

On smoothing estimation for seasonal time series with long cycles

SONG XI CHEN* AND ZHENG XU

We consider a kernel smoothing estimator to the periodic component of seasonal time series which have quite a large periodicity relative to the length of the time series. The estimator is formulated by smoothing the commonly used seasonal-dummy estimator. It combines the neighboring seasonal-dummy estimates of the periodic function so as to reduce the variance of the estimation. We provide some theoretical justifications to the approach as well as simulation evaluations to demonstrate its effectiveness. The proposed approach is used to analyze the return rates of a German electricity price index.

AMS 2000 SUBJECT CLASSIFICATIONS: Primary 62M10, 62G05; secondary 91B84.

KEYWORDS AND PHRASES: Kernel estimator, M-dependent, Seasonal-dummy approach.

1. INTRODUCTION

Many time series contain periodic components due to the seasonal affects on the variable being measured. Existing approaches for estimating the periodic components, when the components are regarded as “fixed” function of time, include the approach of trigonometric functions (Hamilton, 1994), the seasonal-dummy variable approach (Harvey, 1993), non-parametric kernel regression such as the Nadaraya-Watson estimator (Hall, Reimann and Rice, 2000; Hall, 2008), the seasonal ARIMA models and the time-varying splines (Harvey and Koopman, 1993). The estimation of periodic functions, when they are treated as *random*, is typically based on the state space approach which models the time series as the sum of a random trend, seasonal and error components and applies the filtering techniques (Hamilton, 1994). These methods have been further extended by researchers. For example, Harvey, Koopman, and Riani (1997) employed the trigonometric approach for single seasonal time series with multiple sources of errors; and Pedregal and Young (2006) modeled for double-seasonal time series. De Livera, Hyndman and Snyder (2011) employed the Box-Cox transformation to handle the non-linearity of the time series, which allows for multiple nested and non-nested seasonal patterns.

In this paper, we consider estimation of the periodic components when they are viewed as fixed, for which situation

the seasonal-dummy variable approach is a popular method. Comparing with modeling the trend with the trigonometric functions, the seasonal-dummy variable approach has the advantage of being nonparametric and hence avoids the risk of misspecifying the cyclic functions. It is basically a sample average of the time series observations in such a way that one observation per cycle is used in the averaging, in the case of a single cycle. However, often the observational span of the series covers only limited replicates of the cycles, hence a small sample size issue will cause large variability in the seasonal-dummy variable estimation. For example, the Phelix electricity price series, which we will analyze later, spans 10 years and 5 months. Hence, the sample size in the dummy-variable estimation for an annual cycle would be based on 10 observations only, and the variance of the resulting estimates for the seasonality would be large. To improve the seasonal-dummy variable approach, we propose a kernel smoothing estimation approach that conducts a local average of neighboring seasonal-dummy regression estimates so as to achieve a variance reduction, provided the underlying periodic function is smooth enough. By choosing the smoothing bandwidth properly, the bias of the kernel smoothing estimator can be controlled while enjoying the benefits of the variance reduction. Our proposal can be viewed as one that enhances the statistical properties of the seasonal-dummy estimator while enjoying the model-robustness nature of the approach. The proposed approach is particularly applicable for estimation for relatively long periodic functions, for instance estimating the quarterly or annual pattern based on the daily time series.

The rest of the paper is organized as follows. Section 2 analyzes the properties of the seasonal-dummy variable approach and our proposed approach for a seasonal time series with a long cycle. Section 3 extends the approach to the series with a long cycle and a short cycle. Results from simulation experiments are reported in Section 4. Section 5 applies the proposed methodology to empirically study the Phelix electricity price series. Technical details are relegated to the Appendix.

2. TIME SERIES WITH A LONG CYCLE

Suppose we observe the following time series at a regular sampling interval $\delta > 0$

$$(1) \quad Y_{t\delta} = \mu + S(t\delta) + \epsilon_{t\delta} \quad \text{for } t = 1, 2, \dots, n$$

*Corresponding author.

where μ is a constant representing the long term average of the series, $S(\cdot)$ is a periodic function with period $L\delta$ for a positive integer L and $\epsilon_{t\delta}$ are the errors such that $E(\epsilon_{t\delta}) = 0$ and $Var(\epsilon_{t\delta}) = \sigma^2(t\delta)$. We follow the convention regarding the sampling interval so that $\delta = 1$ for yearly observation, which means $\delta = 1/365$ for daily and $\delta = 1/12$ for monthly observed series. With such a convention, $T = n\delta$ represents the total amount of observation time in year of the time series observed.

In order to identify μ and the periodic function, we require $\sum_{j=1}^L S(j\delta) = 0$. We may write $\epsilon_{t\delta} = \sigma(t\delta)e_{t\delta}$ where $e_{t\delta}$ are of zero mean and unit variance. Throughout the paper, we assume $\{e_{t\delta}\}_{t=1}^n$ is an α -mixing process. To simplify the presentation, we assume the variance function $\sigma^2(t\delta)$ has the same periodic feature as periodic component $S(\cdot)$ so that their periodicity is the same. This is to make our analysis less mathematically involved without altering the main results of the paper.

For any $t = 1, 2, \dots, n$, define $t_L = t - L\lfloor t/L \rfloor$ so that $t_L\delta \in \{\delta, 2\delta, \dots, L\delta\}$ where $\lfloor a \rfloor$ denotes the integer strictly less than a , regardless of whether a is an integer or not, which differs from the usual integer truncation. Then Model (1) can be written as

$$(2) \quad Y_{t\delta} = \mu + S(t_L\delta) + \epsilon_{t\delta} \quad \text{for } t = 1, 2, \dots, n$$

Let $\gamma(t\delta) = \mu + S(t\delta)$ which is the elevated periodic component. A popular method for estimating the periodic function is via the seasonal-dummy variable approach (Harvey, 1993; Hamilton, 1994). Define a row vector $A_t = (1_{t_L=1}, 1_{t_L=2}, \dots, 1_{t_L=L})$ for each t , where $1_{t_L=j}$ is the indicator for the j -th day, and $\gamma_L = (\gamma(\delta), \gamma(2\delta), \dots, \gamma(L\delta))$. Then, Model (2) can be written as a linear regression:

$$Y_{t\delta} = A_t \gamma_L + \epsilon_{t\delta} \quad \text{for } t = 1, 2, \dots, n.$$

Let $A = (A'_1, A'_2, \dots, A'_n)'$ and $Y = (Y_1, Y_2, \dots, Y_n)'$, where $'$ denotes the matrix transposition. Then, the least square (LS) estimator of γ_L is

$$\hat{\gamma}_L = (A'A)^{-1}A'Y.$$

Let $J_n = \lfloor n/L \rfloor$ be the number of complete cycles covered by the time series, and $Y_{j,k} = Y_{(jL+k)\delta}$ be the k -th observation in the j -th cycle for $j = 0, 1, 2, \dots, J_n$. It can be shown that the LS estimates of the periodic function have the following forms

$$(3) \quad \hat{\gamma}(t_L\delta) = \begin{cases} \sum_{j=0}^{J_n} Y_{j,t_L} / (J_n + 1), & \text{if } t_L \leq n - J_n L, \\ \sum_{j=0}^{J_n-1} Y_{j,t_L} / J_n, & \text{if } t_L > n - J_n L. \end{cases}$$

Then, the estimated long term average is $\hat{\mu} = \sum_{j=1}^L \hat{\gamma}(j\delta) / L$ and the estimated periodic function $\hat{S}(t_L\delta)$ is $\hat{S}(t_L\delta) = \hat{\gamma}(t_L\delta) - \hat{\mu}$.

We now present some technical conditions needed in our analysis. First we introduce the notion of α -mixing for the standardized residual series $\{e_{t\delta}\}_{t=1}^n$

$$(4) \quad \alpha(k) = \sup_{B \in \mathcal{F}_1^t, C \in \mathcal{F}_{t+k}^n} |P(B \cap C) - P(B)P(C)|, \quad k \geq 1,$$

where $\mathcal{F}_1^t$ and $\mathcal{F}_{t+k}^n$ are the σ -fields generated by $\{e_{s\delta}, s \leq t\}$ and $\{e_{s\delta}, s \geq t+k\}$ respectively. The sequence $\{e_{t\delta}\}_{t=1}^n$ is said to be α -mixing if $\lim_{k \rightarrow \infty} \alpha(k) \rightarrow 0$. Specifically, we need the following condition.

C1: The sequence $\{e_{t\delta}\}_{t=1}^n$ is strictly stationary and α -mixing, $E(|e_{t\delta}|^r) < \infty$ for a $r > 2$ and $\alpha(k) < ak^{-\beta}$ for a positive constant a and $\beta > r/(r-2)$.

Let $Q_n(t_L\delta) = (J_n + 1)^{-1} \sum_{j=0}^{J_n} e_{(jL+t_L)\delta}$ if $t_L \leq n - J_n L$, and $Q_n(t_L\delta) = J_n^{-1} \sum_{j=0}^{J_n-1} e_{(jL+t_L)\delta}$ if $t_L > n - J_n L$. Then, according to (1) and (3),

$$(5) \quad \hat{\gamma}(t_L\delta) = \gamma(t_L\delta) + \sigma(t_L\delta)Q_n(t_L\delta).$$

According to Bosq (1998), under Condition C1, there exists a positive constant $v^2(t_L)$ such that $J_n Var\{Q_n(t_L\delta)\} \rightarrow v^2(t_L)$ as $n \rightarrow \infty$, which means that $Q_n(t_L\delta) = O_p(J_n^{-1/2})$. A detailed derivation on the variance of $Q_n(t_L\delta)$ reveals that

$$v^2(t_L\delta) = \lim_{n \rightarrow \infty} \{1 + 2 \sum_{j=1}^{J_n-1} (1-j/J_n) E\{e_{(L+t_L)\delta} e_{((j+1)L+t_L)\delta}\}\}.$$

If $\{e_{t\delta}\}_{t=1}^n$ is M-dependent, namely $e_{t_1\delta}$ and $e_{t_2\delta}$ are independent if $|t_1 - t_2| > M$, then if $L > M$, we have

$$v^2(t_L\delta) = 1.$$

This is more likely to be true for a series with a long cycle, say an annual cycle.

