UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova

Dipartimento di Fisica

SCUOLA DI DOTTORATO DI RICERCA IN FISICA

XXIV CICLO

Modeling and simulating financial time series over different time scales: from Omori regimes to trading strategies.

Direttore della Scuola: Ch.mo Prof. VITTURI, Andrea

Supervisore: Ch.mo Prof. STELLA, Attilio L.

Co-supervisore: Dott. BALDOVIN, Fulvio

Dottorando: Francesco CAMANA

Abstract

This PhD thesis deals with the modelization of financial time series by means of the application of a non-Markovian non-stationary model, based on the scaling properties of the returns. Some of the theoretical forecasts of the model are compared to historical data of different market indexes and financial instruments and predictive properties of the model are then employed to determine trading strategies, efficient in an intraday context.

After an introductory chapter, where the main stylized facts and financial models are illustrated, in Chap. 2 the model under discussion is described, in its double specialized intra- and inter-day formulation. A particular attention is devoted to the importance of the scaling property in constructing the joint probability distributions for the returns.

In Chap. 3 the model in its interday formulation is applied, with the result of reproducing the statistics of large returns (*aftershocks*) following main shock events in the market (the financial analog of the Omori law for seismic events). It is also discussed whether this analysis may be useful for the calibration of the model's parameters.

In Chap. 4 the model in its intraday version is utilized in applications regarding foreign exchange rates, market indexes, and single stocks quotes. The predictive power of the model is evaluated both with respect to some of the classical correlators (as the volatility autocorrelation) and to the yield obtained by employing trading strategies.

Riassunto

Questa tesi di dottorato riguarda la modellizzazione delle serie temporali finanziarie mediante l'applicazione di un modello non Markoviano e non stazionario basato sulle proprietà di scala dei ritorni/rendimenti. Riferendosi a diversi indici di mercato e a diversi strumenti finanziari, vengono verificate alcune delle previsioni teoriche del modello e si sfruttano le proprietà predittive dello stesso per determinare alfine delle strategie di trading efficaci in abito giornaliero.

Dopo un capitolo introduttivo, nel quale si illustrano i principali modelli utilizzati in ambito finanziario, nel Cap. 2 viene descritto il modello in oggetto, che viene peraltro specializzato nella duplice versione inter- ed intra-giornaliera. Particolare attenzione è dedicata alla centralità dello *scaling* per la costruzione delle distribuzioni di probabilità congiunte dei ritorni.

Nel Cap. 3 il modello viene applicato nella sua versione inter-giornaliera, al fine di riprodurre la statistica dei grandi ritorni (*aftershocks*) susseguenti il verificarsi di eventi di shock nel mercato (l'analogo finanziario della legge di Omori per gli eventi sismici). Viene valutato quanto questa analisi possa essere utile per calibrare i parametri del modello stesso.

Nel Cap. 4 viene utilizzato il modello nella sua versione intra-giornaliera, in applicazioni riguardanti il tasso di cambio, gli indici azionari e il prezzo di singole azioni. Viene valutata la capacità predittiva del modello sia rispetto ad alcuni classici correlatori, quali l'autocorrelazione di volatilità, sia rispetto al rendimento ottenuto dallo sfruttamento di strategie di trading.

To Him, who gives me the strenght.

Contents

1

Introduction

Aus Pflicht.
E. Kant

1.1 Stylized facts

1.1.1 Invitation

The statistical properties of financial indexes have been investigated for more than a century, since the first studies of Bachelier (6). The availability of large, high frequency datasets[1] and the application of computer analysis and simulation methods have recently led to the consolidation of data-based approaches in financial modeling.

The financial time series constitute a rich source of data for statistical studies, especially because these series show non-trivial, but to a certain degree *universal*, statistical properties (see, *e.g.*, 17, 20, 33). One of the key roles of the modeling techniques is to capture these properties, common across different financial instruments, markets and periods, which are usually referred to as *stylized facts*.

Stylized facts emerge as a common denominator among the properties observed in studies of different markets and instruments. They are, to some extent, *qualitative properties* of asset returns and so it is not clear whether they identify in an unique way a financial model.

