Seasonal long memory in intraday volatility and trading volume of Dow Jones stocks

Michelle Voges, Christian Leschinski and Philipp Sibbertsen,

Institute of Statistics, Faculty of Economics and Management,
Leibniz University Hannover, D-30167 Hannover, Germany

This version: June 28, 2017

Abstract

It is well known that intraday volatilities and trading volumes exhibit strong seasonal features. These seasonalities are usually modeled using dummy variables or deterministic functions. Here, we propose a test for seasonal long memory with a known frequency. Using this test, we show that deterministic seasonality is an accurate model for the DJIA index but not for the component stocks. These still exhibit significant and persistent periodicity after seasonal de-meaning so that more evolved seasonal long memory models are required to model their behavior.

JEL-Numbers: C12, C22, C58, G12, G15
Keywords: Intraday Volatility · Trading Volume · Seasonality · Long Memory

---

1 Introduction

The increasing availability of high frequency data poses new challenges for the analysis of seasonality in time series. This is because with the increasing frequency of observations, our datasets contain a higher number of meaningful harmonic oscillations. Harmonic oscillations are those whose period lengths are a multiple of the period with which the observations are sampled. For five-minute returns in US stock markets for example, there are 78 five-minute returns in a trading day. Furthermore, there are 5 trading days per week so that a weekly cycle would have a period of 5×78. Furthermore, there are about 21 trading days per month such that the period of a monthly cycle would be 21×78. In contrast to that, for monthly data, we can at most have a yearly cycle with period 12.

This new prevalence of seasonality requires a careful re-assessment of previous assumptions and practices, especially since the explosion in sample sizes that comes with the availability of high frequency data enables us to specify and estimate models that allow for more complex dynamics.

One area where this issue is particularly important is the intraday dynamics of volatility and trading volume in financial markets. For exchange rates this is documented by Baillie and Bollerslev (1991), Andersen and Bollerslev (1998) and Andersen et al. (2001), among others. Examples from the literature on stock returns include Andersen and Bollerslev (1997), Giot (2005) and Rossi and Fantazzini (2014). The major part of this literature assumes that the seasonality is deterministic. It is common practice to remove deterministic seasonality with seasonal dummy regression or by fitting trigonometric functions. Afterwards, the residuals are assumed to be free from seasonality.

However, in recent years parametric seasonal long memory models for intraday volatility have been proposed by Bordignon et al. (2007) and Rossi and Fantazzini (2014). These treat the seasonal effects as a stochastic process. In contrast to deterministic cycles stochastic seasonal components might change over time. This is why it is important to carefully examine the nature of seasonality.

In the context of seasonally integrated processes, the effect of seasonal demeaning has been studied by Abeysinghe (1991), Abeysinghe (1994), Franses et al. (1995) and da Silva Lopes (1999). They point out that regressing (first-differenced) time series on seasonal dummies might produce spuriously high R^2 if the seasonality originates from a seasonal unit root and refer to this phenomenon as spurious deterministic seasonality. Similar to the seasonal unit root case, ignoring seasonal or periodic long memory, of course, results in misspecified models.

In this paper, we therefore investigate the question whether the seasonality in intraday trading volume and realized volatility is accurately modeled by deterministic dummies

or whether it exhibits seasonal long memory. To do so, we propose a modified version of the G-test that was developed by Leschinski and Sibbertsen (2014) in the context of model specification in a GARMA framework. Here, we take a different perspective, since we are interested in testing rather then model specification. We therefore suggest a semiparametric test for seasonal long memory with a specific periodicity that is robust to the presence of short memory effects. The aim of the test is to assess the presence of seasonal long memory without specifying a model. Hence, we do not assume any parametric structure, e.g. that the seasonal long memory in the series is generated by a GARMA process.

The analysis is carried out for the example of the stocks that are components of the Dow Jones Industrial Average (DJIA) index and the index itself. We find for both the trading volume and the volatility that the majority of the components of the DJIA exhibits seasonal long memory, whereas the index does not. This shows that it is necessary to carefully analyze the nature of the seasonality in the series at hand, before specifying a parametric model.

The following Section 2 summarizes and discusses the existing seasonal long memory models. Afterwards, in Section 3, we introduce our testing procedure. Its finite sample performance is analyzed with help of a Monte Carlo simulation in Section 4. The results of the empirical analysis are presented in Section 5, before Section 6 concludes.

2 Modeling seasonality

Irrespective of its nature, seasonality can be defined as systematic but not necessarily regular behavior that is characterized by spectral peaks at seasonal frequencies and their harmonics (cf. Hylleberg (1992)). Regular periodic patterns can be classified as deterministic seasonality that cause (bounded) peaks in the periodogram. Stochastic cyclical behavior, in contrast, leads to spectral peaks that can be unbounded in the case of seasonal long memory. In general, seasonal data is said to have a period $S \in \mathbb{N}$ that is inferred from the sampling frequency, i.e it gives the number of observations per unit of time (e.g. year or day) so that data which are S observations apart are similar (Box et al. (2013)). In a perfectly deterministic cycle this implies $x_t = x_{t-S}$, i.e. deterministic periodic patterns repeat themselves exactly and can be perfectly forecasted.

It is common to seasonally adjust data by assuming deterministic cyclical patterns and removing them. For example in macroeconomics, a common procedure for seasonal adjustment is the $X-11$ ARIMA and its more advanced versions of the U.S. Census Bureau that uses moving averages in order to adjust trends and seasonal behavior in data such as inflation or unemployment rates.

Another basic procedure that is applied in a wide range of scenarios is to use seasonal

dummies in a linear regression framework

$$X_t = \beta_0 + \sum_{s=1}^{S-1} \beta_s D_{s,t} + Z_t, \tag{1}$$

where $D_{s,t}$ are indicator variables that take the value 1 for $t = s + S(q-1)$ with $q = 1, ..., \lfloor T/S \rfloor$, where $\lfloor \cdot \rfloor$ denotes the greatest integer smaller than the argument and Z_t are the regression residuals. The dummies account for cycles like calender effects (e.g. day of the week) or intradaily cycles and in a purely deterministic framework the regression residuals Z_t are assumed to be free from seasonality.

Slowly varying seasonality is captured with the flexible Fourier form which is a linear combination of sines and cosines. Here, the original data is regressed on sines and cosines that depend on the seasonal frequencies $\omega_s = \frac{2\pi s}{S}$, such that

$$X_t = \sum_{s=1}^{\lfloor S/2 \rfloor} (a_s \cos(\omega_s t) + b_s \sin(\omega_s t)) + Z_t,$$

(cf. Andersen and Bollerslev (1997), Andersen and Bollerslev (1998), Martens et al. (2002), Deo et al. (2006)). It is also conceivable to fit a combination of slowly varying seasonality and basic seasonal dummies. Hereby it is possible to model intraday cycles with more flexibility and account for announcements through dummies at the same time (Andersen and Bollerslev (1998)).

