

Chapter 1

Introduction

Volatility refers to the degree to which financial prices fluctuate. Large volatility means that returns fluctuate over a wide range of outcomes. Figure 1.1 depicts the daily returns of the S&P 500 index over the 21 year period 1988–2008. A value of 0.01 corresponds to a one-day increase in the value of the index by one percent. Two “stylized facts”

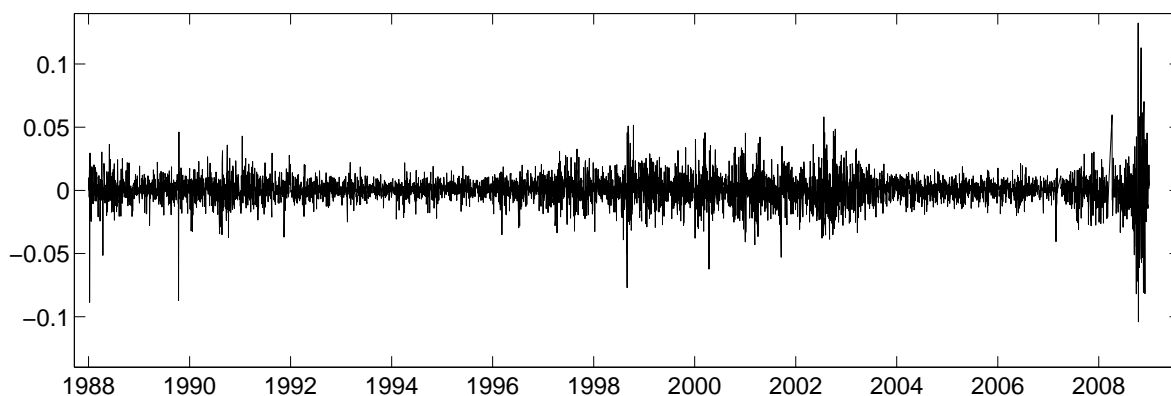


Figure 1.1: Daily returns for the S&P 500 equity index futures over the years 1988–2008, using $N = 5174$ trading days. The returns are the relative changes in daily closing prices. Depicted is the logarithmic change in price. A value 0.01 in the figure corresponds to a 1% increase of the index compared with yesterday’s closing price ($\exp(0.01) - 1 \approx 0.01$ up to a small difference). The S&P 500 index is arguably the most important equity index worldwide. See Chapter 5.A for details on the data.

of daily returns are: (1) returns are unpredictable, that is, it is not possible to forecast whether prices tomorrow will go up or down, and (2) financial markets pass through calm and hectic periods. Compare the calm years 1992–1996 and 2004–2006 with the hectic years 1997–2002 and with the period at the rightmost part of the graph. Zooming in on a period reveals, again, subperiods of relatively large fluctuations and subperiods of

smaller fluctuations; Figure 1.2 zooms in on the hectic period 1998–2001. Volatility is

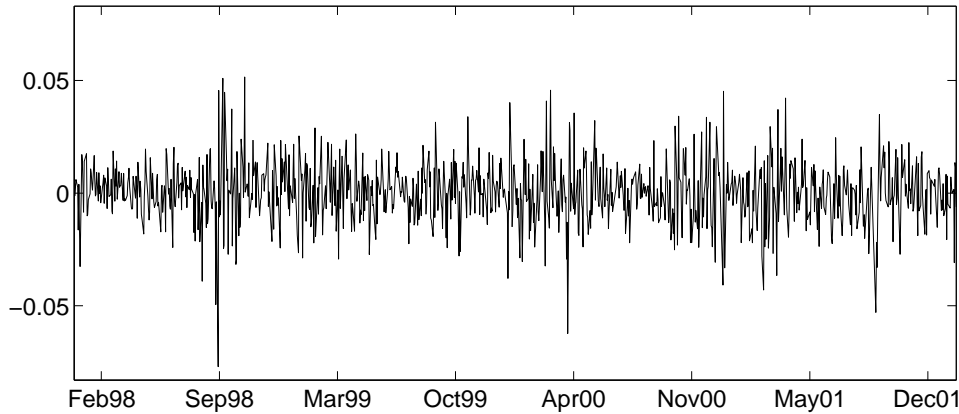


Figure 1.2: Daily S&P 500 returns, zooming in on the days $n = 2500, \dots, 3500$, in our sample.

not constant. One may say that “large changes tend to be followed by large changes – of either sign – and small changes tend to be followed by small changes” (Mandelbrot, 1963). In modern terms, we are dealing with *time varying volatility*. Volatility is persistent and therefore predictable to some extent.

This thesis is concerned with the modelling of daily stock market volatility using time series models. It centres around GARCH models (Generalized Auto Regressive Conditional Heteroscedasticity) and high-frequency data. GARCH is a discrete time model in which the volatility tomorrow is described in terms of the volatility today, and the observed returns. GARCH is a statistical modelling approach for which the initiator, Robert F. Engle, was awarded the 2003 Nobel Prize in economics. It is a standard model for the time varying fluctuations in daily returns.

Many financial data sets include intraday price movements in addition to the daily close-to-close returns. Such high-frequency data sets (containing thousands of transaction prices per day) are a useful source of additional information. This thesis develops a theory that incorporates the intraday price movements into discrete time GARCH models. The approach greatly improves GARCH volatility measurement and forecasting.

1.1 Background

1.1.1 Relation Between Volatility and Economic Fundamentals

Volatility is related to risk and uncertainty. It is an essential ingredient for answering fundamental questions in financial economics. By how much are investors rewarded for

taking risks? (cf. the Capital Asset Pricing Model, Sharpe, 1964, Lintner, 1965). Do prices reflect available information? (the Efficient Market Hypothesis, Fama, 1970). What is the relation between financial market volatility and the real economy?

Anyone with personal investment experience may have more than once wondered why, despite the feeling that fundamentally nothing much has happened over the past twenty four hours, prices are 3% higher or lower today than they were yesterday. Shiller (1981, 1989) posed the question what lies, ultimately, behind the day-to-day movements in prices. Can we trace the source of fluctuations in a logical manner to fundamental shocks affecting the economy, to natural resources, to monetary policy or other instruments of government control? Or are fluctuations also due to changes in opinion or psychology? Investors might become more risk averse when they experienced recent losses, and become more sensitive to news when larger-than-expected price changes occur (McQueen and Vorkink, 2004). Shiller argued that psychology *is* a factor of importance, by showing that stock market prices are much more volatile than is justified by the volatility of economic fundamentals. Investors in the stock market have only a vague idea what a stock is really worth. They may be able to judge whether one stock is overpriced relatively to another, but they have no clear idea how to judge the overall level of prices. As expressed by Bachelier (1900) in the opening sentence of his thesis: “Les influences qui déterminent les mouvements de la Bourse sont innombrables, des événements passés, actuels ou même escomptables, ne présentant souvent aucun rapport apparent avec ses variations, se répercutent sur son cours.”¹

What then, is the relation between aggregate stock market volatility and economic fundamentals? Surprisingly little is known on this matter. Schwert (1989) found only a weak relation between financial volatility and the movements of macro economic fundamentals, such as inflation, money growth, and industrial production. Volatility appears to be higher during recessions (see also Hamilton and Lin, 1996). A recent preprint by Engle, Ghysels, and Sohn (2008) suggests that taking account of macro economic variables when forecasting stock market volatility is useful for long horizons (six months and longer), but hardly improves standard models for forecasting over short horizons (one month). Changes in volatility are generally found to be related to the direction of price movements. Downward price movements go hand in hand with rising volatility, the so-called leverage and volatility feedback effects (Black, 1976, Christie, 1982, French, Schwert, and Stambaugh, 1987). Larger volatility over monthly horizons tends to be followed by larger volatility over weekly and daily horizons, but much less so the other way round (possibly due to the different trading time frames of different kinds of market participants, such as day traders, portfolio managers, and pension funds; Lynch and Zumbach, 2003). Larger

¹“The influences that determine the movements of the Exchange are innumerable; past, current and even anticipated events that often have no obvious connection with its changes have repercussions for the price.” (Davis and Etheridge, 2006).

financial market volatility goes together with larger trading volume (Clark, 1973, Gallant, Rossi, and Tauchen, 1992), with a greater number of trades (Ané and Geman, 2000), with lower liquidity (wider bid-ask spread, a decrease in depth of the order book; Ahn, Bae, and Chan, 2001), and with the arrival of more information (Chang and Taylor, 2003). Overall, there may be many sources for changes in volatility, but as of yet for a large (if not the larger) part there is no economic explanation of the level of volatility.

