Downloaded from UvA-DARE, the institutional repository of the University of Amsterdam (UvA)
http://dare.uva.nl/document/142648

File ID	142648
Filename	Chapter 5: Fit of the scaling model

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	Dissertation
Title	Volatility proxies and GARCH models
Author	M.P. Visser
Faculty	Faculty of Science
Year	2009
Pages	ix, 153
ISBN	978-90-5776-195-9

FULL BIBLIOGRAPHIC DETAILS:
http://dare.uva.nl/record/312511

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)

Chapter 5

Fit of the Scaling Model

Much of the progress in financial asset price modelling in recent years stems from the use of intraday high-frequency data. Compared with the classical situation of having only daily closing prices, researchers nowadays work with data sets consisting per day of thousands of recorded transaction prices and quotes. More data means more information. Advances in volatility measurement and forecasting give rise to improved option pricing models, asset allocation, and risk management procedures.

Despite evident progress, many asset price models are incompatible with high-frequency data. There is a distinguishing property of high-frequency financial data that is generally not taken into account. Intraday price movements display time of day effects. One such effect is a pronounced intraday volatility pattern. On average this pattern takes the form of a U-shape for equity markets, reflecting that price swings during the opening and closing hours tend to be large. In between, the market tends to be less volatile. Moreover, regularly scheduled macroeconomic news announcements, such as consumer confidence figures, are often followed by spikes in volatility. High volatility may be anticipated for certain times of the day. Ignoring such diurnal effects may invalidate inferences drawn from high-frequency data (Andersen and Bollerslev, 1997).

This chapter analyses the *scaling model*, which explicitly takes into account diurnal effects in financial price processes. The aim of the chapter is to gain insight into the intraday price process of the S&P 500 index over the years 1988–2008 using the scaling model, and to assess the fit of the scaling model to the intraday return process.

Before introducing the scaling model, let us briefly discuss the standard setup for financial price processes. Log-prices are widely assumed to follow a semimartingale process; that is, they are the sum of a finite variation process and a local martingale (Protter, 2005); for an econometric discussion, see Andersen, Bollerslev, Diebold, and Labys (2003). The class of semimartingales is large. Brownian motion is a semimartingale. So are Itô processes and Lévy processes. Semimartingales are mathematically interesting since they

form a large class of processes with respect to which stochastic integration is well-defined. They appeal to financial economists as the assumption of no-arbitrage implies a semimartingale (details in Delbaen and Schachermayer, 1994). This is why semimartingales form a cornerstone in derivatives pricing.

For semimartingales the natural measure for variability is quadratic variation. For financial price processes, daily quadratic variation may be approximated by the realized variance (the sum of squared five-minute returns). This observation, combined with the wide availability of high-frequency data in recent years, has spurred econometric research on the measurement and modelling of realized variance (or realized volatility, its square root).

The semimartingale approach does however not come without limitations. The assumption of a semimartingale is so general that it is hard to dispute on empirical grounds. There are no particular assumptions on the path of the price process (as such, it allows for diurnal effects), nor on its dependence structure. One might say that there is no hypothesis to test, hence little to be learned on financial price processes. It is unclear how to forecast volatility, or to price options without additional assumptions. Many particular models have been proposed, often finite-parameter diffusion-type processes with stochastic volatility. Examples are square root processes, Ornstein-Uhlenbeck, and COGARCH. All these models ignore diurnal effects, which makes them unreliable when applied to intraday high-frequency prices.

The scaling model incorporates diurnal effects. For each trading day the high-frequency data determine a continuous time process $R_n(u)$, $0 \leq u \leq 1$, the log-return process for day n. For ease of notation we normalize the trading day to the unit time interval $[0, 1]$. The scaling model describes the intraday return process $R_n(\cdot)$ as the product of a positive scale factor $\sigma_n > 0$ and a stochastic process $\Psi_n(\cdot)$,

$$R_n(u) = \sigma_n \Psi_n(u), \qquad 0 \leq u \leq 1.$$

The *standard processes* $\Psi_n(\cdot)$ are assumed iid. The process $\Psi_n(\cdot)$ is independent of σ_n and of the past given by $\{\sigma_{n-1},\ \Psi_{n-1},\ \sigma_{n-2},\ \Psi_{n-2}, \ldots\}$. In words, *the intraday return process $R_n(\cdot)$ is a scaled standard process.* There are no constraints on the structure of $\Psi_n(\cdot)$; the scaling model may thus accommodate intraday volatility patterns. The process $\Psi_n(\cdot)$ may, but need not, be a semimartingale. By construction the model takes into account diurnal effects.

The scaling model is quite versatile. The daily returns r_n,

$$r_n \equiv R_n(1) = \sigma_n Z_n,$$

where $Z_n \equiv \Psi_n(1)$, satisfy the canonical product structure common to many discrete time volatility models. As such, the scaling model has proven a useful framework for dealing with intraday price movements in discrete time (daily) volatility models, such as GARCH and stochastic volatility models. It has been used for improving the efficiency of parameter estimation for GARCH models, and for increasing forecast accuracy by incorporating volatility proxies into GARCH forecast equations, see Visser (2008b, 2008a).

We test the different assumptions underlying the scaling model. The model is non-parametric (no assumptions on σ_n, nor on the distribution of Ψ_n), so one may expect that it is flexible in matching actual return processes. On the other hand, we have a large data set consisting of more than 5000 trading days and 15 million price ticks. With such a large amount of data even small discrepancies between the model and reality may turn statistically significant. We find almost no dependence between the scale factor σ_n and the process $\Psi_n(\cdot)$, and cannot reject independence for the sequence of standard processes (Ψ_n). The distribution of the process Ψ_n appears quite stable over periods of a year or two, but it does change over the full period of 21 years. So one might think of the sequence (Ψ_n) as locally stationary. Overall, the scaling model gives a good fit to the intraday return process.

The scaling model gives insight into the behaviour of the S&P 500 intraday return processes $R_n(\cdot)$ over the years 1988–2008. There are pronounced intraday volatility patterns. The average level of volatility differs depending on the time of the day. The same holds for volatility persistence; an increase in volatility can have a greater effect on volatility over the subsequent hour depending on the time of the day.

5.1 Exploratory Data Analysis: Intraday Volatility Pattern and Periodicity

In this section, we aim to gain preliminary insights into the diurnal behaviour of the S&P 500 index. We do so without reference to a model. Here we simply use the S&P 500 index tick data to visualize some of the properties of the intraday return process $R_n(\cdot)$. Our S&P 500 futures data consist of 5174 days from 1988–01–04 until 2008–12–31. The S&P 500 futures are traded from 8:30 A.M. to 3:15 P.M. Central Standard Time. See Appendix 5.A for details.

Specifically, we have for each trading day n the sample path of a process $R_n(u)$, $0 \leq u \leq 1$, the continuous time log-return process within that day. Figure 5.1 depicts five realizations of intraday log-return processes $R_n(\cdot)$, for the S&P 500, for the days $n = 2285, \ldots, 2289$. These are the same days as those in Chapter 2, Figure 2.1. Each day starts in the morning with the overnight return (at time $u = 0$). At the end of the day, at time $u = 1$, the graph gives the close-to-close return $R_n(1)$. The second day in the figure

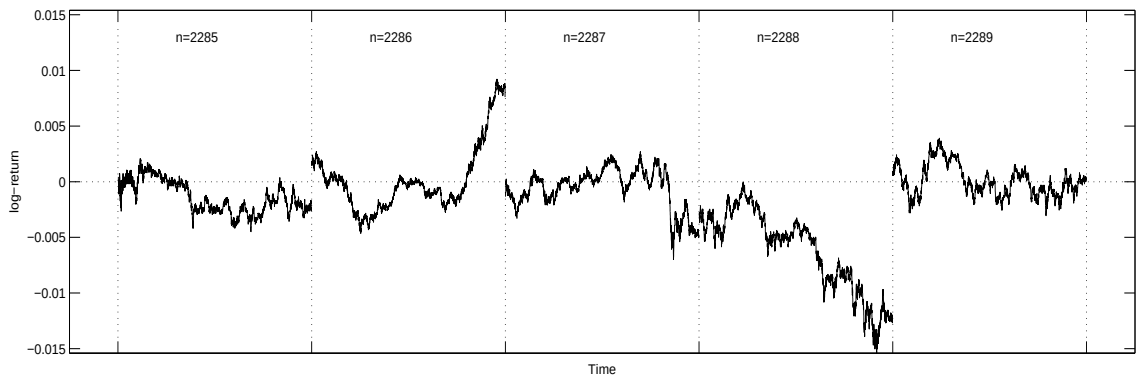


Figure 5.1: Five intraday return processes $R_n(\cdot)$, with respect to the previous day's closing price, $n = 2285, \ldots, 2289$, for the S&P 500. Starting at 1997–02–14.

($n = 2286$ in our sample), for example, starts with a small positive overnight return. The value of the index first decreases, but towards the end of the day increases sharply and finally arrives at a plus of $R_n(1) \approx 0.01$, a 1% increase, at the close of the day, compared with yesterday's closing price.