Although these deterministic models are still popular, they do not always provide a suitable model fit. Therefore, stochastic time series models are proposed. The basic model is a seasonal version of the popular ARIMA model. Seasonal fractionally differenced models (SARFIMA) were introduced by Porter-Hudak (1990) and generalized and extended by Ray (1993). A model of order $(p, d_0, q) \times (P, d_S, Q)$ is given by,

$$\phi(L)\Phi(L^S)(1-L)^{d_0}\left(1-L^S\right)^{d_S} X_t = \theta(L)\Theta(L^S)\varepsilon_t, \tag{2}$$

where L is the lag-operator defined as $LX_t = X_{t-1}$, $|d_S|, |d_0| < \frac{1}{2}$ are the seasonal and non-seasonal fractional orders of integration and ε_t is defined as white noise for the rest of the chapter. As in the deterministic setting, the parameter S determines the period length of the seasonal cycle and the seasonal frequencies are given by $\omega_s = \frac{2\pi s}{S}$ for $s = 1, ..., \lfloor S/2 \rfloor$. The seasonal fractional difference operator is defined in analogy to its non-seasonal counterpart by a binomial expansion such that $\left(1-L^S\right)^{d_S} = \sum_{k=0}^{\infty} \binom{d_S}{k} (-L^S)^k$. Furthermore, $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - ... - \theta_q L^q$ are polynomials of degree p and q in the backshift operator L. The polynomials $\Phi(L^S)$ of degree P and $\Theta(L^S)$ of degree Q are defined analogously and they describe the seasonal short

run dynamics. Hence, SARFIMA models offer a certain flexibility, but by construction such series share the same memory parameter d_S at all seasonal frequencies ω_s. This can be relaxed if several periods S_j are allowed. The flexible ARFISMA model - an alternative but similar version - was introduced by Hassler (1994).

The Gegenbauer ARMA (GARMA) process was introduced by Gray et al. (1989) and it generates one spectral peak at one specific frequency. Woodward et al. (1998) and Giraitis and Leipus (1995) generalize this model and allow for several spectral peaks by introducing the k-factor GARMA process

$$\phi(L)X_t = \Pi_{j=1}^{k}\left(1-2\cos\omega_j L+L^2\right)^{-d_j}\theta(L)\varepsilon_t, \tag{3}$$

where $\left(1-2u_jL+L^2\right)^{-d_j}$ with $u_j=\cos\omega_j$ is the generating function of the Gegenbauer polynomial defined as $\sum_{T=0}^{\infty}C_T^{(d_j)}(u_j)L^T$ with $C_T^{(d_j)}(u_j)=\sum_{k=0}^{\lfloor T/2\rfloor}\frac{(-1)^k(2u_j)^{T-2k}\Gamma(d_j-k+T)}{k!(T-2k)!\Gamma(d_j)}$ and $\omega_j\in[0,\pi]$ is some cyclical frequency. The cyclical frequencies ω_j are not necessarily equal to seasonal frequencies ω_s as in the SARFIMA model and each ω_j has an individual memory parameter d_j so that there may be peaks of different magnitude in the spectrum. This translates, for example, to a two-day cycle with memory parameter d_1 and a daily cycle with memory parameter d_2 ($d_1\neq d_2$). Due to the Gegenbauer polynomial the requirements on the memory parameter(s) for stationarity and invertibility depend on ω_j: for $0<\omega_j<\pi$ this is $|d_j|\in\frac{1}{2}$, whereas for $\lambda=\{0,\pi\}$, $|d_j|\in\frac{1}{4}$ is required. Note that the model is not fractionally integrated in a narrow sense, because integration is related to the fractional differencing operator $(1-L)^d$, but a GARMA model is constructed with a Gegenbauer filter instead. However, the Gegenbauer filter $\left(1-2\cos\omega L+L^2\right)^d$ is equal to $(1-L)^{2d}$ for $\omega=0$ and $(1+L)^{2d}$ for $\omega=\pi$ so that the squared fractional differencing operator is a special case. This is also the reason for frequency-depending stationarity and invertibility requirements discussed above.

Both, the k-factor Gegenbauer model and the rigid SARFIMA model generate spectral poles at one or more frequencies ω that are of the form

$$f(\omega+\lambda)\sim C|\lambda|^{-2d_\omega}\quad\text{as}\quad\lambda\to 0, \tag{4}$$

where C is a positive and finite constant.

Since there is no reason why seasonality should be purely deterministic or purely stochastic, both sources can be considered in one model at the same time. For example Gil-Alana (2005) and Caporale et al. (2012), among others, construct a model with deterministic seasonal means and stochastic seasonal long memory captured with a SARFIMA model, i.e. the regression residuals in (1) are replaced with some version of (2).

Gil-Alana (2005) also suggests a test for seasonality and the seasonal order of integration

d_S in this combined parametric setting. However, this approach has the drawback that the performance of the test depends on the correct specification of the model. In contrast, our test introduced in the next Section 3 circumvents this problem with its semiparametric approach.

3 Testing for Seasonal Long Memory

A seasonal long memory process X_t with period S and seasonal memory parameter d_ω has a pole at frequency $\omega = 2\pi/S$. In its neighborhood, the spectral density is thus given by (4). We are interested in testing the hypothesis that the process does not have seasonal long memory versus the alternative that it has. Thus, our hypotheses are given by

$$H_0 : \ d_\omega = 0 \qquad \text{and} \qquad H_1 : \ d_\omega > 0.$$

Of course seasonal behavior could also be induced by deterministic patterns. We therefore consider the seasonally demeaned series Z_t from (1). It was shown by Ooms and Hassler (1997) that seasonal de-meaning introduces zeros in the periodogram at all Fourier frequencies that coincide with seasonal frequencies, so that $j'T = jS$, where $j' = 1, ..., \lfloor S/2 \rfloor$. However, all other periodogram ordinates remain unaffected (cf. also Arteche (2002)).

Define the periodogram of Z_t by

$$I(\lambda) = (2\pi T)^{-1} \left| \sum_{t=1}^{T} Z_t e^{-\mathbf{i}\lambda t} \right|^2, \quad \text{with} \quad \lambda \in [-\pi, \pi].$$

The periodogram is usually computed at the Fourier frequencies $\lambda_j = 2\pi j/T$, for $j = 1, ..., n$ and $n = \lfloor T/2 \rfloor$.

A semiparametric test for seasonal long memory with period S is obtained by employing a modified version of the G^* test that was suggested by Leschinski and Sibbertsen (2014) in the context of model selection in GARMA models. Their procedure tests for seasonal long memory using the test statistic

$$G_Z^* = \max_j \left\{ \frac{I(\lambda_j)}{\hat{f}_Z(\lambda_j)} \right\} - \log n,$$

where $\hat{f}_Z(\lambda)$ is a consistent estimate of the spectral density under the null hypothesis. To adapt this to our setting, i.e. a simple test for seasonal long memory with a known

period S, we construct a local version of G_Z^*, that only considers m Fourier frequencies to the left and the right of the frequency of interest ω. The bandwidth m has to satisfy the usual condition $1/m + m/T \to 0$ as $T \to \infty$. Thus, the test statistic is given by

$$G = \max_{j \in [0,m]} \left\{ \frac{I(\omega \pm \lambda_j)}{\hat{f}_Z(\omega \pm \lambda_j)} \right\} - \log(2m). \tag{5}$$

For the implementation of this test statistic we require an estimate $\hat{f}_Z(\lambda)$ of the spectral density that is consistent under the null hypothesis. This is usually done with kernel-smoothed versions of the periodogram. However, this has the disadvantage that a single large periodogram ordinate $I(\lambda_j)$ has a significant impact on the spectral density estimate in the neighborhood of λ_j. To avoid this effect, Leschinski and Sibbertsen (2014) adopt the logspline spectral density estimate originally proposed by Cogburn et al. (1974), who showed that this estimator is asymptotically equivalent to a kernel spectral density estimate. A maximum likelihood version of this estimator based on regression splines was proposed by Kooperberg et al. (1995).

Following their notation, define $A_h = [(h-1)\pi/H_T, h\pi/H_T)$ as a subinterval of $[0,\pi]$, for $1 \leq h < H_T$ and set $A_{H_T} = \pi$, where the number (H_T) of subintervals (A_h) is determined according to $H_T = \lfloor 1 + T^c \rfloor$, with $0 < c < 1/2$. Furthermore, let g denote a cubic spline function defined on $[0,\pi]$. That means it is a polynomial of degree three on each subinterval A_h, it is two times continuously differentiable on $[0,\pi]$, and it can be expressed in terms of Basis-splines as

$$g(\lambda,\beta) = \beta_1 B_1(\lambda) + ... + \beta_W B_W(\lambda),$$

where the B_w denote the basis functions, with $1 \leq w \leq 4H_T - 3(H_T - 1) = W$.