In standard economic theory lower stock prices decrease consumer expenditure (by a wealth effect). One may argue that in addition the volatility of financial markets influences the real economy. An increase in volatility means that companies with financial market exposure find themselves in a riskier environment. Faced with a more uncertain economic outlook, companies may experience more difficulties in raising capital. For individual consumers, the uncertainty around the value of pension savings increases. Risk averse consumers may want to postpone investment decisions. Choudry (2003) found that increased stock market volatility is predictive of a decline in the purchase of durable goods. Even consumers without direct stock market exposure may be less willing to spend, if these consumers regard stock market volatility as a leading indicator of income uncertainty. If increased volatility has such effects, it can be used for anticipating the future state of the economy. An account by Fornari and Mele (2008) suggests that financial volatility is predictive of economic activity 12 to 24 months ahead.

1.1.2 A Brief History of Financial Mathematics

The first known account that describes financial prices by a stochastic process is Bachelier (1900). Bachelier described price movements on the Paris stock exchange. He assumed that the price evolves as a continuous, memoryless process, with a fixed probability law for price changes. He observed that the Gaussian probability law fits these assumptions, establishing the stochastic process that is known as Brownian motion. The degree of price fluctuations depends on an unknown constant, the “coefficient d’instabilité,” that nowadays would be called the volatility and denoted by σ . The intuition behind Brownian motion resembles understanding the erratic movements of a pollen particle in water. The pollen particle moves randomly due to many small, independent, collisions with surrounding water molecules.² In finance one can think of many small, independent buyers and sellers pushing prices up and down.

The idea in applying Brownian motion, that financial price changes are unpredictable, fits nicely into economic theory. By the Efficient Market Hypothesis the best estimate of tomorrow’s price is today’s price. Its basic rationale may be described as follows (Bache-

²The Scottish botanist Brown observed such movements by microscope in 1828, hence the name Brownian motion. Einstein (1905), independently of Bachelier, introduced a mathematical model for this phenomenon. The first mathematically rigorous treatment of Brownian motion was given by Wiener (1923). See the historical account in Davis and Etheridge (2006, Chapter 3).

lier, 1900, Fama, 1965, and Samuelson, 1965a). There are many market participants who competitively attempt to incorporate information into asset prices. If prices are expected to rise, a sufficient amount of competition ensures that they would already have risen. Too obvious profit opportunities are arbitrated away. As a result price changes are unpredictable. Empirical research confirms that these claims are broadly true over short time horizons; decades of academic research did so far not yield any convincing evidence of the ability to predict whether prices will go up or down over periods of a day or a month (Cowles, 1933, is an early account).

The theory of stochastic processes permeates modern finance and financial mathematics. By the 50s stochastic process theory had made major progress. Brownian motion was well-defined, there was theory on stochastic integration (developed³ by Itô, 1944, 1951), and there was a well-developed theory on martingales (Doob, 1953). In the past fifty years in financial mathematics two decisive advances have been made: (1) the Black-Scholes option pricing theory, and (2) the empirical insight that, although returns are unpredictable, the absolute returns have positive autocorrelation (a sign that volatility is not constant, and tends to cluster).

Osborne (1959) and Samuelson (1965b) refined the Bachelier model to geometric Brownian motion, ruling out negative prices. Black and Scholes (1973) and Merton (1973), in their landmark theory, applied geometric Brownian motion in deriving the arbitrage-free price of financial derivatives, such as put and call options. They derive a remarkably tractable formula for the value of an option. Their derivations combine results from stochastic integration with the principle of no-arbitrage (which implies that two strategies with identical payoffs have the same value). To calculate a fair option price one needs, besides a few observables, only the value of the unknown constant σ , the volatility! By putting volatility at the centre stage of option valuation, Black-Scholes theory initiated a surge in research effort in volatility estimation.

The Black-Scholes model assumed constant volatility. This was considered a first approximation. Some early comments (Mandelbrot, 1963, Fama, 1965, Praetz, 1969, 1972, Officer, 1973) indicated that volatility may change over time.⁴ Black Monday, the worldwide stock market crash of October 1987, and its hectic aftermath showed that the assumption of constant volatility can be costly. Many financial institutions incurred losses larger than they had ever imagined. Brownian motion (implying that daily log-prices form a random walk with independent Gaussian increments) had proven a not very realistic model for financial return data. Volatility itself is a random process. Financial institutions and regulators evidently needed models that take into account time varying

³The recently discovered work by Doebelin predates Itô's work by a decade and gives the Itô formula in alternative terms.

⁴Nonetheless, as Taylor (2007) wrote, it was generally unknown that absolute or squared returns have positive autocorrelation when Taylor (1986) was published.

volatility.

Over the past twenty years new insights into the behaviour of volatility were gained. Estimates of GARCH models typically imply a contribution of today's price change to tomorrow's volatility in the order of two to ten percent (see Subsection 1.1.3). Over the last decade high-frequency data became widely available. As one insight due to these data, volatility is thought of today as a process that can change dramatically. Daily volatility can double within a few days (see Barndorff-Nielsen and Shephard, 2002b, Figure 2).

Besides in valuing options, volatility nowadays plays an important role in asset allocation (dating back to Markowitz, 1952) and risk management (e.g., McNeil, Frey, and Embrechts, 2005). Financial institutions seek favourable risk-return trade-offs. And, at the end of the trading day they want to have insight into the risk of their portfolio for the next trading day. For these purposes models with time varying volatility are useful.

The next two subsections discuss volatility modelling in more detail. Subsection 1.1.3 is concerned with GARCH models for volatility clustering (these are discrete time models). Subsection 1.1.4 discusses continuous-time semimartingale models, whose widespread use is due to the important role of Black-Scholes theory for pricing derivatives in modern finance.

1.1.3 GARCH Models

ARCH and GARCH models (Generalized Auto Regressive Conditional Heteroscedasticity) were introduced by Engle (1982) and Engle's former student Bollerslev (1986).⁵ Bollerslev (1987) applied the GARCH model to daily financial returns. French, Schwert, and Stambaugh (1987) applied the GARCH model to address the relation between risk and return in the stock market. In subsequent years several hundred papers appeared on the GARCH model; research is ongoing today. GARCH models can take (to some extent) account of three important stylized facts of asset returns:⁶

- returns are uncorrelated,
- returns are heavy tailed (have large extremes),
- extreme returns (as well as moderate ones) tend to cluster.

The widespread use of GARCH models is due to their easy applicability, and to the demand for models with non-constant volatility.

GARCH models are discrete time models. The basic GARCH(1,1) model, applied to financial returns, consists of the two equations (1.1–1.2) below. The first equation

⁵The early working paper by Rosenberg (1972) anticipates ARCH models, see Shephard (2005). The GARCH(1,1) model was independently introduced by Taylor (1986).