Figure 5.2 shows some properties of the variability of the S&P 500 price process, making use of the five-minute returns in $R_n(\cdot)$. Figure 5.2(a) graphs the day-to-day realized volatilities RV_n, based on five-minute intervals. The realized volatility is a widely used indicator for daily volatility. It is obtained by taking the square root of the realized variance $RV_n^2(\Delta)$, which is the sum of the squared returns over intervals of length Δ:

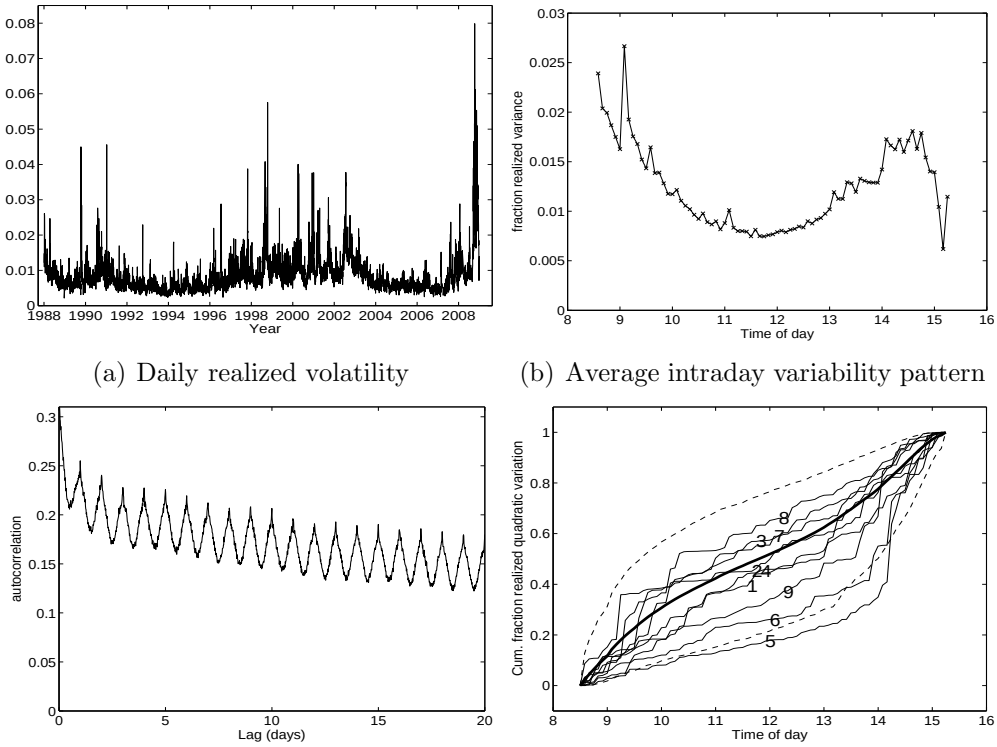
$$RV_n(\Delta) = \left(\sum_{k=1}^{1/\Delta} r_{n,k}^2 \right)^{1/2} . \tag{5.1}$$

We choose Δ such that $1/\Delta$ is an integer. The intraday returns on day n are defined by

$$r_{n,k} = R_n(k\,\Delta) - R_n(\,(k-1)\,\Delta). \tag{5.2}$$

We shall work with the Δ = five-minute time interval, yielding 81 returns per day. Figure 5.2(a) shows that the financial price process exhibits periods of small variation and periods of large variation. Volatility may be high (or low) for several years. In addition, there are spikes, indicating days of hectic trading.

Figure 5.2(b) depicts the U-shaped volatility pattern over the trading day. More precisely, for each five-minute interval the figure displays the average proportion of daily realized variance attributable to this interval. The k-th squared five-minute return, $r_{n,k}^2$, is divided by the realized variance for the day. Then for each interval k the values $r_{n,k}^2/RV_n^2$ are averaged over $n = 1, \ldots, 5174$ days. These averages have been plotted in Figure 5.2(b).



(a) Daily realized volatility

(b) Average intraday variability pattern

(c) One-day periodicity in abs. 5-min. returns

(d) Cumulative intraday proportion of realized variance

Figure 5.2: Properties of the variability of the S&P 500 index price process, over the days $n = 1, \ldots, 5174$ in our full sample. (a) Time series of daily realized volatility, based on five-minute returns, see (5.1). (b) Average intraday variability pattern. For each of the 81 intraday intervals, the curve depicts the squared five-minute return divided by the day's realized variance, averaged over 5174 days. (c) Autocorrelations of absolute five-minute returns, showing one-day periodicity. Lags vary from 5 minutes up to 20 days (the rightmost autocorrelation is lagged by 1620 five-minute intervals). (d) Cumulative intraday proportion of realized variance. Bold line: average of sample (5174 days). Numbered lines: nine consecutive days $n = 2283, \ldots, 2291$ starting at 1997–02–12, including the five days in Figure 5.1.

There is a distinct U-shape: on average a relatively large proportion of realized variance occurs in the period after opening and towards the end of the day. The peak at 9:00 A.M. might be due to the regularly scheduled releases of important macro economic figures, such as Consumer Confidence, Chicago Purchasing Manager, and Leading Indicators. Intraday periodicities and variability patterns have been widely reported, across markets and across time. For references, see McMillan and Speight (2004); for information on scheduled news release times, see Andersen, Bollerslev, Diebold, and Vega (2003).

The U-shape of Figure 5.2(b) suggests that price variability exhibits a pattern that repeats at a one-day frequency. That the basic period really is one day is clear from

Figure 5.2(c), which shows the autocorrelation of the absolute five-minute returns. The lags vary from five minutes to twenty days. The empirical autocorrelation function is a useful tool for detecting periodicities. The empirical autocorrelation of a periodic process is, itself, periodic with the same period. Figure 5.2(c) shows a very pronounced periodicity of one day. This strong periodicity effect may explain why GARCH models applied to daily returns are quite successful. GARCH models applied to five-minute returns do poorly, see Andersen and Bollerslev (1997).

The intraday volatility pattern is very pronounced on average. One may wonder whether this U-shape is visible in each day. To address this question we first construct the cumulative function of the U-shaped function in Figure 5.2(b). This cumulative is given by the bold line in Figure 5.2(d). It tells us how the realized variance is formed over the trading day, on average. At the start of the day the proportion of the day's realized variance is zero, and at the end of the day it is one. The slow increase in the middle part reflects the low values in the graph in Figure 5.2(b) between 10:00 A.M. and 2:00 P.M. In Figure 5.2(d) we also show the cumulative functions for the nine consecutive days $n = 2283, \ldots, 2291$ in our sample. The curves for these days have been numbered one to nine. The numbered curves show that individual daily variability patterns vary substantially: on the fifth day 70% of the day's variation occurs after 2:00 P.M., while on the eighth day 50% of the day's variation has already taken place before 10:30 A.M. The relatively large afternoon variability on day numbered 5 in Figure 5.2(d) corresponds to a sudden downward price movement on day number 2287 in the sample, see Figure 5.1. To make sure that we are not looking at a small set of atypical trading days, Figure 5.2(d) also shows two quantile lines. At the end of each five-minute interval, we calculate the 0.975 sample quantile (over $n = 1, \ldots, 5174$) of the proportion process. The upper dashed line connects these quantiles. Similarly, the lower dashed line connects the 0.025 sample quantiles. At each time point 95% of the curves for the full sample lie in between the dashed lines. The sample of the nine proportion processes in Figure 5.2(d) does not appear to be atypical.

5.2 Diagnostic Checking for the Scaling Model

The previous section shows that the financial price process, certainly for the S&P 500, exhibits daily regularity. In formal terms, one may think of a daily random pattern $\Psi_n(\cdot)$ (a random process), which underlies the periodicity observed in the intraday return processes $R_n(\cdot)$. We shall now introduce a model that takes account of a daily pattern. The scaling model describes $R_n(\cdot)$ as the product of a positive scale factor $\sigma_n > 0$ and a cadlag[1]

[1]The sample paths are right-continuous and have left limits.

process $\Psi_n(\cdot)$,

$$R_n(u) = \sigma_n \Psi_n(u), \qquad 0 \le u \le 1, \tag{5.3}$$

where intraday time u advances from zero to one. We refer to $\Psi_n(\cdot)$ as the *standard process*. We still have to define the dependencies. The sequence (Ψ_n) is iid. Moreover, the scale factor σ_n and the process $\Psi_n(\cdot)$ are independent for each day, and $\Psi_n(\cdot)$ is independent of the past until day $n-1$. It is convenient to express the dependence structure using the σ-algebra $\mathcal{G}_{n-1}$ generated by σ_n and the past, i.e. generated by $\{\sigma_n, \sigma_{n-1}, \Psi_{n-1}, \sigma_{n-2}, \Psi_{n-2}, \ldots\}$. The sequence of standard processes $\Psi_n(\cdot)$ satisfies:

1. for each n, the process $\Psi_n(\cdot)$ is independent of $\mathcal{G}_{n-1}$,
2. the $\Psi_n(\cdot)$ have the same probability distribution for all n.

Equation (5.3) combined with assumptions 1 and 2 define the *scaling model*. By assumptions 1 and 2 the sequence of processes $\Psi_n(\cdot)$ is iid. The scaling model does not impose conditions on the sequence of positive scale factors (σ_n). For financial price processes, the sequence (σ_n) typically depends on its own past, and on the past standard processes. In this case one will typically make the additional assumption that the sequence (σ_n, Ψ_n) is stationary.

The scaling model is consistent with the properties of the S&P 500 index described in the previous section. In this model, persistence in daily volatility corresponds to persistence in the scale factors σ_n. The U-shaped average variability pattern in Figure 5.2(b) is an estimate of the expected variability pattern in $\Psi_n(\cdot)$ (the scale factor σ_n drops out when looking at proportions). By construction the model takes account of diurnal effects.

Assumptions 1 and 2 do not impose constraints on the structure of $\Psi_n(\cdot)$. The process $\Psi_n(\cdot)$ may be a martingale. If it is, then $R_n(\cdot)$ is one too. The scale factor σ_n differs from the square root of quadratic variation which plays an important role in semimartingale models. Quadratic variation and scale factor are related:

$$QV(R_n) = \sigma_n^2 QV(\Psi_n).$$

If σ_n is large, then quadratic variation tends to be large too. For individual days the quadratic variation of $R_n(\cdot)$ can be small while σ_n is large, depending on the variability of the sample path of $\Psi_n(\cdot)$.