The basis for the application of splines in spectral density estimation is the observation that the normalized periodogram ordinates $I(\lambda_j)/f_Z(\lambda_j) = Q_j$ are approximately exponentially distributed with mean one. For the logarithm of $I(\lambda_j)$ follows that $\log I(\lambda_j) = \varphi(\lambda_j) + q_j$, where q_j is the log of the exponential variable Q_j and $\varphi(\lambda_j)$ is the log-spectral density. This linearization allows to apply a spline function $g(\lambda,\beta)$ to estimate $\varphi(\lambda)$. The spectral density estimate $\hat{f}_Z(\lambda) = \exp(g(\lambda,\hat{\beta}))$ is then obtained after reversing the log-transformation.

To estimate $\beta = (\beta_1, ..., \beta_W)'$, we apply the approach suggested by Leschinski and Sibbertsen (2014) and estimate the spectral density $f_Z(\lambda)$ via the OLS estimator

$$\hat{\beta}_{OLS} = \arg\min_{\beta} \sum_{j=-n}^{n} \left[\log(I(\lambda_j)) + \eta - g(\lambda_j,\beta) \right]^2, \tag{6}$$

where η denotes the Euler-Mascheroni constant. Since this OLS estimator has a closed form solution, this approach does not require numerical optimization and is much faster to compute than the ML estimator of Kooperberg et al. (1995).
In direct analogy to Leschinski and Sibbertsen (2014), we then have for the test statistic in (5) that for any $\tilde{z} \in \mathbb{R}$

$$\lim_{T\to\infty} P(G > \tilde{z}) = 1 - \exp(-\exp(-\tilde{z}))$$

under the null hypothesis and under the alternative

$$\lim_{T\to\infty} P(G < \tilde{c}) = 0$$

for any $\tilde{c} < \infty$.
Note that our test is constructed such that it detects seasonal long memory while being robust to the potential presence of seasonal short-run dynamics. Thus, non-rejection of the null hypothesis does not imply the absence of stochastic seasonality.
The test is applied at the first seasonal frequency $\omega = 2\pi/S$. In theory one could also consider the harmonics $\omega_s = 2\pi s/S$, for $s = 2, ..., \lfloor S/2 \rfloor$. However, the periodogram ordinates are proportional to the fraction of the process variance that can be explained by a sinusoidal cycle of the respective frequency. Since by the nature of the Fourier series, cycles with frequencies that are a multiple of this frequency improve the approximation to a non-sinusoidal seasonality, these will not have an effect if the seasonality is indeed of a sinusoidal form. Without additional knowledge about the seasonal properties of the process at hand, it is therefore not clear whether the inclusion of harmonic frequencies will improve or diminish the performance of the test.

4 Monte Carlo

In order to analyze the finite sample performance of our G-test in (5) we conduct a Monte Carlo experiment with 1,000 replications. We use a 1-factor GARMA model as data generating process

$$X_t = \left(1 - 2\cos\left(\frac{2\pi}{13}\right) + L^2\right)^{-d} u_t, \tag{7}$$

where $u_t = \phi u_{t-1} + \varepsilon_t$ is an $AR(1)$-process. We consider $\phi = 0$ for the white noise case with zero mean and unit variance, and $\phi = 0.4$ in order to examine the influence of short-run dynamics. Based on our empirical application in Section 5 we choose $\omega = \frac{2\pi}{13}$, which implies a period of $S = 13$, and we consider memory parameters in the stationary area, $d = \{0,\ 0.1,\ 0.2,\ 0.3,\ 0.4\}$. For $d = 0$ the process is non-seasonal so that the corresponding simulation results display empirical sizes. Our G-test requires two bandwidth choices. First, for the periodogram $m = \lfloor 1 + T^{\delta} \rfloor$ is the number of Fourier frequencies that are considered. Second, in addition to m, a bandwidth that determines the number of knots in the spline-based estimation of the spectral density $H_T = \lfloor 1 + T^{c} \rfloor$ is necessary.

All results are displayed in Table 1. Size and power results both improve with increasing sample size T for all parameter constellations. The size results are already satisfying in small samples and very robust to bandwidth choices and short run dynamics.

The power results depend on the memory parameter d and improve with higher orders of integration. This is intuitive because low seasonal persistence, e.g. $d = 0.1$, is more difficult to detect than higher persistence. However, in large samples the power is already good for $d = 0.2$, and for $d = 0.4$ results are even satisfying in small samples. The bandwidth parameter c influences the performance such that lower values of c improve results significantly. For example, for $d = 0.2$ and $T = 2{,}500$ power is more than twice as high for the lowest value of c than for its highest value. For larger samples the choice of c becomes less important. We find a similar influence of δ. Smaller values lead to much higher power results for low seasonal persistence (about 20 percentage points), but for higher seasonal persistence the influence of δ shrinks. Short-run dynamics of medium size cause no systematic distortions in the results. On the contrary, there are only small and random deviations compared to the white noise case. The semiparametric approach of the test therefore successfully mitigates the impact of short run dynamics.

Overall, our G-test statistic has very good finite sample properties independent of short-run dynamics. One should only be careful with the choice of m and H_T because choosing the bandwidths too large might lead to biased parameter estimates which results in lower power of the test. We therefore recommend to choose low bandwidth parameters.

