat{v}_k \rangle\}.$$

Note that $I_{N,n}^{(k)} I_{N,n+1}^{(k)}$ is positive for a sequence and negative for a reversal.

THEOREM 2.1. *If Assumptions 2.1, 2.2 and 2.3 hold, then for each $1 \leq k \leq p$,*

$$\Lambda^{(k)} := (N-1)^{-1/2} \sum_{i=1}^{N-1} I_{N,n}^{(k)} I_{N,n+1}^{(k)} \xrightarrow{d} N(0, 1).$$

Theorem 2.1 implies that the Cowles–Jones statistic computed using the estimate $\hat{v}_1$ has the same asymptotic distribution as its analogue computed with v_1. To see this, note that

$$\text{number of sequences} = \sum_{i=1}^{N-1} \big(I_{N,n}^{(1)} I_{N,n+1}^{(1)} + 1\big)/2 = (N-1)/2 + (1/2) \sum_{n=1}^{N-1} I_{N,n}^{(1)} I_{N,n+1}^{(1)}.$$

Therefore,

$$(N-1)^{-1/2} \sum_{n=1}^{N-1} I_{N,n}^{(1)} I_{N,n+1}^{(1)} = (N-1)^{-1/2} (2N_s - (N-1)) \xrightarrow{d} N(0, 1),$$

which implies that the number of sequences is asymptotically normal with mean $(N-1)/2$ and variance $1/(4(N-1))$, which agrees with formula (2.2.8) of Campbell et al. (1997) (with $\pi = 1/2$).

We now describe how Theorem 2.1 can be used to determine the predictability of the contribution of the kth FPC. The hypotheses we are testing are, for each j,

$$H_{0,j}: \ E\big[I_1^{(j)} I_2^{(j)}\big] = 0 \quad \text{against} \quad H_{A,j}^- : E\big[I_1^{(j)} I_2^{(j)}\big] < 0 \quad \text{or} \quad H_{A,j}^+ : E\big[I_1^{(j)} I_2^{(j)}\big] > 0,$$

where $I_i^{(j)} = \mathrm{sign}(\xi_{ji})$ (the absence of the subscript N indicates population signs). We break apart the alternative into the case where reversals are more common ($H_{A,j}^-$), and where sequences are more common ($H_{A,j}^+$). For tests with significance level $\alpha \in (0, 1)$, the decision rules are the following:

$$\begin{array}{ll} \text{Accept } H_{0,j} & \text{if } |\Lambda^{(j)}| < z_{1-\alpha/2}, \\ \text{Reject } H_{0,j} \text{ in favour of } H_{A,j}^- & \text{if } \Lambda^{(j)} < -z_{1-\alpha/2}, \\ \text{Reject } H_{0,j} \text{ in favour of } H_{A,j}^+ & \text{if } \Lambda^{(j)} > z_{1-\alpha/2}, \end{array}$$

where $z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution.

The following proposition provides interpretation of the rejection in terms of the conditional probabilities of the scores. It is valid under the assumption that the curves X_n have a symmetric distribution in L^2. This is a much weaker assumption than Assumption 2.3, in particular it says nothing about temporal dependence. The conditional probabilities in Proposition 2.1 can be interpreted as Markov chain transition probabilities similar to those studied for scalar returns in Chapter 2 of Campbell et al. (1997).

PROPOSITION 2.1. *If Assumptions 2.1, 2.2 hold, and the curves X_n have a symmetric distribution in L^2, then we have the following: (a) under $H_{A,j}^+$, $P(\xi_{j,n} > 0 | \xi_{j,n-1} > 0) > 1/2$ and $P(\xi_{j,n} < 0 | \xi_{j,n-1} < 0) > 1/2$, (b) under $H_{A,j}^-$, $P(\xi_{j,n} > 0 | \xi_{j,n-1} > 0) < 1/2$ and $P(\xi_{j,n} < 0 | \xi_{j,n-1} < 0) < 1/2$.*

The consistency of the test based on Theorem 2.1 is established in the following Theorem.

THEOREM 2.2. *Under the assumptions of Proposition 2.1, under $H_{A,j}^-$, $\Lambda^{(j)} \xrightarrow{P} -\infty$, and under $H_{A,j}^+$, $\Lambda^{(j)} \xrightarrow{P} \infty$.*

We conclude this section with a brief discussion of the runs test which has received a great deal of attention in statistical literature, but which reduces to the sequence/reversal test in our setting in which we focus only on signs. Recall that the number of runs in a sequence such as $\{I_{N,1}^{(1)}, \ldots, I_{N,N}^{(1)}\}$ is given by the number of subsequence of consecutive negative ones or positive ones. So, for example, the sequence $(1, 1, 1, 1)$ has only one run while the sequence $(1, -1, 1, 1)$ has three runs. Define

$$N_{runs}^j = \text{the number of runs in} \big\{I_{N,1}^{(j)}, \ldots, I_{N,N}^{(j)}\big\}.$$

Note that for any sequence we have the following relationship:

$$\text{Number of runs} = 1 + \text{number of reversals.}$$

This is trivially true since a new run starts if and only if there is a reversal of the signs. A pair $(I_{N,n}^{(j)}, I_{N,n+1}^{(j)})$ constitutes a reversal if and only if $I_{N,n}^{(j)} I_{N,n+1}^{(j)} = -1$, since they must therefore have different signs. So we have that

$$\frac{1 - I_{N,n}^{(j)} I_{N,n+1}^{(j)}}{2} = \begin{cases} 1 & \text{for a reversal} \\ 0 & \text{for a sequence} \end{cases} .$$

So the number of reversals can be expressed as

$$\text{Number of reversals} = \sum_{i=1}^{N-1} \frac{1 - I_{N,n}^{(j)} I_{N,n+1}^{(j)}}{2},$$

and we have that

$$N_{runs} = 1 + \frac{N-1}{2} - \sum_{i=1}^{N-1} \frac{I_{N,n}^{(j)} I_{N,n+1}^{(j)}}{2}.$$

We can re-express the above as

$$\frac{N_{runs}^{j} - (N+1)/2}{\sqrt{(N-1)/4}} = -\frac{\sum_{i=1}^{N-1} I_{N,n}^{(j)} I_{N,n+1}^{(j)}}{\sqrt{N-1}}.$$

Theorem 2.1 thus implies the following result.

THEOREM 2.3. *If Assumptions 2.1, 2.2 and 2.3 hold, then*

$$\frac{N_{runs}^{j} - (N+1)/2}{\sqrt{(N-1)/4}} \xrightarrow{d} N(0, 1).$$

Note that this again agrees with the univariate results given in Section 2.2 of Campbell et al. (1997) (when taking $\pi = 1/2$, and summing the two types of runs). Due to the equivalence of the tests based on Theorems 2.1 and 2.3, we consider only the the application of Theorem 2.1.

2.2. *Tests based on the first pcomponents*

We now project onto the first p EFPCs to construct a test statistic to determine the predictability of shapes based on all these components. To do it, we first describe the asymptotic joint distribution of the statistics $\Lambda^{(k)}$ defined in Theorem 2.1.

THEOREM 2.4. *Suppose Assumptions 2.1, 2.2 and 2.3 hold. Then, there exists a non-negative definite symmetric matrix* $\mathbf{\Sigma}$ *such that*

$$(\Lambda^{(1)}, \ldots, \Lambda^{(p)})^T \xrightarrow{d} N(\mathbf{0}, \mathbf{\Sigma}),$$

where $N(\mathbf{0}, \mathbf{\Sigma})$ *is a mean zero* p*-dimensional normal random vector with covariance matrix* $\mathbf{\Sigma}$.