The scaling model is of use for incorporating high-frequency intraday price movements in discrete time, daily GARCH models. The daily close-to-close returns, $r_n \equiv R_n(1)$, satisfy

$$r_n = \sigma_n Z_n,$$

where the $Z_n \equiv \Psi_n(1)$ are iid innovations. So, the daily returns r_n satisfy the product structure common to discrete time volatility models, such as GARCH and stochastic volatility models. The use of high-frequency data, in the scaling model, allows one to greatly improve GARCH parameter estimation, compared with estimation based on daily returns (Visser, 2008b). The scaling model also allows one to use high-frequency data to improve volatility forecasts, see Visser (2008a).

The scaling model is nonparametric in the sense that it does not impose constraints on the structure of $\Psi_n(\cdot)$, nor on the sequence of scale factors (σ_n). In applying the model, it is natural to ask: how well does this model fit the behaviour of the S&P 500 over the twenty one year period 1988–2008? In the following, we provide statistical analyses addressing three questions:

- Is the scale factor σ_n independent of the standard process $\Psi_n(\cdot)$, for each day n?
- Are for different days the processes Ψ_k and Ψ_n, $k \neq n$, independent?
- Does the process Ψ_n have the same probability distribution for each day n?

We shall address each of these questions separately in the three subsections below. It is at this point worthwhile to call to mind a general difficulty in hypothesis testing. Hypotheses typically cannot be tested individually. A hypothesis is tested within a larger set of assumptions. Consider for instance the situation of testing independence of a sequence of random variables. Tests for independence usually need random variables that are identically distributed. A rejection of the null hypothesis of independence may be due to dependence; but also to nonstationarity. To come to a specific reason for rejection, one uses tests that have large power against dependence, but low power against nonstationarity. We shall keep these considerations in mind when looking into the three questions above.

5.2.1 Independence of Scale Factor and Standard Process

Does the standard process have the same properties in periods of high volatility as in periods of low volatility? To investigate a possible relation between the scale factor and the standard process, we apply a number of scale free statistics to the intraday return process $R_n(\cdot)$. A *scale free statistic* D is a statistic that is homogeneous of degree zero:

$$D(\alpha R_n) = D(R_n), \qquad \alpha \geq 0.$$

In words, the statistic $D_n \equiv D(R_n)$ does not depend on the scale factor of $R_n(\cdot)$. The identity $R_n(\cdot) = \sigma_n \Psi_n(\cdot)$ implies that

$$D(R_n) = D(\Psi_n).$$

Since the scale free statistic does not depend on σ_n one may use $D(R_n)$ to investigate properties of Ψ_n. By the assumptions on Ψ_n it follows that the D_n form an iid sequence, and for each day n the statistic D_n is not depending on σ_n.

Below we shall examine five scale free statistics, focusing on different aspects of Ψ_n. For each day n the statistics are based on the 81 five-minute returns $(r_{n,k})$, $k = 1, \ldots, 81$, in (5.3):

- $RV5^m/RV5^a$: the ratio of morning variability to afternoon variability, defined as the ratio of the summed squared five-minute returns before and after 12 A.M.
- $RV5^-/RV5^+$: the ratio of the sum of the squared negative five-minute returns to the sum of the squared positive five-minute returns,
- skew: the skewness (standardized third moment) of the five-minute returns,
- kurt: the kurtosis (standardized fourth moment) of the five-minute returns,
- $\rho_1(|r_{n,k}|)$: the first-order autocorrelation coefficient of the sequence of 81 absolute five-minute returns.

We may ask several questions relating these statistics to the size of intraday price fluctuations given by the scale factor σ_n. (1) For $RV5^m/RV5^a$: does σ_n explain asymmetry between morning activity and afternoon activity? (2) Concerning $RV5^-/RV5^+$ and skew: do intraday returns have an increased tendency to show large negative values in periods of large fluctuations? This is measured by $RV5^-$ and skew. (3) On kurt: are intraday returns more heavy tailed during high-volatility periods? (4) For the statistic $\rho_1(|r_{n,k}|)$: is intraday volatility persistence larger during high-volatility periods?

In addressing these questions, one is essentially testing the scaling model. In the scaling model the scale factor σ_n and the standard process $\Psi_n(\cdot)$ are independent. So σ_n and D_n are independent. A relation between σ_n and D_n means a departure from the scaling model.

We investigate whether there is information in D_n that may be explained by the scale factor $\hat{\sigma}_n$ by applying a regression for the five statistics D_n:

$$D_n = c + b\,\hat{\sigma}_n + \varepsilon_n, \qquad n = 1, \ldots, N. \tag{5.4}$$

In the scaling model $b \equiv 0$, $c = \mathbb{E}D_n$, and (ε_n) is a mean-zero iid sequence. For $\hat{\sigma}_n$ we take the log-GARCH estimate based on Visser (2008a), see Appendix 5.C. We obtain

the coefficient of determination R^2 for the regression and calculate the t-value for the null hypothesis that b is zero. For an account of tests of independence for pairs of random variables, see the classical works by Kendall and Stuart (1973) and Hollander and Wolfe (1973). We purely focus on regressions of the type (5.4) as the statistics D_n permit a sufficient amount of freedom in zooming in on different aspects of Ψ_n.

Table 5.1 provides the R^2s and t-values for $b = 0$. As expected, given the large amount

D_n	R^2	t_b
$RV5^m/RV5^a$	0.002	-2.86
$RV5^-/RV5^+$	0.001	-2.72
skew $(r_{n,k})_{k=1,\ldots,81}$	0.003	4.58
kurt $(r_{n,k})_{k=1,\ldots,81}$	0.004	-4.69
ρ_1 $(\lvert r_{n,k}\rvert)_{k=1,\ldots,81}$	0.014	-6.13

Table 5.1: R^2s and t-values for the regression (5.4), for $n = 1, \ldots, 5174$. Newey-West t-values (t_b) for the hypothesis $b = 0$. A t-value outside $(-2, 2)$ statistically indicates dependence between $\hat{\sigma}_n$ and Ψ_n. For each day n the scale free statistics D_n are based on the 81 intraday five-minute returns $r_{n,k},\ k = 1, \ldots, 81$ of $R_n(\cdot)$.

of data, the t-values lie outside the 95% confidence interval (-2,2). The R^2s of order 0.01 (and smaller) reflect minor effects. The largest effect is for ρ_1, the autocorrelation of the absolute five-minute returns, with a t-value equal to -6.13 and a modest R^2. The larger part of this t-value may be explained by a structural break in the mean of ρ_1. Breaking the sample into parts takes away most of the effect, though some significance remains.[2] So the autocorrelation of the standardized absolute five-minute returns is slightly smaller during periods of high volatility. The small R^2s in the table indicate almost no departure from the model. An obvious place to look for reasons for values $R^2 > 0$ are market microstructure effects, such as the bid-ask bounce and price discretization, but such is not the aim of this chapter.

5.2.2 Independence of the Sequence of Standard Processes

Each sequence of scale free statistics (D_n), defined in the previous subsection, is iid in the scaling model since the sequence (Ψ_n) is iid. If (D_n) is a dependent sequence, the standard processes (Ψ_n) are dependent too. We may therefore examine independence for the sequence (Ψ_n) by applying standard tests of independence to the five scale free statistics introduced in Section 5.2.1. We apply a turning point test, difference-sign test and two kinds of Ljung-Box (portmanteau) tests, see Appendix 5.B for details.

[2]Inspection of the time series of ρ_1 suggests a downward shift in the mean during the year 1997. Calculations on the subsample before 1997 (days 1–2253) and the subsample following 1997 (days 2502–5174) yield (R^2, t_b) equal to $(0.001,\ -1.24)$ and $(0.005,\ -3.24)$.

	turning point		diff.-sign		Ljung-Box		Ljung-Box-sub	
D_n	$\bar{T}$	p	$\bar{S}$	p	$Q_{k=20}$	p	$Q_{k=20}^{\text{sub}=50}$	p
$RV5^m/RV5^a$	-1.286	0.198	0.746	0.455	1090.4	0.000	22.6	0.003
$RV5^-/RV5^+$	0.857	0.391	-0.120	0.904	36.3	0.014	21.4	0.053
skew $(r_{n,k})_{k=1,\ldots,81}$	0.660	0.510	1.902	0.057	25.6	0.180	19.2	0.826
kurt $(r_{n,k})_{k=1,\ldots,81}$	1.253	0.210	-0.072	0.942	119.1	0.000	17.6	0.997
ρ_1 $(\lvert r_{n,k}\rvert)_{k=1,\ldots,81}$	-1.319	0.187	-0.602	0.547	473.3	0.000	19.4	0.733

Table 5.2: Tests of randomness (iid) of scale free statistics, for $n = 1, \ldots, 5174$. The statistics $\bar{T}$, $\bar{S}$, and $Q_{k=20}^{\text{sub}=50}$ may be regarded primarily tests for independence. The statistic Q_k tends to reject both in the case of dependence and in the case of nonstationarity. A p-value below 0.05 reflects statistical significance. The scale free statistics D_n are described in Section 5.2.1.

Table 5.2 provides for each of the five scale free statistics D the values of these four test-statistics, and in each case a p-value. A p-value smaller than 0.05 signifies individual statistical significance. Given the number of tests one may expect a few p-values smaller than 0.05 if the D_n are iid. The turning point and difference sign tests do not reject the null hypothesis of independence for any of the statistics D, quite an accomplishment given the large number of observations. In contrast, the standard Ljung-Box test detects some structure for most statistics D, most notably in the case of the ratio morning to afternoon variability. This might indicate dependence within the first 20 days of history. It may also indicate nonstationarity. The turning point test and difference-sign test do *not* detect dependence within runs of two to three successive days. It would therefore be remarkable if there is dependence between the processes Ψ that are more than two days apart. To gain more insight into this matter, we include the average Ljung-Box statistic over fifty subsamples of length approximately ninety days (four months), $Q_{k=20}^{\text{sub}=50}$. Over such a short period stationarity is a reasonable assumption. If a break occurs it will only affect one of the 50 subsamples. The rightmost column shows that only one of the subsampled Ljung-Box statistics yields a significant p-value. This suggests that the standard Ljung-Box rejections are due to nonstationarity. We shall further address the question of stationarity in the next subsection. Table 5.2 is therefore in line with independence of the sequence (Ψ_n).