⁶For overviews of stylized facts of asset returns, see Cont (2001) Dacorogna et al. (2001), and Taylor (2005).

describes the return r_n as the product of the *scale factor* $\sigma_n > 0$ and a random noise term Z_n independent of σ_n ,

$$r_n = \sigma_n Z_n. \quad (1.1)$$

The sequence (Z_n) is assumed iid, mean zero, unit variance. The returns (r_n) fluctuate randomly (because Z_n is random noise). If the scale factor σ_n is large, then the return r_n tends to be large (in absolute value). So in this model the variables (σ_n) determine the sizes of the return fluctuations. The scale factors (σ_n) are *latent*, i.e. they are not observable; the data contain the returns r_n in equation (1.1), but not the model variables σ_n and Z_n . The GARCH(1,1) model⁷ assumes a simple recursion for the scale factor σ_n ,

$$\sigma_n^2 = \kappa + \alpha r_{n-1}^2 + \beta \sigma_{n-1}^2. \quad (1.2)$$

The scale factor is not a constant parameter: the variables σ_n form the terms of a discrete time stochastic process. In first instance one may think of the Z_n as independent standard Gaussian variables, and of the scale factor σ_n as the volatility underlying the return r_n of day n . If σ_0 is known, then (1.2) determines the day-by-day sequence $\sigma_1, \sigma_2, \dots, \sigma_n$ up to today (or tomorrow if today's return r_n is known). We can easily simulate the future behaviour of the sequence (σ_n) from a sequence of iid standard Gaussian variables Z_n . The parameters κ, α, β are positive. One often assumes $\alpha + \beta < 1$ for stationarity (e.g., for $\beta \geq 1$ the recursion (1.2) is explosive, Nelson, 1990b). The scale factors are determined by past fluctuations, they show persistence. The persistence is strong if $\alpha + \beta$ is close to one; the k -th order autocorrelations of σ_n^2 and r_n^2 can be shown to have a geometric decay as a constant times $(\alpha + \beta)^k$. The larger σ_{n-1} was yesterday, the larger σ_n will be today; the more extreme the price movement was yesterday (r_{n-1}), the larger σ_n will be today. GARCH is a simple model that takes account of volatility clustering.

In practice, the GARCH parameters κ, α, β may be estimated using standard estimation procedures. The parameter κ is generally found to be small. The parameter α is typically of order 0.05 and the sum $\alpha + \beta$ is large (close to one; larger than 0.95). The estimated recursion based on daily close-to-close returns for our full S&P 500 index sample reads (standard errors in parentheses⁸)

$$\hat{\sigma}_n^2 = \begin{array}{r} 8.9 \text{ e-}7 \\ (2.82\text{e-}7) \end{array} + \begin{array}{r} 0.054 \\ (0.008) \end{array} r_{n-1}^2 + \begin{array}{r} 0.939 \\ (0.008) \end{array} \hat{\sigma}_{n-1}^2. \quad (1.3)$$

Note that $\hat{\alpha} + \hat{\beta} = 0.993$ indicates strong persistence. Since $\mathbb{E}Z_{n-1}^2$ equals 1 the expected contribution of $r_{n-1}^2 = \sigma_{n-1}^2 Z_{n-1}^2$ to the squared scale factor σ_n^2 is roughly 5% (κ is of no

⁷A GARCH(p, q) specification would extend the recursion (1.2) to include p returns and q scale factors.

⁸Standard quasi-maximum likelihood estimates, with Bollerslev-Wooldridge (1992) standard errors.

importance here, since σ_n is of order 0.01). Most of the time the value of Z_{n-1}^2 is smaller than 4 (for the standard Gaussian distribution $Z_{n-1}^2 < 4$ more than 95% of the time). The contribution of yesterday's return to today's squared scale factor σ_n^2 is therefore in the order of five to twenty percent; to today's scale factor σ_n this is of order two to ten percent.

In the GARCH literature one commonly refers to σ_n as volatility. We use the more neutral "scale factor," as it is nowadays common in financial mathematics to use the word volatility in relation to the quadratic variation of the return process, as will be discussed in Subsection 1.1.4.

The unit of time for the returns is one day. This is the natural sampling frequency to be used in GARCH modelling as is clear when one examines autocorrelation functions. Figure 1.3 shows the empirical autocorrelations of the absolute five-minute returns for the S&P 500 index futures data. The lags vary from five minutes to twenty days. There

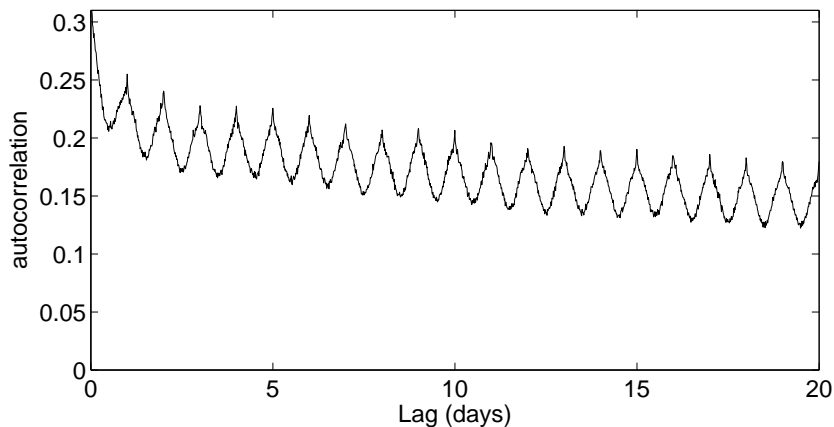


Figure 1.3: Empirical autocorrelations of absolute five-minute returns. Lags vary from 5 minutes up to 20 days (given 81 five-minute intervals per trading day, the rightmost autocorrelation is lagged by 1620 intervals). S&P 500 index futures data over 1988–2008 (5174 trading days). See also Chapter 5.1.

is a very pronounced periodicity of one day (for a detailed account see Andersen and Bollerslev, 1997). The GARCH model (1.1–1.2) does not take account of periodicities. To avoid the periodicity in the data the model is applied to daily returns. Intuitively the daily periodicity results from the 24 hour day schedule of everyday life. The trading day evolves during a window of a fixed number of hours; stock markets close at the end of the afternoon. Financial markets show pronounced intraday activity patterns (see Chapter 5.1). GARCH models applied to daily returns do surprisingly well at describing daily volatility, see Andersen and Bollerslev (1998a).

The GARCH model is a purely statistical model. There are no economic fundamentals in equations (1.1–1.2). The model does not include inflation rates, industrial production

figures, or GDP growth, only past returns. The fit of daily GARCH models hardly improves upon including macro economic data (Andersen and Bollerslev, 1998b). For portfolio managers and options traders it is attractive that to forecast volatility one can do with a simple model, based on historical prices only. The model does not exclude that economic fundamentals may be related to volatility. It assumes that all relevant information on fundamentals is already contained in the price process.

The term *conditional heteroscedasticity* is used since GARCH models have time varying conditional variances. Since σ_n is available from yesterday's information, one has the relations $\text{var}(\sigma_n Z_n | \mathcal{F}_{n-1}) = \sigma_n^2 \text{var}(Z_n | \mathcal{F}_{n-1}) = \sigma_n^2$. Here \mathcal{F}_{n-1} denotes yesterday's information, given by the σ -algebra generated by the returns $\{r_{n-1}, r_{n-2}, \dots\}$. So, the squared scale factor σ_n^2 is the conditional variance of the return:

$$\sigma_n^2 = \text{var}(r_n | \mathcal{F}_{n-1}). \quad (1.4)$$

There are many GARCH models around, see for instance the glossary to ARCH (Bollerslev, 2009) containing approximately 150 acronyms pertaining to different ARCH and GARCH models. These models typically share the product structure (1.1), but differ in specifying the GARCH recursion (1.2). We shall not provide an overview of the GARCH literature here. Overviews may be found in Bollerslev, Engle, Nelson (1994), and Taylor (2005). We also mention a related class of models, stochastic volatility models, in which the recursion for the scale factor is driven by a second source of randomness separate from the (Z_n) . These models date back to Taylor (1982); for overviews see Shephard (1996, 2005).