From (5) and under C1,

$$(6) \quad \hat{\gamma}(t_L\delta) \xrightarrow{P} \gamma(t_L\delta) \quad \text{as } n \rightarrow \infty.$$

Furthermore, according to the central limit theorem for stationary α -mixing sequences as given in Theorem 1.7 of Bosq (1998), we have under conditions C1 and C2,

$$(7) \quad \sqrt{J_n} \{\hat{\gamma}(t_L\delta) - \gamma(t_L\delta)\} \xrightarrow{d} N(0, \sigma^2(t_L\delta)).$$

To show the consistency and the asymptotic normality of $\hat{\gamma}(t_L\delta)$, we need to consider the asymptotic normality of

$$\sqrt{J_n} (\hat{\gamma}(\delta) - \gamma(\delta), \dots, \hat{\gamma}(L\delta) - \gamma(L\delta))^T.$$

To this end, we need to define the asymptotic covariance matrix. For $i, j \in \{1, \dots, L\}$, let

$$\rho(i\delta, j\delta) = \lim_{n \rightarrow \infty} J_n Cov\{Q_n(i\delta), Q_n(j\delta)\},$$

which exist under condition C1, and $\nu(i\delta, j\delta) = \sigma(i\delta)\sigma(j\delta)\rho(i\delta, j\delta)$. Then, define a $L \times L$ matrix $V(\delta) =$

$(\nu(i\delta, j\delta))$. If $\{e_{t\delta}\}_{t=1}^n$ is M -dependent with $M < L$, then $\nu(i\delta, i\delta) = 1$ for $i = 1, \dots, L$ and $\nu(i\delta, i\delta) = 0$ if $|i - j| > M$. Hence, $V(\delta)$ will be a banded matrix with bandwidth M , which is the maximal number of non-zero sub-diagonals.

By the central limit theorem for stationary α -mixing sequences, for instance Theorem 1.8 of Bosq (1998), and the Crámer-Wold device, it can be shown that

$$(8) \quad \sqrt{J_n}(\hat{\gamma}(\delta) - \gamma(\delta), \dots, \hat{\gamma}(L\delta) - \gamma(L\delta))^T \xrightarrow{d} N_L(0, V(\delta))$$

where $N_L(0, V(\delta))$ denotes the L -dimensional multivariate normal distribution with zero mean and asymptotic covariance matrix V .

The seasonal-dummy variable estimation encounters a large variance if $J_n = \lfloor n/L \rfloor$ is not large, which is very likely to happen if the periodicity L is large. For example, the Phelex electricity price series, which we will analyze later, covers 11 years and 5 months and may have an annual cycle. Hence, J_n is 10, which means that the dummy-variable estimates for each $\gamma(t_L\delta)$ will be based on 11 observations only and the variance of the dummy-variable estimates would be large.

We propose a kernel smoothing approach that combines neighboring periodic functional estimates so as to increase the “sample size” used in estimating each $\gamma(t_L\delta)$. We need the following smoothness condition regarding the periodic function and those involved with the covariances.

C.2: The periodic $\gamma(t_L\delta)$ and the covariance related quantities $\sigma^2(i\delta)$ and $\nu(i\delta, j\delta)$ are realizations at the discrete points $\{i\delta\}_{i=1}^L$ of the same named functions $\gamma(s)$, $\sigma^2(i\delta)$ and $\nu(s_1, s_2)$, respectively, which have continuous second derivatives over $[0, L\delta]$. Both $\sigma^2(t)$ and $\rho(t_1, t_2)$ are periodic with period L .

Let K be a continuous symmetric unimodal probability density with compact support, its maximum at the origin and finite second moment. Define the kernel weights $K_h(t) = K(t/h)/h$ where h is a smoothing bandwidth that controls the bias and variance of the smoothed estimate. Let $m = (L - 1)/2$ for odd m and $m = (L - 2)/2$ for even m . The kernel smoothing estimator of $\gamma(t_L\delta)$ is

$$(9) \quad \begin{aligned} \tilde{\gamma}_h(t_L\delta) &= \frac{K_h(0)\hat{\gamma}(t_L\delta) + \sum_{j=1}^m K_h(j\delta)[\hat{\gamma}\{(t_L+j)\delta\} + \hat{\gamma}\{(t_L-j)\delta\}]}{K_h(0) + 2\sum_{j=1}^m K_h(j\delta)}, \end{aligned}$$

which is a weighted average of the seasonal-dummy estimates around $t_L\delta$.

Proposition 2.1. *Under Conditions C.1 and C.2,*

$$\text{Bias}\{\tilde{\gamma}_h(t_L\delta)\} = \frac{\sum_{j=1}^m K_h(j\delta)j^2\delta^2}{K_h(0) + 2\sum_{j=1}^m K_h(j\delta)}\gamma''(t_L\delta)$$

and

$$\begin{aligned} & \text{Var}\{\tilde{\gamma}(t_L\delta)\} \\ &= [J_n\{\sum_{j=-m}^m K_h(j\delta)\}^2]^{-1} \sum_{i=-m}^m \sum_{j=-m}^m K_h(i\delta)K_h(j\delta) \\ & \quad \nu\{(i+t_L)\delta, (j+t_L)\delta\}. \end{aligned}$$

We now outline some benefit of the smoothing if $\nu(i\delta, i\delta) = \nu(j\delta, j\delta)$ for all $i \neq j$, namely the residuals are homoscedastic and the covariance matrix $V(\delta)$ has identical diagonal elements. In this case, $\nu\{(i+t_L)\delta, (j+t_L)\delta\} \leq \nu(t_L\delta, t_L\delta)$ for all i, j . This together with the fact that all the kernel weights are non-negative leads to

$$\begin{aligned} & J_n\{\sum_{j=-m}^m K_h(j\delta)\}^2 \text{Var}\{\tilde{\gamma}(t_L\delta)\} - \text{Var}\{\hat{\gamma}(t_L\delta)\} \\ &= \sum_{i=-m}^m \sum_{j=-m}^m K_h(i\delta)K_h(j\delta) \\ & \quad \{\nu\{(i+t_L)\delta, (j+t_L)\delta\} - \nu(t_L\delta, t_L\delta)\} \leq 0. \end{aligned}$$

Hence, the kernel smoothing reduces the variability of the dummy variable estimator $\hat{\gamma}(t_L\delta)$.

When the variances $\{\nu(i\delta, i\delta)\}$ are not the same, smoothing at a low variance location say $t_L\delta$ may incur larger variance due to its neighbors having larger variance. However, this can be controlled for M -dependent time series with $M < L$, if the ratio $\max_t \sigma^2(t\delta)/\min_t \sigma^2(t\delta)$ satisfies

$$(10) \quad \begin{aligned} & 2\sum_{k=1}^m \sum_{i=-m}^{m-k} K_h(i\delta)K_h((i+k)\delta)\rho\{(t_L+i)\delta, (t_L+i+k)\delta\} \\ & \leq \frac{\min_t \sigma^2(t\delta)}{\max_t \sigma^2(t\delta)} \left\{ \sum_{j=-m}^{j=m} K_h(j\delta)\}^2 - \sum_{j=-m}^m K_h^2(j\delta) \right\}. \end{aligned}$$

If the residuals are homoscedastic, then $\max_t \sigma^2(t\delta) = \min_t \sigma^2(t\delta)$ and (10) is valid. To appreciate the role of (10), let $U = \max_t \sigma^2(t\delta)/\min_t \sigma^2(t\delta)$, then (10) implies

$$\sigma_{(t_L+i)\delta}^2 \leq U\sigma^2(t_L\delta) \text{ and } \sigma_{(t_L+i)\delta}\sigma_{(t_L+i+k)\delta} \leq U\sigma^2(t_L\delta).$$

Then, from the variance in Proposition 2.1,

$$\begin{aligned} & \text{Var}\{\tilde{\gamma}(t_L\delta)\} \\ & \leq [J_n\{\sum_{j=-m}^m K_h(j\delta)\}^2]^{-1} \left[\sum_{i=-m}^m K_h^2(i\delta)U\sigma_{t_L\delta}^2 \right. \\ & \quad + 2\sum_{k=1}^M \sum_{i=-m}^{m-k} K_h(i\delta)K_h((i+k)\delta) \\ & \quad \left. U\sigma_{t_L\delta}^2\rho\{(t_L+i)\delta, (t_L+i+k)\delta\} \right] \\ & \leq [J_n\{\sum_{j=-m}^m K_h(j\delta)\}^2]^{-1} \sigma_{t_L\delta}^2 U \left[\sum_{j=-m}^m K_h^2(j\delta) \right] \end{aligned}$$

$$\begin{aligned}
& +2 \sum_{k=1}^M \sum_{i=-m}^{m-k} K_h(i\delta) K_h((i+k)\delta) \\
& \rho\{(t_L + i)\delta, (t_L + i + k)\delta\} \\
& \leq [J_n \{ \sum_{j=-m}^m K_h(j\delta) \}^2]^{-1} \sigma_{t_L\delta}^2 \{ \sum_{j=-m}^m K_h(j\delta) \}^2 \\
& = \text{Var}\{\hat{\gamma}(t_L\delta)\},
\end{aligned}$$

where the second last inequality is based on (10) and the last equality is based on M -dependence with $M < L$.