For completeness, we state the following result from the runs test. It follows from the arguments leading to Theorem 2.3 and from Theorem 2.4.

THEOREM 2.5. *Under the assumptions of Theorem 2.4,*

$$\left(\frac{N_{runs}^{(1)} - (N+1)/2}{\sqrt{(N-1)/4}}, \ldots, \frac{N_{runs}^{(p)} - (N+1)/2}{\sqrt{(N-1)/4}}\right)^T \xrightarrow{d} N(\mathbf{0}, \mathbf{\Sigma}),$$

where $\mathbf{\Sigma}$ *is the same matrix as in Theorem 2.4.*

The diagonal entries of $\mathbf{\Sigma}$ are all 1. One cannot guarantee that the off diagonal entries are zero, even if X is assumed symmetric. However, under a stronger assumption that the $\{\langle X_1, v_j\rangle\}$ are jointly symmetric, i.e. $(\langle X_1, v_1\rangle, \ldots, \langle X_1, v_p\rangle) \underset{=}{\mathcal{L}} (c_1\langle X_1, v_1\rangle, \ldots, c_p\langle X_1, v_p\rangle)$ for any deterministic sequence $\{c_j\}$ taking values 1 or -1, the off-diagonal entries of $\mathbf{\Sigma}$ are zero. This follows from the symmetry argument used at the beginning of the proof of Theorem 2.1. Regardless, that the off diagonals are zero is not essential.

Using Theorem 2.4 one can construct *overall* tests for predictability. The one that we will use in Section 3 is based on the convergence

$$\Lambda_p = \sum_{j=1}^{p} \hat{\lambda}_j \Lambda^{(j)} \xrightarrow{d} N(0, \lambda^T \mathbf{\Sigma} \lambda), \tag{2.1}$$

where $\lambda = (\lambda_1, \ldots, \lambda_p)$. The matrix $\mathbf{\Sigma}$ can be estimated using the sample covariance matrix of the $\{I_{n,N}^{(k)}\}$. The goal of the test statistic Λ_p is to pool the information from the coordinate wise statistics to determine if sequences or reversals are more likely in the process overall. We weight the components by λ_j since those correspond to the amount of variability explained by the corresponding component. An analogous pooled statistic can be constructed for the runs test, but due to the equivalence explained in the previous section, it is not applied in the following.

We close this section by noting that, under additional assumptions on the rate of decay of the population eigenvalues λ_j, we could modify the convergence in (2.1) by allowing p to tend to infinity with N at a controlled rate. Such an extension would require a fair deal of additional mathematical arguments, and is not pursued here. An example of such an approach is given in Cardot et al. (2003).

The approach developed here uses perhaps the simplest representation of curves. Fourier and wavelet expansions could potentially be used as well, and more sophisticated tests based on the work of Fan (1996) and Fan and Lin (1998) could be explored.

3. EMPIRICAL EVIDENCE

3.1. Application to stock CIDRs

We applied our tests to 10 stock CIDRs taken from December 1999 to April 2007, adjusted for splits but not for dividends. This gives a total of $N = 1843$ days. The original data for each stock is available as one minute averages (390 observations per day). These averages were mildly smoothed as shown in Figure 2, merely to create functional objects in the `R` package `fda`. The application of a test similar to the test for the equality of mean functions described in Section 5.1 of Horváth and Kokoszka (2012) shows that for all stocks with the exception of BOA the mean function can be assumed to be zero. Even for BOA the mean function is several orders of

Table 2. P-values for the Λ_p test based on the statistic.

p	1	2	3	4
BOA	0.255	0.243	0.232	0.240
CITI	0.194	0.213	0.229	0.223
COCA	*0.053	*0.062	*0.063	*0.060
CVX	0.106	0.093	0.086	0.085
DIS	0.181	0.226	0.204	0.193
IBM	0.590	0.656	0.653	0.642
MCD	0.239	0.218	0.212	0.209
MSFT	0.952	0.841	0.887	0.896
WMT	0.984	0.927	0.916	0.897
XOM	0.016	0.015	0.015	0.015

Notes: * denotes the P-values under 10% favouring sequences.

Table 3. P-values of the $\Lambda^{(j)}$ tests.

j	1	2	3	4
BOA	0.255	0.646	0.309	0.194
CITI	0.194	0.618	0.223	0.290
COCA	*0.053	0.413	0.857	0.168
CVX	0.106	0.419	0.196	0.628
DIS	0.181	0.098	*0.040	*0.033
IBM	0.590	0.290	0.857	0.369
MCD	0.239	0.391	0.536	0.536
MSFT	0.952	0.069	*0.040	0.413
WMT	0.984	0.194	0.646	0.115
XOM	0.016	0.592	0.834	0.625

Notes: * denotes the P-values under 10% favouring sequences.

magnitude smaller than the CIDRs so assuming that it is zero is reasonable. We note that the assumption of the zero mean is important only for the interpretation of the results: if the mean function $\mu(t)$ is not zero, our conclusions pertain to the predictability of the shapes of centred CIDRs $r_n(t) - \mu(t)$.

We applied the test based on (2.1) using the first four principal components, which explain approximately 96% of the variability in each stock. The first principal component explains about 85% of the variance. The increments in the amount of variability explained are very similar for all stocks and equal to about 7%, 2.5% and 1.3%. The P-values are summarised in Table 2. In contrast to Table 1, for every stock the conclusion does not depend on p. For eight stocks we find no evidence for the predictability of the shapes. For COCA we find a marginal, not significant at the usual 5% level, evidence that sequences of similar shapes are more likely to occur. For XOM we find evidence that reversals are significantly more likely than allowed by the unpredictable pattern of the signs of scores. These conclusions are supported by the P-values in Table 3. These are the P-values for the tests based on the statistics $\Lambda^{(j)}$ for $j = 1, 2, 3, 4$. These statistics look at the predictability of the signs of the individual v_j separately. As intuitively expected, the marginal evidence for sequences in COCA is found only in the first FPC; the other components are unpredictable. Analogous observation applies to XOM, so for this stock a decreasing pattern of the intrayday price curve is more likely to be followed by an increasing pattern than by another decreasing pattern. There are some marginal rejections in Table 3 (DIS and MSFT, $j = 3$ and 4). These could be spurious due to multiple testing, Table 3 presents the P-values of forty tests, but even if they are not, these P-values are not highly significant, and they deal with principal components that together account for less than 4% of variance.

Our very limited empirical study points out that the shapes of the intraday price curves are not predictable. There might be exceptions, but we do not undertake in this paper a more detailed analysis of why predictability might exist in some stocks. Our goal has been to develop a statistical tool for testing this hypothesis, and we hope that our contribution will be received with some interest by applied researchers.

Our small empirical study has been intended to highlight the differences between the portmanteau statistic of of Gabrys and Kokoszka (2007) and the new statistic Λ_p which is not sensitive to the choice of p. This is a very important issue in FDA. Many methods of optimally choosing p have been proposed, including cross-validation and information criteria, see, e.g. Horváth and Kokoszka (2012). The most common criterion is however to choose p which explains about 90% of the variance. For our data, all values of p would thus be admissible because they all explain 90% $\pm$ 5% of the variance. Yet the portmanteau test is very sensitive to the choice of p. In Section 3.3, we investigate this issue further by means of a simulation study. First, we propose functional time series models which are suitable for such a study.