1.1.4 Semimartingales, Quadratic Variation, and Realized Volatility

Many financial data sets include the price movements over the course of the trading day. Our S&P 500 index future data set, for instance, contains on average more approximately three thousand price ticks per day (or 7 price ticks per minute). Continuous time models are naturally suited to take account of the informational content of these high-frequency data. Since the start of the century much research attention has been devoted to applying semimartingale models to high-frequency data. Two results from this research area are (1) high-frequency data may be used for estimating volatility using so-called realized volatility, and (2) AR(FI)MA-type models for realized volatility give impressive volatility forecasts. Let us now provide a more detailed discussion.

In spirit, a semimartingale is a process consisting of signal plus noise. In financial models the noise is usually large, i.e. the signal is locally of no importance compared

with the noise. Let t denote a real-valued time parameter. The continuous-time log-price process (Y_t) (also called log-return process) is a semimartingale if it is the sum of two processes (A_t) , the signal, and (M_t) , the noise,

$$Y_t = A_t + M_t,$$

where the process (A_t) has relatively smooth sample paths, and the process (M_t) is an erratic, unpredictable process. The increments in A_t may be thought of as rewards for investing in the risky process Y_t . In formal terms the process A_t is assumed to be of bounded variation, which means that, for finite time intervals, sums of the absolute values of the increments over subintervals are bounded; the process (M_t) is a local martingale (Protter, 2005). For econometric discussions, see Barndorff-Nielsen and Shephard (2002b), and Andersen, Bollerslev, Diebold, and Labys (2003). Semimartingales are mathematically interesting since they constitute the large class of processes with respect to which stochastic integration is well defined. They appeal to economic intuition as the assumption of no-arbitrage implies a semimartingale (for details, see Delbaen and Schachermayer, 1994). This is why they form a cornerstone in derivatives pricing.⁹

The class of semimartingales is large. Models in continuous-time finance typically fall into this class. Examples are Brownian motion, Itô processes, and Lévy processes (allowing for jumps). Thorough analyses based on S&P 500 index tick data do not reject the semimartingale hypothesis, see Peters and de Vilder (2006), and Andersen, Bollerslev, and Dobrev (2007).

For semimartingales, the natural measure for variability is *quadratic variation*. The quadratic variation of (Y_t) over the unit time interval, t in $[0, 1]$, may be obtained as follows. Divide the unit interval into m subintervals of equal length $\Delta = 1/m$. First define the *realized variance* $RV^2(\Delta)$ as the sum of the squared increments in Y :

$$RV^2(\Delta) = \sum_{k=1}^m |Y_{k\Delta} - Y_{(k-1)\Delta}|^2. \quad (1.5)$$

The square root of the realized variance is called the *realized volatility* RV . Quadratic variation¹⁰ QV may be defined as the limit (in probability) of realized variance as the sampling intervals approach zero,

$$QV = \text{p-}\lim_{\Delta \downarrow 0} RV^2(\Delta).$$

⁹There is recent interest in derivatives valuation with non-semimartingale models, e.g. Bender, Sottinen, and Valkeila (2008), and Jarrow, Protter, and Sayit (2009). An example of a non-semimartingale is fractional Brownian motion.

¹⁰In technical terms *optional quadratic variation*.

Realized variance with five and ten-minute intervals will turn up regularly in this thesis.

There is a natural reason for summing increments to the power of $p = 2$. In general for continuous semimartingales, the paths of Y_t are so irregular that for powers $p < 2$ the sum tends to infinity. Brownian motion, for instance, has paths of unbounded variation that are nowhere differentiable. Proceeding at a fixed speed it would take infinitely long to follow the Brownian sample path over any time interval. On the other hand, for powers $p > 2$ the sums tend to zero for all continuous semimartingales. Taking $p = 2$ yields a well-defined limit for all semimartingales. Here we have a nonparametric (model-free) definition of variability, with enclosed a practical prescription for obtaining it! The volatility for day n is nowadays interpreted as the square root of QV_n , the quadratic variation over day n ,

$$\text{volatility day } n = \sqrt{QV_n}. \quad (1.6)$$

In this way volatility is purely a sample path property. For Brownian motion the quadratic variation is deterministic. For standard Brownian motion QV_n equals 1 for all n . In Subsection 1.2.2 we shall discuss the relation between $\sqrt{QV_n}$ and the GARCH scale factor σ_n .

High-frequency data do not allow one to calculate quadratic variation to any desired degree of precision. Per day one may use the five-minute returns for determining the day's realized variance. In an ideal world, with a continuously observed asset price in a frictionless market, one could simply increase the sampling frequency, and obtain the quadratic variation. In real life it turns out that the use of intervals below a certain time unit (in the order of five-minutes for liquid equity markets) does not improve the estimates. On small time scales the price process shows a square saw-tooth effect that cannot be interpreted as a semimartingale. See Hansen and Lunde (2006b). It may be tempting to posit that quadratic variation is observed, up to "small" noise. Barndorff-Nielsen and Shephard (2002b) showed that realized variance may actually have quite large noise in practical situations (with confidence intervals for volatility larger than the level of volatility itself). So in practice volatility may not be treated as an observable.

There is now an active research area on estimating quadratic variation and related measures, such as bipower variation, absolute power variation, and realized range. Much of this research is concerned with the effects of market microstructure noise (bid-ask bounce, price discreteness, non-synchronous trading) and with the task of separating jumps from the continuous part of the sample paths. For overviews see McAleer and Medeiros (2008) and Andersen, Bollerslev, and Diebold (2009).

How does one set about forecasting tomorrow's volatility? It is unclear how to derive, only assuming a semimartingale, the optimal forecast of tomorrow's quadratic variation. Do we only need past quadratic variations for such a forecast? Does it make sense to

write

$$QV_n = f(QV_{n-1}, QV_{n-2}, \dots) + \varepsilon_n. \quad (1.7)$$

Equation (1.7) is essentially a reduced form of the log-price process Y . Researchers face two challenges in such forecasting questions: (1) formally, one has to deduce from the continuous time process (Y_t) the optimal forecast function f , and the time series properties of the errors (ε_n) , and (2) the daily quadratic variation is not observed; one can merely approximate it by realized variance. These questions have been addressed for a particular class of so-called eigenfunction stochastic volatility models. In this setup, a good approximation to the optimal forecast is obtained by applying ARMA models based on realized variance, see Meddahi (2003) and Andersen, Bollerslev, and Meddahi (2004).

Common practice is to take a pragmatic approach. Researchers replace the volatility \sqrt{QV} by realized volatility RV . Realized volatility is then used in regression type approaches, such as (log)AR(FI)MA models based on realized volatility, MIDAS regressions (Ghysels, Santa-Clara, and Valkanov, 2006), and HAR-RV (Andersen, Bollerslev, and Diebold, 2007). These forecasting schemes have shown impressive empirical success. As one may expect (since they use intraday data), such simple ARMA-type models outperform the GARCH-type models (based on daily returns) in out-of-sample forecasting, see Andersen, Bollerslev, Diebold, and Labys (2003), and Koopman, Jungbacker, and Hol (2005).

1.2 The Scaling Model and Volatility Proxies

The theory of semimartingales and quadratic variation offers a natural way of dealing with high-frequency data by using realized volatilities. We now have two measures for the day-to-day fluctuations of financial processes: the scale factor σ_n and the square root of the quadratic variation, $\sqrt{QV_n}$. In order to avoid confusion we henceforth use the term volatility only in the meaning it has in semimartingale theory, the square root of the increase in quadratic variation:

$\begin{aligned} \text{volatility} &= \sqrt{QV_n} && \text{(semimartingale)} \\ \text{scale factor} &= \sigma_n && \text{(GARCH)} \end{aligned}$
--

Both σ_n and $\sqrt{QV_n}$ measure the level of the fluctuations in log-prices. Both are model constructs. They can be approximated by statistics based on high-frequency data. Re-

alized volatility RV_n is commonly used to estimate $\sqrt{QV_n}$. In Subsection 1.2.2 we will introduce a class of statistics H , so-called proxies, to estimate σ_n .