To gain more insight on the bias and variance expressions given in Proposition 2.1, we consider a high frequency scenario where $\delta \rightarrow 0$, $\delta/h \rightarrow 0$ but $m\delta$ remains fixed to be $T/2$, the half length of the cycle. This allows approximations of the sums in the bias and variance by integrals under Condition C.2. Let $\sigma_K^2 = \int z^2 K(z) dz$.

Proposition 2.2. *Under Conditions C.1 and C.2, as $\delta \rightarrow 0$, $h \rightarrow 0$, $\delta/h \rightarrow 0$ but $m\delta$ is fixed,*

$$\begin{aligned}
\text{Bias}\{\tilde{\gamma}(t_L\delta)\} & \approx \frac{1}{2} h^2 \sigma_K^2 \gamma''(t\delta); \\
\text{Var}\{\tilde{\gamma}(t_L\delta)\} & \approx \delta J_n^{-1} h^{-1} \{ \nu(t_L\delta, t_L\delta) \int K^2(t) dt \\
& + 2 \sum_{j=1}^{2m} \nu(t_L\delta, (t_L + j)\delta) \\
& \int K(t) K(j\frac{\delta}{h} + t) dt \}.
\end{aligned}$$

Let $R = 2 \sum_{j=1}^{2m} \nu\{t_L\delta, (t_L + j)\delta\} \int K(t) K(j\frac{\delta}{h} + t) dt$. The mean square error (MSE) of $\tilde{\gamma}(t_L\delta)$ is approximately

$$\begin{aligned}
(11) \quad \text{MSE}\{\tilde{\gamma}(t_L\delta)\} & \approx \frac{1}{4} h^4 \sigma_K^4 \{ |\gamma''(t_L\delta)| \}^2 \\
& + \frac{[\nu(t_L\delta, t_L\delta) R(K) + R] \delta}{J_n h}.
\end{aligned}$$

We can compare the MSEs of the seasonal-dummy estimator $\hat{\gamma}(t_L\delta)$ and the kernel smoothed estimator under the M -dependence. Specifically,

$$\begin{aligned}
(12) \quad \text{MSE}\{\hat{\gamma}(t_L\delta)\} - \text{MSE}\{\tilde{\gamma}(t_L\delta)\} \\
= \frac{\nu(t_L\delta, t_L\delta)}{J_n} \{ 1 - \frac{R(K)\delta}{h} \} - \frac{R\delta}{J_n h} - \frac{1}{4} h^4 \sigma_K^4 |\gamma''(t_L\delta)|^2.
\end{aligned}$$

Hence if h is chosen such that $(R(K) + \frac{\mathbf{R}}{\nu(t_L\delta, t_L\delta)})\delta/h < 1$ which is equivalent to $(R(K) + \frac{\mathbf{R}}{\nu(t_L\delta, t_L\delta)})\delta < h$, smoothing will reduce the variance of estimation. Intuitively speaking, the smaller δ is, the approach is more likely to reduce the variance.

The optimal bandwidth h^* that minimizes the approximate mean square error (MSE) of $\tilde{\gamma}(t_L\delta)$ is

$$(13) \quad h^* = \frac{[\nu(t_L\delta, t_L\delta)\{R(K)\} + R]^{\frac{1}{5}} \delta^{\frac{1}{5}}}{J_n^{\frac{1}{5}} \sigma_K^{\frac{4}{5}} |\gamma''(t_L\delta)|^{\frac{2}{5}}}.$$

In practice the cross validation (CV) method (Härdle, 1990; Fan and Gijbels, 1996) can be used to select the smoothing bandwidth h . In the simulation and empirical study reported in Sections 4 and 5, the CV method is shown to be effective in producing quality estimation of a long periodic component.

3. TIME SERIES WITH A LONG CYCLE AND A SHORT CYCLE

We consider in this section extending our analysis to the case of the time series having a long cycle and a short cycle. We are to show that the same form of results we have established in the last section for the case of a single long cycle can be attained. A key intermediate result is that there is a kind of orthogonality between the short cycle and the long cycle estimates for the seasonal-dummy approach.

Consider the following time series model:

$$(14) \quad Y_{t\delta} = \mu + S_w(t\delta) + S_a(t\delta) + \epsilon_{t\delta} \quad \text{for } t = 1, 2, \dots, n,$$

where $S_w(\cdot)$ denotes a short periodic function with periodicity L_w and $S_a(\cdot)$ denotes a longer periodic function with periodicity L_a , and $\epsilon_{t\delta}$ are the residuals which satisfied the same assumptions as the previous section. The subscripts w and a had their origins in denoting the weekly and annual cycles in the study of the Phelix electricity time series, which motivated our analysis. We assume that the periodicities L_a and L_w ($L_w < L_a$) are integers with 1 being their greatest common divisor. For identification purposes, it is required that $\sum_{j=1}^{L_a} S_a(j\delta) = 0$ and $\sum_{j=1}^{L_w} S_w(j\delta) = 0$.

Similar to the notation in Section 2, we denote $\gamma(t\delta) \equiv c + S_a(t\delta)$, $t_a \equiv t - L_a \lfloor t/L_a \rfloor$ and $t_w \equiv t - L_w \lfloor t/L_w \rfloor$. Thus, $t_a \in \{1, 2, \dots, L_a\}$ and $t_w \in \{1, 2, \dots, L_w\}$. Then Model (14) can be written as

$$(15) \quad Y_{t\delta} = S_w(t_w\delta) + \gamma(t_a\delta) + \epsilon_{t\delta} \quad \text{for } t = 1, 2, \dots, n.$$

Define $W_t = (1_{t_w=1} - 1_{t_w=L_w}, 1_{t_w=2} - 1_{t_w=L_w}, \dots, 1_{t_w=L_w} - 1_{t_w=L_w})$ and $A_t = (1_{t_a=1}, 1_{t_a=2}, \dots, 1_{t_a=L_a})$ be the day indicator vectors for the short and the long cycles respectively. With these notations, Model (15) can be expressed as the following linear regression model:

$$(16) \quad Y_{t\delta} = W_t \beta_w + A_t \beta_a + \epsilon_{t\delta} \quad (t = 1, 2, \dots, n),$$

where $\beta_w = (S_w(\delta), \dots, S_w(L_w - 1\delta))'$ and $\beta_a = (\gamma(\delta), \gamma(2\delta), \dots, \gamma(L_a\delta))'$.

Let $W = (W_1', W_2', \dots, W_n')'$, $A \equiv (A_1', A_2', \dots, A_n')'$ and $Y \equiv (Y_\delta, Y_{2\delta}, \dots, Y_{n\delta})'$, $e \equiv (e_\delta, e_{2\delta}, \dots, e_{n\delta})'$, then (16) becomes

$$Y = W\beta_w + A\beta_a + e.$$

Furthermore, define $X \equiv (W, A)$ and $\beta \equiv (\beta_w', \beta_a')'$, then

$$(17) \quad Y = X\beta + e.$$

The LS estimator of β is

$$(18) \quad \hat{\beta} = (X'X)^{-1}X'Y$$

Our interest is in estimating β_a which corresponds to $\gamma(t_a\delta)$ of the long cycle. A key fact we use is that the columns of A and columns of W are nearly orthogonal such that

$$(19) \quad n^{-1}W'A = O(n^{-1}).$$

Let us first consider the case where the length of the series n is an integer multiple of L_wL_a . Denote (i, j) -th element of $W'A$ as $(W'A)_{i,j}$. Define vectors

$$\begin{aligned} \eta_{t_w=i} &= (1_{1-L_w \lfloor \frac{1}{L_w} \rfloor = i}, 1_{2-L_w \lfloor \frac{2}{L_w} \rfloor = i}, \dots, 1_{n-L_w \lfloor \frac{n}{L_w} \rfloor = i})' \in \mathcal{R}^n, \end{aligned}$$

and

$$\begin{aligned} \eta_{t_a=i} &= (1_{1-L_a \lfloor \frac{1}{L_a} \rfloor = i}, 1_{2-L_a \lfloor \frac{2}{L_a} \rfloor = i}, \dots, 1_{n-L_a \lfloor \frac{n}{L_a} \rfloor = i})' \in \mathcal{R}^n. \end{aligned}$$

As the greatest common divisor of L_w and L_a is one, it can be checked that

$$(20) \quad \begin{aligned} (W'A)_{i,j} &= (\eta_{t_w=i} - \eta_{t_w=L_w})' \eta_{t_a=j} \\ &= \lfloor n/(L_wL_a) \rfloor - \lfloor n/(L_wL_a) \rfloor = 0. \end{aligned}$$

This indicates “orthogonality” between the column vectors of W and A . Let 1_{L_a} denote a vector of length L_a with all its entries being 1 and 0_{L_w-1} denote a vector of length $L_w - 1$ with all its entries being 0. In this case, the least square estimator of β_a is

$$\begin{aligned} \hat{\beta}_a &= [0'_{L_w-1}, 1'_{L_a}] \begin{pmatrix} (W'W)^{-1} & O \\ O & (A'A)^{-1} \end{pmatrix} X'Y \\ &= (A'A)^{-1}A'Y. \end{aligned}$$

This means that under the orthogonality, the seasonal-dummy estimator of β_a is the same as the one that ignores the short cycle. Similarly, the estimation of the short cycle can ignore the long cycle as well under the orthogonality.

