Chapter 5
Process Overview

So far, the focus was on extracting as much information as possible from the financial time series, with a special concern on generic features. This overview of the "stylized facts" sets a benchmark against which various processes can be gauged. The next goal is to construct mathematical processes that can reproduce some, possibly all, stylized facts. This is of primary concern for many areas in finance, for example, portfolio construction, derivative pricing, or risk control.

This chapter gives a first "grand tour" of the existing processes used to model financial time series. The goal is to set the notations and basic ideas, as well as to present the *structures* for the main models used in finance. Then, subsequent chapters delve into more details and extensions of the basic processes. The challenge then is to be realistic, namely to reproduce as many stylized facts as possible. This will require more complex structures, but possibly some empirical features cannot be reproduced at all inside a given structure. The "mug shots" give a convenient summary of the main properties, either of the empirical data or of the processes.

5.1 Why Using a Finite Time Increment for the Processes?

The process equations are given in this book in a discrete form, with a finite time increment δt. There are several reasons for this choice.

- The empirical data are known with a finite time increment, mostly of one day, possibly at a shorter time increment with tick-by-tick data. The processes should reproduce the empirical time series at this time scale and longer, and the behavior of the data and processes at shorter time horizon than δt is irrelevant, as unaccessible empirically.
- A continuous-time process needs to be discretized in order to perform Monte Carlo simulations, to compute a forecast, or to evaluate a figure of merits like a log-likelihood. The discretization of a continuous-time stochastic equation is a nontrivial business. Because in general stochastic equations involve terms of

G. Zumbach, *Discrete Time Series, Processes, and Applications in Finance*,
Springer Finance, DOI 10.1007/978-3-642-31742-2_5,

order $\sqrt{\delta t}$ and δt, inequivalent discretizations can be constructed. A good example is the Heston (continuous-)time process which can be discretized in different ways, with different properties with respect to positivity [73].

- The continuous-time limit of some discretized process may not exist, or be quite different to a "naive" limit. An example of a process with a nonobvious limit is the GARCH(1, 1) process [45, 115]. At a deeper level, the description of financial time series at the tick-by-tick time horizons involves a new phenomenology, related to order queue, limit orders and market orders, and to the exchange of information [46, 144, 162]. This new phenomenology is important at time horizons ranging from a few seconds to a few hours. Only at longer time scales, the details of the price formation mechanism are small enough to be safely neglected. This very short term effects are not included, and are not needed, at a mesoscopic scales ranging from a few hours to months. Therefore, the continuous limit of a mesoscopic model is anyway missing some part of the price formation mechanism.
- The comparison of the process properties with empirical data shows that for *all* processes, fat-tailed innovations should be used. The most common choice is to use Student innovations with a number of degrees of freedom in the range from 3 to 7, albeit other distributions are possible. The Student distribution is not a stable distribution, namely the sum of independent Student variables is not distributed according to a Student distribution. For this reason, it is not possible to construct a process where the increments over several time intervals are distributed according to a Student distribution. The Student distribution has a weaker property, namely is infinitely divisible. An infinitely divisible distribution $P(y)$ is such that for any integer n, a probability distribution $p_n(x)$ can be found so that the sum $y = \sum x_i$ of n independent variables $x_i \sim p_n(x)$ is distributed according to the distribution P. This condition is needed to construct Lévy processes (see, e.g., [44]). Therefore, a Lévy process can be constructed so that the innovations over one selected time interval δt are distributed according to a Student distribution. But the independence condition for the increments leads to a central limit type convergence toward a normal distribution for longer time intervals.

In the continuum formulation, the Itô calculus is based on Wiener processes. The Itô calculus is very important to price derivatives and to formulate rigorously the associated replication strategy. In order to price contingent claims in a Black–Scholes scheme, the process equations can be fairly general, but the source(s) of randomness has to be infinitely divisible. The scheme can be generalized to accommodate random volatility, processes with jumps or a broader class of distributions (but still with infinite divisibility) [44]. Yet, stochastic calculus cannot accommodate dependency involving the process history like in a GARCH model.