One wonders how the GARCH scale factors σ_n compare to the realized volatilities RV_n . To this end we determined for the S&P 500 data the five-minute based realized volatilities $RV5 = RV(\Delta = 5 \text{ min.})$ based on 81 time intervals per trading day, and we determined the GARCH(1,1) scale factors $\hat{\sigma}_n$, using open-to-close returns¹¹ with parameter estimates over this subperiod $\hat{\kappa} = 4.5e-6$, $\hat{\alpha} = 0.082$, $\hat{\beta} = 0.890$. Figure 1.4(a) shows both $\hat{\sigma}_n$ and $RV5_n$ for the 1001 days numbered 2500 to 3500 in our sample. This comprises a relatively hectic period, so that differences between the two quantities are clearly visible. We show the values of $\hat{\sigma}_n$ and $RV5_n$ on a logarithmic scale. The two series

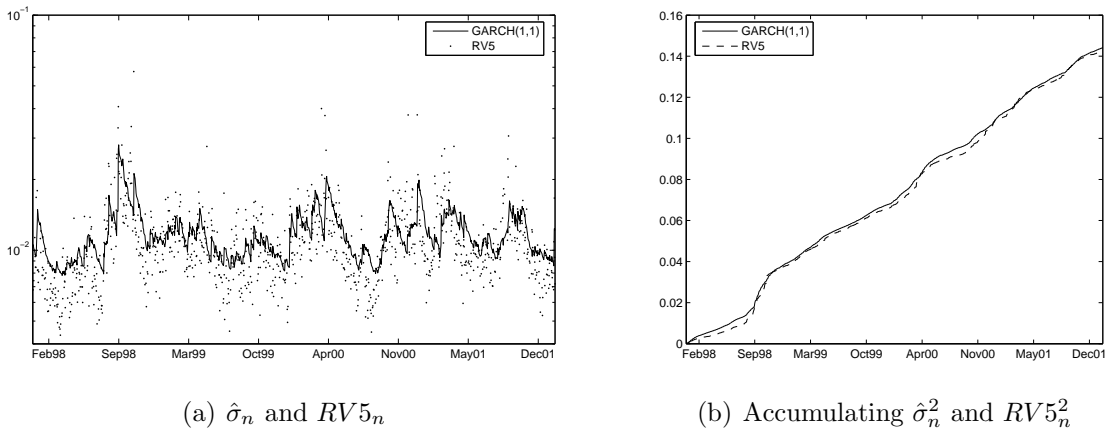


Figure 1.4: (a) The GARCH(1,1) scale factors $\hat{\sigma}_n$ (eq. (1.2)) and the five-minute realized volatilities $RV5$ (based on (1.5) using open-to-close returns and $\hat{\kappa} = 4.5e-7$, $\hat{\alpha} = 0.082$, $\hat{\beta} = 0.890$) for the S&P 500 over the days $n = 2500, \dots, 3500$, in our sample. The vertical axis has a logarithmic scale. The GARCH parameters κ, α, β were estimated using standard quasi-maximum likelihood over this subsample. (b) The cumulatives of $\hat{\sigma}_n^2$ and $RV5_n^2$.

show some resemblance. When $\hat{\sigma}_n$ is larger, then $RV5_n$ tends to be larger too. This is further stressed by Figure 1.4(b), which accumulates $\hat{\sigma}_n^2$ and $RV5_n^2$. This second figure shows that the longer term behaviour of both series is essentially the same. For those interested in the volatility over a longer period there is almost no difference. Recall that $RV5_n$ estimates the square root of the quadratic variation. The GARCH model uses only daily returns. GARCH does quite well, considering that in the semimartingale approach through quadratic variation the complete set of tick data is available for every day. One wonders how well the GARCH approach would do if high-frequency data would be incorporated. That is the question addressed in this thesis.

¹¹For r_n we took the open-to-close return so that the $\hat{\sigma}_n$ may be compared with the five-minute realized volatility, which is based on 81 five-minute returns during trading hours only.

This thesis introduces a continuous time asset price model that incorporates the intraday price movements into the discrete time GARCH model. Contrary to the semimartingale approach the quadratic variation now is no longer a key objective, but rather the discrete time scale factor σ_n is our object of interest. The approach is quite versatile. It does *not* make use of semimartingale theory. It allows one to address a variety of questions;

- Do high-frequency data give rise to improved parameter estimators for the GARCH model?
- Given a day's worth of data, what is the optimal estimator for the scale factor? (what is the optimal *proxy*?).
- How does one use the intraday price movements for improving standard volatility forecasts?

Each of these questions is the topic of a separate chapter of the thesis. The next subsection introduces the model. We shall also introduce a class of statistics based on intraday data, so-called volatility proxies. These statistics give an indication of the size of the fluctuations of the price process over the course of the trading day.

1.2.1 The Scaling Model

For each trading day, $n = 1, \dots, N$, the high-frequency data determine a continuous time process $R_n(u)$, $0 \leq u \leq 1$, the intraday 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 scale factor $\sigma_n > 0$ and a cadlag¹² process $\Psi_n(\cdot)$,

$$R_n(u) = \sigma_n \Psi_n(u), \quad 0 \leq u \leq 1. \quad (1.8)$$

One should compare this to equation (1.1). The *standard processes* $\Psi_n(\cdot)$ are assumed iid, like the Z_n in (1.1). As in GARCH models the scale factor σ_n is *latent*. We only observe the processes $R_n(\cdot)$, as expressed by the information set \mathcal{F}_n generated by $\{R_n, R_{n-1}, \dots\}$. In simulations the log-return process $R_n(\cdot)$ is determined in two steps. At first the scale factor σ_n is appointed. There are no constraints on the way in which the scale factors are generated. The scale factor is a positive variable. It may depend on the course of the intraday price process over the past days; it may also be purely random. We often assume stationarity of the sequence (σ_n, Ψ_n) . In the second step, the process $\Psi_n(\cdot)$ is determined, independent of the system so far. More precisely, the process $\Psi_n(\cdot)$ is independent of the

¹²The sample paths are right-continuous and have left limits.

σ -field \mathcal{G}_{n-1} generated by $\{\sigma_n, \sigma_{n-1}, \Psi_{n-1}, \sigma_{n-2}, \Psi_{n-2}, \dots\}$. In particular, for different days the standard processes Ψ_k and Ψ_n , $k \neq n$, are independent. There are no constraints on the probability distribution of $\Psi_n(\cdot)$. It may accommodate intraday volatility patterns. The process $\Psi_n(\cdot)$ may, but need not, be a semimartingale. By construction diurnal effects are taken into account in the scaling model.

The scaling model has a direct relation to the GARCH model. The scaling model yields daily close-to-close returns, $r_n \equiv R_n(1)$, that satisfy

$$r_n = \sigma_n Z_n, \quad (1.9)$$

by setting $Z_n \equiv \Psi_n(1)$. So the daily returns satisfy the canonical product structure (1.1) common to many discrete time volatility models. If the scale factors are generated according to the GARCH(1,1) recursion (1.2), then the full model reads

$$\begin{aligned} R_n(u) &= \sigma_n \Psi_n(u), & 0 \leq u \leq 1, \\ \sigma_n^2 &= \kappa + \alpha r_{n-1}^2 + \beta \sigma_{n-1}^2, \end{aligned}$$

where $R_n(1) \equiv r_n$. So we now have a way to incorporate the intraday price movements into discrete time models. The model is simple and practical. The dependence between different days is determined by the discrete time scale factors σ_n . So, even though we have a continuous time model, we may apply the well-developed and easy-to-apply theory on time series models to model the day-to-day dependencies.