Another way to examine dependence between $\Psi_n(\cdot)$ and $\Psi_{n+1}(\cdot)$ is to check for persistence in fluctuations from one five-minute interval to the other. Let us first briefly look into the persistence *within* the day. Does the fluctuation over today's 9:10–9:15 A.M. trading interval for $\Psi_n(\cdot)$ affect the 11:50–11:55 A.M. interval? Figure 5.3(a) reveals such dependencies by showing correlations between five-minute high-low ranges (these high-low ranges give per five-minute interval the supremum minus the infinum in R_n). We

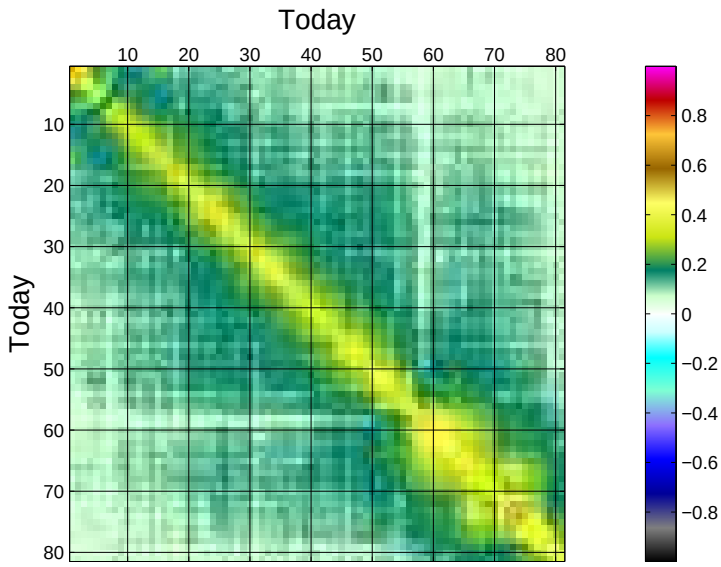
substitute for Ψ_n its estimate $\hat{\Psi}_n$ defined as the standardized process

$$\hat{\Psi}_n(\cdot) = R_n(\cdot)/\hat{\sigma}_n. \tag{5.5}$$

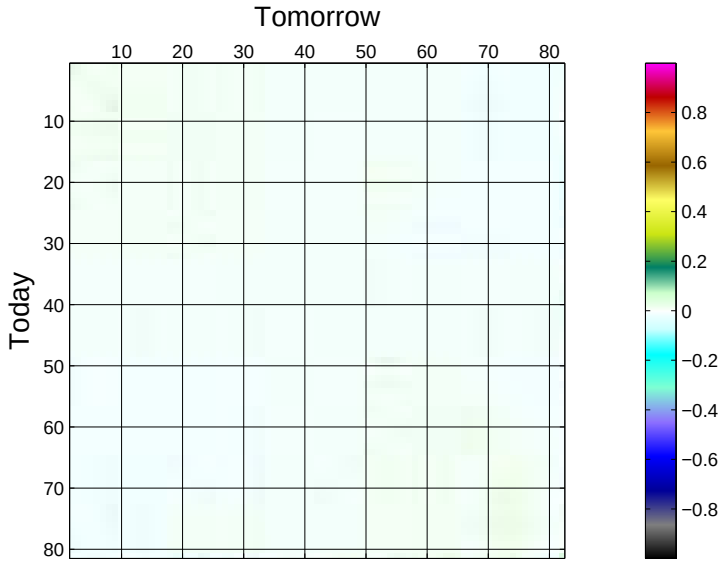
For the daily scale factor $\hat{\sigma}_n$ we take the log-GARCH estimate obtained in Visser (2008a), see Appendix 5.C.

Correlations close to the diagonal are large in Figure 5.3(a); these concern time spans that are close together. Apparently certain hot spots of large correlation exist. Such hot spots occur at the opening, and after interval 56 (just after 1:00 P.M.). The speed of autocorrelation decay is not constant throughout the day: the persistence at the opening largely fades out within fifteen minutes; the persistence after 1:00 P.M. lasts longer. The overall picture is one of volatility persistence in $\hat{\Psi}_n(\cdot)$ that decays fast within the first thirty minutes and decreases to levels close to zero within a couple of hours.

Let us now examine persistence of intraday fluctuations in the standard process from one day to the next. Figure 5.3(b) shows the correlations between the high-low ranges of today's five-minute intervals in $\hat{\Psi}_n(\cdot)$ and those of tomorrow's intervals in $\hat{\Psi}_{n+1}(\cdot)$. One might expect that a large volatility during today's late afternoon induces larger volatility in the early morning tomorrow. Such dependencies result in positive correlations around the lower left corner of the figure. The figure shows that these correlations are practically zero, as are all other correlations. The average of these 81×81 correlations equals -0.003, which is of no practical relevance. This property of zero correlation agrees with the scaling model assumption of independent standard processes.



(a) Persistence in fluctuations within standard process



(b) Persistence in fluctuations between standard processes

Figure 5.3: Correlations between five-minute high-low ranges within the day for $\hat{\Psi}_n(\cdot)$ (part a), and between days for $\hat{\Psi}_n$ (today) and $\hat{\Psi}_{n+1}$ (tomorrow), part (b), based on 5174 data points (days). The process $\hat{\Psi}_n(\cdot)$ is defined in (5.5). Each day is divided into 81 five-minute intervals. The yellow-coloured block at (61, 62) in part (a) represents a correlation coefficient of approximately 0.45 between the high-low ranges of the time interval 1:00–1:05 P.M. and the time interval 1:05–1:10 P.M. within $\hat{\Psi}_n(\cdot)$. The maximal correlation in part (a) equals 0.54 (block (72,73)); the minimal correlation in part (a) equals 0.016 (block (7,74)). The correlations on the diagonal (identically 1) were forced to have colours similar to their neighbours.

5.2.3 Stationarity of the Sequence of Standard Processes

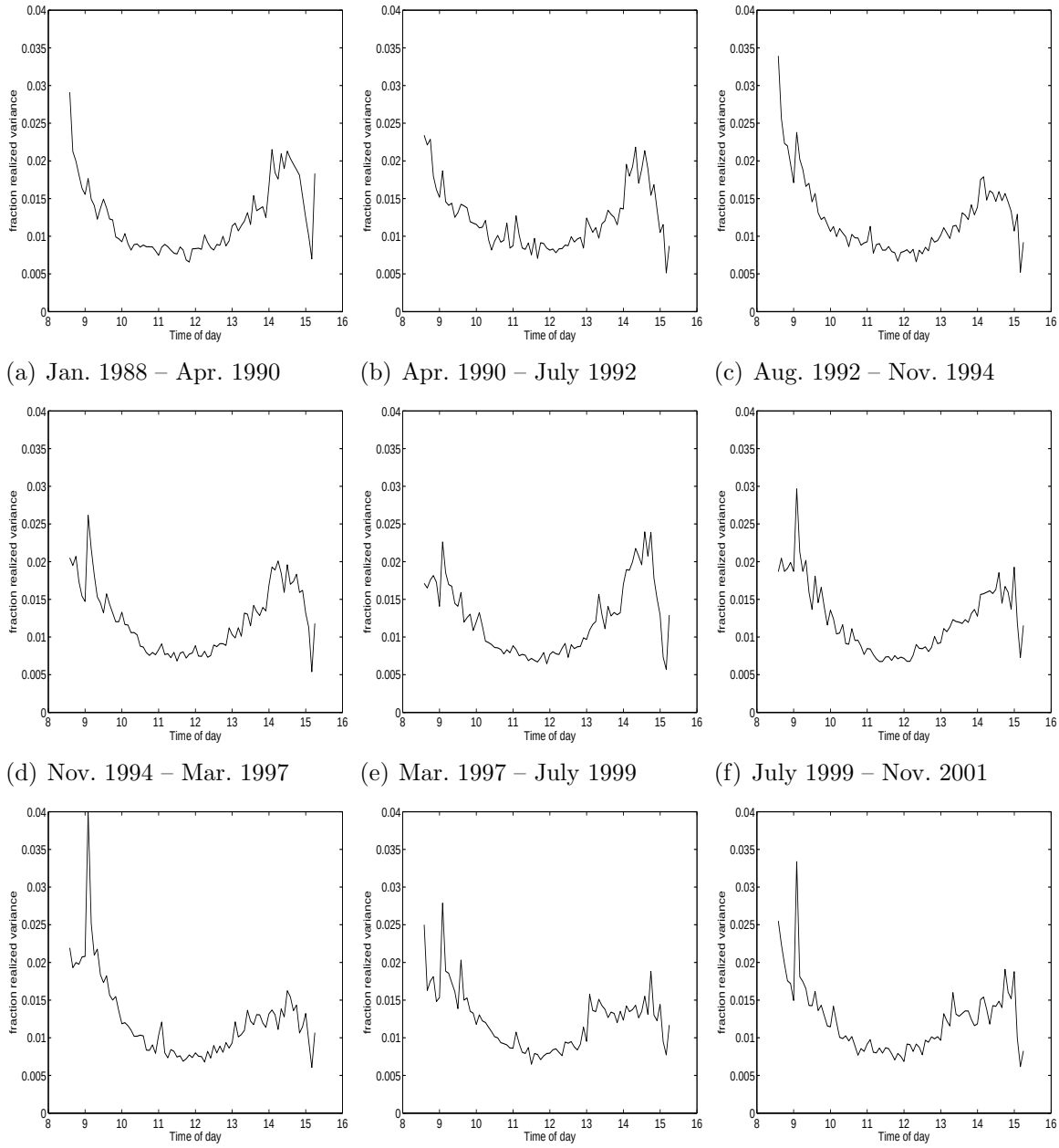
Does the standard process have the same properties in 2008 as 20 years earlier in 1988? Let us start by a visual examination of the volatility U-shape and the intraday volatility persistence in subsamples.