In a discrete-time formulation, there is no rigorous Itô calculus, and this creates a problem with respect to option pricing as an important computational tool is now missing. Yet, if the time increment is sufficiently small, a "physicist" approach can be used in algebraic manipulation by using the simple rule that $E[dz^2] = dt$ and neglecting higher-order terms. Therefore, the key issue is not the continuum limit *per se*, but that many interesting process equations do not admit a straightforward continuum extension. Yet, it would be wrong to select a particular process structure because of its mathematical tractability in an idealized continuum limit. It is to the mathematics and our models to adapt to the empirical properties of the financial time series, as clearly the financial world will not change just to conform to an elegant mathematical idealization. The formulation of the process equations directly in a discretized framework avoid these difficulties. Recent progresses for option pricing has been made in order to accommodate a broader class of volatility processes like GARCH [40] and fat-tailed innovations [116]. These formulations are based on a discrete-time increment for the underlying processes and are presented in Chap. 16.

5.2 The Definition of the Returns

The vast majority of the time series models in finance are derived from a random walk for the logarithm of the price

$$\log\bigl(p(t+\delta t)\bigr) = \log\bigl(p(t)\bigr) + r(t+\delta t). \tag{5.1}$$

The change of variable $x = \log(p)$ makes this equation slightly simpler. The (logarithmic) return is then

$$r(t+\delta t) = \mu(t) - \frac{1}{2}\sigma_{\mathrm{eff}}^2(t) + \sigma_{\mathrm{eff}}(t)\epsilon(t+\delta t) \tag{5.2}$$

with μ fixing the drift and σ_{eff} the effective volatility. The Itô term $-1/2\sigma_{\mathrm{eff}}^2$ is such that the mean expected price is μ. In general, the innovations ϵ are i.i.d. random variables with

$$E[\epsilon(t)] = 0 \tag{5.3}$$

$$E\bigl[\epsilon^2(t)\bigr] = 1 \tag{5.4}$$

but the distribution is otherwise not specified. Both conditions are needed so that the drift μ and volatility σ_{eff} can be identified.

This is not the only way to write the basic random walk equation for a finite discretization step δt. A slightly different discretization is given by

$$p(t+\delta t) = p(t)\bigl(1 + r(t+\delta t)\bigr) \tag{5.5}$$

with the (relative) return

$$r(t+\delta t) = \mu(t) + \sigma_{\mathrm{eff}}(t)\epsilon(t+\delta t). \tag{5.6}$$

Up to higher-order terms in δt, both discretizations are equivalent, and therefore they share the same continuum limit. Yet, they have different properties for fat-tailed innovations, and in particular the discretization in term of the logarithmic return leads to diverging expectations. This issue is explored in more details in Chap. 6.

The Gaussian random walk corresponds to no drift $\mu = 0$, a constant volatility $\sigma_{\text{eff}}(t) = \sigma_\infty$, and to a normal distribution for the innovations $p(\epsilon) = N(0, 1)$. An example of returns for a Gaussian random walk is given in Fig. 5.4 in the top panel. This random walk model was written by Bachelier in 1900 [12], but for the price (i.e., with p instead of $x = \ln(p)$ in Eq. (5.1)). It is only in the 1960s that it becomes clear that the model should be written for the logarithm of the price. This change of variable can be deduced from a change of numeraire argument, namely the change $p \to \alpha p$ for any positive α corresponds to a change of the currency unit, and this change should not affect the economy. The logarithm transformation changes the scaling by α into an additive transformation, and this leaves the logarithmic price differences $r(t) = x(t) - x(t - \delta t)$ invariant for any α. This argument can be used for any contract which basically trades the numeraire, like stocks, stock indexes, or FX rates. For other time series, like interest rates, implied volatilities or spreads, the change of numeraire argument cannot be used (even if denominated in a currency unit, a bond price is trading the underlying interest rates). For these other time series, the change of variable that transforms the time series into a random walk is open as the change of numeraire argument cannot be used. For example, with interest rates, [154] uses the condition that the volatility $\sigma = \sigma(x)$ computed with the transformed price should be uncorrelated with the price, namely $\rho(p, \sigma) \simeq 0$.