Far from being a fully solved problem (15), the characterization of the underlying process for the financial indexes has received contributions both from economics and mathematical finance, but recently also from statistical physics. Physicists' natural tendency to focus on universal laws is often regarded with skepticism by some economists,

[1] Since 1993 financial and market data have been recorded 'tick-by-tick', transaction-by-transaction, which implies that the average time between two successive records can be as short as a few seconds.

but the development of the theory of complex systems in the last decades have brought a significant number of results, that are now relevant in financial and market contexts, too. Achievements made by physicists in various fields of statistical mechanics, like disordered systems or phase transitions, have thus found applications in the financial context, too, where the concepts of unpredictable time series, scaling, self-organized systems and power laws are today of common knowledge.

As a matter of fact, the modeling of the random fluctuations of asset prices is of obvious importance in financial applications as risk assessment and derivative pricing: correct predictions and estimates on indexes can only be achieved through an as precise as possible characterization of the stochastic process related to the observed indicator.

Indeed, the classic and widely applied *Brownian walk* model, dating back to Bachelier (6), has proven to be unfit and has been ultimately invalidated, as anticipated by Mandelbrot (31), after intensive statistical studies, *e.g.* on the recently available high frequency time series.

1.1.2 Universal features of financial indicators

Let us primarily fix some notations. Denoting with $S(t)$ the value at time t of a certain financial asset (a stock price, an exchange rate or a market index), following (39) it is usual to define the *return* $x_\tau(t)$ at time t and at scale τ as the difference of the logarithms of the prices at times $t+\tau$ and t:

$$x_\tau(t) = \ln S(t+\tau) - \ln S(t). \tag{1.1}$$

Most of the times, the value of $S(t)$ is assumed to be *detrended*, so that

$$\ln S(t) \mapsto \ln S(t) - \rho t, \tag{1.2}$$

where ρ is the average linear growth of the logarithm of the values of the asset over the whole time series. This exponential trend is related to the exponential growth followed by an amount stored in a bank account, usually used as a reference in finance; this trend is actually a first example of stylized fact, which is commonly detected in stock prices and other market indexes (see, *e.g.*, Fig.1.1).

We summarize as follows the most relevant stylized facts:

- **Absence of (linear) autocorrelations**:

  Given a discrete-time series of successive returns $x_\tau(0), x_\tau(\tau), \dots x_\tau(n\tau)$ and a time interval $T = k \cdot \tau$, with $k \in \mathbb{N}$ the empirical linear autocorrelation function,

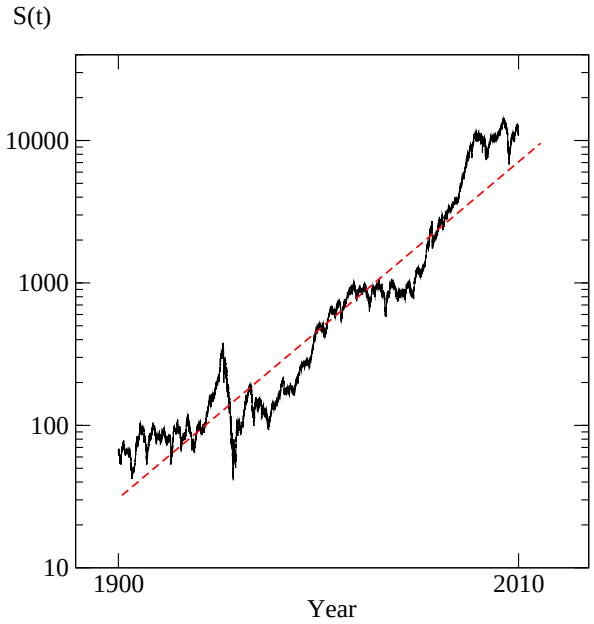


Figure 1.1: DJI index values from 1900 to 2010. Notice the log scale in the vertical axis and the quasi-linear trend in this scale.

defined as

$$\langle x_\tau(t)x_\tau(t+T)\rangle_e \equiv \frac{1}{n-k+1}\sum_{i=0}^{n-k} x_\tau(i\tau)x_\tau\left(i\tau+T\right), \tag{1.3}$$

vanishes if T is longer than 5-10 minutes[1]. This has proven to be a direct consequence of the efficiency of a market, which quickly suppresses any arbitrage opportunity (23).