To the best of our knowledge the scaling model (1.8) is a new asset price model. Only the special case of scaled Brownian motion (take $\Psi_n(\cdot) = W_n(\cdot)$ the standard Brownian motion) has been used as a financial model (Lildholdt, 2002, Brandt and Jones, 2006).

The scaling model is nonparametric in the sense that it does not impose constraints on the distribution of Ψ_n . It can therefore be a truly realistic model for the intraday log-price process. In principle one could use the data to obtain the empirical distribution of the processes Ψ_n . This thesis does not pursue this route. The theoretical results and the empirical applications in the thesis are obtained without having to know the distribution of the standard processes (Ψ_n).

Identification of σ_n and Ψ_n in (1.8) requires additional assumptions. Suppose that the scale factors satisfy the GARCH(1,1) recursion. The scale factor σ_n is then uniquely determined by the usual GARCH standardization,

$$\mathbb{E}\Psi_n^2(1) = \mathbb{E}Z_n^2 = 1.$$

The process Ψ_n can be obtained by descaling R_n ,

$$\Psi_n(\cdot) = R_n(\cdot)/\sigma_n.$$

In Chapter 2, on optimizing proxies for σ_n , we do not impose model assumptions on the scale factors σ_n . There the assumption of independence in (1.8) suffices to optimize proxies. So for some purposes identification is not needed.

As an alternative standardization it might be tempting to standardize the standard processes Ψ_n to have quadratic variation 1 over the time interval $[0, 1]$. Let us briefly indicate that this would be an unfortunate standardization. The process $\tilde{\Psi}_n$ with unit quadratic variation may be obtained from the original standard process by $\tilde{\Psi}_n = \Psi_n/\sqrt{QV(\Psi_n)}$. The quadratic variation then becomes part of the alternative scale factor $\tilde{\sigma}_n = \sigma_n\sqrt{QV(\Psi_n)}$. If $QV(\Psi_n)$ and $\Psi_n/\sqrt{QV(\Psi_n)}$ are dependent, then in general so are $\tilde{\sigma}_n$ and $\tilde{\Psi}_n$. This results in the representation $R_n = \tilde{\sigma}_n\tilde{\Psi}_n$ which in general does not satisfy the scaling model; it violates the independence in (1.8).

1.2.2 Proxies

The scale factors σ_n are latent. They have to be estimated from the data. The square root of quadratic variation $\sqrt{QV_n}$, or rather the realized volatility RV_n , is an obvious candidate. If Ψ is a standard Brownian motion then $QV(\Psi) = 1$, so $\sqrt{QV_n} = \sigma_n$. In this case the square root of quadratic is a perfect estimator. In general $\sqrt{QV_n}$ is not an optimal estimator; if the standard process has continuous sample paths of bounded variation, then $\sqrt{QV_n} \equiv 0$ is clearly not a good estimator of σ_n . For another example with $\sqrt{QV_n} > 0$, see Example 2.D.2.1 in Chapter 2.

What then is the relation between the scale factor σ_n and the volatility $\sqrt{QV_n}$? In the scaling model the scale factor σ_n determines $\sqrt{QV_n}$ up to multiplicative noise,

$$\sqrt{QV_n} = \sigma_n \sqrt{QV(\Psi_n)}.$$

This may be an explanation for their common time series behaviour as depicted in Figure 1.4. We can say more. Suppose $\mathbb{E}QV(\Psi_n) = 1$.¹³ Take for σ_n a GARCH scale factor (which is \mathcal{F}_{n-1} -measurable). The relation $QV_n = \sigma_n^2 QV(\Psi_n)$ now yields

$$\mathbb{E}(QV_n|\mathcal{F}_{n-1}) = \sigma_n^2.$$

A common interpretation (e.g., Andersen, Bollerslev, Diebold, and Labys, following Corollary 1, 2003) is that quadratic variation is a natural (ex-post) estimator for the squared

¹³One could achieve this standardization through dividing Ψ_n by $\sqrt{\mathbb{E}QV(\Psi_n)}$ (if $0 < \mathbb{E}QV(\Psi_n) < \infty$). Also, if Ψ_n is a martingale with $\mathbb{E}\Psi_n^2(1) = 1$, then $\mathbb{E}QV(\Psi_n) = 1$.

scale factor (or conditional return variance, cf. equation (1.4)). Another insightful interpretation is that the *squared scale factor is the best forecaster (in terms of mean squared error) of quadratic variation*. Quadratic variation and scale factor are two different concepts for price fluctuations, both having their own merits. Quadratic variation is a measure pertaining to the *sample path* fluctuations during the trading day. The scale factor is a *latent variable* that influences the size of the fluctuations during the trading day, and that predicts volatility $\sqrt{QV_n}$ (if QV_n exists). A forecast of a scale factor may therefore be regarded as a volatility forecast, even though the scale factor itself is not volatility. For this reason we shall speak of *volatility forecasting* when using the scaling model for prediction (Chapter 4).

As we have seen the square root of the quadratic variation is not necessarily an optimal estimator of σ_n , so it is natural to consider other estimators. We shall consider a general class of statistics H based on intraday data. Per day these statistics are applied to the sample path of the intraday return process. These statistics act as stand-ins for σ_n , and are called proxies. They have a particular form. Proxies may be seen as estimators. We prefer to use the term “proxy” rather than “estimator”: the proxies do not estimate a fixed population parameter, but instead approximate a sequence of unobservable random variables.

There are a number of intraday based statistics around to measure price fluctuations: the realized volatility (e.g. Barndorff-Nielsen and Shephard, 2002b, and Andersen, Bollerslev, Diebold, and Labys, 2003), and absolute power variation (Barndorff-Nielsen and Shephard, 2004), the intraday high-low range (Parkinson, 1980). The most basic example is the absolute daily return $|r_n|$. Many authors have motivated these statistics with a Brownian motion or a semimartingale in mind. These statistics all are *positive* and have the property of *positive homogeneity*: if the process $R_n(\cdot)$ is multiplied by a factor $\alpha \geq 0$, then so is the statistic H :

$$H(\alpha R_n) = \alpha H(R_n), \quad \alpha \geq 0. \quad (1.10)$$

In this thesis a *proxy* is a positive statistic that satisfies the positive homogeneity (1.10). We write $H_n \equiv H(R_n)$, and refer to both the statistic H and the variable H_n as proxies. The property of positive homogeneity fits naturally into the scaling model. Homogeneity implies $H(\sigma_n \Psi_n) = \sigma_n H(\Psi_n)$, and hence

$$H_n = \sigma_n H(\Psi_n). \quad (1.11)$$

Here, for each day n , $H(\Psi_n)$ is independent of σ_n . So in the scaling model a proxy equals σ_n up to multiplicative noise. The sequence of random variables $H(\Psi_n)$ is iid.

Equation (1.11) shows that a proxy provides information on the scale factor σ_n . If σ_n is large the proxy H_n tends to be large too. The proxy H_n has the product structure typical of discrete time models, cf. equation (1.1). Equation (1.11) underlies many of the results in the thesis.

The class of proxies defined by (1.10) is new. There are many positively homogeneous statistics. If one has two proxies, it is easy to create additional ones. The sum of two proxies is a proxy. So is the minimum or the maximum of two proxies. For two real-valued coefficients $w = (w_1, w_2)$, the geometric combination of two strictly positive proxies also is a proxy:

$$H_n^{(w)} = (H_n^{(1)})^{w_1} (H_n^{(2)})^{w_2}, \quad w_1 + w_2 = 1. \quad (1.12)$$

In this thesis we shall use proxies for two distinct purposes. First, as σ_n is not observed we need proxies to give us information on the value of σ_n . Proxies are used as stand-ins for the scale factors, i.e. as *measurement* variables. It is crucial to have good measurement variables to evaluate model fit, and to obtain sharp parameter estimates in GARCH models. The second purpose of proxies is in *forecasting*. The size of the fluctuations tomorrow may be influenced by specific aspects of today's price behaviour. One may use proxies that focus on different aspects of intraday price movements. We apply these proxies to improve forecasts.