5.3 The Most Important Stylized Facts

As a daily return is the sum of many intra-day transactions, the central limit theorem justifies to model the returns with a Gaussian distribution, at least at first glance. In the other direction, as a Gaussian random variable can be decomposed as a sum of random variables distributed according to a Gaussian law, a continuum limit can be constructed where the infinitesimal random variables are Gaussian. This construction is such that, at any time scale, the corresponding random variable is Gaussian. This construction can be made rigorous, leading to Wiener process. Yet, as explained in the later sections, the relationship between return distribution and the central limit theorem is quite complex because the returns are not independent. Independence is one of the prerequisites for the central limit theorem to hold, and as the dependency between returns is subtle, the return distribution is converging to a Gaussian but at a much slower pace.

The basic Gaussian random walk is very successful at giving a first good description of the behavior of the prices, and today a large part of finance is relying on this model. However, a simple visual comparison of returns time series shows that this model is deficient on two important aspects.

1. The volatility is constant, namely the time series is homoscedastic.
2. The return distribution is Gaussian instead of the observed fat-tailed distribution.

Both points can be seen directly on plots of the returns as function of time, like in Figs. 5.1, 5.2, and 5.3 for empirical time series and on Fig. 5.4 for some popular

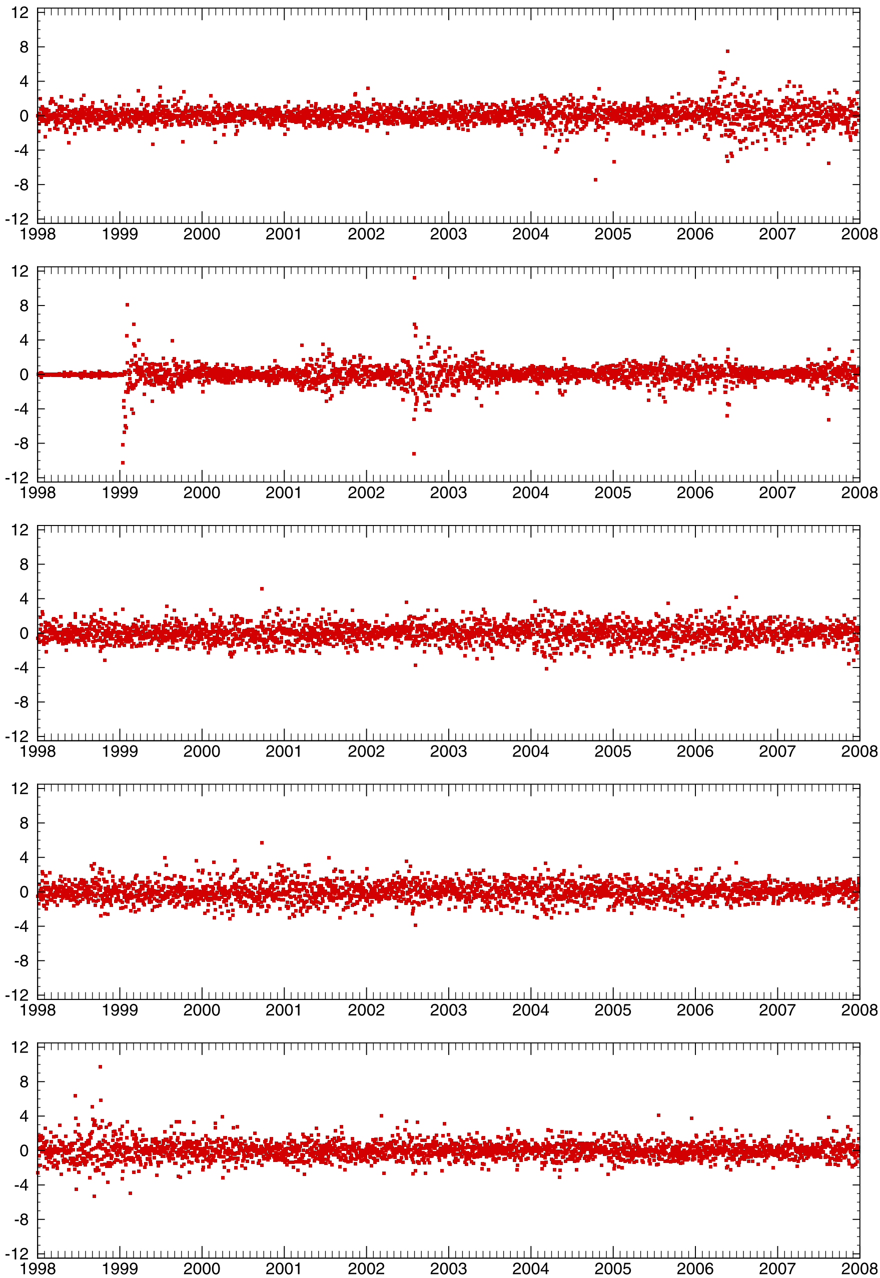

Fig. 5.1 The standardized return for one commodity and some FX against the USD (from top to bottom): 3 month future on copper, BRL/USD (Brazil), GBP/USD (GB), EUR/USD (Euroland), and JPY/USD (Japan)