Figure 5.4 shows for nine equally sized subsamples of length two years and four months the average proportion of daily realized variance per five-minute interval. The subsamples span the full sample. For the full-sample pattern, see Figure 5.2(b). The 9:00 A.M. peak is hardly there in the first three graphs. The curve in part (g) has a larger peak at 9:00 A.M. The overall picture suggests subtle differences in the variability pattern over this 21 year period. Shifts appear to take place gradually: successive U-shapes are similar. Below we shall examine the stability of morning variability to afternoon variability in a formal test for stationarity.

Let us first visually examine whether the intraday volatility persistence patterns have changed over time. Figure 5.5 shows the correlations between the five-minute high-low ranges within the day, as in Figure 5.3(a), but now for the nine subsamples of two years. The data tell an interesting story. The first two graphs for the years 1988–1992 (Figure 5.5(a-b)) show similar intraday volatility persistence. Persistence did not last long (as may be seen from the narrow yellow coloured diagonal). There were certain hot spots with correlations around 0.7, at the opening and after 2:15 P.M. For the following two years we see persistence started spreading out over the day (part (c)), reaching the highly persistent pattern for the years 1996–1999 (part (e)). Here, fluctuations in Ψ had an effect on the entire remaining trading day: one may observe an increased width of the yellow diagonal and an increase in the size of the dark green area. During the years 1999–2001 (part (f)) intraday volatility persistence started to weaken. One may also observe a change of structure around 1:00 P.M. The four and a halve years 2002–2006 (parts (g-h)), clearly show a structural break around 1:00 P.M during the day: fluctuations before this time of day appear of little influence to the fluctuations after this time. Intraday volatility persistence was strong again during the years 2006–2008 (part (i)). The graph for these recent years resembles the graph in part (e) for the years 1996–1999.

The persistence patterns in Figure 5.5 change over time. Figures 5.4 and 5.5 suggest that most of the time there are either hardly changes, or that changes take place gradually; looking from one subgraph to the next the patterns appear similar.

One may wonder whether these changes are related to the level of the scale factor $\hat{\sigma}_n$. In Section 5.2.2 we saw that the data are in line with independence of scale factor and standard process, so there should be no relation. We can perform an additional check here. Let us calculate for each day the first, sixth, and twelfth autocorrelations (lags five, thirty, and sixty minutes) of the 81 five-minute high-low ranges per day. These statistics are refinements of the statistic $\rho_1(|r_{n,k}|)$ in Table 5.1. The linear regressions of these three



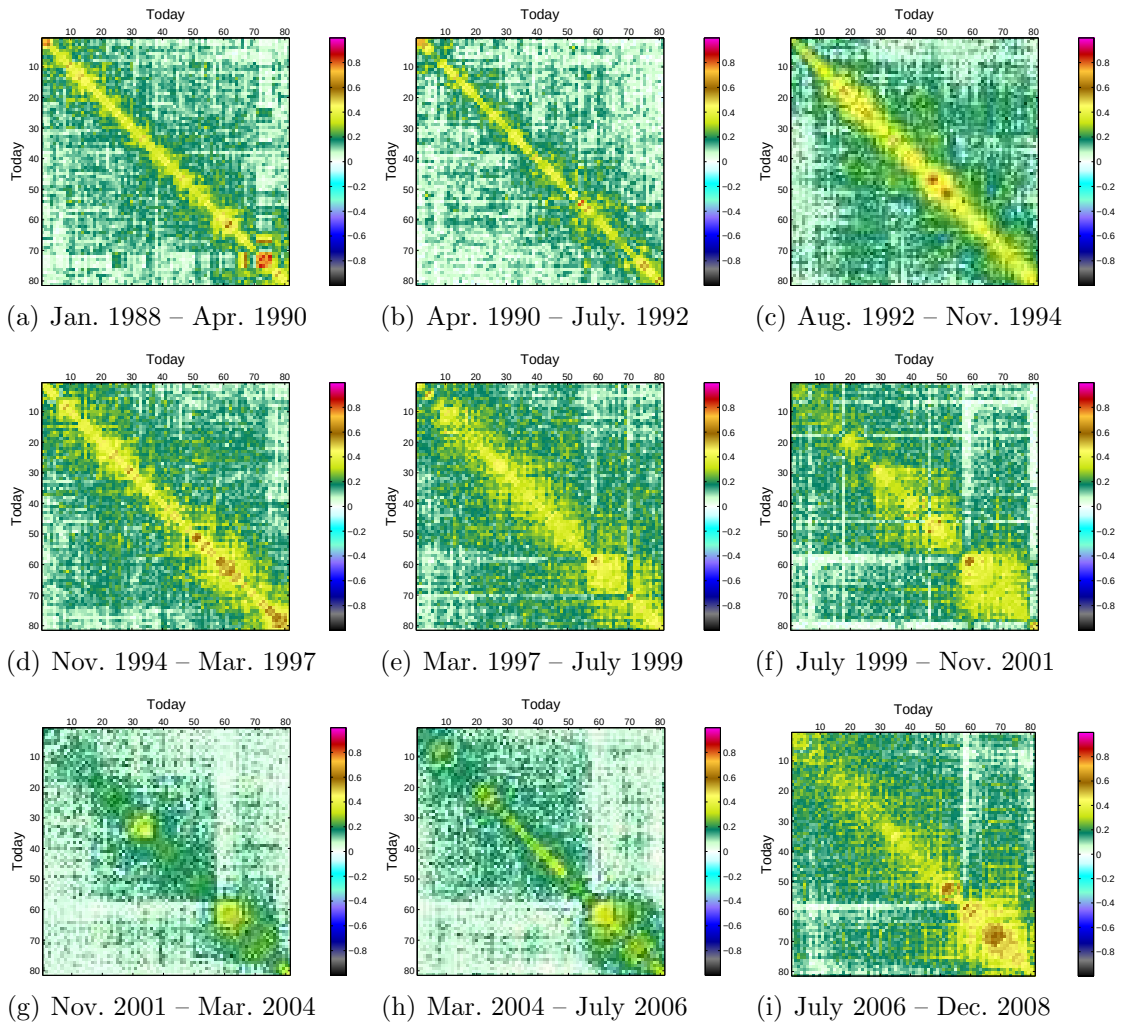
(a) Jan. 1988 – Apr. 1990 (b) Apr. 1990 – July 1992 (c) Aug. 1992 – Nov. 1994

(d) Nov. 1994 – Mar. 1997 (e) Mar. 1997 – July 1999 (f) July 1999 – Nov. 2001

(g) Nov. 2001 – Mar. 2004 (h) Mar. 2004 – July 2006 (i) July 2006 – Dec. 2008

Figure 5.4: Intraday variability pattern, for nine subsamples. For the full-sample pattern, see Figure 5.2(b). The subsamples are days 1–574, 575–1149, 1150–1724, 1725–2299, 2300–2874, 2875–3449, 3450–4024, 4025–4599, 4600–5174.

series of autocorrelations with the scale factor $\hat{\sigma}_n$ (like regression (5.4)) yield a maximal $R^2 = 0.006$, which is even smaller than the effect for ρ_1 in Table 5.1. There is no sign of a relation between the level of the scale factor and the intraday volatility persistence pattern.



(a) Jan. 1988 – Apr. 1990 (b) Apr. 1990 – July. 1992 (c) Aug. 1992 – Nov. 1994

(d) Nov. 1994 – Mar. 1997 (e) Mar. 1997 – July 1999 (f) July 1999 – Nov. 2001

(g) Nov. 2001 – Mar. 2004 (h) Mar. 2004 – July 2006 (i) July 2006 – Dec. 2008

Figure 5.5: Persistence in intraday fluctuations, for nine subsamples. For the full sample, see Figure 5.3. The nine subsamples are given at Figure 5.4.

The exploratory visualizations of this section suggest that the distribution of the standard process changes over time. We shall now formally test stationarity of the sequence (Ψ_n) by examining the stationarity of the time series of the five scale free statistics $RV5^-/RV5^+$, $RV5^m/RV5^a$, skew, kurt, and ρ_1 from Table 5.2. As discussed in Subsection 5.2.2, in the scaling model these are iid, hence identically distributed, random variables. Here, we test for changes in distribution by comparing the nine subsamples.

Table 5.3 provides for each of the five variables the two-sample Kolmogorov-Smirnov test statistics with p-values in parentheses. Consider the subtable for the daily statistic $RV5^-/RV5^+$. The first entry .058 (.274) in this subtable gives the two-sample Kolmogorov-Smirnov test statistic for equality in distribution in the first and second subsamples. The two-sample Kolmogorov-Smirnov test statistic is the supremum of the

absolute difference between the empirical cumulative distribution functions. A large value indicates a difference in distribution. Only 4 out of the 36 p-values in this subtable is smaller than 0.05, suggesting that the ratio of downward to upward volatility has (almost) the same distribution in different subsamples.

Now consider the first subtable, concerning the ratio of morning variability to afternoon variability $RV5^m/RV5^a$. This is the variable whose distribution appears to change most clearly from one subsample to the other. Equality in distribution is rejected for seven out of eight pairs of adjacent intervals (seven p-values on the diagonal smaller than 0.05). This means that the variability patterns depicted in Figure 5.4 are statistically different. To gain insight into these differences in distribution we provide Figure 5.6. It depicts for every subsample a kernel density estimate for the log of the variability ratio, $\log(RV5^m/RV5^a)$. By applying the logarithm to this strictly positive ratio we obtain more symmetric densities. All densities are centred near zero, i.e. the ratio is centred around 1. There are subtle shifts in the location of the peak. As an example, for the years 1988–1990 the modus is negative, whereas for the years 2002–2004 (part (g)) the modus is positive. This reflects a shift from afternoon activity to morning activity. The graphs for the years 1999–2006 show a higher peak near zero. A higher peak means that the distribution of the ratio morning to afternoon activity is more concentrated around 1. Overall, the density plots suggest that there are shifts in location. Are the shifts in location significant? Figure 5.7 shows the cumulative of the variables $\log(RV5_n^m/RV5_n^a)$. If these variables form a stationary sequence then their cumulative would show a random walk with constant drift. For comparison the figure also shows the sample path of a random walk with the same drift and variance as the variables $\log(RV5_n^m/RV5_n^a)$. The curve for the variables $\log(RV5_n^m/RV5_n^a)$ shows nonstationarity. In some periods afternoon activity shows a relative increase, in other periods the morning activity gains importance. These effects are clear, but modest: from Figure 5.6 one observes that from one graph to the next the densities for $\log(RV5_n^m/RV5_n^a)$ over subsamples are similar.