- **Presence of correlation on absolute returns and higher order long-ranged correlations**:

  The autocorrelation function of the absolute returns

$$\langle |x_\tau(t)||x_\tau(t+T)|\rangle_e \equiv \frac{1}{n-k+1}\sum_{i=0}^{n-k} |x_\tau(i\tau)||x_\tau\left(i\tau+T\right)|, \tag{1.4}$$

  is positive and slowly decaying, as a function of T. The decay is well reproduced by a power law, $\langle |x_\tau(t)||x_\tau(t+T)|\rangle_e \simeq \frac{A}{T^\beta}$, with $\beta \in [0.2, 0.4]$ (see, *e.g.*, 33).

  Moreover, different measures of volatility (*e.g.*, the squared returns), display similar positive autocorrelation (*persistence*) over several days. The effect is called

[1] In the following, we will use the notation $\langle\cdot\rangle_e$ for the empirical averages, $\langle\cdot\rangle_p$ for the averages calculated on the basis of a probability density function p and $\langle\cdot\rangle$ when addressing to both.

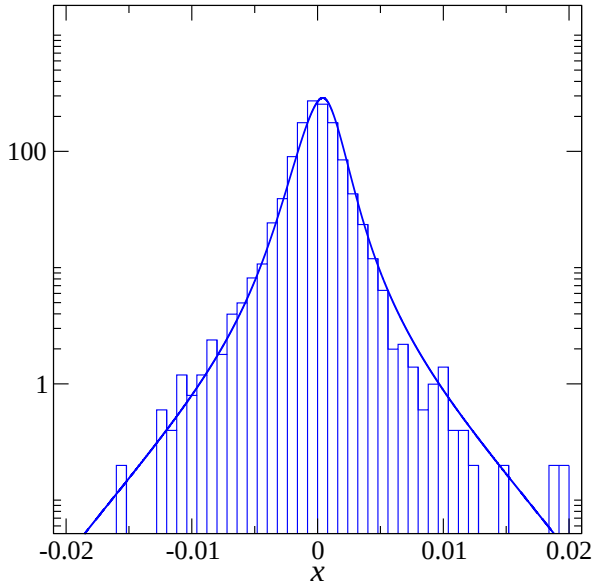


Figure 1.2: An example of histogram of returns, taken from the high frequency S&P500 index (1985-2010). A tentative fitting function (PDF) is given, too.

volatility clustering and measures the tendency of the market to cluster large returns (see, *e.g.*, 17, 33). Interestingly, the volatility has a power law decay in time (30) after very large returns (shocks).

- **Heavy-tailed distributions**:

  For sufficiently long historical time series, the probability density function (PDF) of the returns $p_{X_\tau}(x)$ can be estimated[1]. This function seems to show fat tails, with a decay index approximately in the range $2 \div 5$, for most assets (an example is given in Fig. 1.2). However, the fitting of the tails, being related to extreme events in limited time series, usually gives a non-reliable result.

- **Skewed distributions**:

  The PDF of the returns $p_{X_\tau}(x)$ displays an asymmetry, which privileges large negative returns rather than positive equally large movements (see, *e.g.*, 44).

- **Leverage effect**:

  A negative correlation is detected between returns and volatilities: the leverage correlation function, defined as $L_\tau(T) = \frac{1}{K}\langle |x_\tau(t+T)|^2 x_\tau(t)\rangle$ (K is an appropriate

[1] In the following, as a common rule, we will denote with the capital letter X the stochastic variable associated to the return x.

normalization factor), is very small, indistinguishable from noise, for $T \leq 0$, negative and asymptotically increasing toward zero for $T > 0$. In a nutshell, large negative returns are usually associated to a following period of high volatility (see, *e.g.*, 16).

- **Scaling properties**:
  - **Simple scaling**: For τ's in the range from one day to a few months, the PDF p_τ approximately obeys a simple-scaling symmetry: $p_{X_\tau}(x) = (\frac{1}{\tau^D})g(\frac{x}{\tau^D})$, where the scaling exponent D is close to $1/2$ for the indexes of the most developed markets. The scaling function $g(x)$ is non-Gaussian and has power-law, Pareto tails at large $|x|$ (see, *e.g.*, 34).
  - **Aggregational Gaussianity**: The above stated simple scaling property breaks for longer time lags. As one increases the τ's, the distribution of the returns progressively deforms from an heavy tailed distribution to a nearly Gaussian shape. This has proven to be a consequence of the Central Limit Theorem (CLT) (9), valid on a time-scale in which all correlations vanish.
  - **Multiscaling**: A more detailed analysis of the non-linear empirical moments

$$M_\tau(q) \equiv \langle |x_\tau(t)|^q \rangle_e \equiv \frac{1}{n+1} \sum_{i=0}^{n} |x_\tau(i\tau)|^q \tag{1.5}$$

    invalidates the simple scaling assumption and shows the presence of a *multiscaling* behaviour: $M_\tau(q) \propto \tau^{f(q)}$, being $f(q)$ a non-linear function of q. More details on scaling will be discussed in Sec. 3.1.