		c	0.1			0.15			0.2			0.25		
ϕ	d	T/δ	0.5	0.6	0.7	0.5	0.6	0.7	0.5	0.6	0.7	0.5	0.6	0.7
0	0	250	0.06	0.06	0.05	0.07	0.06	0.06	0.09	0.07	0.07	0.08	0.07	0.06
		500	0.08	0.07	0.06	0.07	0.07	0.05	0.06	0.06	0.07	0.07	0.08	0.06
		1000	0.05	0.04	0.05	0.06	0.06	0.06	0.05	0.06	0.04	0.08	0.07	0.05
		2500	0.05	0.05	0.05	0.05	0.06	0.06	0.04	0.06	0.04	0.06	0.06	0.05
	0.1	250	0.10	0.11	0.08	0.12	0.09	0.09	0.13	0.12	0.09	0.14	0.10	0.08
		500	0.17	0.13	0.09	0.14	0.11	0.09	0.16	0.11	0.10	0.13	0.11	0.09
		1000	0.24	0.16	0.12	0.22	0.16	0.11	0.22	0.16	0.14	0.14	0.10	0.08
		2500	0.35	0.24	0.18	0.35	0.26	0.16	0.31	0.21	0.15	0.15	0.10	0.09
	0.2	250	0.33	0.24	0.23	0.32	0.29	0.23	0.29	0.27	0.23	0.33	0.25	0.22
		500	0.48	0.41	0.30	0.51	0.40	0.35	0.49	0.42	0.36	0.41	0.36	0.27
		1000	0.68	0.56	0.54	0.71	0.60	0.50	0.66	0.59	0.50	0.48	0.38	0.34
		2500	0.92	0.85	0.78	0.91	0.82	0.73	0.84	0.76	0.64	0.43	0.36	0.28
	0.3	250	0.62	0.56	0.47	0.60	0.55	0.50	0.58	0.53	0.51	0.59	0.55	0.51
		500	0.83	0.78	0.71	0.84	0.79	0.73	0.81	0.78	0.70	0.70	0.69	0.63
		1000	0.97	0.93	0.91	0.96	0.94	0.92	0.96	0.94	0.88	0.82	0.76	0.71
		2500	1.00	1.00	0.99	1.00	1.00	0.99	1.00	0.99	0.98	0.83	0.76	0.72
	0.4	250	0.84	0.80	0.76	0.83	0.81	0.75	0.83	0.82	0.76	0.85	0.77	0.76
		500	0.97	0.96	0.94	0.97	0.96	0.94	0.97	0.95	0.93	0.95	0.91	0.89
		1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.96	0.95	0.94
		2500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	0.96	0.95
0.4	0	250	0.06	0.07	0.05	0.06	0.07	0.06	0.09	0.07	0.06	0.09	0.08	0.08
		500	0.07	0.07	0.04	0.06	0.06	0.05	0.07	0.06	0.07	0.06	0.08	0.06
		1000	0.05	0.06	0.05	0.04	0.05	0.04	0.06	0.05	0.04	0.06	0.07	0.06
		2500	0.05	0.05	0.05	0.05	0.07	0.07	0.06	0.06	0.06	0.05	0.07	0.06
	0.1	250	0.12	0.09	0.08	0.10	0.11	0.08	0.11	0.12	0.09	0.12	0.08	0.09
		500	0.14	0.11	0.08	0.15	0.12	0.10	0.16	0.12	0.09	0.15	0.11	0.09
		1000	0.20	0.18	0.11	0.22	0.13	0.12	0.21	0.17	0.11	0.14	0.09	0.09
		2500	0.33	0.24	0.17	0.33	0.24	0.15	0.26	0.21	0.15	0.14	0.11	0.08
	0.2	250	0.27	0.26	0.22	0.33	0.26	0.19	0.31	0.25	0.22	0.29	0.25	0.24
		500	0.46	0.38	0.33	0.51	0.42	0.34	0.47	0.39	0.33	0.38	0.32	0.28
		1000	0.66	0.58	0.48	0.67	0.60	0.50	0.65	0.61	0.50	0.43	0.37	0.29
		2500	0.89	0.84	0.75	0.92	0.85	0.74	0.83	0.71	0.66	0.45	0.33	0.29
	0.3	250	0.58	0.53	0.47	0.59	0.53	0.48	0.59	0.53	0.51	0.59	0.52	0.51
		500	0.82	0.78	0.73	0.83	0.78	0.74	0.81	0.75	0.70	0.71	0.67	0.64
		1000	0.96	0.91	0.90	0.97	0.94	0.91	0.95	0.93	0.89	0.81	0.76	0.70
		2500	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.98	0.83	0.78	0.68
	0.4	250	0.80	0.77	0.75	0.83	0.80	0.75	0.82	0.82	0.77	0.81	0.80	0.78
		500	0.97	0.96	0.94	0.97	0.96	0.93	0.97	0.95	0.94	0.93	0.91	0.87
		1000	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.99	0.97	0.96	0.93
		2500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	0.98	0.95

Table 1: Rejection frequency of our G-test statistic for a nominal significance level of $\alpha = 0.05$ and the DGP given in (7).

5 Empirical analysis of Dow Jones stocks

We analyze the intraday log-realized volatility and the log-trading volume of the 30 Dow Jones Industrial Average (DJIA) stocks and the index itself. To do so, we use five-minute data and aggregate them to half-hourly observations for the time span from January 2011 to December 2015 which makes about 16,300 observations for each stock. They are all traded on the New York Stock Exchange and on the Nasdaq stock market, during the trading hours from 09:30 to 16:00 local time. Consequently, we obtain 13 half-hourly observations per day so that our period is $S = 13$ for a daily cycle and the seasonal frequency of interest is $\omega = \frac{2\pi}{13} = 0.48$.

Let $r_{i,t}$ be the ith return in the tth 30-minute interval so that realized volatilities are calculated via $\hat{\sigma}_t^2 = \sum_{i=1}^{6} r_{i,t}^2$ for each half-hour interval $t = 1, ..., T$. For trading volume, we proceed analogously, i.e. we sum six five-minute observations up in order to obtain half-hourly volumes. Finally, we take the logarithm of both series. As customary, we add 0.001 to the volatility series before applying the logarithm to avoid infinite values in case of constant prices (cf. for example Lu and Perron (2010) and Xu and Perron (2014). The transformed series are approximately Gaussian in contrast to the original series that are usually right-skewed (cf. Andersen et al. (2003), among others) and they exhibit less outliers which is desirable from a data analytic point of view.

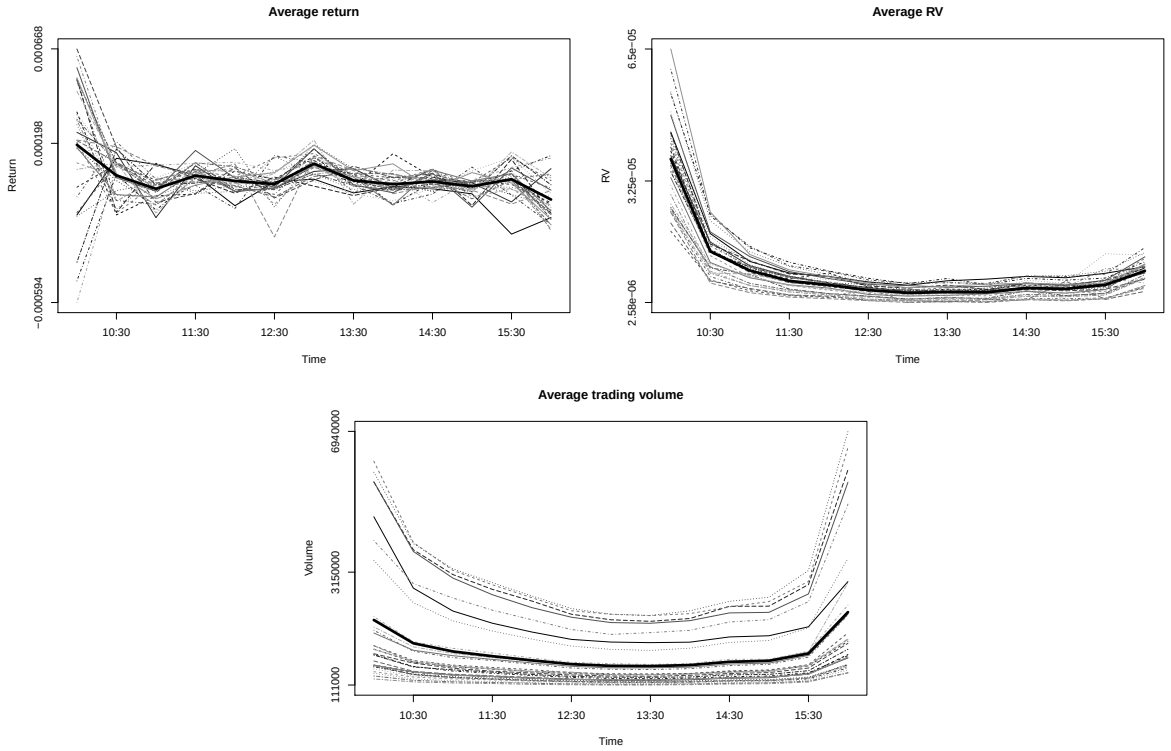

Figure 1: Intraday averaged return, realized volatility and trading volume for all stocks. The bold lines indicate the overall average.