For the remaining three scale free statistics, skew, kurt, and ρ_1, the table shows statistical differences in distribution for a substantial number of subsamples. However, many of the adjacent subsamples (the ones on the diagonal) are not statistically different. This confirms the impression given by the U-shapes (Figure 5.4) and the coloured correlation plots (Figure 5.5) that the structure of the standard process changes over time, but that changes tend to take place gradually. We may therefore think of the standard processes as a locally stationary sequence of random processes.

$RV5^m/RV5^a$

	2	3	4	5	6	7	8	9
1	.144 (.000)	.244 (.000)	.131 (.000)	.057 (.308)	.248 (.000)	.354 (.000)	.243 (.000)	.197 (.000)
2		.134 (.000)	.052 (.405)	.106 (.003)	.108 (.002)	.230 (.000)	.113 (.001)	.082 (.041)
3			.130 (.000)	.224 (.000)	.111 (.001)	.123 (.000)	.092 (.014)	.087 (.024)
4				.122 (.000)	.125 (.000)	.226 (.000)	.123 (.000)	.094 (.012)
5					.200 (.000)	.317 (.000)	.198 (.000)	.179 (.000)
6						.157 (.000)	.035 (.872)	.059 (.259)
7							.155 (.000)	.169 (.000)
8								.063 (.203)

$RV5^-/RV5^+$

	2	3	4	5	6	7	8	9
1	.058 (.274)	.056 (.313)	.040 (.727)	.044 (.632)	.051 (.437)	.064 (.188)	.042 (.671)	.056 (.318)
2		.063 (.203)	.073 (.089)	.057 (.292)	.054 (.364)	.061 (.230)	.092 (.014)	.057 (.292)
3			.052 (.405)	.038 (.786)	.040 (.739)	.078 (.056)	.080 (.048)	.028 (.978)
4				.040 (.739)	.054 (.364)	.087 (.024)	.043 (.639)	.042 (.689)
5					.040 (.739)	.075 (.076)	.056 (.327)	.045 (.590)
6						.054 (.364)	.075 (.076)	.043 (.639)
7							.097 (.008)	.068 (.136)
8								.064 (.178)

skew $(r_{n,k})_{k=1,\ldots,81}$

	2	3	4	5	6	7	8	9
1	.086 (.026)	.057 (.297)	.051 (.443)	.065 (.175)	.060 (.249)	.055 (.338)	.109 (.002)	.074 (.085)
2		.063 (.203)	.070 (.119)	.094 (.012)	.113 (.001)	.049 (.493)	.150 (.000)	.118 (.001)
3			.089 (.020)	.106 (.003)	.085 (.029)	.099 (.006)	.106 (.003)	.101 (.005)
4				.070 (.119)	.087 (.024)	.043 (.639)	.127 (.000)	.087 (.024)
5					.068 (.136)	.087 (.024)	.137 (.000)	.045 (.590)
6						.101 (.005)	.083 (.034)	.042 (.689)
7							.141 (.000)	.117 (.001)
8								.117 (.001)

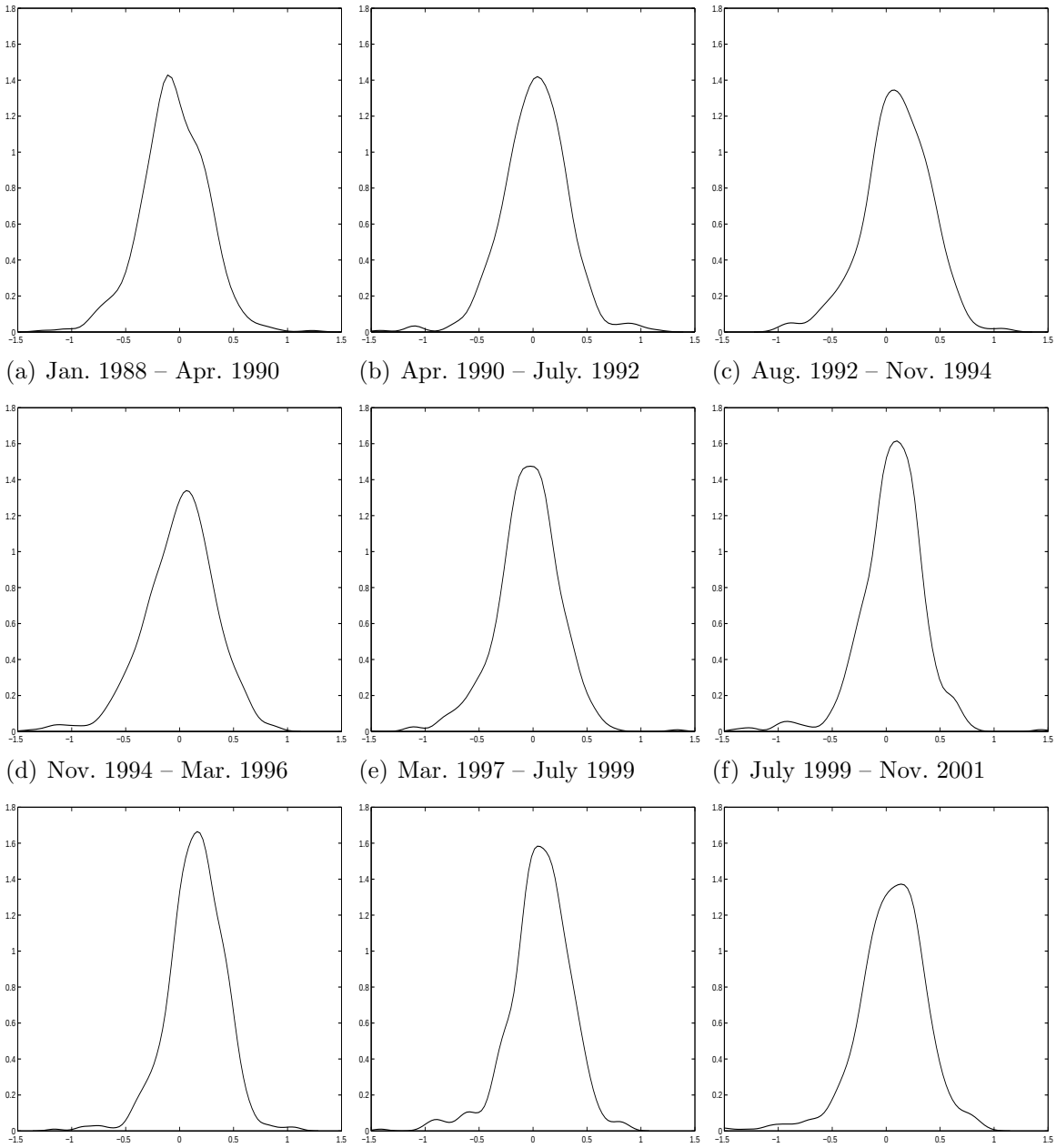
kurt $(r_{n,k})_{k=1,\ldots,81}$

	2	3	4	5	6	7	8	9
1	.052 (.418)	.097 (.008)	.071 (.110)	.145 (.000)	.158 (.000)	.104 (.004)	.043 (.663)	.097 (.009)
2		.092 (.014)	.080 (.048)	.155 (.000)	.163 (.000)	.106 (.003)	.064 (.178)	.132 (.000)
3			.162 (.000)	.226 (.000)	.238 (.000)	.179 (.000)	.129 (.000)	.183 (.000)
4				.089 (.020)	.103 (.004)	.054 (.364)	.042 (.689)	.064 (.178)
5					.033 (.908)	.057 (.292)	.118 (.001)	.070 (.119)
6						.078 (.056)	.129 (.000)	.085 (.029)
7							.078 (.056)	.049 (.493)
8								.077 (.066)

ρ_1 $(|r_{n,k}|)_{k=1,\ldots,81}$

	2	3	4	5	6	7	8	9
1	.054 (.357)	.034 (.886)	.036 (.853)	.156 (.000)	.211 (.000)	.231 (.000)	.199 (.000)	.220 (.000)
2		.049 (.493)	.042 (.689)	.143 (.000)	.195 (.000)	.210 (.000)	.174 (.000)	.207 (.000)
3			.047 (.541)	.155 (.000)	.195 (.000)	.217 (.000)	.183 (.000)	.203 (.000)
4				.137 (.000)	.191 (.000)	.212 (.000)	.170 (.000)	.202 (.000)
5					.089 (.020)	.103 (.004)	.071 (.103)	.097 (.008)
6						.047 (.541)	.054 (.364)	.038 (.786)
7							.078 (.056)	.031 (.937)
8								.077 (.066)

Table 5.3: Testing equality in distribution for standard processes among nine subsamples. For the five scale free statistics described in Section 5.2.1, the two-sample Kolmorogov-Smirnov test is applied to pairs of subsamples. The tables give the values of the test statistic and in parentheses p-values. A p-value smaller than 0.05 indicates statistical significance. The nine subsamples are given at Figure 5.4.