- **Intraday seasonalities**:

  Both for foreign exchange rates (11), stock markets (3) and market indexes (4) a clear intraday pattern for the volatility and, more in general, for the even moments of the PDF of the returns can be detected. One may calculate, for instance, the volatility of stock prices at a scale of a few minutes or even as low as a few seconds: the result, depicted in a daily scale, shows a clear U-shaped pattern (1), with a decay in the first two hours of trading as a power law with an exponent around 0.3 (3). In addition, cross correlation between stock values increases throughout the day (3).

  For foreign exchange rates the detectable peaks in the volatility pattern coincide with the opening and closure time of the main world stock exchange markets (11).

1.2 Financial models: classification

After the publication of the Black and Scholes option pricing model (13), back in 1973, which was essentially based on the *geometric Brownian motion* (see Sec. 1.2.2), theorists and practitioners developed analyses and studies that clearly demonstrated the need of corrections to the model. So, alternative models started being considered and built: up to now the problem of the identification of the stochastic process describing the financial returns remains open.

As already stated, the problem is of utmost importance, for it involves delicate matters as the fair price of the derivatives and the estimate of the portfolio-associated risks, which are crucial in an economic and financial perspective.

One of the key aspects, somehow characterizing the analysis of the financial and economic data, is that the empirical analyses which can be performed on these data are not equivalent to those used in the usual experimental investigations in physical sciences. Large scale and repeated experiments are not possible, and one single, not so long, history is usually available[1]. So this is a problem from the point of view of the validation of the models. As an example, we may cite the fact that, after the studies of Mandelbrot (31), a question about the finiteness of the variance of the data arose. It may seem a simple check to do, but the community has finally completely agreed on its finiteness only after dedicated studies on high frequency datasets (17, 26).

Since the core argument of this thesis is the study of the properties and the predictive power of a novel financial model, I think it relevant to spend a few words about the zoo of the so far proposed models, with a focus on their assumptions and on their mathematical and statistical characteristics.

As a general overview, we can identify several studies focusing on different aspects of the analyzed stochastic processes: the shape of the PDF of the returns (17, 24, 33, 40), the memory of the process (35), the higher-order statistical properties (37). Moreover, other studies have proposed models attempting to describe the features of the real markets, and even if many of them are able to reproduce many of these features (as fat-tailed or non-Gaussian shape of the PDF), they all ultimately fail somewhere else (*e.g.* time invariance or leverage effect). Finally other studies consider the analogies and differences between the dynamics of prices and that of ecological or social systems (40), or even turbulent regimes in fluidynamics (24, 35).

We will focus on the most common models.

[1] Actually, this limitation is also affecting other developed areas of physics as astrophysics, geophysics, atmospheric physics; ultimately, this analogy justifies the task.

1.2.1 Brownian motion (simple)

The very first financial model, proposed by Bachelier (6) in 1900, described the evolution of the time-series of the returns as a stochastic process that is currently named *Brownian motion*. Actually, Bachelier defined the *return* just as the difference of subsequent prices, not as the difference of the logarithm of the prices. So, just for this section, we define $x_\tau(t) \equiv S(t+\tau) - S(t)$. The properties of the Brownian motion are:

(i) the returns are independent random variables;

(ii) the returns are stationary (*e.g.* their distribution $p_{X_\tau(t)}$ is t-independent);

(iii) each return calculated over a time interval of duration τ follows a Gaussian distribution with zero mean and variance $\sigma^2\tau$.