3.2. Models for the CIDRs

While high frequency intraday price data have been the focus of extensive research, the interest has been in their micro-structure. The motivation has been to exploit the properties of this microstructure and to model it in order to devise trading strategies on scales of several minutes. Our objectives are different, and we treat the intraday price curves as complete statistical objects. We are not aware of any empirical research attempting to fit functional time series models to intraday price curves. We therefore discuss several possibilities and demonstrate in Section 3.3 that some of them come very close to emulating the dynamics of the CIDRs in respects that relevant to our purpose.

The simplest data generating process is formed by i.i.d. curves. However, we can also allow more sophisticated processes that are suitable for financial data. In our simulations we will focus on processes that fall under the second example below and which seem to mimic the patterns in the data quite nicely.

1. *Independent and identically distributed.* If $X_1, \ldots, X_n$ are i.i.d., then they are ergodic. Furthermore, if they are symmetric then the independence trivially implies that they are conditionally symmetric.
2. *ARCH-type processes with univariate volatilities*. Suppose that the X_i are given by

$$X_i = \sigma_i W_i,$$

where $\{W_i\}$ are i.i.d. and symmetric L^2 processes. We assume the volatility process, σ_i, is univariate and given by

$$\sigma_i = f(X_{i-1}, X_{i-2}, \ldots, \delta_i, \delta_{i-1} \ldots),$$

where δ_i are i.i.d. and univariate. To be clear, $\{W_i\}$ and $\{\delta_i\}$ are assumed to be totally independent. Clearly the $\{X_i\}$ are symmetric and ergodic. Using the symmetry of W_i and ε_i, they are also conditionally symmetric. If σ_i, W_i and ε_i all have finite second moments, then Assumption 2.1 is also satisfied.

3. *ARCH-type processes with operator valued volatilities.* A more detailed theoretical development of a functional version of the ARCH model can be found in Hörmann et al. (2013). Suppose that

$$X_i = \boldsymbol{\sigma}_i(W_i),$$

where again $\{W_i\}$ are i.i.d. and symmetric L^2 processes. However, now assume that $\boldsymbol{\sigma}_i$ is a random linear bounded operator acting on L^2. Once again assume that

$$\boldsymbol{\sigma}_i = \mathbf{f}(X_{i-1}, X_{i-2} \ldots, Z_i, Z_{i-1}, \ldots),$$

where $\{Z_i\}$ is now an i.i.d. L^2 valued sequence. Again, since $\boldsymbol{\sigma}_i$ is linear and independent of W_i, X_i is ergodic, symmetric, and conditionally symmetric. If $\boldsymbol{\sigma}_i$ and W_i have finite second moments, then again Assumption 2.1 is satisfied. (We emphasise that $\boldsymbol{\sigma}_i$ is a linear operator in L^2, the function $\mathbf{f}$ need not be linear.)

3.3. A simulation study

To better understand the differing behaviour of our procedure and the portmanteau test in the stock CIDRs, we constructed a relevant simulation study. We generated heteroscedastic functional time series designed to emulate the volatility patterns found in stock data. To that end, we simulated from what we call a *functional EARCH* model, which is a special case of Model 2 discussed in Section 3.2. Define

$$X_i(t) = \sigma_i W_i(t),$$

where W_i is a standard Brownian motion and σ_i is a univariate process defined recursively as

$$\log(\sigma_i) = a \log(\sigma_{i-1}) + 0.5\delta_i,$$

and the δ_i are i.i.d. standard normal random variables. Since the δ_i are normal, σ_i has a stationary solution (when $|a| < 1$) with a finite second moment. Thus X_i satisfies Assumptions 2.1 and 2.3. The X_i are thus not i.i.d., but they are not predictable either. In particular, the cross-covariance operators between the observations are zero. Therefore, one would think that neither our procedure nor the portmanteau test would detect any sort of dependence, which for these simulated data would be a desirable outcome.

To determine the calibration of the tests, we simulated data for sample sizes $N = 500$ and 1000, with 1000 replications of each scenario. The curves are evenly sampled at a temporal resolution of 0.001 (on the $[0, 1]$ interval) and we use a burn-in sample of 100 curves. We simulated under $a = 0$ (no heteroscedasticity), $a = 0.25$ and $a = 0.5$. The results are summarised in Table 4. We report the P-values for $p = 1, 2, 3, 4$.

We see that the heteroscedasticity does not have a notable effect on our signed procedure while the rejection rates for the Portmanteau tests progressively increase with the degree of heteroscedasticity. Furthermore, the effect becomes magnified as the number p of the FPCs used in the test increases. This helps explain the small portmanteau P-values in the CIDRs (cf. Table 1). We computed several additional simulation scenarios (not given here) and the

Table 4. Rejection levels for the Λ_p and portmanteau statistics.

$N = 500, a = 0$					$N = 1000, a = 0$				
		1	2	3	4			1	2	3	4
10%	PM	0.113	0.101	0.098	0.094	10%	PM	0.099	0.108	0.102	0.105
	Λ_p	0.111	0.106	0.107	0.108		Λ_p	0.102	0.098	0.102	0.100
5%	PM	0.055	0.049	0.043	0.048	5%	PM	0.041	0.057	0.053	0.052
	Λ_p	0.057	0.061	0.059	0.06		Λ_p	0.044	0.043	0.046	0.047
1%	PM	0.010	0.016	0.012	0.016	1%	PM	0.010	0.011	0.018	0.019
	Λ_p	0.012	0.014	0.013	0.013		Λ_p	0.015	0.013	0.012	0.014
$N = 500, a = 0.25$						$N = 1000, a = 0.25$					
		1	2	3	4			1	2	3	4
10%	PM	0.163	0.189	0.237	0.318	10%	PM	0.146	0.203	0.26	0.294
	Λ_p	0.102	0.095	0.096	0.094		Λ_p	0.113	0.114	0.112	0.111
5%	PM	0.103	0.120	0.156	0.208	5%	PM	0.085	0.125	0.153	0.209
	Λ_p	0.051	0.054	0.057	0.057		Λ_p	0.045	0.050	0.051	0.052
1%	PM	0.031	0.041	0.053	0.068	1%	PM	0.022	0.039	0.054	0.082
	Λ_p	0.012	0.011	0.012	0.012		Λ_p	0.011	0.011	0.011	0.012
$N = 500, a = 0.5$						$N = 1000, a = 0.5$					
		1	2	3	4			1	2	3	4
10%	PM	0.217	0.36	0.507	0.652	10%	PM	0.235	0.384	0.515	0.667
	Λ_p	0.113	0.103	0.105	0.107		Λ_p	0.102	0.103	0.098	0.103
5%	PM	0.139	0.256	0.393	0.537	5%	PM	0.165	0.273	0.417	0.569
	Λ_p	0.052	0.053	0.051	0.055		Λ_p	0.043	0.046	0.045	0.045
1%	PM	0.054	0.121	0.219	0.354	1%	PM	0.064	0.130	0.227	0.385
	Λ_p	0.008	0.009	0.007	0.006		Λ_p	0.006	0.007	0.007	0.007

apparent pattern in the portmanteau rejection rates is that while the sample size has little effect, they increase steadily with the degree of heteroscedasticity. In particular, for $a \geq 0.7$, the rejection rates for the portmanteau test were over 90% at a 10% significance level. In our simulations, increasing a by 0.25 resulted, approximately, in a 25–30 point increase in the rejection percentage. This indicates that the portmanteau test severely overrejects in the presence of heteroscedasticity. This is not entirely unexpected as this test was derived to test the null hypothesis of i.i.d. curves, but, on the other hand, the test statistics checks only the temporal cross-correlations, which are zero for functional EARCH data. As the analysis that follows reveals, heteroscedasticity generates heavy tailed score processes. The connection between heteroscedasticity and heavy tails has been extensively studied in ARCH-type models for scalar returns, see e.g. Embrechts et al. (1997) and Mikosch and Stărică (2000). No such theoretical results are available for functional models, but we found using simulations that similar properties hold for the score processes. Our procedure is based on signs, and so it is not thrown out by large observations. In addition, our procedure employs an appropriate weighting scheme that weighs down less important FPCs v_k.