We now consider the more general case where n is not an integer multiple of L_aL_w so that there is a residual block toward the end of the series. It can be shown that the (i, j) -th element of $W'A$ satisfy

$$|(W'A)_{i,j}| \leq 1$$

with the non-zero entries largely contributed by the residual fraction. Hence, as $n \rightarrow \infty$

$$\begin{aligned} W'W &= \frac{1}{n} \lfloor n/L_w \rfloor (I_{L_w-1, L_w-1} + 1_{L_w-1} 1'_{L_w-1}) + O(n^{-1}) \\ &\rightarrow \frac{1}{L_w} (I_{L_w-1, L_w-1} + 1_{L_w-1} 1'_{L_w-1}), \\ W'A &= 0 + O(n^{-1}) \rightarrow 0, \end{aligned}$$

$$A'W = 0 + O(n^{-1}) \rightarrow 0,$$

$$\begin{aligned} A'A &= \frac{1}{n} \lfloor n/L_a \rfloor I_{L_a, L_a} + O(n^{-1}) \\ &\rightarrow \frac{1}{L_a} I_{L_a, L_a}, \end{aligned}$$

where $O(n^{-1})$ refers to a matrix with all its elements being at the magnitude n^{-1} . So as $n \rightarrow \infty$

$$\begin{aligned} &n^{-1}X'X \\ &= n^{-1} \begin{pmatrix} W'W & W'A \\ A'W & A'A \end{pmatrix} \\ &\rightarrow \begin{pmatrix} \frac{1}{L_w} (I_{L_w-1, L_w-1} + 1_{L_w-1} 1'_{L_w-1}) & 0 \\ 0 & \frac{1}{L_a} I_{L_a, L_a} \end{pmatrix}. \end{aligned}$$

This implies that asymptotically $(X'X)$ is block diagonal, hence the least square estimation of one periodic component can be carried out by ignoring the other cyclic component as far as the estimation is concerned.

Denote $J_n = \lfloor n/L_a \rfloor$. Hence, the seasonal dummy estimator of the long cycle

$$(21) \quad \hat{\gamma}(t_L\delta) = \begin{cases} \sum_{j=0}^{J_n} Y_{j, t_L} / (J_n + 1), & \text{if } t_L \leq n - J_n L, \\ \sum_{j=0}^{J_n-1} Y_{j, t_L} / J_n, & \text{if } t_L > n - J_n L; \end{cases}$$

which assumes a similar form to that in Section 2 when there is a single cycle.

The properties of the seasonal-dummy estimator can be evaluated by noticing that

$$(22) \quad \hat{\gamma}(t_L\delta) = \gamma(t_L\delta) + B_{1n}(t_L\delta) + \sigma(t_L\delta)Q_n(t_L\delta),$$

where, by utilizing the fact that the sum of consecutive L_w terms in S_w is 0 and the fact that the greatest common divisor of L_w and L_a is 1,

$$\begin{aligned} B_{1n}(t_L\delta) &= (J_n + 1)^{-1} \sum_{j=L_w \lfloor J_n/L_w \rfloor}^{J_n} S_w((jL_a + t_L)\delta) \\ &\quad \text{if } t_L \leq n - J_n L_a, \\ B_{1n}(t_L\delta) &= J_n^{-1} \sum_{j=L_w \lfloor J_n/L_w \rfloor}^{J_n-1} S_w((jL_a + t_L)\delta) \\ &\quad \text{if } t_L > n - J_n L_a \end{aligned}$$

and

$$\begin{aligned} Q_n(t_L\delta) &= (J_n + 1)^{-1} \sum_{j=0}^{J_n} e_{(jL_a + t_L)\delta} \text{ if } t_L \leq n - J_n L_a, \\ Q_n(t_L\delta) &= J_n^{-1} \sum_{j=0}^{J_n-1} e_{(jL_a + t_L)\delta} \text{ if } t_L > n - J_n L_a. \end{aligned}$$

We note that $B_{1n}(t_L\delta) = O(J_n^{-1})$ since it effectively captures less than L_w terms of $S_w(\cdot)$. This implies that the

Table 1. Average squared bias, variance and MSE (multiplied by 10^3) of the seasonal-dummy estimator ($\hat{S}_{a,1}(t_L\delta)$), the seasonal-dummy estimator that ignores the short cycle ($\hat{S}_{a,2}(t_L\delta)$), and the kernel estimator ($\hat{S}_{a,3}(t_L\delta)$) with cross-validated bandwidth for the four simulation settings under Model (23). The unit for the sample size n is 10^3

Model	n	h_{cv}	Bias ²			Var			MSE		
			$\hat{S}_{a,1}$	$\hat{S}_{a,2}$	$\hat{S}_{a,3}$	$\hat{S}_{a,1}$	$\hat{S}_{a,2}$	$\hat{S}_{a,3}$	$\hat{S}_{a,1}$	$\hat{S}_{a,2}$	$\hat{S}_{a,3}$
$L_a = 90$ Homos.	1	4.5	0.13	2.07	1.35	88.45	88.36	5.88	88.58	90.43	7.23
	2	3.9	0.07	0.22	0.81	44.33	44.32	3.27	44.40	44.54	4.08
	4	3.4	0.04	0.14	0.48	22.24	22.23	1.84	22.28	22.37	2.32
	6	3.1	0.03	0.07	0.33	14.68	14.68	1.33	14.71	14.75	1.67
	8	2.9	0.02	0.04	0.24	11.03	11.03	1.11	11.05	11.07	1.34
$L_a = 90$ Heteros.	1	4.5	0.18	1.92	1.24	100.42	100.27	7.09	100.59	102.19	8.33
	2	4.0	0.08	0.27	0.80	49.33	49.33	3.82	49.40	49.59	4.62
	4	3.5	0.05	0.15	0.45	25.28	25.28	2.27	25.32	25.42	2.73
	6	3.2	0.03	0.08	0.33	16.73	16.73	1.65	16.76	16.81	1.98
	8	3.0	0.02	0.05	0.26	12.51	12.51	1.27	12.53	12.56	1.52
$L_a = 365$ Homos.	1	18.3	0.75	30.90	1.19	378.92	375.15	5.95	379.66	406.05	7.15
	2	15.9	0.39	4.74	0.61	183.40	183.23	3.30	183.79	187.97	3.90
	4	13.9	0.17	2.09	0.37	91.34	91.32	1.83	91.51	93.40	2.20
	6	12.9	0.12	0.86	0.29	60.58	60.57	1.28	60.70	61.43	1.58
	8	12.2	0.08	0.22	0.22	45.53	45.53	1.05	45.61	45.74	1.27
$L_a = 365$ Heteros.	1	19.0	0.87	30.59	1.44	431.24	426.99	6.27	432.10	457.58	7.70
	2	16.4	0.43	4.98	0.87	205.75	205.57	3.55	206.18	210.55	4.41
	4	14.1	0.20	2.13	0.42	101.90	101.86	1.97	102.09	103.99	2.40
	6	13.1	0.15	0.90	0.33	67.84	67.83	1.46	67.99	68.74	1.79
	8	12.4	0.10	0.24	0.27	51.32	51.32	1.15	51.42	51.56	1.42

central limit theorem we establish in Section 2 can be also made here even in the presence of another cycle.

Because the form and the properties of $\hat{\gamma}_a(t\delta)$ in the presence of another cycle are the same with those in Section 2 in the absence of another cycle, we can propose the same kernel estimator for the long cycle as given in Equation (9) that smooths the seasonal-dummy estimates $\hat{\gamma}_a$ in (18). The kernel estimator for the long cycle is denoted as $\tilde{\gamma}_a(t\delta)$.

4. SIMULATION STUDIES

We conducted simulation experiments which were designed to evaluate the estimators for the long cycle proposed in the previous section. The model used in the data generating process was

$$(23) \quad Y_{t\delta} = 2 + S_w(t_w\delta) + \sin(2\pi t/L_a) + \epsilon_{t\delta},$$

where the short cycle $S_w(t_w\delta)$ represents a weekly cycle with $S_w(t_w\delta) = 0.4, 0.2, 0.15, 0.05, -0.1, -0.3, -0.4$ for $t_w = 1, 2, \dots, 7$ respectively and $\epsilon_{t\delta} \stackrel{d}{\sim} N(0, V^2)$ with $V = 1$ and $1 + 0.5 \sin(2\pi t/30)$ respectively. We note here that the conditional variance has different cycles from the cycles in the mean, which deviates from our assumptions in the previous sections. This was designed to illustrate that the assumption is not essential for our approach. The value of $S_{w\delta}$ given above was the estimates from the empirical study on the return rates of Phelix electricity prices, which we will analyze in the next section. The long cycle was $\sin(2\pi t/L_a)$

with $L_a = 90$ and 365 respectively. We considered *four* settings that cross the homoscedastic and heteroscedastic errors with the length of the long cycle, respectively. Specifically, Setting 1: $L_a = 90$ and $V = 1$; Setting 2: $L_a = 90$ and $V = 1 + 0.5 \sin(2\pi t/30)$; Setting 3: $L_a = 365$ and $V = 1$ and Setting 4: $L_a = 365$ and $V = 1 + 0.5 \sin(2\pi t/30)$. In each setting, we chose the sample size $n = 1,000, 2,000, 4,000, 8,000$ and $16,000$, respectively.

We considered three estimators of the long cycle: (i) the full seasonal-dummy (the least square) estimator $\hat{S}_{a,1}(t_L\delta)$, (ii) the seasonal-dummy estimator that ignores the short cycle $\hat{S}_{a,2}(t_L\delta)$, and (iii) the kernel smoothing estimator $\hat{S}_{a,3}(t_L\delta)$ which smooths the full seasonal-dummy estimator $\hat{S}_{a,1}(t_L\delta)$.