(a) Jan. 1988 – Apr. 1990 (b) Apr. 1990 – July. 1992 (c) Aug. 1992 – Nov. 1994

(d) Nov. 1994 – Mar. 1996 (e) Mar. 1997 – July 1999 (f) July 1999 – Nov. 2001

(g) Nov. 2001 – Mar. 2004 (h) Mar. 2004 – July 2006 (i) July 2006 – Dec. 2008

Figure 5.6: Kernel density estimates for $\log(RV5^m/RV5^a)$, for nine subsamples. Equality of distribution for $RV5_n^m/RV5_n^a$ is tested in the first subtable of Table 5.3. The applied kernel is the Gaussian kernel. The nine subsamples are defined in Figure 5.4.

5.3 Discussion

The present chapter shows that the scaling model gives a good fit to the data. We shall here reflect upon the use of the scaling model, and proxies, in the earlier chapters.

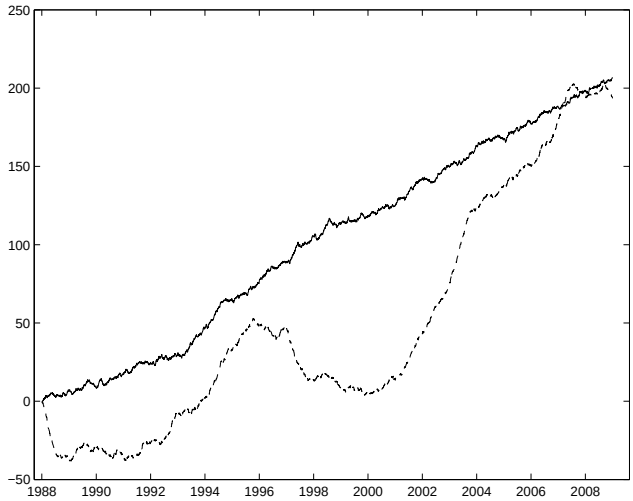


Figure 5.7: The cumulative of the morning/afternoon ratios $\log(RV5_n^m/RV5_n^a)$ (dashed curve). The full-sample mean and variance of $\log(RV5_n^m/RV5_n^a)$ are mean=0.037 and var=0.096. The solid curve shows a sample path of a Gaussian random walk with increments that have mean=0.037 and variance=0.096.

The fit of the scaling model is not perfect. The analysis in the previous section suggests that the sequence of processes $\Psi_n(\cdot)$ is not stationary, but changes over time. Given the time span of twenty one years one can think of many causes for changes in structure of the intraday return process. A technological revolution has taken place; connection speeds have increased, telephones have become less important. The Chicago Mercantile Exchange installed a fully electronic platform for futures trading, Globex, in 1992, and a second generation Globex in September 1998. There are institutional changes: exchange boards adjust trading rules and the cost of trading. In 1994 the market opened for overnight trading in the S&P futures. As of November 1997 the tick size for the S&P 500 future, the price series used in our analyses, increased from \$0.05 to \$0.10. The release-time schedule for important macro economic figures is not fixed in practice over such a long time period. For instance, in December 1993 the personal consumption expenditures announcement time moved from 9:00 A.M. to 7:30 A.M. Central Standard Time. The same happened to the business inventory announcement time in January 1997. The amount of capital available to different market participants changes. The invested amount of pension funds, day traders, market makers, and individual investors change depending on their past performance, or because of changes in regulation. Demography influences the amount of pension money invested in stocks. The share of algorithmic trading, with trading decisions and order execution driven by computer programs, has increased. Derivatives trading has increased, giving rise to more hedging trades in underlying values. In view of these considerations, it is remarkable that properties of the $\Psi_n(\cdot)$

are relatively stable over time. The relative stability is expressed well in the nine intraday variability patterns of Figure 5.4.

The scaling model has proven fruitful for dealing with high-frequency data in GARCH-type discrete time models. The model yields a tractable structure for volatility proxies. Underlying many results is that the proxy $H_n = H(R_n)$ can be written as

$$H_n = \sigma_n Z_{H,n}, \tag{5.6}$$

where the scale factor σ_n and the innovation $Z_{H,n}$ are independent. Moreover, the $Z_{H,n}$ are iid. Strictly speaking, the scaling model, $R_n = \sigma_n \Psi_n$, assumes more than we need to obtain (5.6). We do not need (5.6) to hold for *all* possible proxies H, only for the ones that we use. It may be worthwhile to verify the independence assumptions for the $Z_{H,n}$ of the proxies (5.6) directly, rather than the independence assumptions for the process Ψ_n.

Let us check equation (5.6) for some frequently applied proxies, using the S&P 500 data. Use the log-GARCH scale factor $\hat{\sigma}_n$ to define the estimated proxy innovation

$$\hat{Z}_{H,n} = H_n / \hat{\sigma}_n.$$

A key proxy in the thesis is the combined proxy $H_n^{(\hat{w})}$. It is the result of the optimized combination in Chapter 2; in later chapters it is used for parameter estimation and to measure weekly and monthly price fluctuations for forecasting. Other proxies that have gained special attention are the absolute daily return $|r_n|$, the daily high-low range hl_n, the absolute power variation RAV_n and its upward and downward decompositions RAV^+ and RAV^-.

Let us apply the diagnostic checks of the previous sections to the proxy innovations of $H_n^{(\hat{w})}$. Figure 5.8 shows the time series of the innovations $\hat{Z}_{H,n}$ for $H = H^{(\hat{w})}$. We shall now examine to what extent this time series may be regarded as a (non-zero mean) white noise sequence. The regression

$$\hat{Z}_{H,n} = c + b\hat{\sigma}_n + \varepsilon_n, \qquad 1, \ldots, 5174,$$

checks for a linear relationship between the proxy innovation and the scale factor (cf. Section 5.2.1). It yields an insignificant t-value for $b = 0$ equal to 1.58. Taking squares of either the regressor or the regressant also yields (R^2, t_b) equal to $(0.001, 2.03)$ and $(0.001, 2.02)$. These same exercises for $|r|$, hl, $RAV5$, $RAV5^+$, and $RAV5^-$ result in either insignificant t-values or a significant t-value with a modest R^2. The largest R^2 is smaller than 0.01 (for $RAV5$ the regression displayed above yields $(R^2, t_b) = (0.008, 4.96)$). In line with the assumption of independence between proxy innovation and scale factor, we find almost no relation between $\hat{\sigma}_n$ and the proxy innovations.

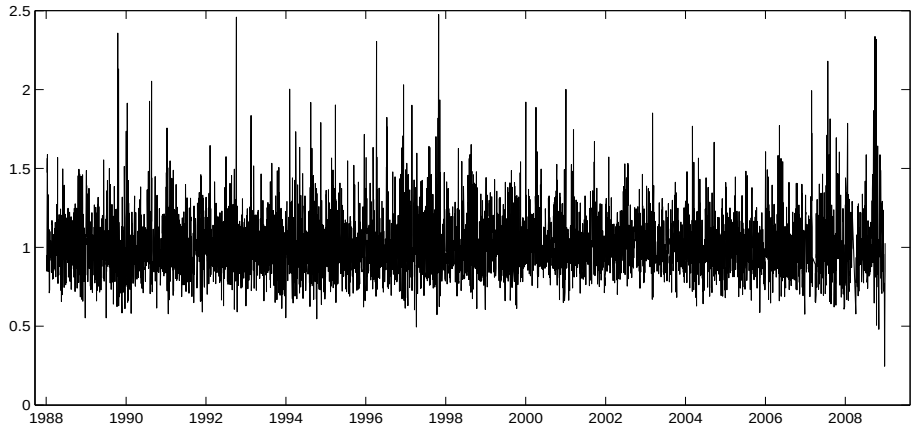


Figure 5.8: Time series of the proxy innovations $\hat{Z}_{H,n}$, for $H = H^{(\hat{w})}$.

	turning point		diff.-sign		Ljung-Box-sub	
H	$\bar{T}$	p-value	$\bar{S}$	p-value	$Q_{k=20}^{\text{sub}=50}$	p-value
$H^{(\hat{w})}$	0.660	0.510	2.528	0.011	20.9	0.162
$\|r\|$	2.045	0.041	0.457	0.647	20.6	0.254
$\|hl\|$	3.990	0.000	-0.072	0.942	22.2	0.010
$RAV5$	1.748	0.080	2.721	0.007	21.0	0.134
$RAV5^+$	0.396	0.692	-0.795	0.427	19.8	0.599
$RAV5^-$	1.451	0.147	-1.132	0.258	20.8	0.182

Table 5.4: Tests of independence (cf. Table 5.2) for the sequence of proxy innovations $\hat{Z}_{H,n}$, for several H.

Table 5.4 applies the same tests for independence as in Table 5.2 to the proxy innovations. We leave out the regular Ljung-Box values, since these are sensitive to nonstationarities as previously indicated in the discussion on Table 5.2. The turning point test for the high-low range has a p-value close to zero. This may be due to the small but significant first order autocorrelation of -0.05. This small autocorrelation cannot be expected to have practical relevance. As in Table 5.2, the overall picture is in line with the assumption of independence.

Table 5.5 addresses the question of stationarity of the proxy innovations for the same nine subperiods as in Table 5.3, for the four proxies $H^{(\hat{w})}$, $|r|$, hl, and $RAV5^-$ (these proxies make up the log-GARCH recursion). Most p-values are insignificant. The nonstationarity present in the sequence of standard processes $\Psi_n(\cdot)$ does not appear to translate into clear nonstationarities in the sequences of proxy innovations. As one explanation, our proxies place equal weight on each intraday time interval. This gives some robustness. For instance, a shift in price variability from the morning to the afternoon does not affect

these proxies.

Overall, we find that Equation (5.6) gives a good description of the proxies used in this thesis.