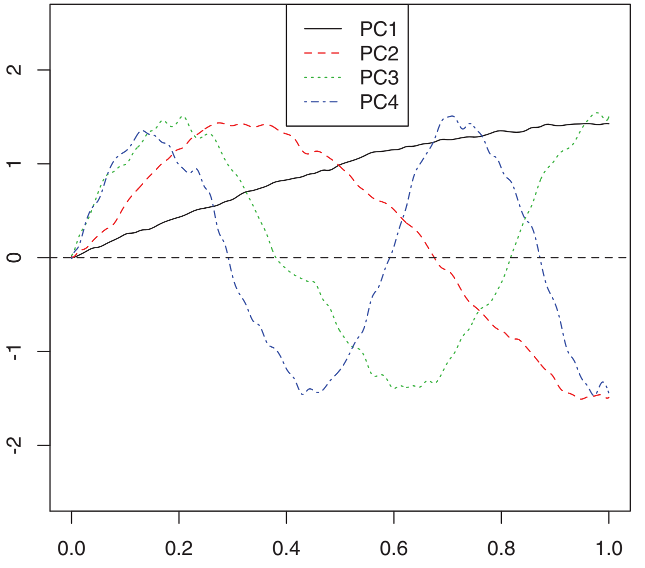

Figure 4. First four EFPCs of a simulated EARCH repetition.

Before presenting a power study, we thus first investigate if the simulated data resembles the stock CIDRs. Figure 4 shows the first four EFPCs $\hat{v}_k$ for one replicate of the functional EARCH process of length $N = 1000$ with $a = 0.5$. The plots are very similar to the ones given in Figure 3. Figure 5 shows the time series of the PC scores $\hat{\xi}_{1i}$ and $\hat{\xi}_{2i}$, $1 \leq i \leq N$, for real and simulated data; the top panels plot the first (left-hand panel) and second (right-hand panel) PC scores for WMT, the middle panels plot the first two PC scores for the simulated process in Section 3.3 with $a = 0.5$, and the bottom panels plot the first two PC scores for the simulated process in Section 3.3 with $a = 0.5$, and a 50% drop in variance after the 1250th observation. The appearance of these series for real data is very similar to the daily returns series: there are occasional 'extremal' scores and long term volatility shifts. By experimenting with various values of a, we think that the functional EARCH process with $a = 0.5$ has score processes that exhibit similar tail behaviour to that of real data. This model does not however exhibit volatility shifts. These are easy to introduce by multiplying σ_i by a constant for a specified range of i. We ran a simulation with a volatility shift in the middle of the sample and found that our test is not effected by it, but the size distortion of the portmanteau test increases by a further 10–15%. This can again be explained by the fact that our test is based on signs. While our assumptions postulate stationarity of the function valued process, they could in fact be weakened, but we do not pursue such technical modifications. We merely note that the stationarity in Assumption 2.1 can be dropped and replaced with an assumption that the signs of $\langle X_i, \hat{v}_j \rangle$ are strongly stationary, and Theorem 2.1 will still hold. The symmetry assumptions on the X_i will guarantee that the martingale arguments will still hold.

We conclude this section by analysing the power of our procedure in the presence of varying levels of heteroscedasticity. The curves under the alternative follow the functional autoregressive model with EARCH errors:

$$X_i(t) = \Phi(X_{i-1})(t) + \sigma_i W_i(t),$$

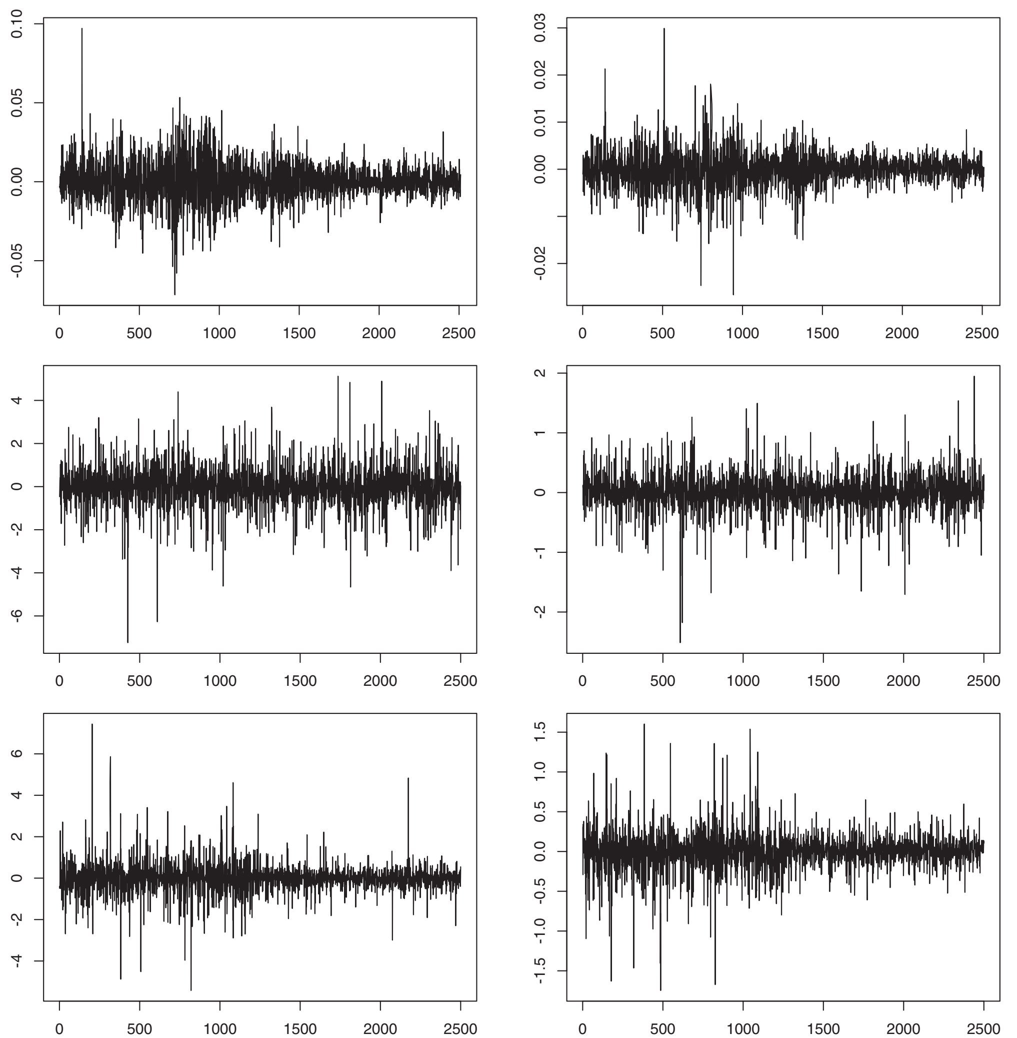

Figure 5. PC scores.

where W_i and σ_i are defined as before. The operator Φ is an integral Hilbert–Schmidt operator with the kernel

$$\phi(t, s) = c\frac{\exp(-(t-s)^2)}{0.8739}.$$

This means that for any function $x \in L^2$, $\Phi(x) = \int \phi(t, s)x(s)ds$. Note the denominator above is there so that $\|\Phi\| \approx c$. We use differing values of c to control for the predictability of the curves, and we use differing values of a to control the level of conditional heteroscedasticity. The properties of functional autoregressive processes are well known, see Bosq (2000). In particular, a stationary solution exists if the Hilbert–Schmidt norm of Φ is less than unity. In our setting, this is equivalent to $|c| < 1$.