For the kernel estimator, the fixed bandwidths of 3, 7 and 15 days were used together with the data driven cross-validation bandwidth.

Simulation results for each of the four settings based on 500 replications are reported in Tables 1 and 2, respectively. Table 1 reports the performance of different estimators $\hat{S}_{a,1}(t_a\delta)$, $\hat{S}_{a,2}(t_a\delta)$ and $\hat{S}_{a,3}(t_a\delta)$ with the bandwidth selected by cross-validation. Table 2 reports the performance of $\hat{S}_{a,3}(t_a\delta)$ using different bandwidths. Table 2 also reports the average of the cross-validation bandwidth $\bar{h}_{cv}$. The main observation of Table 1 is that the kernel smoothing estimator $\hat{S}_{a,3}(t_a\delta)$ was always better than the other two estimators $\hat{S}_{a,1}(t_a\delta)$ and $\hat{S}_{a,2}(t_a\delta)$, mainly due to a significant reduction in the variance. In almost all the settings, the variance of the

Table 2. Average squared bias and variance (multiplied by 10^3) of the kernel smoothing estimator with different bandwidths 3, 7, 15 and the cross validation (CV) under Model (23). The unit for the sample size n is 10^3

Model Bandwidth	n	h_{CV}	$Bias^2$				Var			
			CV	3	7	15	CV	3	7	15
$L_a = 90$ Homos.	1	4.5	1.35	0.28	6.53	87.84	5.88	7.41	2.64	0.71
	2	3.9	0.81	0.30	6.60	87.99	3.27	3.72	1.30	0.34
	4	3.4	0.48	0.28	6.53	87.83	1.84	1.85	0.66	0.18
	6	3.1	0.33	0.27	6.49	87.73	1.33	1.24	0.43	0.11
	8	2.9	0.24	0.24	6.35	87.39	1.11	0.95	0.34	0.09
$L_a = 90$ Heteros.	1	4.5	1.24	0.22	6.16	86.92	7.09	8.30	2.95	0.79
	2	4.0	0.80	0.24	6.33	87.33	3.82	4.05	1.42	0.37
	4	3.5	0.45	0.24	6.34	87.37	2.27	2.13	0.76	0.21
	6	3.2	0.33	0.25	6.40	87.51	1.65	1.44	0.53	0.14
	8	3.0	0.26	0.24	6.35	87.39	1.27	1.07	0.39	0.11
$L_a = 365$ Homos.	1	18.3	1.19	0.05	0.04	0.50	5.95	35.00	14.47	6.09
	2	15.9	0.61	0.04	0.02	0.44	3.30	16.90	6.98	2.98
	4	13.9	0.37	0.02	0.02	0.49	1.83	8.41	3.47	1.46
	6	12.9	0.29	0.01	0.03	0.51	1.28	5.52	2.26	0.96
	8	12.2	0.22	0.01	0.02	0.49	1.05	4.18	1.74	0.75
$L_a = 365$ Heteros.	1	19.0	1.44	0.11	0.06	0.54	6.27	39.26	16.15	6.76
	2	16.4	0.87	0.04	0.05	0.59	3.55	18.90	7.77	3.32
	4	14.1	0.42	0.02	0.03	0.51	1.97	9.35	3.79	1.58
	6	13.1	0.33	0.01	0.03	0.54	1.46	6.28	2.58	1.09
	8	12.4	0.27	0.01	0.03	0.54	1.15	4.74	1.95	0.83

kernel estimator was less than 10% of its seasonal-dummy counterparts, while the bias was only slightly larger. This spells out nicely the benefits of the kernel smoothing in estimating long periodic component. Table 2 shows that among the kernel estimator with different bandwidths, the cross-validation bandwidth offered satisfactory performance. The best performing bandwidth under each model setting was the one that was the closest to the cross-validation bandwidth, which lends further support to the cross-validation approach.

To gain more insight on the benefits of the smoothing, we plot in Figure 1 the seasonal-dummy and the kernel smoothed estimates for the long cycle based on two simulated time series under Model (23). The volatility of the seasonal-dummy estimates was the most visible while the kernel estimates were quite nice and smoothed, and were quite close to the underlying truth.

5. EMPIRICAL STUDY

We analyzed the Physical Electricity Index (Phelix) in the German-based European Exchange (EEX). This index is the average hourly price for the 24 hour period of a day. We referred to it as Phelix in our paper. The average prices, denoted as P_t , are the reference price for swap, futures and options, and have drawn much attention in many empirical studies (e.g. Benth, Benth and Koekebakker, 2008; Benth, Kiesel and Nazarova, 2012). Electricity spot prices tend to have seasonality, which can be explained by the commodity's characteristics. Seasonality of the electricity spot prices

is widely observed in many markets including the EEX. It is caused by the working week pattern and the annual natural seasons. There has been a set of literature on modeling electricity spot prices including Benth, Benth and Koekebakker (2008), Rebennack et al. (2010), Benth, Kiesel and Nazarova (2012). We studied the seasonality in the return rates of electricity prices. We intended to answer the question whether the return rate of electricity prices has an annual and/or a weekly cycle.

The period of the Phelix index we considered ranges from July 7, 2000 to December 31, 2011, consisting of 11 years and 6 months with 4,213 daily prices $P_{t\delta}$ at $\delta = 1/365$. There were two negative prices on October 4, 2009 and October 26, 2009, which can happen and reflect the fact that electricity power companies sometime have to pay to discharge a big surplus supply in order to maintain the order of the power production. In the calculation of the log prices $\ln(P_{t\delta})$, these two negative prices were discarded. Thus, the time series of log-prices has a sample size $n = 4,211$. We then calculated the log returns $Y_{t\delta} = \ln(P_{t\delta}) - \ln(P_{(t-1)\delta})$. The Phelix price $P_{t\delta}$, its log prices $\ln(P_{t\delta})$ and the log returns $Y_{t\delta}$ are plotted in Figure 2, which show that the log returns were largely stationary.

To explore the seasonality, we calculated the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF), the widely used tools in exploring the seasonal patterns of a time series (Hamilton 1994), which are displayed in Figure 3 at lags values ranging from 1 to 40 and 1 to 400 respectively. Both the sample ACF and the PACF showed spikes at lag values in multiples of 7, which in-

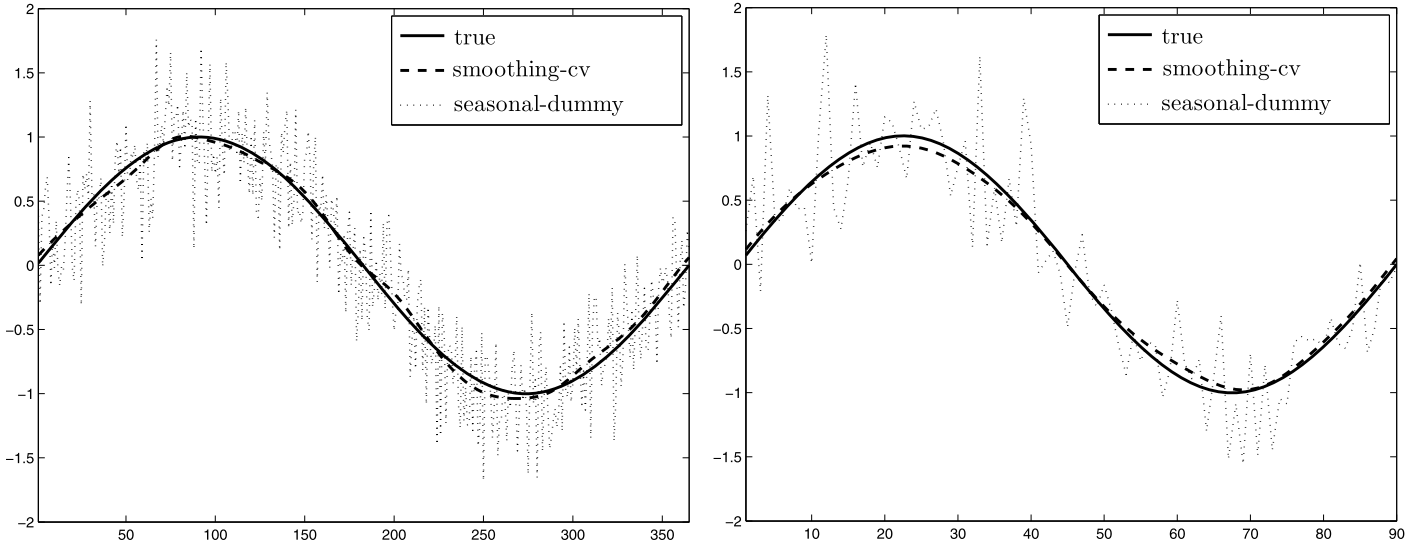


Figure 1. Estimated long-cycle by the seasonal-dummy estimator and the kernel estimator with the cross-validation bandwidth based on two simulated series from Model (23). Left panel: $L_a = 365$ with $n = 4,000$; Right Panel: $L_a = 90$ with $n = 1,000$.

indicated a weekly cycle in the log returns. However, the ACF and PACF did not provide much support for an annual cycle as no spikes around the lag 365 in the ACF and PACF were observed. To gain empirical support for the weekly pattern, we plotted the 601 weekly return rates in Figure 4 by connecting each weekday's returns by lines. The figure showed a consistent monotone decreasing pattern that peaked on Mondays, and decreased from Tuesday to Sunday.