Therefore, because of (ii) and (iii), the PDF of the single returns over a time τ is:

$$p_{X_\tau(t)}(x) = \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left(-\frac{x^2}{2\sigma^2\tau}\right), \tag{1.6}$$

which is independent of t. Moreover, from (i), the joint PDF for n consecutive returns follows:

$$p_{X_\tau(0)X_\tau(\tau)\ldots X_\tau((n-1)\tau)}(x_0, x_1, \ldots x_{(n-1)}) = \prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left(-\frac{x_i^2}{2\sigma^2\tau}\right). \tag{1.7}$$

However the description of the evolution of the prices according to a Brownian motion generates some inconsistencies with respect to the observed asset behaviour:

- the Brownian assumption states that the fluctuations of the prices are independent of the size of the price itself, while empirical evidence shows that the amplitude of the fluctuations is related to the values of the asset;

- the price of the asset, can go below zero, whereas most financial time series are strictly positive.

Thus the model required refinements, especially for long-time analyses, where market values make large changes. The first result of this refinement was the *Geometric Brownain motion* model, that has been for a long time the most common and widely used model for financial price dynamics.

1.2.2 Geometric Brownian motion

In this model, differently from the simple Brownian motion, the *logarithm* of the price, rather than the price itself, follows a simple Brownian motion. It's quite common in finance to say that the returns are Log-normal distributed.

Log-normal distributions forbid negative prices, by construction. Moreover, the size of the fluctuations is dependent on the value of the asset. So, both of the problems encountered in the simple Brownian motion are overcome. The Eqs. 1.6 and 1.7 are still valid, but with the returns re-defined as difference of the logarithms of the prices, as in Sec. 1.1.2.

Given the gaussianity of the PDF, an important property, called *self-similarity*, can be easily proven. A process is said to be *self-similar* if, for any $\tau > 0$, the PDF of the return $x_\tau(t)$ satisfies the following rule

$$p_{X_\tau}(x) = \frac{1}{\tau^D} g\left(\frac{x}{\tau^D}\right). \tag{1.8}$$

The formula indicates that, up to a rescaling factor dependent on τ, the returns X_τ are all identically distributed according to a function $g(x)$, which is called *scaling function*. The parameter D is called *self similarity* or *scaling* exponent. By direct inspection, it's clear that the Brownian and the geometric Brownian motion are self-similar processes with $D = 1/2$ and with $g(x)$ Gaussian.

The model accounts for many stylized facts of financial data, but it provides just a first approximation, although still widely used in finance. Actually, systematic deviations from the model prediction are observed: the empirical distributions are more leptokurtic than Gaussian distributions, with fatter tails than in the Gaussian case.

1.2.3 Lévy flights

The first model to account for the leptokurtic shape of the PDF, originally for cotton prices, was proposed by Mandelbrot back in 1963 (31). In his work the logarithmic return of the cotton price is stated to follow a stochastic process with Lévy stable non-Gaussian increments. *Lévy processes* are self-similar processes with independent increments. The difference respect to the previous models is that the distribution of the returns is no longer a Gaussian, but a Lévy distribution. The characteristic function of a (symmetric) Lévy distribution is

$$\hat{L}_\mu(k) = \exp\left(-a_\mu |k|^\mu\right), \tag{1.9}$$

while an analytical form for the distribution itself is not generally available. An important property of these symmetrical distributions is the power-law behaviour of the tails (the so called *Pareto tails*):

$$L_\mu(x) \sim \frac{\mu A^\mu}{|x|^{1+\mu}} \quad \text{for } x \to \pm\infty, \tag{1.10}$$

where a simple relation $A^\mu = k(\mu)a_\mu$, between a_μ and A^μ, holds (see, *e.g.*, 17).

Mandelbrot assumed the following:

(i) the returns are independent random variables;

(ii) the returns are stationary (*i.e.* their distribution is time-independent);

(iii) each return obeys a Lévy distribution described by the following characteristic function:

$$\hat{p}_1(k) = \exp\left(-a|k|^\mu\right), \tag{1.11}$$

with $0 < \mu < 2$ and $a > 0$. These rules define the *Lévy flights* stochastic process.

It's easy to prove that the Lévy flights are self-similar processes with $D = 1/\mu$ and that the relative distribution is stable, *i.e.* the Lévy distributions share with the Gaussians the property of being an attractor in probability space in the sense that, under repeated convolution, a large class of distributions converges towards it (29).

It is evident from the tail decay that, for $\mu \leq 2$, the variance of the distribution[1] is formally infinite. As previously noted in Sec. 1.2, the problem of the finiteness of the variance is an important topic.