Table 5. Rejection rates for Λ_p.

$N = 500, a = 0.25, c = 0.05$					$N = 2500, a = 0.25, c = 0.05$			
	1	2	3	4		1	2	3	4
10%	0.190	0.183	0.179	0.184	10%	0.483	0.495	0.494	0.492
5%	0.107	0.104	0.106	0.108	5%	0.372	0.380	0.387	0.385
1%	0.031	0.031	0.030	0.030	1%	0.195	0.201	0.203	0.201
$N = 500, a = 0.25, c = 0.10$					$N = 2500, a = 0.25, c = 0.10$				
	1	2	3	4		1	2	3	4
10%	0.442	0.437	0.442	0.442	10%	0.941	0.948	0.947	0.948
5%	0.309	0.309	0.315	0.320	5%	0.891	0.905	0.906	0.906
1%	0.124	0.134	0.130	0.132	1%	0.747	0.763	0.764	0.762
$N = 500, a = 0.25, c = 0.15$					$N = 2500, a = 0.25, c = 0.15$				
	1	2	3	4		1	2	3	4
10%	0.723	0.721	0.726	0.724	10%	1.000	1.000	1.000	1.000
5%	0.580	0.596	0.604	0.607	5%	0.998	1.000	1.000	1.000
1%	0.306	0.324	0.329	0.328	1%	0.992	0.992	0.993	0.993
$N = 500, a = 0.50, c = 0.05$					$N = 2500, a = 0.50, c = 0.05$				
	1	2	3	4		1	2	3	4
10%	0.195	0.187	0.192	0.194	10%	0.470	0.479	0.480	0.481
5%	0.105	0.105	0.108	0.108	5%	0.339	0.359	0.364	0.360
1%	0.032	0.032	0.035	0.035	1%	0.156	0.163	0.163	0.165
$N = 500, a = 0.50, c = 0.10$					$N = 2500, a = 0.50, c = 0.10$				
	1	2	3	4		1	2	3	4
10%	0.435	0.439	0.439	0.441	10%	0.943	0.946	0.948	0.948
5%	0.288	0.301	0.301	0.301	5%	0.894	0.907	0.905	0.905
1%	0.126	0.132	0.135	0.134	1%	0.723	0.734	0.738	0.738
$N = 500, a = 0.50, c = 0.15$					$N = 2500, a = 0.50, c = 0.15$				
	1	2	3	4		1	2	3	4
10%	0.689	0.690	0.691	0.690	10%	0.999	0.998	0.999	0.998
5%	0.562	0.578	0.579	0.578	5%	0.996	0.998	0.998	0.998
1%	0.329	0.364	0.367	0.363	1%	0.981	0.985	0.984	0.984

We simulate data for sample sizes $N = 500$ and 2500, with 1000 replications of each scenario. The curves are evenly sampled at a temporal resolution of 0.001 (on the [0, 1] interval), the integral of $\Phi(X_i)$ is approximated on grid with resolution 0.001×0.001, and we use a burn-in sample of 100 curves. We simulate under $a = 0.25$ and 0.5. The values of c are taken as $c = 0.05$, 0.10 and 0.15. The rejection percentages are given in Table 5. As we can see, the procedure enjoys a reasonable level of power even for small amounts of predictability ($c = 0.05$). One nice feature

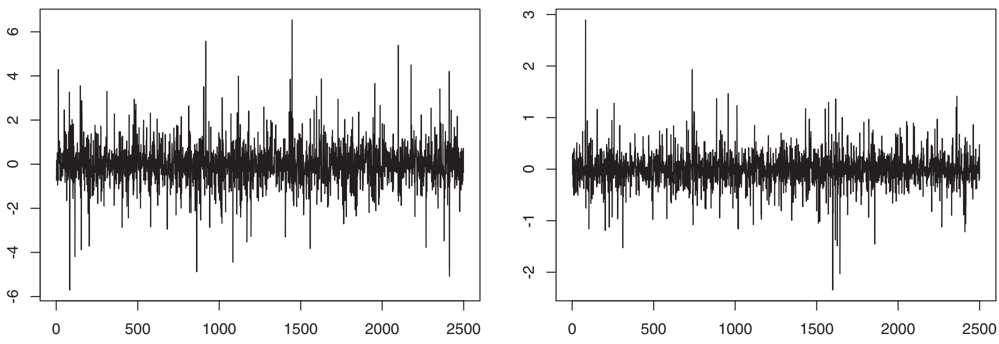

Figure 6. Simulated PC scores.

of our approach is that the results seem relatively robust against the number of components, p, used and the level of conditional heteroscedasticity. Such a property is not typical among FDA procedures, the portmanteau test can give quite different results depending on p. We performed additional simulations, not shown here, with more values of p whose eigenvalues decay at different rates and this robustness property remained. The plot of the PC score time series $\hat{\xi}_{1i}$ (left-hand panel) and $\hat{\xi}_{2i}$ (right-hand panel), $1 \leq i \leq N$, is given in Figure 6 for the simulated process in Section 3.3 with $a = 0.5$ and $c = 0.15$. If we compare the plot to the ones given in Figure 5, they seem very similar. By eye, it seems difficult to notice much of a difference, even though our procedure has a power of almost 1 under such a setting.

4. SUMMARY AND CONCLUSIONS

The shape of daily price curves may be used to determine timing of investment decisions. A natural question to ask is whether the shape of these curves on previous days provides any hint to the shape on the following day. We have attempted to develop a statistical framework for testing the null hypothesis that that such predictions are not possible. Our limited empirical study hints that the null hypothesis is true for most broadly traded US stocks. We hope that our statistical methodology will be of interest to applied researchers with access to large data bases.

Our methodology rests on the following steps. The price curves are suitably normalised to eliminate trends and scale changes, and leave only the fundamental shape. The curves so transformed are decomposed into a number of ‘principal shapes’, the FPCs. The shape is then quantitatively summarised by the weights (scores) with which these principal shapes contribute to the curves. The predictability of shapes is thus reduced to the predictability of the score processes. A test statistic is constructed which is based on the signs of the scores and which takes their relative importance into account. Focusing on signs is natural because it is interpretable (at least for the first principal component) as focusing on increasing or decreasing patterns. It has the additional advantage that the test is robust to extreme data, which do not impact the predictability, but which severely distort the size of an existing test.

The research presented in this paper leads to some interesting questions. In addition to more extensive empirical studies, we hope it will motivate an effort to find suitable functional time series models for the intraday price curves. Only an initial investigation in this direction is presented here. The testing framework we developed is fully non-parametric and suitable only for testing the basic hypothesis of predictability of with respect to the information set generated by the past observations of the curves of the same asset. Appropriate models, perhaps with a parametric structure, might be needed to test the predictability with respect to larger information sets. These models will depend on the classes of assets to be investigated. We hope that our modest contribution will generate some interest in such issues.