1.3 Overview of the Thesis

The results of the thesis are presented in Chapters 2 to 5.

Chapter 2 addresses the problem of finding good proxies for σ_n , in the scaling model. A good proxy H_n is “close” to σ_n . To determine the quality of a proxy we examine the relative measurement errors U_n ,

$$U_n = \log(H_n/\sigma_n).$$

If $H_n = \sigma_n$ then $U_n = 0$. We compare the quality of different proxies by comparing their measurement variances λ^2 ,

$$\lambda^2 = \text{var}(U_n).$$

The smaller the measurement variance the better the proxy. It will be shown that there is no bias-variance trade-off: the class of proxies considered here ensures that proxies with

minimal MSE have minimal measurement variance. Proxies with small measurement variance constitute a time series that resembles the time series of scale factors (σ_n). Smaller measurement variances lead to more efficient GARCH parameter estimators (as will be shown in Chapter 3). We have seen that $\sqrt{QV_n}$ is in general not an optimal proxy. Chapter 2 shows that an optimal proxy exists, though we do not know how it should be constructed. How then do we know what constitutes a good proxy?

Suppose that we are given the data of the intraday log-return processes $R_n(\cdot)$. The observed processes $R_n(\cdot)$ are all we have. Suppose we are not allowed to make additional model assumptions for σ_n , nor are we allowed to assume a particular process for $\Psi_n(\cdot)$. Is it now possible to say which of the two proxies $H^{(1)}$ and $H^{(2)}$ is the better one? The remarkable answer of Chapter 2 to this question is: Yes! Moreover, it is possible to use the data to combine proxies in an optimal way into a single better proxy, by applying the idea of geometric combinations given by (1.12). This results in a nonparametric and easy-to-apply theory for ranking and combining volatility proxies. Empirical analysis based on the S&P 500 index tick data yields a proxy that outperforms five-minute realized volatility RV5 by 40%. The time series of this optimized proxy is plotted in Figure 2.2(d), page 43.

Chapter 3 deals with GARCH parameter estimation. For ease of notation we focus on the GARCH(1,1) recursion,

$$\sigma_n^2 = \kappa + \alpha r_{n-1}^2 + \beta \sigma_{n-1}^2,$$

as previously given in equation (1.2). Standard practice is to estimate the GARCH parameters (κ, α, β) by quasi maximum likelihood: specify a Gaussian likelihood for the daily returns r_n , and maximize the likelihood by varying the parameter values. It is well known that this procedure yields a quasi maximum likelihood estimator (QMLE) that is consistent and asymptotically normal, even if the distribution of the returns r_n is not Gaussian.

The chapter makes use of the likelihood theory in Straumann and Mikosch (2006). Their theory treats the QMLE based on the close-to-close returns r_n . The chapter generalizes the classical QMLE to a QMLE that takes account of the intraday price movements. The idea is to derive the likelihood for the proxy H_n . Since H_n extracts information from the intraday return process one may expect that it yields improved parameter estimators.

These ideas yield an easy-to-apply and powerful estimation theory. The asymptotic relative efficiency of the GARCH parameter estimators is determined by the quality of the proxy (see λ^2 above). We compare our generalized QMLE with the classical GARCH(1,1) QMLE for the S&P 500 data. There is a large efficiency gain by a factor 20. Simulations confirm that large gains may be expected.

Chapter 4 is concerned with the role of downward price pressure in volatility forecasting. Given a good proxy H_n it is natural to use this proxy for forecasting the value of σ_{n+1} . We introduce a log-GARCH recursion. In its most basic form our recursion reads

$$\log(\sigma_n) = \kappa + \alpha \log(H_{n-1}) + \beta \log(\sigma_{n-1}). \quad (1.13)$$

One advantage of the log-recursion is that it ensures positivity for σ_n . No positivity constraints on the parameters are needed; here κ , α , β are real-valued parameters. One may easily include multiple proxies in one equation. This gives great flexibility in specifying the forecast equation. Notice the difference compared with the standard practice of using only daily close-to-close returns, as in the GARCH(1,1) equation (1.2). One may now simultaneously include proxies that focus on different aspects of intraday price movements. It is natural to include a proxy H_{n-1} that gives a good measurement of σ_{n-1} . But other proxies, focusing on jumps, or on the downward price movements, will be shown to do better.

On the theoretical side we derive conditions for stationarity and invertibility using a theorem on the stationarity of autoregressive processes from Bougerol and Picard (1992). That is, we derive for what parameter values the stochastic system defined by the log-GARCH equations is well-behaved.

GARCH research based on daily close-to-close returns has widely documented the important *leverage effect*: downward price movements significantly increase future volatility (e.g. Nelson, 1991). One wonders how strong the effect of downward price movements is when high-frequency data are included. For this purpose Chapter 4 introduces a proxy that measures downward price pressure. The proxy is given by the sum of the downward five-minute returns over the trading day, and is called *downward absolute power variation*. Including the downward absolute power variation in a log-GARCH specification with several other proxies, we find that it is the most important variable for forecasting S&P 500 daily volatility one day ahead. Our analysis based on intraday high-frequency data shows how pronounced the effect actually is (it has a coefficient $\hat{\alpha} > 0.2$). The log-GARCH model using downward absolute power variation does well at forecasting. Evaluating more than three thousand out-of-sample forecasts using as a benchmark the optimized proxy obtained in Chapter 2, we find a coefficient of determination R^2 larger than 0.8 (the GARCH(1,1) forecasts have $R^2 \approx 0.7$). Figure 1.5 provides an analogue of Figure 1.4, but now the scale factors $\hat{\sigma}_n$ are estimated using the log-GARCH model. As one difference compared with Figure 1.4, the log-GARCH scale factors (using high-frequency data) capture changes in volatility more quickly than the classical GARCH(1,1) scale factors (which are based on daily returns).

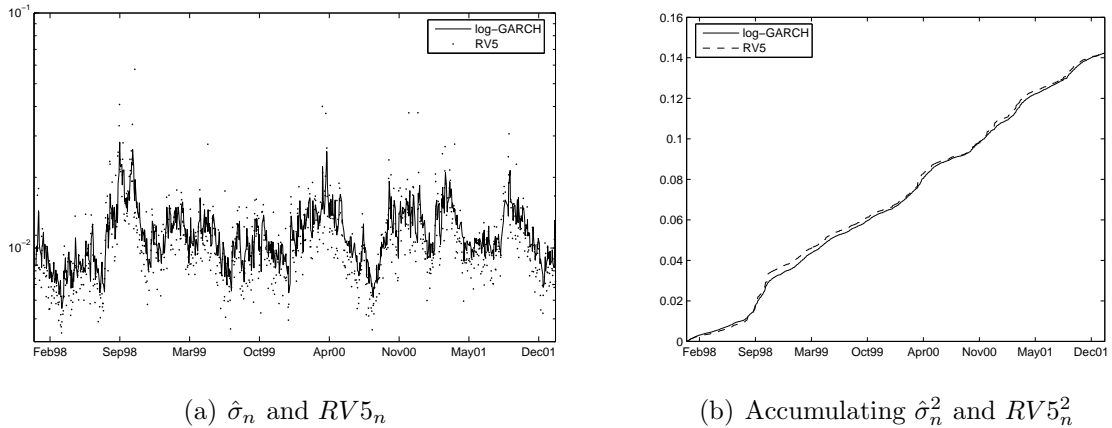


Figure 1.5: The same as Figure 1.4, but now with $\hat{\sigma}_n$ estimated by the log-GARCH model including downward absolute power variation. The log-GARCH specification is given in equation (4.20), page 96. The parameters for the subperiod considered here were estimated by log-Gaussian quasi-maximum likelihood: $\hat{\alpha}^{(1)} = 0.150$, $\hat{\alpha}^{(2)} = 0.114$, $\hat{\alpha}^{(3)} = 0.092$, $\hat{\alpha}^{(4)} = 0.274$, $\hat{\beta} = 0.320$. The effect of downward absolute power variation is estimated by $\hat{\alpha}^{(4)}$. The scale factors $\hat{\sigma}_n$ were rescaled such that the two curves in (b) have the same end points.