Although accounting for some stylized facts, this model cannot explain:

- the multiscaling behaviour;
- the scaling breakdown at large τ's;
- the finiteness of the variance;
- higher order, long range correlations (and particularly the volatility clustering)

[1] The variance σ_ρ of a (zero-mean) distribution $\rho(x)$ is defined as follows: $\sigma_\rho = \int_{-\infty}^{\infty} x^2 \rho(x) dx$.

1.2.4 Fractional Brownian motion

One way to explain the long range correlations is that of relaxing the request for independence of the returns. An example of model which combines self-similarity and stationarity, but at the same time imposes non-independence, is given by the *Fractional Brownian motion*. Without entering the details of the model, originally proposed again by Mandelbrot together with van Ness (32), it suffices here to state that the model explains the long range correlations and the all-important property of volatility clustering.

On the other hand the model poses serious problems of ergodicity[1] and does not show the vanishing of linear correlations.

1.2.5 Truncated Lévy flights

The problem of the finiteness of the variance, generated by the studies of Mandelbrot and posed by the Lévy flights has led to the definition of a new stochastic process with finite variance: the truncated Lévy flight (TLF) process. Its distribution is simply defined as a truncation of the Lévy symmetric distribution:

$$p(x) = \begin{cases} 0 & x < \ell \\ cL_\mu(x) & -\ell \leq x \leq \ell \\ 0 & x > \ell \end{cases} \tag{1.12}$$

The TLF is a non-stable stochastic process but, possessing finite variance, it converges, in a CLT sense, to a Gaussian process (36). So it can bypass the limit of the (simple) Lévy flight model in that it accounts for the scaling breakdown at large τ's.

Still based on the assumption of *i.i.d.* (independent, identically distributed) returns, the model fails to describe in a proper way the time-dependent volatility observed in market data (33).

1.2.6 Mixture of Gaussian distributions

One alternative model to take into account the leptokurtic features of the empirical PDF of the returns was originally proposed in 1973 by Clark (19). Starting from the observation that trading operations are inhomogeneous in time and volume, he specifies and applies the concept of *subordinate stochastic process* to justify the inapplicability of the standard CLT to financial series (*e.g.* cotton price). He selects the stochastic process coming from the, possibly inhomogeneous, time series and further defines a

[1] The *ergodicity* property guarantees that the sample averages, calculated along a single realization of the process, tend to the ensemble averages as the length of the sample tends to $+\infty$.

new process, which turns to be a weighted mixture of Gaussians, which has stationary independent increments, zero mean and finite variance. The time-independent PDF for the return x can be expressed in a Gaussian mixture form as follows:

$$p(x) = \sum_{k=1}^{K} w_k \mathcal{N}(x; \mu_k, \sigma_k) \tag{1.13}$$

with $w_1, \ldots, w_K$ normalized weights and $\mathcal{N}(x; \mu, \sigma)$ a Gaussian of mean μ and variance σ^2. In its continuous form, Eq. 1.13 reads:

$$p(x) = \int_{-\infty}^{\infty} d\mu \int_{0}^{\infty} d\sigma \; \psi(\mu, \sigma) \mathcal{N}(x; \mu, \sigma). \tag{1.14}$$

Generally speaking, the model has no definite scaling properties.

1.2.7 ARCH/GARCH processes

In section 1.1.2 we pointed out the time dependence of the volatility, which takes high levels in periods of uncertainty or after *large* returns, and reverts to lower values in calmer periods. This phenomena is called *heteroskedasticity* and suggests that the variance itself can be then modeled as a stochastic process.

Some models have been proposed to explain and to describe such a behaviour. Generally speaking, in these models the random part of the return can be written as:

$$x_k = \epsilon_k \sigma_k, \tag{1.15}$$

where ϵ_k are i.i.d. random variables with zero mean and unit variance, and σ_k are deterministic constants or random processes (a process generally correlated in time and, possibly, with the variables ϵ_j with $j < k$) corresponding to the local size of the fluctuations. To reproduce the market behaviour, one may want, for example, that a large σ_i is followed with high probability by a large σ_{i+1}, although with an arbitrary sign for the return.

In this context of volatility modelization, a rich class is represented by the ARCH models, introduced by Engle in 1982 (22). The acronym stands for *autoregressive conditional heteroskedasicity* and well describes its characteristic: the variance at time t depends, conditionally, on some past values of the square value of the random signal itself. The different choice for the PDF of the process generating the random variable at time t, and the number of conditioning terms (the memory p), span a class of stochastic models that have found wide application in finance and are, to some extent, a reference for testing other models, too.