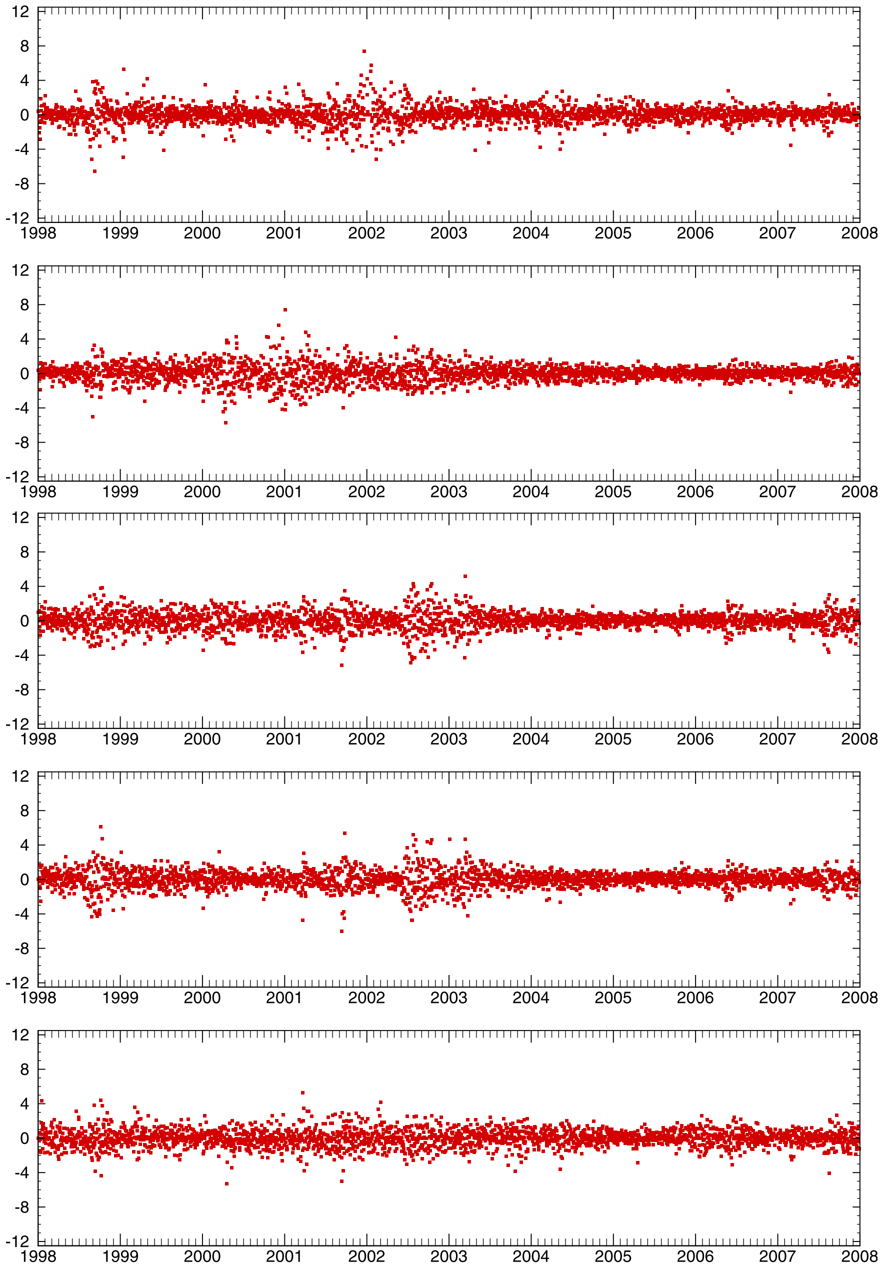

Fig. 5.2 The standardized return for several stock indices (from top to bottom): Merval (Argentina), Nasdaq (USA), FTSE-100 (GB), SMI (Switzerland), and Nikkei-225 (Japan)