Chapter 5 is devoted to assessing the fit of the scaling model to the intraday return process. Since the model is nonparametric (no assumptions on σ_n , nor on the distribution of Ψ_n), 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 become statistically significant.

This chapter concludes that the scaling model gives an accurate description of the S&P 500 intraday return processes over the years 1988–2008. There is almost no dependence between the scale factor σ_n and the standard process Ψ_n . The assumption of independence cannot be rejected for the sequence of processes (Ψ_n) . We do find nonstationarity in the distribution of the Ψ_n . That is, the generating mechanism underlying the standard process Ψ_n is not the same in 2008 as it was in 1988. Further analysis shows that this nonstationarity has only a minor effect on the proxies used for obtaining empirical results in previous chapters.

Chapter 5 makes clear that the scaling model is a powerful tool for understanding financial price processes. It provides an opportunity to examine for instance intraday activity patterns, the persistence in intraday volatility, and the factors driving daily volatility.

1.4 Discussion

This thesis incorporates high-frequency data into GARCH models. There is almost no other research that incorporates high-frequency data into discrete time, daily GARCH models. Lildholdt (2002), and Brandt and Jones (2006) used the daily high-low range for estimating and forecasting GARCH volatility, assuming an intraday Brownian motion. Their model is therefore a special case of our scaling model. In fact Brownian motion is a deceptive special case. Here σ_n equals $\sqrt{QV_n}$, and the intraday quadratic variation process is deterministic. For Brownian motion, σ_n can be read off by looking at the first ten minutes of the sample path of the intraday return process. This is not the way we like to think of the scaling model. In general the scale factor σ_n is latent, and the intraday quadratic variation process evolves randomly.

There is other research on GARCH models and high-frequency data, but not on models that have GARCH returns at the daily frequency. Drost and Nijman (1993) showed that the GARCH model is not closed under temporal aggregation. One can estimate the parameters of a GARCH process with a five-minute time unit and derive the implied daily process, but the daily process is then no longer GARCH. If one applies GARCH models with time units that approach zero, one may obtain so-called GARCH and COGARCH diffusions (Nelson, 1990a, Klüppelberg, Lindner, and Maller, 2004, Kallsen and Vesenmayer, 2009). These diffusions have stationary increments. They are empirically unattractive as they do not allow for daily periodicity (cf. Figure 1.3). In general, modelling intraday high-frequency data by stationary processes leads to unreliable conclusions (Andersen and Bollerslev, 1997). Alternatively, the direct modelling of intraday activity patterns is unattractive since it soon involves hard-to-estimate models with many parameters. These insights seemed to end the hope of incorporating high-frequency data into GARCH models. Andersen, Bollerslev, Diebold, and Labys (2003) wrote:

“It has become apparent that standard volatility models used for forecasting at the daily level cannot readily accommodate the information in intraday data.”

Chapters 2 to 5 show that these words were too pessimistic. Intraday price movements can be easily, effectively, and realistically incorporated into daily GARCH models by using the scaling model presented in this thesis.

Subsection 1.1.4 described the current practice of dealing with high-frequency data in terms of semimartingales. Semimartingales are natural models for return processes. They conform to the no-arbitrage assumption, they describe the data in continuous time, they come with the elegant theory of quadratic variation, and they allow for many particular models (also processes with diurnal effects).

Semimartingale theory does have its limitations. The appealing theoretical property of observable volatility (\sqrt{QV}), does not carry over into practice since volatility is merely observed up to large noise (hence the large amount of research on realized volatility measures). The assumption of no-arbitrage may be troublesome. For equity tick data there seems to be a five minute barrier below which the data do not satisfy a semimartingale model (over short time intervals a semimartingale would resemble a martingale). It is hard to imagine fully efficient markets. Grossman and Stiglitz (1980) noted that “prices cannot perfectly reflect the information which is available, since if it did, those who spent resources to obtain it would receive no compensation.” Arbitrage is possible; it is what traders live by. As another inconvenience, the insight gained over the past twenty years that volatility itself is an extremely volatile process complicates option pricing seriously. The quadratic variation process, which is the monitor of the degree of price fluctuations, appears to fluctuate more heavily than the prices themselves on an intraday basis. Jumps also irreversibly complicate modern option pricing theory.

One serious drawback of the general semimartingale assumption is a lack of structure. High-frequency data show a clear periodicity in intraday volatility, but the general semimartingale assumption does not incorporate this explicitly. It does not take advantage of this structure. In principle semimartingale theory allows one to calculate weekly or hourly quadratic variation. Nonetheless, common practice is to estimate the *daily* quadratic variation. As Shephard (2005) wrote: “daily realised variations are somewhat robust to these types of intraday patterns, in the same way as yearly inflation is somewhat insensitive to seasonal fluctuations in the price level.” The semimartingale assumption enables the definition of quadratic variation, but for parameter estimation or volatility forecasting one has to turn to specific models. One such model is presented in this thesis. If $\Psi_n(\cdot)$ is a semimartingale, then so is the intraday return process.

Consider now the scaling model. The few parameters underlying the GARCH scale factors are easily estimated, and volatility forecasts are easily obtained by applying the GARCH recursion. The model tells one to work with *daily* volatility proxies. Procedures for ranking and combining proxies are available. The great efficiency increase in estimating parameters enables the estimation over a shorter history of the financial price process (this may be of value if there is nonstationarity). The scaling model comes without the cost of imposing model assumptions on the intraday price process.

We believe that the theory developed in this thesis may set the stage for additional interesting research. The improved GARCH measurements and forecasts allow us to gain deeper insights into the driving forces of asset prices. Improved volatility forecasts increase the quality of risk management, portfolio allocation, and option valuation. One interesting open question is how to derive theoretical option prices assuming the scaling model.

1.5 Reading Guide

Chapters 2–4 in the thesis each consist of a separate paper. Chapter 2 is based on the preprint de Vilder and Visser (2008), and Chapters 3 and 4 are based on the preprints (Visser, 2008b, 2008a). This has the disadvantage that readers may experience a certain amount of repetition, as in each chapter the scaling model and the definition of proxy are reintroduced. It has the advantage that each chapter is self contained. Readers having a particular interest in one of the chapters can simply start there. We have done our best to achieve a high level of readability. Derivations of technical results have been placed in appendices. In each chapter the theoretical results are also applied in an empirical analysis. The empirical analyses in Chapters 2–4 are based on the S&P 500 index data over the years 1988–2006. The analysis in the more recent Chapter 5 uses the data over 1988–2008. We hope that the main text is accessible to readers having a basic background in probability, stochastic processes, and time series analysis. Chapter 2 is in spirit rather abstract. It is concerned with ranking and combining proxies. Readers who prefer to start with a chapter that applies proxies in a practical situation are advised to start by having a look at Chapter 3.

As background material one may consult the GARCH overview paper by Bollerslev, Engle, and Nelson (1994), the book by Taylor (2005), the overview of volatility measurement and modelling by Andersen, Bollerslev, and Diebold (2009), and the Handbook of Financial Time Series (Andersen, Davis, Kreiss, and Mikosch, 2009).