We then analyzed the log return data by considering the following model:

$$(24) \quad Y_{t\delta} = \mu + S_w(t_w\delta) + S_a(t_a\delta) + \epsilon_{t\delta},$$

with a weekly cycle and an annual cycle, where $t_w = t - 7\lfloor t/7 \rfloor$, $t_a = t - 365\lfloor t/365 \rfloor$, and $\sum_{j=1}^7 S_w(j\delta) = 0$ and $\sum_{j=1}^{365} S_a(j\delta) = 0$ for the identification purpose. The residuals were assumed to be α -mixing as assumed in Section 2.

We first estimated the weekly cycle and the annual cycle by the seasonal-dummy variable approach, followed by carrying out the kernel smoothing using the bandwidth of 3 days and 15 days. These estimates are displayed in Figure 5. The seasonal-dummy estimate of the long cycle given in Panel (2) of the figure displayed strong volatility which reflected the fact that each estimate was only based on 11 or 12 observations. The kernel smoothed estimates of the annual cycle were quite flat regardless of the bandwidths used. The bandwidth selected by the cross-validation was very large, which may be viewed as an indication that the annual cycle was not significant. Panel (1) of Figure 5 also displayed the seasonal-dummy estimate (in dashed line) of the short cycle assuming no annual cycle, which was almost identical to the estimate assuming the presence of the long annual cycle. This is another indication that the annual cycle was not significant.

We observe that the estimates of the weekly cycle were monotone decreasing from Monday with the lowest values during the weekend. This was consistent with the plot of the weekly returns in Figure 4. The presence of the weekly cycle is tested by comparing the model that prescribes only the long cycle as specified in Model (1) with the model that has both the weekly and the annual cycles as specified in Model (14). The standard F test is used to carry out the testing with the F statistic = 858.93, which corresponds to a p-value of almost 0. Thus, we reject the hypothesis of no weekly cycle at an overwhelming significance. The significance of the weekly cycle is consistent with our previous findings on the sample ACF and PACF in Figure 3 and the plot of the weekly return rates in Figure 4.

We then carried out a goodness-of-fit test for the presence of the annual pattern in the log returns $Y_{t\delta}$. Based on Model (14), the null hypothesis is

$$H_0 : \gamma_a(i\delta) = k \quad \text{for all } i \text{ and an arbitrary constant } k$$

versus the alternative hypothesis

$$H_1 : \gamma_a(i\delta) \neq \gamma_a(j\delta) \quad \text{for some } i \neq j.$$

The test statistic is

$$(25) \quad Q_n^a = L_a^{-1} \sum_{j=1}^{L_a} \{\tilde{\gamma}_a(j\delta) - \bar{\tilde{\gamma}}_a\}^2,$$

where $\bar{\tilde{\gamma}}_a = \sum_{j=1}^{L_a} \tilde{\gamma}_a(j\delta)$ and $\tilde{\gamma}_a(j\delta)$ are the kernel smoothed estimates of the long cycle with the bandwidth being a pre-determined fixed bandwidth or the cross-validation bandwidth. The test statistic $Q_n^a = 2.3975 \times 10^{-5}$.

The wild bootstrap method (Härdle and Mammen, 1993) was used to approximate the distribution of Q_n^a , which had

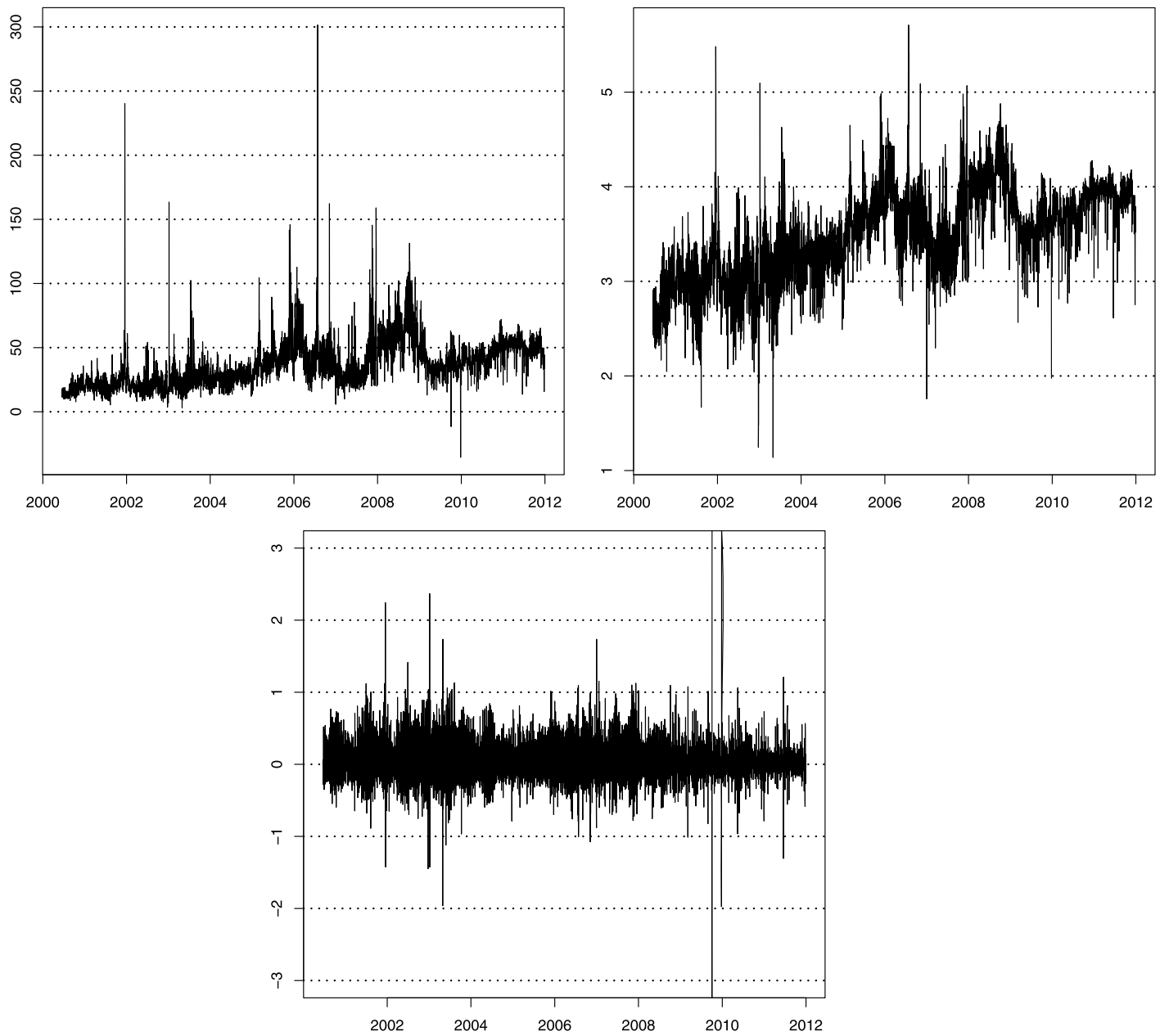


Figure 2. *Phelix Electricity Index (upper left), the log prices (upper right), and the log returns (lower) between July 7, 2000 and December 31, 2011.*

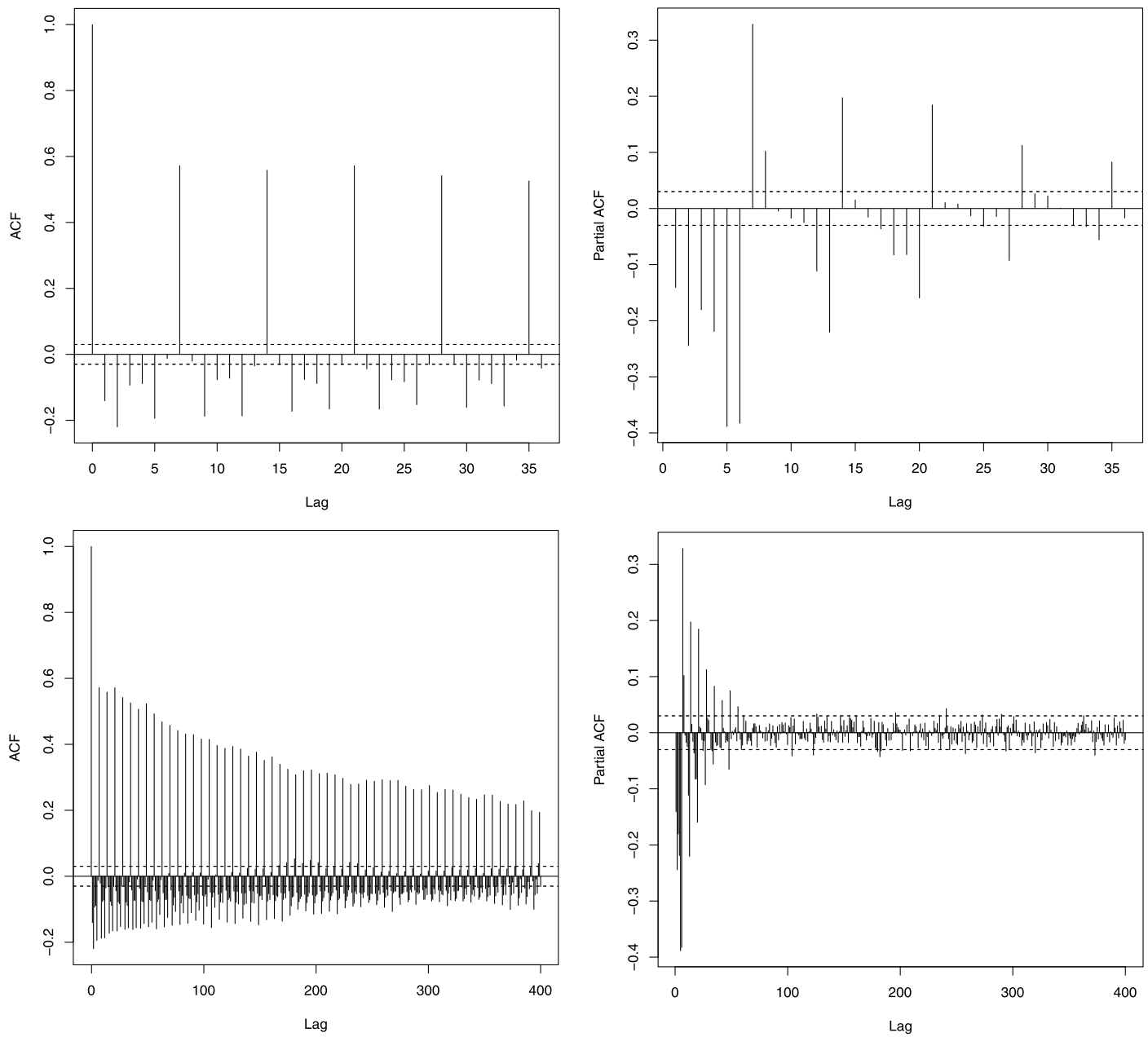


Figure 3. Sample auto-correlation (left panels) and partial autocorrelation (right panels) with lags up to 40 (upper panels) and lags up to 400 (lower panels).