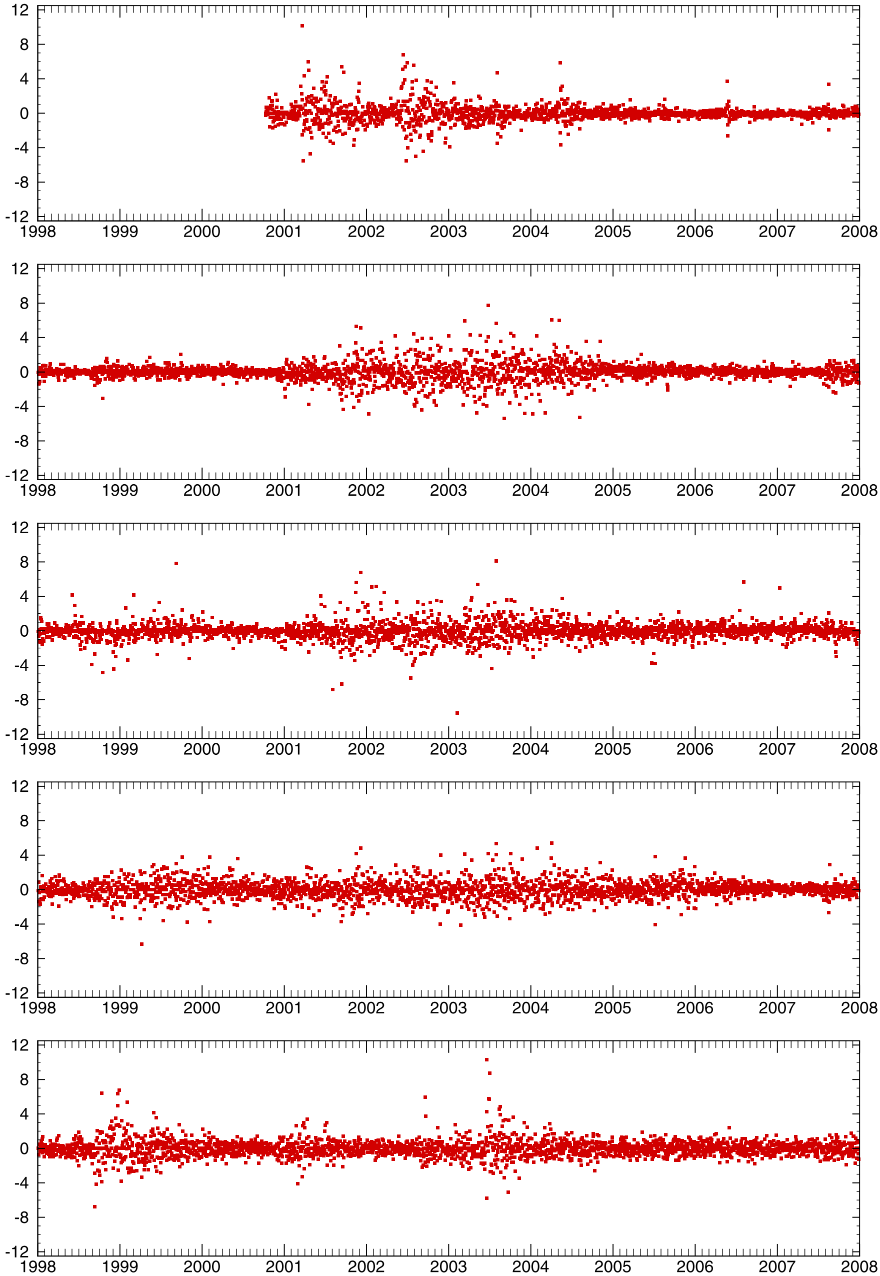

Fig. 5.3 The standardized return for several interest rates (from top to bottom): Brazil at 1 year, US at 1 year (interest rate swap), UK at 1 year (interest rate), Euro at 1 year (Government debt), and Japan at 10 years (Government debt; this period had very (very) low rate in Japan for IR up to 1 year)