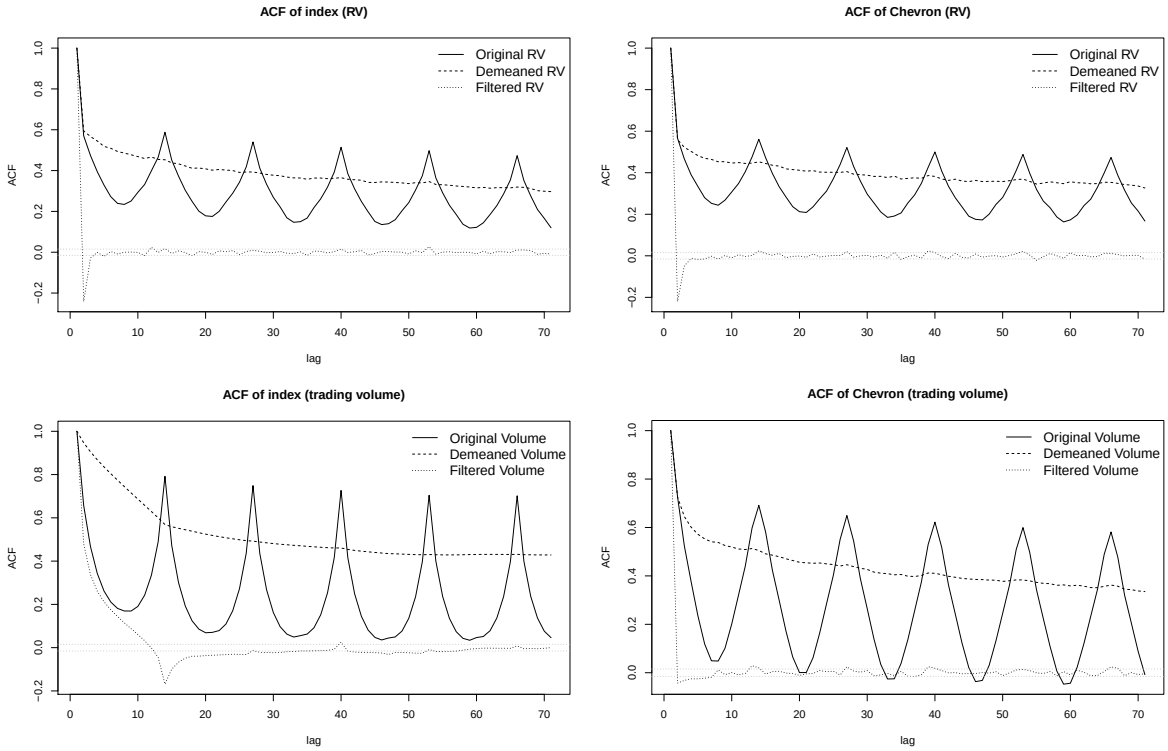

Figure 2: Autocorrelation functions for *logarithms of realized volatility* and *logarithms of trading volume* of the index (left panel) and Chevron (right panel).

As expected, the intraday average return is almost constant for all stocks and slightly positive, see Figure 1. Only overnight returns have a higher variance. We also find the typical time of the day phenomena. Market opening and closing times are characterized by higher market activity that leads to higher trading volume and volatility than at lunch time which causes a pronounced U-shape in intraday trading volume. For realized volatility we find an inverse J-shape instead of an U-shape. Hence, on average volatility does not increase much at the end of trading days. As discussed in the introduction, these results are well known in the literature, cf. Wood et al. (1985) for early evidence of these phenomena and Bordignon et al. (2008) and Bollerslev et al. (2016) for more recent analyses.

Since we are interested in seasonal patterns, we take a look at the autocorrelation function (ACF) (Figure 2) and the periodogram (Figures 3 and 4). The left side of all three figures displays the index, i.e. the average behavior, and the right side displays a single constituent of the index (Chevron). Clear cyclical behavior can be seen from the solid lines in all four graphs in Figure 2 which are the ACFs of original data. The same characteristic is also present in the periodograms in the upper panels of Figures 3 and 4 where we find peaks at seasonal frequencies. This observations holds equally for realized volatility and trading volume of both the index and Chrevron. In order to

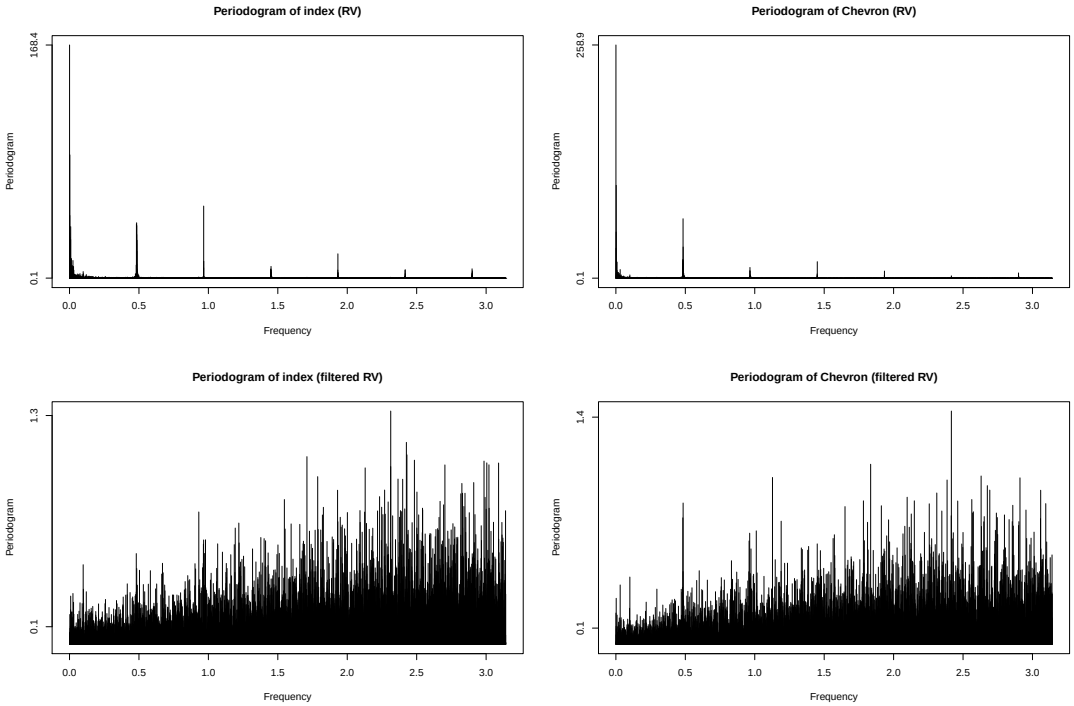

Figure 3: Periodogram of original data, and after dummy regression and filtering at zero frequency for *logarithms of realized volatility* of the index (left panel) and of Chevron (right panel).