$H^{(\hat{w})}$

	2	3	4	5	6	7	8	9
1	.033 (.904)	.057 (.308)	.071 (.108)	.061 (.234)	.059 (.256)	.095 (.011)	.044 (.621)	.079 (.051)
2		.056 (.327)	.063 (.203)	.064 (.178)	.050 (.448)	.083 (.034)	.038 (.786)	.089 (.020)
3			.033 (.908)	.061 (.230)	.080 (.048)	.125 (.000)	.045 (.590)	.054 (.364)
4				.080 (.048)	.087 (.024)	.130 (.000)	.052 (.405)	.066 (.156)
5					.057 (.292)	.089 (.020)	.089 (.020)	.080 (.048)
6						.066 (.156)	.066 (.156)	.113 (.001)
7							.097 (.008)	.146 (.000)
8								.089 (.020)

$|r|$

	2	3	4	5	6	7	8	9
1	.046 (.572)	.042 (.682)	.037 (.825)	.082 (.038)	.098 (.007)	.091 (.016)	.053 (.381)	.079 (.051)
2		.064 (.178)	.037 (.831)	.054 (.364)	.066 (.156)	.059 (.259)	.038 (.786)	.052 (.405)
3			.047 (.541)	.089 (.020)	.101 (.005)	.090 (.017)	.064 (.178)	.087 (.024)
4				.064 (.178)	.082 (.041)	.077 (.066)	.050 (.448)	.073 (.089)
5					.033 (.908)	.038 (.786)	.063 (.203)	.033 (.908)
6						.033 (.908)	.080 (.048)	.042 (.689)
7							.071 (.103)	.052 (.405)
8								.056 (.327)

hl

	2	3	4	5	6	7	8	9
1	.044 (.615)	.051 (.429)	.040 (.749)	.037 (.828)	.048 (.525)	.088 (.023)	.060 (.252)	.057 (.310)
2		.068 (.136)	.056 (.327)	.042 (.689)	.052 (.405)	.071 (.103)	.066 (.156)	.035 (.872)
3			.035 (.872)	.064 (.178)	.057 (.292)	.094 (.012)	.035 (.872)	.087 (.024)
4				.054 (.364)	.057 (.292)	.094 (.012)	.045 (.590)	.063 (.203)
5					.049 (.493)	.075 (.076)	.056 (.327)	.040 (.739)
6						.050 (.448)	.059 (.259)	.070 (.119)
7							.097 (.008)	.059 (.259)
8								.082 (.041)

$RAV5^{-}$

	2	3	4	5	6	7	8	9
1	.034 (.899)	.057 (.306)	.041 (.720)	.076 (.070)	.102 (.005)	.089 (.019)	.043 (.665)	.086 (.027)
2		.073 (.089)	.056 (.327)	.070 (.119)	.090 (.017)	.073 (.089)	.040 (.739)	.082 (.041)
3			.043 (.639)	.117 (.001)	.146 (.000)	.129 (.000)	.094 (.012)	.132 (.000)
4				.101 (.005)	.132 (.000)	.101 (.005)	.077 (.066)	.110 (.002)
5					.063 (.203)	.085 (.029)	.082 (.041)	.047 (.541)
6						.082 (.041)	.080 (.048)	.064 (.178)
7							.059 (.259)	.078 (.056)
8								.070 (.119)

Table 5.5: Testing equality in distribution for proxy innovations $\hat{Z}_{H,n}$ among nine subsamples. For the proxy innovations of each of the four proxies $H^{(\hat{w})}$, $|r|$, hl, and $RAV5^{-}$ a two-sample Kolmogorov-Smirnov test is applied to pairs of subsamples. The tables give the values of the test statistic and in parentheses p-values. A p-value smaller than 0.05 indicates statistical significance. The nine subsamples are given at Figure 5.4.

Appendices

5.A Data

Our data set is the U.S. Standard & Poor's 500 equity index future, traded at the Chicago Mercantile Exchange (CME), for the period 1st of January, 1988 until December 31st, 2008. The data were obtained from Nexa Technologies Inc. (www.tickdata.com). The futures trade from 8:30 A.M. until 15:15 P.M. Central Standard Time. Each record in the set contains a timestamp (with one second precision) and a transaction price. The tick size is \$0.05 for the first part of the data and \$0.10 from 1997–11–01. The data set consists of 5322 trading days. We removed one hundred and thirty days for which the closing hour is 12:15 P.M. (early closing hours occur on days before a holiday). Eighteen more days were removed, either because of too late first ticks, too early last ticks, or a suspiciously long intraday no-tick period. These removals leave us with a data set of 5174 days with more than 15 million price ticks, on average approximately 3 thousand price ticks per day, or more than 7 price ticks per minute.

There are four expiration months: March, June, September, and December. We use the most actively-traded contract: we roll to a next expiration as soon as the tick volume for the next expiration is larger than for the current expiration.

5.B Turning Point, Difference-Sign, and Ljung-Box Tests

The turning point test, difference sign test, and Ljung-Box test can each be used for examining the assumption that a sequence of random variables is iid. These tests are usually interpreted in first place as tests for independence. Let us briefly discuss the tests, see also Brockwell and Davis (1991, § 9.4).

The turning point statistic for a series y_n is obtained by counting turning points T: one registers a count if either y_n is larger than both y_{n-1} and y_{n+1} or if it is smaller than both these values. The statistic $\bar{T} = (T - \mu_T)/\sigma_T$ denotes the normalized number of turning points. Here $\mu_T = 2(N-2)/3$ and $\sigma_T^2 = (16N - 29)/90$. A large positive value of $\bar{T}$ indicates that the series is fluctuating more rapidly than expected for an iid sequence, whereas a negative value may be due to positive correlation between neighbouring observations. The statistic $\bar{T}$ is asymptotically standard Gaussian distributed under the assumptions that per triple (y_{n-1}, y_n, y_{n+1}) its three elements are identically distributed, independent, and have a continuous distribution. Notice that the statistic aggregates a purely local property of the data; a turning point count is determined locally, i.e. only by

(y_{n-1}, y_n, y_{n+1}), without reference to the other values of the series (y_n). The turning point test is therefore insensitive to a structural break or a gradual change in the distribution underlying the observations.

The difference-sign statistic $\bar{S}$ is obtained by normalization of the number of times S that $y_n > y_{n-1}$. Set $\bar{S} = (S - \mu_S)/\sigma_S$ with $\mu_S = 1/2(N-1)$ and $\sigma_S^2 = (N+1)/12$. If the series (y_n) is iid then the statistic $\bar{S}$ is asymptotically standard Gaussian distributed. A large positive value may be due to an increasing trend in the data. As the turning point test, this statistic aggregates a purely local property. The test statistic is insensitive to the presence of certain cyclic components, or a structural break.

The standard Ljung-Box Q_k statistic is obtained by summing (and normalizing) the first k-th order squared autocorrelations,

$$Q_k = N(N+1) \sum_{i=1}^{k} \hat{\rho}_i^2(y_n)/(N-i),$$

where N denotes the sample size. A large value indicates significant autocorrelations. Under the null hypothesis that the y_n are identically distributed, and independent, the statistic Q_k has asymptotically a chi-square distribution with k degrees of freedom. The Ljung-Box statistic requires calculation of the sample mean and sample autocovariances. It therefore cannot be regarded an aggregate of solely local properties of the data. In contrast to the turning point test and the difference sign test, the Ljung-Box test is sensitive to nonstationarities: it tends to be large if there is a trend in the data, or a structural break. To minimize the effect of such slow nonstationarities we also calculate the Q_k statistics after splitting the sample into equally sized, adjacent subsamples. The sum of these Q_k statistics has asymptotically a chi-square distribution with degrees of freedom equal to k times the number of subsamples, for N tending to infinity, keeping the number of subsamples fixed. The statistic Q_k^{sub} denotes the average of the subsampled Q_k statistics.

5.C Log-GARCH Estimates

The estimated scale factor $\hat{\sigma}_n$ is given by a log-GARCH equation based on the indicators: high-low range hl, five-minute based downward absolute power variation $RAV5^-$ (the sum of downward absolute five-minute returns), and the combined proxy $H^{(\hat{w})}$. The proxy $H^{(\hat{w})}$ combines the sum of the ten-minute highs, the sum of the ten-minute lows, and the sum of the ten-minute absolute returns. The ten-minute high is obtained by the difference of the maximum of $R_n(\cdot)$ and the starting value of $R_n(\cdot)$ over the ten-minute interval in question. The lows are obtained similarly, and made positive by taking absolute values.

The proxy $H^{(\hat{w})}$ reads

$$H_n^{(\hat{w})} = (RAV10HIGH_n)^{1.04}(RAV10LOW_n)^{0.72}(RAV10_n)^{-0.76}, \tag{5.7}$$

which is a good volatility proxy for the S&P 500 index. The coefficients $\hat{w}$ are obtained by minimizing the measurement variance of the proxy, see de Vilder and Visser (2008). The indicators $H_{n-1,\text{Week}}^{(\hat{w}),\log}$ and $H_{n-1,\text{Month}}^{(\hat{w}),\log}$ are moving averages of $H^{(\hat{w})}$ representing week and month price fluctuations. The log-GARCH equation is now given by

$$\begin{aligned}\log(\sigma_n) &= \kappa + \alpha^{(1)}\, H_{n-1,\text{Week}}^{(\hat{w}),\log} + \alpha^{(2)}\, H_{n-1,\text{Month}}^{(\hat{w}),\log} + \alpha^{(3)} \log(hl_{n-1}) + \alpha^{(4)} \log(RAV5_{n-1}^{-}) \\ &\quad + \beta \log(\sigma_{n-1}).\end{aligned}$$

We estimate the full-sample parameter values by log-Gaussian QML in the same way as in Visser (2008a). For the S&P 500 index data over the years 1988–2008 we obtain the parameter values $\hat{\kappa} = 0.054$, $\hat{\alpha}^{(1)} = 0.210$, $\hat{\alpha}^{(2)} = 0.113$, $\hat{\alpha}^{(3)} = 0.110$, $\hat{\alpha}^{(4)} = 0.229$, $\hat{\beta} = 0.297$. These parameter values are used in determining the sequence $(\hat{\sigma}_n)$.