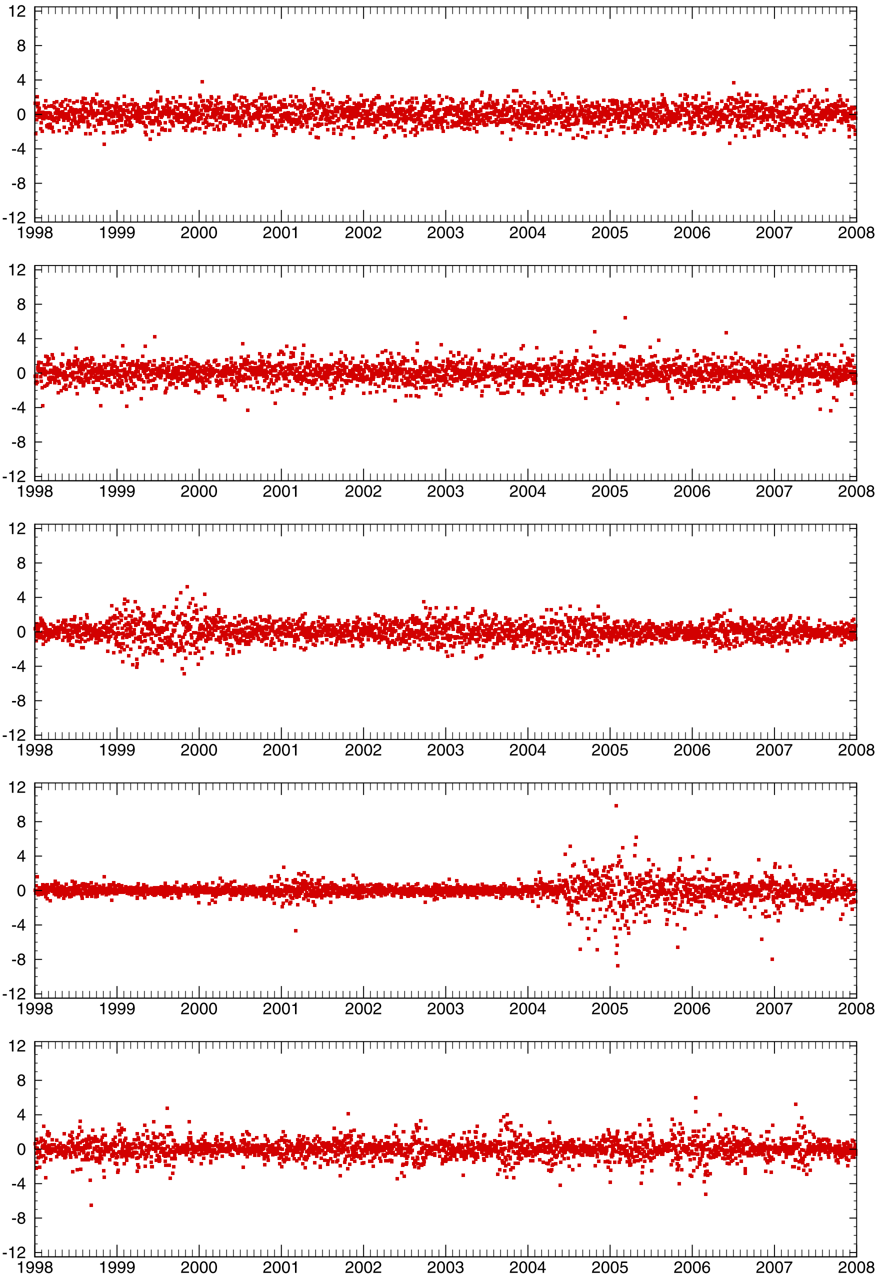

Fig. 5.4 The standardized return for a few processes (from top to bottom): constant volatility with normal innovations, constant volatility with Student innovations, GARCH(1, 1) with normal innovations, GARCH(1, 1) with Student innovation, and Heston process with normal innovations. Student innovations have 5 degrees of freedom, and the characteristic time of the GARCH and Heston processes is of one month (21 business day)

theoretical processes. For all plots, the returns are normalized using the empirical standard deviation computed on the visualized sample. In this way, the returns are displayed directly in terms of the sample empirical standard deviation. The empirical time series have been chosen to display many geographic areas and many instruments types over the same 5 year period, but have not been selected specifically otherwise. When visually comparing a Gaussian constant volatility random walk (Fig. 5.4, top panel) with the empirical time series, the deficiencies can be directly seen without any statistical analysis. Let us discuss both points in turn.

The magnitude for the returns is not constant, but there are periods of low volatility and periods of large volatility. The figure is at a time scale of half a decade, and the cluster of volatility with a characteristic length between months and years are clearly visible. If only one year of data is used for the figure, a similar picture emerges, but with a characteristic length for the clusters between weeks and months. This intuitively shows that there is not a single characteristic time for the volatility clusters, but that many time scales are needed to describe their dynamic. The multiple time scales structure is quantified by the lagged correlations for the volatilities, which show a slow decay (and not an exponential decay). This important empirical stylized fact was discussed in Chap. 3 and is illustrated by two graphs in the mug shots.

The range for the y-axes $(-12, +12)$ has been chosen so that the largest standardized daily return appears on the graphs. This range is very wide, and many daily returns are beyond a 6 sigma level. On the other hand, the Gaussian random walk does not display any point above a 5 sigma level. This simple comparison shows clearly that the returns have a fat-tailed distribution.

Both features need to be quantified by a serious statistical analysis of empirical data, yet the dominant characteristics of any financial time series can be seen with naked eyes on most graphs of the time series of returns. The next question is how to incorporate them in a process used to describe the evolution of prices. The easiest part is the fat tail, as it is enough to use fat-tailed innovations. More complex mechanism can be used, as, for example, in a GARCH process, but in practice all realistic models need to include fat-tailed innovations. The more complex part is the heteroscedasticity: to capture this feature, some memory from the past volatility needs to be included in the price process. Essentially three paths can be taken that are introduced in the next three paragraphs. A full discussion of each process classes are presented in the subsequent chapters, with Monte Carlo simulations of the key processes and the corresponding mug shots.