An ARCH(p) model is defined by:

$$\sigma_k^2 = \alpha_0 + \alpha_1 \epsilon_{k-1}^2 + \ldots + \alpha_p \epsilon_{k-p}^2, \tag{1.16}$$

with $\alpha_0, \ldots, \alpha_p$ positive variables and ϵ_k random variables with zero mean and variance σ_k. The PDF for the ϵ's is usually taken to be a Gaussian, but that's not restrictive. The choice of the parameters α determines the shape of the PDF of the process, and particularly its kurtosis. Even in case of Gaussian distributed ϵ's, the resulting process can have a leptokurtic PDF, with fat tails resulting from the volatility fluctuations and correlations (17). However, it is still an open question what is the asymptotic PDF of the ARCH process, given the PDF for the ϵ's.

An extension of the ARCH models was introduced in 1986 by Bollerslev (14), to overcome the problem of the optimization of the α parameters. In these generalized ARCH models, the so called GARCH(p,q), the following stochastic process is defined:

$$\sigma_k^2 = \alpha_0 + \alpha_1 \epsilon_{k-1}^2 + \ldots + \alpha_q \epsilon_{k-q}^2 + \beta_1 \sigma_{k-1}^2 + \ldots + \beta_p \sigma_{k-p}^2, \tag{1.17}$$

with $\alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p$ control parameters. A quite large zoo of GARCH models is used in finance. One of these models, the asymmetric GJR-GARCH by Glosten, Jagannathan and Runkle (25) will be used as reference and competitor model in Sec. 4.2.3.

A crucial limitation of ARCH/GARCH models is that they generate histories for which the scaling properties of the sampled PDF are not guaranteed (33).

2

A model based on scaling

All models are wrong, but some are useful.
G.E.P. Box

2.1 The scaling properties of financial data

The idea of studying and applying the concept of *scaling* in finance dates back to the last half of the past century, following the ideas generated by the works of Osborne and Mandelbrot (see, *e.g.*, 34, 37, and references therein). This key concept plays a central role in the analysis of a wide class of complex systems, where a large number of degrees of freedom are strongly correlated with each other.

The scaling laws were first introduced during the 1960s in independent works by Kadanoff, Fisher and Widom on the thermodynamics of systems around critical points (in proximity of phase transitions). Their assumptions of scaling in the critical region found a consistent interpretation and foundation during the 1970s, when the ideas of *renormalization group* were defined and took appropriate form.

The renormalization method and the self-similarity approach, originally developed for the critical phenomena in thermodynamics, has found wide application in different fields of natural and social sciences (from geology to ecology, from fluidynamics to human geography), especially in systems where consistent dynamic equations are not available or known. In these fields the study of the scaling transformation properties can be of help.

Effectively, after some pioneering, isolated contributions, physicists entered the scene of finance around 1995, with the ingenious invention of the provocative brand name of *econophysics*, right with the application of scaling concepts to market series (for an introduction on the subject see, *e.g.* (45)).

2.2 The interday formulation

2.2.1 The stationarity and self-similarity assumptions

The model presented in this chapter, which is the ground of the entire thesis work, bases its development on the self-similarity assumption and on the concept of scaling of financial data over different time lags. The foundations of the model have been laid down in 2007 by Baldovin and Stella (8, 48), whose reasoning will be hereafter reported.

As formerly stated, the concept of self-similarity for the PDF of the financial returns is summarized by the Eq. 1.8, which characterizes a process which has, upon rescaling the variables with a proper normalization, the same statistical distribution over different (but limited) times.

To introduce the model, we consider a discrete series of equally spaced in time returns $\{x_1(t)\}$, for $t = 1, \ldots, T$, defined as in section 1.1.2. Taking advantage of the additive properties of the definition of the returns, different series of returns can be built, over different time lags. For example, being $x_2(1) = x_1(2) + x_1(1)$, the series $\{x_2(t)\}$ of non-overlapping returns can be extracted, for suitable values of t.

To start with, we suppose that the series is generated by a stochastic process with the following properties:

(i) the returns $\{x_1(t)\}$ are identically distributed[1];

(ii) the self-similarity property (Eq. 1.8) holds, with a given scaling exponent D (simple scaling).