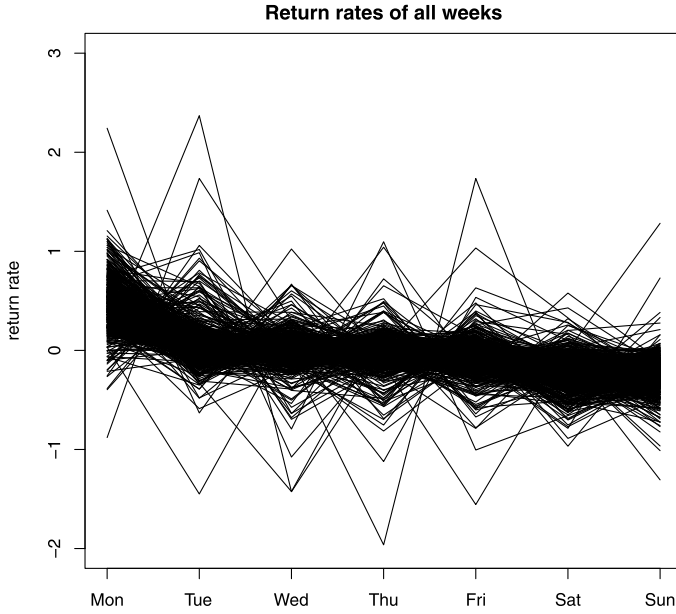


Figure 4. Weekly Return rates of the log electricity prices for all the 601 weeks.

the following steps in its implementation. We first obtained the estimated residuals

$$(26) \quad \hat{\epsilon}_{t\delta} = Y_{t\delta} - \hat{S}_w(t_w\delta) - \tilde{\gamma}_a(t_a\delta).$$

Then we generated the wild bootstrap residuals $\epsilon_{t\delta}^*$ from a two point discrete distribution with the probability of $\epsilon_{t\delta}^* = (\sqrt{5}+1)\epsilon_{t\delta}/2$ being $(5-\sqrt{5})/10$ and the probability of $\epsilon_{t\delta}^* = (1-\sqrt{5})\epsilon_{t\delta}/2$ being $(5+\sqrt{5})/10$. The wild bootstrapped resample was

$$Y_{t\delta}^* = \tilde{\gamma}_a + \hat{S}_w(t_w\delta) + \epsilon_{t\delta}^*,$$

which respect the H_0 of no annual cycle. We repeated the above procedures independently 500 times to obtain 500 bootstrap statistics Q_n^{a*} . We found that all the wild bootstrap statistics Q_n^{a*} were bigger than Q_n^a , which indicated that there was insufficient evidence to reject the null hypothesis of no annual cycle. Our finding was that there was no annual cycle in the log return rates of the Phelix index series.

APPENDIX

Proof of Proposition 2.1. We first consider the bias. Note that

$$\begin{aligned} \text{Bias}\{\tilde{\gamma}(t_L\delta)\} &= E\{\tilde{\gamma}(t_L\delta)\} - \gamma(t_L\delta) \\ &= \frac{\sum_{j=1}^m K_h(j\delta)[\gamma\{(t_L+j)\delta\} + \gamma\{(t_L-j)\delta\} - 2\gamma(t_L\delta)]}{K_h(0) + 2\sum_{j=1}^m K_h(j\delta)}. \end{aligned}$$

By Taylor expansion, we have

$$\begin{aligned} \gamma\{(t_L+j)\delta\} &= \gamma(t_L\delta) + j\delta\gamma'(t_L\delta) \\ &\quad + \frac{j^2}{2}\delta^2\gamma''(t_L\delta) + \frac{j^3}{6}\delta^3\gamma'''(\eta_1), \\ \gamma\{(t_L-j)\delta\} &= \gamma(t_L\delta) - j\delta\gamma'(t_L\delta) \\ &\quad + \frac{j^2}{2}\delta^2\gamma''(t_L\delta) - \frac{j^3}{6}\delta^3\gamma'''(\eta_2), \end{aligned}$$

with $\eta_1 \in [t_L\delta, (t_L+j)\delta]$ and $\eta_2 \in [(t_L-j)\delta, t_L\delta]$. Then

$$\begin{aligned} &\gamma\{(t_L+j)\delta\} + \gamma\{(t_L-j)\delta\} - 2\gamma(t_L\delta) \\ &= j^2\delta^2\gamma''(t_L\delta) + \frac{j^3}{6}\delta^3\{\gamma'''(\eta_1) - \gamma'''(\eta_2)\} \\ &\approx j^2\delta^2\gamma''(t_L\delta). \end{aligned}$$

So

$$\text{Bias}\{\tilde{\gamma}(t_L\delta)\} = \frac{\sum_{j=1}^m K_h(j\delta)j^2\delta^2}{K_h(0) + 2\sum_{j=1}^m K_h(j\delta)}\gamma''(t_L\delta),$$

as stated in the first part of the proposition.

For the variance,

$$\begin{aligned} &\text{Var}\{\tilde{\gamma}(t_L\delta)\} \\ &= \{J_n(\sum_{j=-m}^m K_h(j\delta))^2\}^{-1} \\ &\quad \sum_{-m \leq i, j \leq m} K_h(i\delta)K_h(j\delta)\nu\{(t_L+i)\delta, (t_L+j)\delta\} \\ &= \{J_n(\sum_{j=-m}^m K_h(j\delta))^2\}^{-1} \\ &\quad [\sum_{i=-m}^m K_h^2(i\delta)\nu\{(t_L+i)\delta, (t_L+i)\delta\} \\ &\quad + 2\sum_{k=1}^{2m} \sum_{i=-m}^{m-k} K_h(i\delta)K_h((i+k)\delta) \\ &\quad \nu\{(t_L+i)\delta, (t_L+i+k)\delta\}]. \end{aligned}$$

This completes the proof of Proposition 2.1. $\square$

Proof of Proposition 2.2. On the bias, we note that

$$\begin{aligned} \text{Bias}\{\tilde{\gamma}(t_L\delta)\} &= \frac{\sum_{j=1}^m K_h(j\delta)j^2\delta^2}{K_h(0) + 2\sum_{j=1}^m K_h(j\delta)}\gamma''(t_L\delta) \\ &= \frac{1}{2} \frac{h^2 \frac{\delta}{h} \sum_{j=-m}^m K(\frac{j\delta}{h})(\frac{j\delta}{h})^2}{\frac{\delta}{h} \sum_{j=-m}^m K(\frac{j\delta}{h})} \gamma''(t_L\delta) \\ &\approx \frac{1}{2} h^2 \frac{\int_{-m\delta/h}^{m\delta/h} z^2 K(z) dz}{\int_{-m\delta/h}^{m\delta/h} K(z) dz} \gamma''(t_L\delta). \end{aligned}$$

As $[-m\delta/h, m\delta/h]$ covers the support of K as $h \rightarrow 0$,

$$\int_{-m\delta/h}^{m\delta/h} K(z) dz = 1, \quad \int_{-m\delta/h}^{m\delta/h} z^2 K(z) dz = \sigma_K^2.$$