REFERENCES

Aue, A., S. Hörmann, L. Horváth and M. Reimherr (2009). Break detection in the covariance structure of multivariate time series models. *Annals of Statistics 37*, 4046–87.

Bosq, D. (2000). *Linear Processes in Function Spaces*. New York, NY: Springer.

Box, G. E. P. and D. A. Pierce (1970). Distribution of residual autocorrelations in autoregressive moving average time series models. *Journal of the American Statistical Association 65*, 1509–26.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Campbell, J. Y. and M. Yogo (2006). Efficient tests of stock return predictability. *Journal of Financial Economics 81*, 27–60.

Cardot, H., F. Ferraty, A. Mas, and P. Sarda (2003). Testing hypothesis in the functional linear model. *Scandinavian Journal of Statistics 30*, 241–55.

Coles, A. and D. G. Hawkins (2011). *MIDAS Technical Analysis: A VWAP Approach to Trading and Investing in Today's Markets*. New York, NY: Bloomberg Press.

DasGupta, A. (2008). *Asymptotic Theory of Statistics and Probability*. New York, NY: Springer.

Dauxois, J., A. Pousse, and Y. Romain (1982). Asymptotic theory for principal component analysis of a vector random function. *Journal of Multivariate Analysis 12*, 136–54.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer.

Fama, E. F. (1991). Efficient markets. *Journal of Finance 46*, 1575–618.

Fan, J. (1996). Test of significance based on wavelet thresholding and Neyman's truncation. *Journal of the American Statistical Association 91*, 674–88.

Fan, J. and S.-K. Lin (1998). Test of significance when data are curves. *Journal of the American Statistical Association 93*, 1007–21.

Ferraty, F. and Y. Romain (Eds.) (2011). *The Oxford Handbook of Functional Data Analysis*. Oxford: Oxford University Press.

Folger, J. and L. Leibfarth (2007). *Make Money Trading: How to Build a Winning Trading Business*. Columbia, MD: Marketplace Books.

Gabrys, R. and P. Kokoszka (2007). Portmanteau test of independence for functional observations. *Journal of the American Statistical Association 102*, 1338–48.

Hörmann, S., L. Horváth, and R. Reeder (2013). A functional version of the ARCH model. *Econometric Theory*. Forthcoming.

Hörmann, S. and P. Kokoszka (2010). Weakly dependent functional data. *Annals of Statistics 38*, 1845–84.

Hörmann, S. and P. Kokoszka (2012). Functional time series. In T. S. Rao, S. S. Rao and C. R. Rao (Eds.), *Handbook of Statistics, Time Series: Methods and Applications*, *Volume 30*, 157–86. Amsterdam: Elsevier.

Horváth, L. and P. Kokoszka (2012). *Inference for Functional Data with Applications*. New York, NY: Springer.

Kallenberg, O. (1997). *Foundations of Modern Probability*. New York, NY: Springer.

Ljung, G. and G. Box (1978). On a measure of lack of fit in time series models. *Biometrika 66*, 67–72.
McLeish, D. L. (1974). Dependent central limit theorems and invariance principles. *Annals of Probability 2*, 620–8.
Mikosch, T. and C. Stărică (2000). Limit theory for the sample autocorrelation and extremes of a GARCH(1, 1) process. *Annals of Statistics 28*, 1427–51.
Pesaran, M. H. and A. Timmermann (1995). Predictability of stock returns: robustness and economic significance. *Journal of Finance 50*, 1201–28.
Pettenuzzo, D. and A. Timmermann (2011). Predictability of stock returns and asset allocation under structural breaks. *Journal of Econometrics 164*, 60–78.
Ramsay, J. O. and B. W. Silverman (2005). *Functional Data Analysis*. New York, NY: Springer.
Shao, X. and W. B. Wu (2007). Asymptotic spectral theory for nonlinear time series. *Annals of Statistics 35*, 1773–801.

APPENDIX A: ASYMPTOTIC THEORY

An important ingredient of our asymptotic arguments is the following martingale central limit theorem established by McLeish (1974).

THEOREM A.1. *Let $Y_{n,N}$ be a triangular array of martingale differences such that(a) $\max_{1\le n\le N}|Y_{n,N}|$ is uniformly bounded in $L_2(P)$ norm, (b) $\max_{1\le n\le N}|Y_{n,N}|\to 0$, in probability, (c) $\sum_{1\le n\le N}Y_{n,N}^2\to 1$, in probability. Then, $\sum_{1\le n\le N}Y_{n,N}\xrightarrow{\mathcal{L}} N(0,1)$.*

Theorem A.1 is well suited to the work with the random variables $I_{N,n}^{(1)}$ because they depend on the whole sample through the $\hat{v}_1$, so the computation of the moments conditional on the past is difficult. Such computations are required in the more popular martingale CLTs, see e.g. Section 9.4 of DasGupta (2008). Theorem A.1 allows us to overcome the difficulty discussed in Section 1.3, namely the transition from the v_k to the $\hat{v}_k$ without using bound (1.3).

The next ingredient is the following lemma, which is used repeatedly throughout the proofs.

LEMMA A.1. *If Assumptions 2.1, 2.2 and 2.3 hold, then*

$$(c_1X_1,\ldots,c_NX_N,\hat{v}_1,\ldots,\hat{v}_p)\stackrel{\mathcal{L}}{=}(X_1,\ldots,X_N,\hat{v}_1,\ldots,\hat{v}_p),$$

where $c_1,\ldots,c_N$ is any (deterministic) sequence taking values -1 or 1.

Proof: Note that the sample covariance operator is invariant with respect to the $c_1,\ldots,c_N$:

$$\frac{1}{N}\sum_{n=1}^{N}c_iX_n(s)c_iX_n(t)=\frac{1}{N}\sum_{n=1}^{N}c_i^2X_n(s)X_n(t)=\frac{1}{N}\sum_{n=1}^{N}X_n(s)X_n(t).$$

Thus the lemma follows since $\hat{v}_1,\ldots,\hat{v}_p$ are functions of $(X_1,\ldots,X_N)$ (through the estimated covariance operator) which are symmetric and conditionally symmetric. □

LEMMA A.2. *Under the assumptions of Theorem 2.4,*

$$\lim_{N\to\infty} Var\left[\frac{\sum_{i=1}^{N-1}Z_{N,i}^2}{N-1}\right]=0.$$

Proof: To show that the variance tends to zero, consider

$$E\left[\frac{\sum_{i=1}^{N-1} Z_{N,i}^2}{N-1}\right]^2 = \frac{\sum_{i,j=1}^{N-1} E\left[Z_{N,i}^2 Z_{N,j}^2\right]}{(N-1)^2}.$$

We will argue that due to the strong stationarity of the $\mathbf{I}_{N,i}$, the above summands only take three different values: when $i = j$, $|i - j| = 1$, or $|i - j| \geq 2$. To see this, we write out the corresponding cases. To lighten the notation we suppress the subscript N. Observe that

$$\begin{aligned} E\left[Z_i^2 Z_j^2\right] &= \sum_{k,k'} a_k a_{k'} \sum_{\ell,\ell'} a_\ell a_{\ell'} E\left[I_i^{(k)} I_{i+1}^{(k)} I_i^{(k')} I_{i+1}^{(k')} I_j^{(\ell)} I_{j+1}^{(\ell)} I_j^{(\ell')} I_{j+1}^{(\ell')}\right] \\ &= \sum_{k,k'} a_k a_{k'} \sum_{\ell,\ell'} a_\ell a_{\ell'} E\left[I_i^{(k)} I_i^{(k')} I_{i+1}^{(k)} I_{i+1}^{(k')} I_j^{(\ell)} I_j^{(\ell')} I_{j+1}^{(\ell)} I_{j+1}^{(\ell')}\right]. \end{aligned}$$