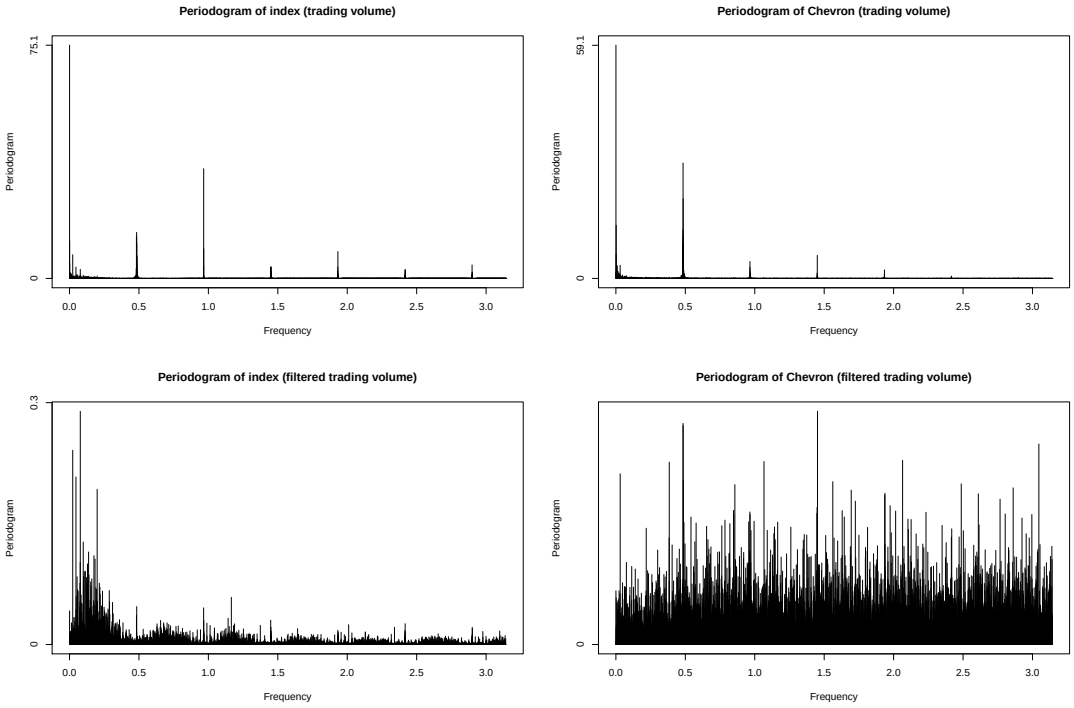

Figure 4: Periodogram of original data, and after dummy regression and filtering at zero frequency for *logarithms of trading volume* of the index (left panel) and of Chevron (right panel).

.DJI	Dow Jones Index	KO	Coca-Cola
AAPL.O	Apple	MCD	McDonald's
AXP	American Express	MMM	3M
BA	Boeing	MRK	Merck
CAT	Caterpillar	MSFT.O	Microsoft
CSCO.O	Cisco	NKE	Nike
CVX	Chevron	PFE	Pfizer
DD	E.I. du Pont de Nemours & Company	PG	Procter & Gamble
DIS	Disney	TRV	Travelers Companies Inc
GE	General Electric	UNH	UnitedHealth
GS	Goldman Sachs	UTX	United Technologies
HD	Home Depot	V	Visa
IBM	IBM	VZ	Verizon
INTC.O	Intel	WMT	Wal-Mart
JNJ	Johnson & Johnson	XOM	Exxon Mobil
JPM	JPMorgan Chase		

Table 2: List of rics and corresponding companies.

assess the influence of seasonal means we remove them with help of the dummy regression from (1). Again, the ACFs of all four seasonally demeaned series considered in Figure 2 show a similar behavior, that is the dashed lines decay slowly at a hyperbolic rate. The corresponding periodograms are omitted for reasons of space but they all have a clear singularity at zero frequency. Hence, both volatility and trading volume series exhibit long memory. Therefore, we estimate the order of integration with the local Whittle estimator (cf. Kuensch (1987), Robinson (1995)) or more specifically with the exact local Whittle estimator (Shimotsu et al. (2005)) because trading volume is known to be nonstationary in some cases. After that we filter the seasonally demeaned series with their respective memory estimates. The dotted lines in Figure 2 display the ACFs of these series. For the index there are only very few significant autocorrelations left. This is similar for Chevron, but when taking a closer look we find more significant autocorrelations especially at seasonal lags. The lower panels of Figures 3 and 4 show the periodograms of demeaned and differenced data. These are already close to being constant like the periodogram of a white noise process. However, one can conjecture small peaks at the seasonal frequencies, e.g. 0.48. This is more pronounced for Chevron than for the index - especially for the realized volatilities in Figure 3.

In order to examine the question whether the cyclical behavior in the data is already sufficiently accounted for by removing deterministic seasonality with seasonal dummies we apply our test for no seasonal long memory to the individual stock data and the index.[1]

[1]We find that the flexible Fourier form does not have a significant influence on the elimination of seasonality because seasonal dummies already eliminate all deterministic periodicity.

c		0.1			0.15			0.2	
RICS\ δ	0.5	0.6	0.7	0.5	0.6	0.7	0.5	0.6	0.7
.DJI	0.053	0.138	0.354	0.080	0.200	0.462	0.127	0.300	0.605
AAPL.O	**0.031**	0.063	0.066	**0.031**	0.071	0.117	**0.020**	0.055	0.153
AXP	**0.009**	**0.027**	0.074	**0.013**	**0.037**	0.098	**0.025**	0.063	0.156
BA	0.084	0.135	0.421	0.145	0.159	0.456	0.324	0.263	0.591
CAT	**0.002**	**0.009**	**0.027**	**0.005**	**0.018**	**0.050**	**0.008**	**0.019**	**0.048**
CSCO.O	0.310	0.324	0.579	0.339	0.325	0.589	0.400	0.335	0.633
CVX	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**	**0.003**	**0.001**	**0.004**	**0.009**
DD	**0.020**	0.062	0.222	**0.037**	0.070	0.239	0.068	0.087	0.250
DIS	**0.044**	0.143	0.113	0.081	0.224	0.233	0.139	0.314	0.206
GE	**0.047**	0.163	0.417	0.118	0.328	0.676	0.196	0.420	0.755
GS	0.059	0.178	0.395	0.085	0.227	0.481	0.113	0.262	0.493
HD	0.111	0.296	0.160	0.201	0.467	0.339	0.316	0.626	0.362
IBM	**0.008**	**0.034**	0.128	**0.014**	**0.046**	0.139	**0.026**	0.063	0.149
INTC.O	0.490	0.750	0.187	0.623	0.801	0.128	0.607	0.798	0.139
JNJ	**0.036**	0.123	0.094	0.078	0.220	0.147	0.195	0.424	0.150
JPM	0.077	0.206	0.064	0.116	0.287	0.105	0.109	0.259	0.113
KO	**0.000**	**0.001**	**0.005**	**0.001**	**0.003**	**0.008**	**0.001**	**0.003**	**0.009**
MCD	**0.025**	0.092	0.334	0.058	0.171	0.461	0.189	0.410	0.411
MMM	**0.045**	0.077	0.217	0.063	0.084	0.232	0.163	0.155	0.375
MRK	0.119	**0.025**	0.086	0.214	0.062	0.160	0.388	0.115	0.244
MSFT.O	0.104	0.263	0.606	0.116	0.284	0.619	0.162	0.372	0.708
NKE	0.085	0.246	0.548	0.107	0.280	0.592	0.189	0.416	0.756
PFE	**0.006**	**0.025**	0.132	**0.019**	0.063	0.214	**0.050**	0.119	0.261
PG	**0.035**	0.101	0.321	0.061	0.163	0.421	0.108	0.256	0.529
TRV	**0.003**	**0.012**	**0.049**	**0.008**	**0.024**	0.077	**0.017**	**0.043**	0.103
UNH	0.172	0.210	0.517	0.271	0.254	0.593	0.469	0.382	0.619
UTX	**0.021**	0.089	0.245	**0.047**	0.142	0.348	0.063	0.142	0.324
V	0.583	0.653	0.917	0.710	0.705	0.949	0.791	0.734	0.967
VZ	**0.002**	**0.004**	**0.004**	**0.003**	**0.009**	**0.012**	**0.006**	**0.016**	**0.012**
WMT	0.060	0.236	0.224	0.121	0.347	0.228	0.169	0.366	0.234
XOM	**0.000**	**0.001**	**0.005**	**0.001**	**0.003**	**0.010**	**0.002**	**0.004**	**0.010**

Table 3: p-values of the G-test at frequency $\omega = 2\pi/13$ for *logarithms of realized volatility*. The data is seasonally demeaned and long memory at frequency zero has been removed.

Note that the index can be interpreted as the average of individual stock data because it is calculated as the sum of single stock prices scaled with the so called Dow Divisor, i.e. the DJIA index is a weighted average of its components.

Tables 3 (realized volatility) and 4 (volume) show p-values of our G-test from equation (5) for all Dow Jones stocks and the index for the same bandwidth choices as in our Monte Carlo experiment.