5.4 ARCH Processes

In a (G)ARCH process, the volatility is a function of the past return magnitudes, and this dependency on the past brings in the volatility clustering. The simplest model is to have the squared volatility (i.e., the variance) to be a weighted sum of past squared returns. For most ARCH processes, this dependency is induced by exponential moving averages of the squared returns. This structure provides for a efficient mean to construct estimators of the past volatilities at various horizons, but the dependency of the past returns can be more complex. Ultimately in a GARCH process, the volatility is a function of the past (squared) returns. The key point is that the volatility can be removed completely from the equations and the magnitude of the next return is a function of the past returns only. In this family of processes, the volatility is a parsimonious way to construct the equations and to interpret them, but is nothing more than a convenient intermediate variable that could be eliminated from the equations.

5.5 Stochastic Volatility Processes

In a stochastic volatility process, the variance has a higher status as it becomes an integral part of the process with its own source of randomness. The heteroscedasticity is introduced by a dependency of the volatility on its past, typically by a moving average. As the volatility has to be mean reverting, a natural candidate is an Ornstein–Uhlenbeck process. Then, the return appears as a slave process, with the magnitude given by the volatility process. In particular, there is no feed-back from returns to the volatility. Many variations can be constructed around this theme, while preserving the core feature of a volatility depending on its past, but not on the returns.

The process for the volatility has to preserve the positivity of the volatility. One simple way to enforce this restriction is to write a process for the logarithm of the volatility. Another way is to modify the amplitude of the random term by a power of the volatility, so that when the volatility decreases, the random term decreases, leaving the mean reverting term to "pull up" the volatility. Both families will be explored.

A closely related family of processes is based on the idea of subordinating the return to a random time. The time is viewed as progressing randomly, for example, increasing according to the events on the market. The process for the time can depend on its own past, in order to introduce heteroscedasticity. Then, the return follows a simple random walk with constant volatility but is indexed by the random time. This setup is in fact similar to a stochastic volatility process, with the volatility acting like the "speed of time". This equivalence is explained in Sect. 8.1.1.

5.6 Regime-Switching Processes

In a regime-switching process, a new hidden stochastic variable is introduced. This variable define the "state of the word". The state of the world follows its own process, typically given by a Markov chain over a discrete space with only a few states, for example, with two states taken as "quiet" and "crisis". The volatility takes values depending only of the state of the world, and the return follows a simple random walk with the given state dependent volatility. In this family, the heteroscedasticity is introduced by the Markov chain for the state of the world, while the volatility and return are simple slave processes, without feed-back on the state of the world.

5.7 The Plan for the Forthcoming Chapters

The next four chapters discuss in details the possible processes inside the different structures. First, the discretization of price process is investigated, using either the logarithmic returns or the relative returns. This part investigates whether (5.1) or (5.5) should be used. Then, the heteroscedasticity is added to the price random walk in the next three chapters which investigate in details respectively the ARCH, stochastic volatility, and regime switching structures. This exploration gives many mug shots which can be compared to the empirical ones. The systematic process investigation allows us to select the relevant mathematical structures to be used in our mathematical description of the price dynamics. Some graphs in the mug shots acquire a more important status, as they allow one to decide clearly in favor of some process structures. In particular, the time reversal invariance, the precise form for the heteroscedasticity, and the leverage effect for stocks are very important and deserve further chapters in their own. Finally, the distribution of the innovations can be analyzed. With a clear view on the process structure, the applications to market risk evaluations and to option pricing are presented. This completes the discussion of the univariate case, and the subsequent chapters move to the analysis of multivariate time series.

http://www.springer.com/978-3-642-31741-5

Discrete Time Series, Processes, and Applications in Finance
Zumbach, G.
2013, XXII, 322 p., Hardcover
ISBN: 978-3-642-31741-5