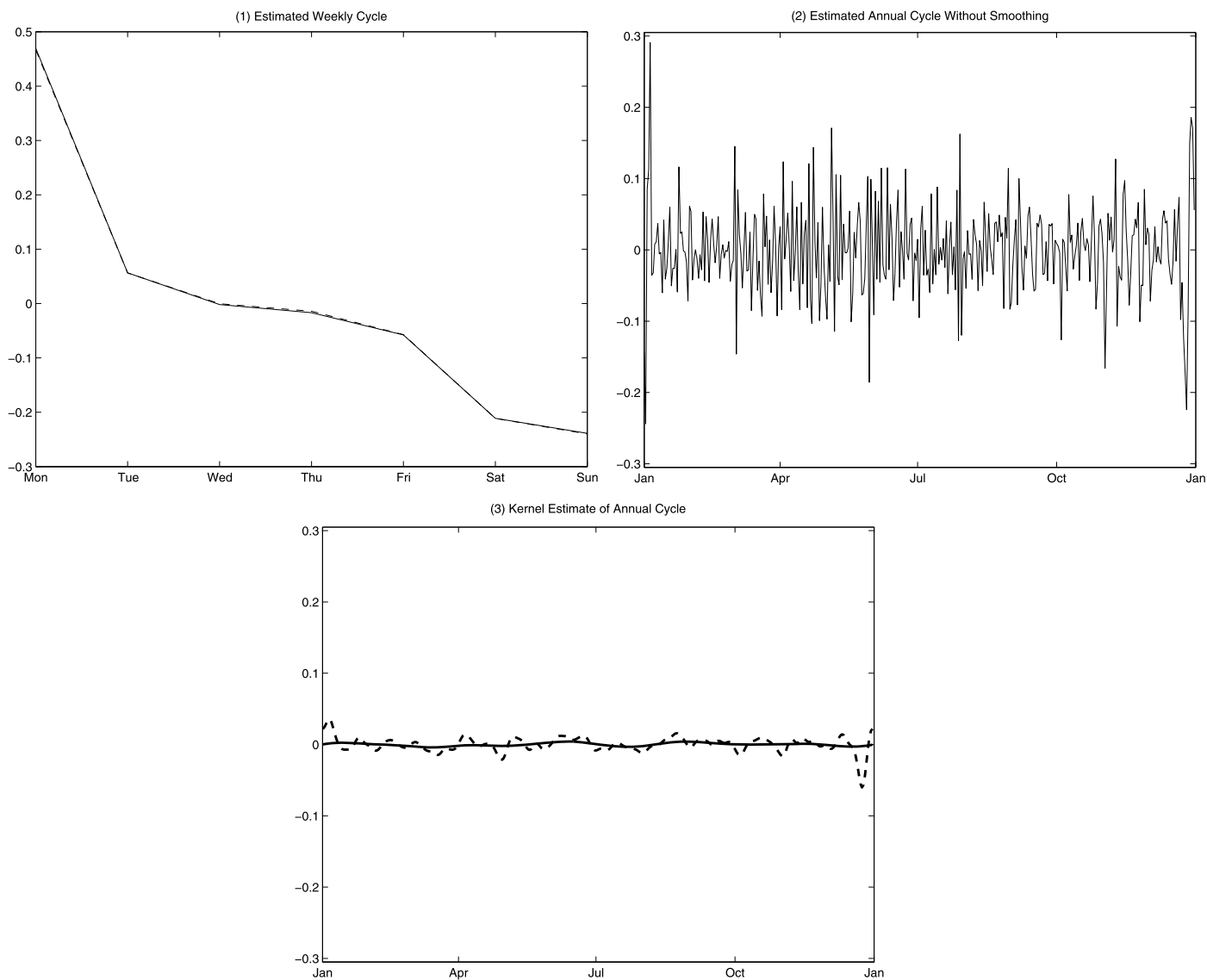


Figure 5. Estimated weekly cycles (upper left panel) and the estimated annual patterns by seasonal-dummy variable approach (upper right panel) and by the kernel estimator (bottom panel) for the electricity price data. In the top panel, the estimated weekly pattern by seasonal-dummy variable approach estimator (solid line) and ignore-long-cycle-regression estimator (dashed line) are nearly identical. In the bottom panel, the lines represent the kernel estimator with the bandwidth of 3 days (dashed line) and 15 days (solid line).

Hence, approximately $Bias\{\tilde{\gamma}(t_L\delta)\} \approx \frac{1}{2}h^2\sigma_K^2\gamma''(t_L\delta)$.

Under Condition C.2,

$$\left\{ \sum_{j=-m}^m K_h(j\delta) \right\}^2 \approx \delta^{-2} \left[\int_{-\frac{m\delta}{h}}^{\frac{m\delta}{h}} K(t)dt \right]^2 \approx \delta^{-2}$$

and

$$\begin{aligned} & \sum_{j=-m}^m K_h^2(j\delta) \nu\{(t_L + j)\delta, (t_L + j)\delta\} \\ &= \frac{1}{h\delta} \frac{\delta}{h} \sum_{j=-m}^m K^2\left(\frac{j\delta}{h}\right) \nu(t_L\delta + h\frac{j\delta}{h}, t_L\delta + h\frac{j\delta}{h}) \\ &\approx \frac{1}{h\delta} \int_{-\frac{m\delta}{h}}^{\frac{m\delta}{h}} K^2(t) \nu(t_L\delta + ht, t_L\delta + ht) dt \\ &\approx \frac{1}{h\delta} \left\{ \nu(t_L\delta, t_L\delta) \int_{-\infty}^{\infty} K^2(t) dt + O(h^2) \right\}. \end{aligned}$$

By applying the standard technique in the analysis of kernel smoothing,

$$\begin{aligned} & \sum_{j=-m}^{m-k} K_h(j\delta) K_h((j+k)\delta) \nu((t_L + j)\delta, (t_L + j+k)\delta) \\ &= \frac{1}{h\delta} \sum_{j=-m}^{m-k} \frac{\delta}{h} K\left(\frac{j\delta}{h}\right) K\left(\frac{k\delta}{h} + \frac{j\delta}{h}\right) \\ & \quad \times \nu(t_L\delta + h\frac{j\delta}{h}, t_L\delta + k\delta + h\frac{j\delta}{h}) \\ &= \frac{1}{h\delta} \int_{-\frac{m\delta}{h}}^{\frac{(m-k)\delta}{h}} K(t) K\left(k\frac{\delta}{h} + t\right) \\ & \quad \times \nu(t_L\delta + ht, t_L\delta + k\delta + ht) dt \\ &\approx \frac{1}{h\delta} \int_{-\frac{m\delta}{h}}^{\frac{(m-k)\delta}{h}} K(t) K\left(k\frac{\delta}{h} + t\right) \\ & \quad [\nu(t_L\delta, t_L\delta + k\delta) + h t \nu'_1(\epsilon_1, t_L\delta + k\delta) \\ & \quad + h t \nu'_2(t_L\delta, \epsilon_2)] dt \\ &\leq \frac{1}{h\delta} \int_{-\infty}^{\infty} K(t) K\left(k\frac{\delta}{h} + t\right) \\ & \quad [\nu(t_L\delta, t_L\delta + k\delta) + h t \nu'_1(\epsilon_1, t_L\delta + k\delta) \\ & \quad + h t \nu'_2(t_L\delta + \epsilon_2)] dt \\ &\approx \frac{1}{h\delta} \nu(t_L\delta, t_L\delta + k\delta) \int_{-\infty}^{\infty} K(t) K\left(k\frac{\delta}{h} + t\right) dt \end{aligned}$$

Then, we have the approximate variance for $\tilde{\gamma}(t_L\delta)$. $\square$

Received 13 September 2013

REFERENCES

- BENTH, F. E., BENTH, J. S. and KOEKEBAKKER, S. (2008). *Stochastic Modelling of Electricity and Related Markets*, World Scientific Publishing Co, Singapore. [MR2416686](#)
- BENTH, F. E., KIESEL, R. and NAZAROVA, A. (2012). A critical empirical study of three electricity spot price models. *Energy Economics*, online version.
- BOSQ, D. (1998). *Nonparametric Statistics For Stochastic Processes, Estimation and Prediction*, 2nd ed., Springer-Verlag, New York, USA. [MR1640691](#)
- DE LIVERA, A. M., HYNDMAN, R. J. and SNYDER, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, **106**, 1513–1527. [MR2896853](#)
- FAN, J. and GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications*, Chapman and Hall, London, UK. [MR1383587](#)
- GREENE, W. H. (2012). *Econometric Analysis*, 7th ed., Prentice Hall.
- HALL, P. (2008). Nonparametric methods for estimating periodic functions, with applications in Astronomy. *COMSTAT 2008*, Part I, 3–18. [MR2509576](#)
- HALL, P., REIMANN, J. and RICE, H. (2000). Nonparametric estimation of a periodic function. *Biometrika*, **87**, 545–557. [MR1789808](#)
- HAMILTON, J. D. (1994). *Time Series Analysis*, Princeton University Press, New Jersey, USA. [MR1278033](#)
- HÄRDLE, W. (1990). *Applied Nonparametric Regression*, Econometric Society Monographs 19, Cambridge University Press, Melbourne, Australia. [MR1161622](#)
- HÄRDLE, W. and MAMMEN, E. (1993). Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, **21**, 1926–1947. [MR1245774](#)
- HARVEY, A. (1993). *Time Series Models*, 2nd ed., MIT Press, Cambridge, Massachusetts, USA. [MR1230848](#)
- HARVEY, A. and KOOPMAN, S. J. (1993). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association*, **88**, 1228–1236.
- HARVEY, A., KOOPMAN, S. J. and RIANI, M. (1997). The modeling and seasonal adjustment of weekly observations. *Journal of Business and Economic Statistics*, **15**, 354–368.
- NADARAYA, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, **9**, 141–142.
- PEDREGAL, D. J. and YOUNG, P. C. (2006). Modulated cycles, an approach to modelling periodic components from rapidly sampled data. *International Journal of Forecasting*, **22**, 181–194.
- REBENNACK, S., PARDALOS, P. M., PEREIRA, M. V. F. and ILIADIS, N. A. (2010). *Handbook of Power Systems II*, Springer-Verlag, New York, USA.
- WATSON, G. S. (1964). Smooth regression analysis. *Sankhy: The Indian Journal of Statistics, Series A*, **26**, 359–372. [MR0185765](#)
- Song Xi Chen
Department of Statistics
Iowa State University
Ames, IA, 50010
USA
Guanghua School of Management and
Center for Statistical Science
Peking University, Beijing, 100871
China
E-mail address: songchen@iastate.edu
- Zheng Xu
Department of Biostatistics and
Department of Genetics
University of North Carolina at Chapel Hill
Chapel Hill, NC, 27599
USA
E-mail address: xuzheng@email.unc.edu