As expected from the graphical investigation we find no significant seasonal long memory in the volatility of the index at frequency $\omega = 2\pi/13$ at the 5% level, and for the index'

c		0.1			0.15			0.2	
RICS\ δ	0.5	0.6	0.7	0.5	0.6	0.7	0.5	0.6	0.7
.DJI	0.245	0.571	**0.000**	0.092	0.240	**0.000**	**0.016**	**0.040**	**0.000**
AAPL.O	**0.000**	**0.002**	**0.007**	**0.006**	**0.023**	0.068	**0.026**	0.061	0.149
AXP	0.136	0.289	0.402	0.104	0.238	0.417	0.090	0.223	0.513
BA	0.212	0.192	0.190	0.136	0.176	0.161	0.085	0.139	0.163
CAT	0.067	0.180	0.302	0.053	0.140	0.279	**0.036**	0.090	0.230
CSCO.O	0.181	0.460	0.527	0.157	0.387	0.580	0.124	0.288	0.560
CVX	**0.002**	**0.003**	**0.010**	**0.004**	**0.009**	**0.025**	**0.004**	**0.011**	**0.028**
DD	**0.006**	**0.004**	**0.008**	**0.003**	**0.004**	**0.008**	**0.002**	**0.005**	**0.013**
DIS	0.247	0.377	0.674	0.215	0.398	0.719	0.172	0.413	0.760
GE	0.062	0.147	0.256	0.062	0.151	0.304	**0.032**	0.082	0.212
GS	**0.003**	**0.007**	**0.012**	**0.003**	**0.008**	**0.017**	**0.002**	**0.006**	**0.016**
HD	**0.020**	**0.043**	0.077	**0.020**	**0.046**	0.097	**0.026**	0.069	0.177
IBM	0.298	**0.049**	0.073	0.222	**0.047**	0.075	0.186	**0.045**	0.093
INTC.O	**0.015**	**0.034**	**0.050**	**0.014**	**0.033**	0.066	**0.006**	**0.017**	**0.048**
JNJ	0.083	0.127	0.229	0.061	0.117	0.242	**0.049**	0.129	0.313
JPM	**0.005**	**0.009**	**0.021**	**0.005**	**0.010**	**0.025**	**0.003**	**0.007**	**0.019**
KO	0.181	0.081	0.151	0.158	0.077	0.148	0.129	0.062	0.135
MCD	0.080	0.141	0.130	**0.048**	0.096	0.127	**0.018**	**0.045**	0.110
MMM	0.381	0.293	0.473	0.348	0.290	0.476	0.310	0.281	0.511
MRK	**0.004**	**0.008**	**0.007**	**0.002**	**0.004**	**0.006**	**0.001**	**0.002**	**0.007**
MSFT.O	0.573	0.847	0.777	0.552	0.855	0.859	0.415	0.764	0.866
NKE	**0.017**	**0.017**	**0.026**	**0.012**	**0.018**	**0.037**	**0.006**	**0.018**	**0.049**
PFE	**0.006**	**0.011**	**0.030**	**0.009**	**0.020**	0.052	**0.007**	**0.019**	**0.048**
PG	**0.018**	**0.024**	**0.025**	**0.009**	**0.017**	**0.028**	**0.002**	**0.007**	**0.021**
TRV	0.098	0.132	0.300	0.076	0.140	0.321	**0.042**	0.117	0.281
UNH	0.111	0.165	0.253	0.078	0.142	0.236	**0.040**	0.104	0.239
UTX	0.168	0.348	0.495	0.147	0.326	0.553	0.079	0.197	0.460
V	0.113	0.144	0.228	0.113	0.144	0.241	0.094	0.125	0.253
VZ	**0.007**	**0.007**	**0.014**	**0.007**	**0.011**	**0.025**	**0.005**	**0.014**	**0.039**
WMT	0.782	0.911	0.739	0.725	0.909	0.719	0.576	0.866	0.729
XOM	**0.011**	**0.022**	0.083	**0.020**	**0.047**	0.141	**0.029**	0.078	0.184

Table 4: p-values of the G-test at frequency $\omega = 2\pi/13$ for *logarithms of volume*. The data is seasonally demeaned and long memory at frequency zero has been removed.

volume data the results are only significant for large bandwidths. However, we find significant seasonal long memory in a large proportion of our single stock data because we reject the null hypothesis in almost two third of both the volatility and the volume series. There are a few stocks where we have significant results in both series but most often we find it in only one of them for the same stock. This suggests that the seasonal components in volume and volatility might be driven by different factors. Nonetheless, the results clearly prove that there is stochastic seasonality in the shape of seasonal long memory in addition to deterministic cycles in individual stock data.

RICS\ δ	0.5	0.6	0.7	RICS\ δ	0.5	0.6	0.7
.DJI				KO	0.111 (0.031)	0.072 (0.019)	0.046 (0.012)
AAPL.O	0.147 (0.031)			MCD	0.063 (0.031)		
AXP	0.120 (0.031)	0.058 (0.019)		MMM	0.112 (0.031)		
BA				MRK		0.056 (0.019)	
CAT	0.113 (0.031)	0.059 (0.019)	0.025 (0.012)	MSFT.O			
CSCO.O				NKE			
CVX	0.179 (0.031)	0.102 (0.019)	0.086 (0.012)	PFE	0.002 (0.031)	0.021 (0.019)	
DD	0.038 (0.031)			PG	0.114 (0.031)		
DIS	0.138 (0.031)			TRV	0.163 (0.031)	0.101 (0.019)	0.063 (0.012)
GE	0.101 (0.031)			UNH			
GS				UTX	0.123 (0.031)		
HD				V			
IBM	0.067 (0.031)	0.046 (0.019)		VZ	0.146 (0.031)	0.093 (0.019)	0.058 (0.012)
INTC.O				WMT			
JNJ	0.089 (0.031)			XOM	0.149 (0.031)	0.101 (0.019)	0.066 (0.012)
JPM							

Table 5: Empty fields indicate no rejection of G-test applied at frequency $\omega = 2\pi/13$ at the 5%-level for *logarithms of realized volatility*. The numbers give estimates of seasonal long memory at frequency ω and standard errors in brackets.

RICS\ δ	0.5	0.6	0.7	RICS\ δ	0.5	0.6	0.7
.DJI	0.113 (0.031)	-0.012 (0.019)	-0.125 (0.012)	KO			
AAPL.O	0.114 (0.031)	0.059 (0.019)	0.063 (0.012)	MCD	0.107 (0.031)	0.063 (0.019)	
AXP				MMM			
BA				MRK	0.160 (0.031)	0.075 (0.019)	0.042 (0.012)
CAT	0.108 (0.031)			MSFT.O			
CSCO.O				NKE	0.194 (0.031)	0.109 (0.019)	0.069 (0.012)
CVX	0.131 (0.031)	0.087 (0.019)	0.063 (0.012)	PFE	0.162 (0.031)	0.084 (0.019)	0.043 (0.012)
DD	0.117 (0.031)	0.065 (0.019)	0.046 (0.012)	PG	0.169 (0.031)	0.080 (0.019)	0.059 (0.012)
DIS				TRV	0.187 (0.031)		
GE	0.099 (0.031)			UNH	0.068 (0.031)		
GS	0.088 (0.031)	0.074 (0.019)	0.018 (0.012)	UTX			
HD	0.170 (0.031)	0.071 (0.019)		V			
IBM		0.107 (0.019)		VZ	0.190 (0.031)	0.076 (0.019)	0.061 (0.012)
INTC.O	0.089 (0.031)	0.064 (0.019)	0.036 (0.012)	WMT			
JNJ				XOM	0.124 (0.031)	0.087 (0.019)	
JPM	0.173 (0.031)	0.078 (0.019)	0.039 (0.012)				

Table 6: Empty fields indicate no rejection of G-test applied at frequency $\omega = 2\pi/13$ at the 5%-level for *logarithms of volume*. The numbers give estimates of seasonal long memory at frequency ω and standard errors in brackets.