If $i = j$, since $(\mathbf{I}_i, \mathbf{I}_{i+1})$ has the same distribution as $(\mathbf{I}_1, \mathbf{I}_2)$,

$$\begin{aligned} E\left[Z_i^4\right] &= \sum_{k,k'} a_k a_{k'} \sum_{\ell,\ell'} a_\ell a_{\ell'} E\left[I_i^{(k)} I_i^{(k')} I_{i+1}^{(k)} I_{i+1}^{(k')} I_i^{(\ell)} I_i^{(\ell')} I_{i+1}^{(\ell)} I_{i+1}^{(\ell')}\right] \\ &= \sum_{k,k'} a_k a_{k'} \sum_{\ell,\ell'} a_\ell a_{\ell'} E\left[I_1^{(k)} I_1^{(k')} I_1^{(\ell)} I_1^{(\ell')} I_2^{(k)} I_2^{(k')} I_2^{(\ell)} I_2^{(\ell')}\right], \end{aligned}$$

so

$$E\left[Z_i^4\right] = E\left[Z_1^4\right], \quad 1 \leq i \leq N-1.$$

If $j = i + 1$ or $i = j + 1$, using the fact that $(\mathbf{I}_i, \mathbf{I}_{i+1}, \mathbf{I}_{i+2})$ has the same distribution as $(\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3)$, one can similarly verify that

$$E\left[Z_i^2 Z_{i+1}^2\right] = E\left[Z_1^2 Z_2^2\right], \quad 1 \leq i \leq N-2. \tag{A.1}$$

Turning to the case where $|i - j| \geq 2$, observe that

$$\begin{aligned} E\left[Z_1^2 Z_3^2\right] &= \sum_{k,k'} a_k a_{k'} \sum_{\ell,\ell'} a_\ell a_{\ell'} E\left[I_1^{(k)} I_1^{(k')} I_2^{(k)} I_2^{(k')} I_3^{(\ell)} I_3^{(\ell')} I_4^{(\ell)} I_4^{(\ell')}\right] \\ &= \sum_{k,k'} a_k a_{k'} \sum_{\ell,\ell'} a_\ell a_{\ell'} E\left[I_i^{(k)} I_i^{(k')} I_{i+1}^{(k)} I_{i+1}^{(k')} I_j^{(\ell)} I_j^{(\ell')} I_{j+1}^{(\ell)} I_{j+1}^{(\ell')}\right], \end{aligned}$$

for any indices i and j such that $|i - j| \geq 2$. Thus

$$E\left[Z_i^2 Z_j^2\right] = E\left[Z_1^2 Z_3^2\right], \quad \text{if } |i - j| \geq 2.$$

There are $N - 1$ terms where $i = j$, $2(N - 2)$ terms where $|i - j| = 1$, and, finally, $(N - 1)^2 - (N - 1) - 2(N - 2) = (N - 1)^2 + o((N - 1)^2)$ terms where $|i - j| \geq 2$. Thus, since $Z_{N,1}$ and $Z_{N,3}$ are asymptotically independent by Assumption 2.3, we have

$$\frac{\sum_{i,j=1}^{N-1} E\left[Z_{N,i}^2 Z_{N,j}^2\right]}{(N-1)^2} = E\left[Z_{N,1}^2 Z_{N,3}^2\right] + o(1) \to \left(\sum_{j,k}^{p} a_j a_k (E[\mathrm{sign}(\langle X_1, v_j\rangle \langle X_1, v_k\rangle)])^2\right)^2,$$

So the second moment of $(N-1)^{-1} \sum_{i=1}^{N-1} Z_{N,i}^2$ is asymptotically equal to the first moment squared. Therefore the variance tends to zero. □

Proof of Theorem 2.1: Set

$$Y_{n,N} = (N-1)^{-1/2} I^{(j)}_{N,n} I^{(j)}_{N,n+1}, \quad \mathcal{F}^{(j)}_n = \sigma\left\{ I^{(j)}_{N,1}, \ldots, I^{(j)}_{N,n} \right\}.$$

Using Lemma A.1, we have

$$\begin{aligned}(N-1)^{1/2} E\left[Y_{n,N} | \mathcal{F}^{(j)}_n \right] &= E\left[\mathrm{sign}(\langle X_n, \hat{v}_j \rangle \langle X_{n+1}, \hat{v}_j \rangle) | \mathcal{F}^{(j)}_n \right] \\ &= E\left[\mathrm{sign}(\langle X_n, \hat{v}_j \rangle \langle -X_{n+1}, \hat{v}_j \rangle) | \mathcal{F}^{(j)}_n \right] \\ &= -E\left[\mathrm{sign}(\langle X_n, \hat{v}_j \rangle \langle X_{n+1}, \hat{v}_j \rangle) | \mathcal{F}^{(j)}_n \right] \\ &= -(N-1)^{1/2} E\left[Y_{n,N} | \mathcal{F}^{(j)}_n \right]\end{aligned}$$

which implies that

$$E\left[Y_{n,N} | \mathcal{F}^{(j)}_n \right] = 0.$$

Therefore, for each N, $\{Y_{n,N}\}$ is a martingale difference sequence with respect to $\{\mathcal{F}^{(j)}_i\}$.

We now verify conditions (a)–(c) of Theorem A.1. Since the $I^{(j)}_{N,n} I^{(j)}_{N,n+1}$ are bounded,

$$\max_{1 \le i \le N-1} \frac{\left| I^{(j)}_{N,n} I^{(j)}_{N,n+1} \right|}{\sqrt{N-1}} a.s.\!\to 0, \quad \text{as } N \to \infty.$$

Since $(I^{(j)}_{N,n})^2 (I^{(j)}_{N,n+1})^2 = 1$, we have that

$$(N-1)^{-1} \sum_{n=1}^{N-1} \left(I^{(j)}_{N,n} \right)^2 \left(I^{(j)}_{N,n+1} \right)^2 = 1. \qquad \square$$

Proof of Proposition 2.1: Since the proofs under $H^-_{A,j}$ and $H^+_{A,j}$ are identical, assume that $H^+_{A,j}$ is true. Note that $(I^{(j)}_n + 1)/2$ is a Bernoulli random variable whose expectation is $P(\xi_{j,n} > 0)$. Similarly

$$\begin{aligned}P(\xi_{j,n} > 0, \xi_{j,n-1} > 0) &= E\left[\left(\frac{I^{(j)}_2 + 1}{2} \right) \left(\frac{I^{(j)}_1 + 1}{2} \right) \right] = \frac{E\left[I^{(j)}_2 I^{(j)}_1 \right]}{4} + \frac{E\left[I^{(j)}_1 \right]}{4} + \frac{E\left[I^{(j)}_2 \right]}{4} \\ &\quad + \frac{1}{4} = \frac{1}{4} + \frac{E\left[I^{(j)}_2 I^{(j)}_1 \right]}{4},\end{aligned}$$

since, by our symmetry assumption, $E[I^{(j)}_1] = E[I^{(j)}_2] = 0$. Therefore, we have that $P(\xi_{j,n} > 0, \xi_{j,n-1} > 0) > 1/4$ under $H^+_{A,j}$. This implies that

$$P(\xi_{j,n} > 0 | \xi_{j,n-1} > 0) = \frac{P(\xi_{j,n} > 0, \xi_{j,n-1} > 0)}{P(\xi_{j,n-1} > 0)} > \frac{1/4}{1/2} = \frac{1}{2}.$$