Tables 5 (realized volatility) and 6 (volume) display seasonal memory estimates in the cases where we find seasonal long memory according to our G-test for any choice of c. The estimates are calculated with the generalized local Whittle estimator of Arteche and Robinson (2000) at frequency $\omega = 2\pi/13$ for different bandwidth choices m. They vary from 0.05 to 0.2 and are significant with only three exceptions. Overall, the volume series exhibit slightly higher seasonal persistence than volatility. Since our test has

better power properties for higher seasonal memory parameters, even more seasonal long memory would be found if the series were more persistent.
All in all it seems to be enough to account for deterministic seasonality in averaged data like the DJIA index. Thus, aggregation in an index eliminates the stochastic seasonality present in individual stocks. However, we show that for a large proportion of individual stock data there is seasonal long memory that has to be considered after removing deterministic cycles. This is in line with Bordignon et al. (2007) who show that dummy regression is not sufficient in the context of volatility modeling.

6 Conclusion

In recent years the availability of larger data sets from using intraday data has become important from an econometric perspective, and by now it is established that this intraday data exhibits a strong seasonal structure that has to be considered.
In this chapter we therefore examine intraday seasonality and show that it is not always well characterized by deterministic models. To do so we introduce a semiparametric test for seasonal long memory and prove its good finite sample performance in a Monte Carlo experiment. Due to its semiparametric nature we do not encounter the problem of a potentially misspecified model and are robust to short run dynamics.
This procedure is applied to intraday realized volatility and trading volume data of the DJIA index and its constituents. We find that for the index seasonality is deterministic, but the inspection of individual stocks indicates that there is seasonal long memory in both realized volatility and trading volume. Hence, such data is characterized by "normal" long memory, deterministic cycles and seasonal long memory. In contrast, for the averaged behavior we only find long memory and deterministic cycles. We therefore conclude that the nature of intraday seasonality in the index and in individual stock data is not identical so that they should be treated accordingly.

References

Abeysinghe, T. (1991). Inappropriate use of seasonal dummies in regression. *Economics Letters*, 36(2):175–179.

Abeysinghe, T. (1994). Deterministic seasonal models and spurious regressions. *Journal of Econometrics*, 61(2):259–272.

Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2):115–158.

Andersen, T. G. and Bollerslev, T. (1998). Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *The Journal of Finance*, 53(1):219–265.

Andersen, T. G., Bollerslev, T., and Das, A. (2001). Variance-ratio statistics and high-frequency data: Testing for changes in intraday volatility patterns. *The Journal of Finance*, 56(1):305–327.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Arteche, J. (2002). Semiparametric robust tests on seasonal or cyclical long memory time series. *Journal of Time Series Analysis*, 23(3):251–285.

Arteche, J. and Robinson, P. M. (2000). Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, 21(1):1–25.

Baillie, R. T. and Bollerslev, T. (1991). Intra-day and inter-market volatility in foreign exchange rates. *The Review of Economic Studies*, 58(3):565–585.

Bollerslev, T., Li, J., and Xue, Y. (2016). Volume, volatility and public news announcements. Technical report, Technical report, CREATES Research Paper 2016-19.

Bordignon, S., Caporin, M., and Lisi, F. (2007). Generalised long-memory garch models for intra-daily volatility. *Computational Statistics & Data Analysis*, 51(12):5900–5912.

Bordignon, S., Caporin, M., and Lisi, F. (2008). Periodic long-memory garch models. *Econometric Reviews*, 28(1-3):60–82.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2013). *Time series analysis: forecasting and control*. John Wiley & Sons.

Caporale, G. M., Cunado, J., and Gil-Alana, L. A. (2012). Deterministic versus stochastic seasonal fractional integration and structural breaks. *Statistics and Computing*, 22(2):349–358.

Cogburn, R., Davis, H. T., et al. (1974). Periodic splines and spectral estimation. *The Annals of Statistics*, 2(6):1108–1126.

da Silva Lopes, A. C. (1999). Spurious deterministic seasonality and autocorrelation corrections with quarterly data: Further monte carlo results. *Empirical Economics*, 24(2):341–359.

Deo, R., Hurvich, C., and Lu, Y. (2006). Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment. *Journal of Econometrics*, 131(1):29–58.

Franses, P. H., Hylleberg, S., and Lee, H. S. (1995). Spurious deterministic seasonality. *Economics Letters*, 48(3):249–256.

Gil-Alana, L. A. (2005). Deterministic seasonality versus seasonal fractional integration. *Journal of Statistical Planning and Inference*, 134(2):445–461.

Giot, P. (2005). Market risk models for intraday data. *The European Journal of Finance*, 11(4):309–324.

Giraitis, L. and Leipus, R. (1995). A generalized fractionally differencing approach in long-memory modeling. *Lithuanian Mathematical Journal*, 35(1):53–65.

Gray, H. L., Zhang, N.-F., and Woodward, W. A. (1989). On generalized fractional processes. *Journal of Time Series Analysis*, 10(3):233–257.

Hassler, U. (1994). (mis) specification of long memory in seasonal time series. *Journal of Time Series Analysis*, 15(1):19–30.

Hylleberg, S. (1992). *Modelling seasonality*. Oxford University Press.

Kooperberg, C., Stone, C. J., and Truong, Y. K. (1995). Rate of convergence for logspline spectral density estimation. *Journal of Time Series Analysis*, 16(4):389–401.

Kuensch, H. R. (1987). Statistical aspects of self-similar processes. In *Proceedings of the first World Congress of the Bernoulli Society*, volume 1, pages 67–74. VNU Science Press Utrecht.

Leschinski, C. and Sibbertsen, P. (2014). Model order selection in seasonal/cyclical long memory models. Technical report, Discussion Paper, Wirtschaftswissenschaftliche Fakultät, Leibniz University of Hannover.

Lu, Y. K. and Perron, P. (2010). Modeling and forecasting stock return volatility using a random level shift model. *Journal of Empirical Finance*, 17(1):138–156.

Martens, M., Chang, Y.-C., and Taylor, S. J. (2002). A comparison of seasonal adjustment methods when forecasting intraday volatility. *Journal of Financial Research*, 25(2):283–299.

Ooms, M. and Hassler, U. (1997). On the effect of seasonal adjustment on the log-periodogram regression. *Economics Letters*, 56(2):135–141.

Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. *Journal of the American Statistical Association*, 85(410):338–344.

Ray, B. K. (1993). Long-range forecasting of ibm product revenues using a seasonal fractionally differenced arma model. *International Journal of Forecasting*, 9(2):255–269.

Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of statistics*, pages 1630–1661.

Rossi, E. and Fantazzini, D. (2014). Long memory and periodicity in intraday volatility. *Journal of Financial Econometrics*, 13(4):922–961.

Shimotsu, K., Phillips, P. C., et al. (2005). Exact local whittle estimation of fractional integration. *The Annals of Statistics*, 33(4):1890–1933.

Wood, R. A., McInish, T. H., and Ord, J. K. (1985). An investigation of transactions data for nyse stocks. *The Journal of Finance*, 40(3):723–739.

Woodward, W. A., Cheng, Q. C., and Gray, H. L. (1998). A k-factor garma long-memory model. *Journal of Time Series Analysis*, 19(4):485–504.

Xu, J. and Perron, P. (2014). Forecasting return volatility: Level shifts with varying jump probability and mean reversion. *International Journal of Forecasting*, 30(3):449–463.