Finally, note that

$$\begin{aligned}P(\xi_{j,n} < 0, \xi_{j,n-1} < 0) &= E\left[\left(\frac{1 - I^{(j)}_2}{2} \right) \left(\frac{1 - I^{(j)}_1}{2} \right) \right] = \frac{E\left[I^{(j)}_2 I^{(j)}_1 \right]}{4} - \frac{E\left[I^{(j)}_1 \right]}{2} + \frac{1}{4} \\ &= \frac{1}{4} + \frac{E\left[I^{(j)}_2 I^{(j)}_1 \right]}{4},\end{aligned}$$

and using the same arguments as before we obtain that $P(\xi_{j,n} < 0|\xi_{j,n-1} < 0) > 1/2$, which concludes the proof. $\square$

Proof of Theorem 2.2: The proof under $H_{A,j}^-$ is the same as under $H_{A,j}^+$, therefore we will assume without loss of generality that $H_{A,j}^+$ is true.

The ergodicity of $\{X_n\}$ implies that $\xi_{n,j}$, $I_n^{(j)}$ and $I_n^{(j)} I_{n+1}^{(j)}$ also form ergodic sequences. Since the $I_n^{(j)}$ are also bounded, we have by the ergodic theorem

$$(N-1)^{-1} \sum_{n=1}^{N-1} I_n^{(j)} I_{n+1}^{(j)} \xrightarrow{P} c > 0, \quad \left(c = E\left[I_1^{(j)} I_2^{(j)}\right]\right).$$

This implies that $(N-1)^{-1/2} \sum_{n=1}^{N-1} I_n^{(j)} I_{n+1}^{(j)} \approx N^{1/2} c \to \infty$.

Next we show that the same limit holds when $I_n^{(j)}$ is replaced with $I_{N,n}^{(j)}$. Consider the difference

$$E\left|(N-1)^{-1} \sum_{n=1}^{N-1} I_n^{(j)} I_{n+1}^{(j)} - (N-1)^{-1} \sum_{n=1}^{N-1} I_{N,n}^{(j)} I_{N,n+1}^{(j)}\right|$$
$$\leq \frac{\sum_{n=1}^{N-1} E\left|I_n^{(j)} I_{n+1}^{(j)} - I_{N,n}^{(j)} I_{N,n+1}^{(j)}\right|}{N-1}.$$

By Assumption 2.1 the above reduces to

$$\frac{\sum_{n=1}^{N-1} E\left|I_n^{(j)} I_{n+1}^{(j)} - I_{N,n}^{(j)} I_{N,n+1}^{(j)}\right|}{N-1} = E\left|I_1^{(j)} I_2^{(j)} - I_{N,1}^{(j)} I_{N,2}^{(j)}\right|.$$

By Assumption 2.1 and the assumed ergodicity, we have that $\|\hat{v}_j - v_j\| P \to 0$, see Lemma 3.2 of Hörmann and Kokoszka (2010), and therefore

$$I_{N,1}^{(j)} I_{N,2}^{(j)} \xrightarrow{\mathcal{L}} I_1^{(j)} I_2^{(j)}.$$

Since the variables are bounded we have that

$$E\left|I_1^{(j)} I_2^{(j)} - I_{N,1}^{(j)} I_{N,2}^{(j)}\right| \to 0.$$

So the difference between the two sums tends to zero and we can conclude that $\Lambda^{(j)} P \to \infty$. $\square$

Proof of Theorem 2.4: Define

$$\mathbf{I}_{N,n}^T = \left(I_{N,n}^{(1)}, \ldots, I_{N,n}^{(p)}\right), \quad \mathcal{F}_n = \sigma\{\mathbf{I}_{N,1}, \ldots, \mathbf{I}_{N,n}\}.$$

By the symmetry argument used in the proof of Theorem 2.1, for $j = 1, \ldots, p$, we have

$$E\left[I_{N,n}^{(j)} I_{N,n+1}^{(j)} | \mathcal{F}_n\right] = 0.$$

Therefore

$$\mathbf{I}_{N,n} \circ \mathbf{I}_{N,n+1},$$

is a p-dimensional martingale difference (here $\circ$ refers to the Hadamard product, i.e. component wise multiplication) with respect to $\mathcal{F}_n$.

We will now make use of the Cramer–Wold device. Define

$$Z_{N,n} = \sum_{j=1}^{p} a_k I_{N,n}^{(j)} I_{N,n+1}^{(j)},$$

for $a_j \in \mathbb{R}$. Since $\{\mathbf{I}_{N,n} \circ \mathbf{I}_{N,n+1}\}$ is a p-dimensional martingale difference, $\{Z_{N,n}\}$ is a one dimensional martingale difference array. All of the $Z_{N,n}$ are bounded so that

$$\left| \max_{1 \le n \le N-1} \frac{Z_{N,n}}{\sqrt{N-1}} \right| \overset{a.s}{\le} \sum_{j=1}^{p} \frac{|a_j|}{\sqrt{N-1}} \to 0.$$

Using the stationarity of the $\{\mathbf{I}_{N,i}\}$, we have that

$$E\left[\frac{\sum_{i=1}^{N-1} Z_{N,i}^2}{N-1}\right] = \sum_{i=1}^{N-1} \sum_{j,k}^{p} \frac{a_j a_k E\left[I_{N,i}^{(j)} I_{N,i+1}^{(j)} I_{N,i}^{(k)} I_{N,i+1}^{(k)}\right]}{N-1}$$
$$= \sum_{j,k}^{p} a_j a_k E\left[I_{N,1}^{(j)} I_{N,2}^{(j)} I_{N,1}^{(k)} I_{N,2}^{(k)}\right].$$

As $N \to \infty$, $\hat{v}_j P{\to} v_j$, and by Assumption 2.3, we have

$$E\left[\frac{\sum_{i=1}^{N-1} Z_{N,i}^2}{N-1}\right] \to \sum_{j,k}^{p} a_j a_k \left(E\left[\operatorname{sign}(\langle X_1, v_j\rangle \langle X_1, v_k\rangle)\right]\right)^2.$$

Finally, if we let

$$\sum_{j,k} = \left(E\left[\operatorname{sign}(\langle X_1, v_j\rangle \langle X_1, v_k\rangle)\right]\right)^2,$$

then we need to show that

$$\frac{\sum_{i=1}^{N-1} Z_{N,i}^2}{N-1} \xrightarrow{P} \sum_{j,k}^{p} a_j a_k \left(E\left[\operatorname{sign}(\langle X_1, v_j\rangle \langle X_1, v_k\rangle)\right]\right)^2 = a^T \mathbf{\Sigma} a. \tag{A.2}$$

By Lemma A.2,

$$\lim_{N\to\infty} Var\left[\frac{\sum_{n=1}^{N-1} Z_{N,n}^2}{N-1}\right] = 0,$$

and so (A.2), and thus Theorem A.1, holds. By the Cramer–Wold device we can conclude

$$\frac{\sum_{n=1}^{N-1} \mathbf{I}_{N,i} \circ \mathbf{I}_{N,n+1}}{\sqrt{N-1}} \xrightarrow{\mathcal{L}} N(\mathbf{0}, \mathbf{\Sigma}). \qquad \square$$