The (i) property implies that the marginal ensemble PDF of the returns is the same for every return of the series, $p_{X_1(1)}(x) = \ldots = p_{X_1(T)}(x) \equiv p_1(x)$. The same property holds for the series constructed for different τ's: $p_{X_\tau(1)}(x) = p_{X_\tau(1+\tau)}(x) = p_{X_\tau(1+2\tau)}(x) = \ldots \equiv p_\tau(x)$. These PDF's can be empirically built, given a sufficiently long series of market data.

It is useful to switch into a Fourier transformed space, where the *characteristic function* $\tilde{p}_{X_1,\ldots,X_n}(q_1, \ldots, q_n)$ is defined, starting from a joint probability density function $p_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$, by the formula:

$$\tilde{p}_{X_1,\ldots,X_n}(q_1, \ldots, q_n) = \left(\prod_{i=1}^{n} \int dx_i e^{iq_i x_i} \right) p_{X_1,\ldots,X_n}(x_1, \ldots, x_n). \tag{2.1}$$

From (i) one has:

$$\tilde{p}_{X_\tau(t)}(q) = \tilde{p}_{X_\tau(1)}(q), \quad \text{for } \tau = 1, 2, \ldots \quad \text{and } t = 1 + \tau, 1 + 2\tau, \ldots \ . \tag{2.2}$$

[1] This assumption will be relaxed in Sec. 2.2.3

Upon setting $p_1(q) \equiv g(q)$, Eq. 1.8 can be rewritten in terms of characteristic function:

$$\begin{aligned}\tilde{p}_\tau(q) &= \int dx e^{iqx} p_\tau(x) \\ &= \int dx e^{iqx} \frac{1}{\tau^D} p_1\left(\frac{x}{\tau^D}\right) \\ &= \int dx e^{iq\tau^D x} p_1(x) = \tilde{p}_1(\tau^D q) = \tilde{g}(\tau^D q). \end{aligned} \tag{2.3}$$

The expression in Eq. 1.8 for the scaling of the PDF can be extended to the case of a joint PDF, obtaining the most complete definition of self-similarity, in the following way:

$$p_{X_\tau(1),\ldots,X_\tau(1+(n-1)\tau)}(x_1,\ldots,x_n) = \frac{1}{\tau^{nD}} p_{X_1(1),\ldots,X_1(1+(n-1)\tau)}\left(\frac{x_1}{\tau^D},\ldots,\frac{x_n}{\tau^D}\right), \tag{2.4}$$

or equivalently, in the Fourier space,

$$\tilde{p}_{X_\tau(1),\ldots,X_\tau(1+(n-1)\tau)}(q_1,\ldots,q_n) = \tilde{p}_{X_1(1),\ldots,X_1(1+(n-1)\tau)}\left(\tau^D q_1,\ldots,\tau^D q_n\right). \tag{2.5}$$

Where clear, the following easier notations will be in use:

$$\begin{aligned} p_{X_\tau(1)} &\equiv p_\tau, \\ \tilde{p}_{X_\tau(1)} &\equiv \tilde{p}_\tau, \\ p_{X_\tau(1),\ldots,X_\tau(1+(n-1)\tau)} &\equiv p_\tau^{(n)}, \\ \tilde{p}_{X_\tau(1),\ldots,X_\tau(1+(n-1)\tau)} &\equiv \tilde{p}_\tau^{(n)}. \end{aligned}$$

2.2.2 The stationary model with D=$\frac{1}{2}$

An interesting key observation is that the joint PDF is constrained by the stationarity and the simple scaling assumptions of Sec. 2.2.1. Let us consider two successive returns $x_1 \equiv x_\tau(1)$ and $x_2 \equiv x_\tau(1+\tau)$ and the total return $x \equiv x_{2\tau}(1) = x_\tau(1) + x_\tau(1+\tau) = x_1 + x_2$. The following relations naturally follow:

$$\int dx_1 dx_2 p_\tau^{(2)}(x_1,x_2)\delta(x - x_1 - x_2) = p_{2\tau}(x), \tag{2.6}$$

$$\int dx_2 p_\tau^{(2)}(x_1,x_2) = p_\tau(x_1), \tag{2.7}$$

$$\int dx_1 p_\tau^{(2)}(x_1,x_2) = p_\tau(x_2). \tag{2.8}$$

where the results in Eq. 2.7 and Eq. 2.8 follow from the fact that the joint PDF $p_\tau^{(2)}$ is sampled from a sequence of intervals of length 2τ, whose first and second halves

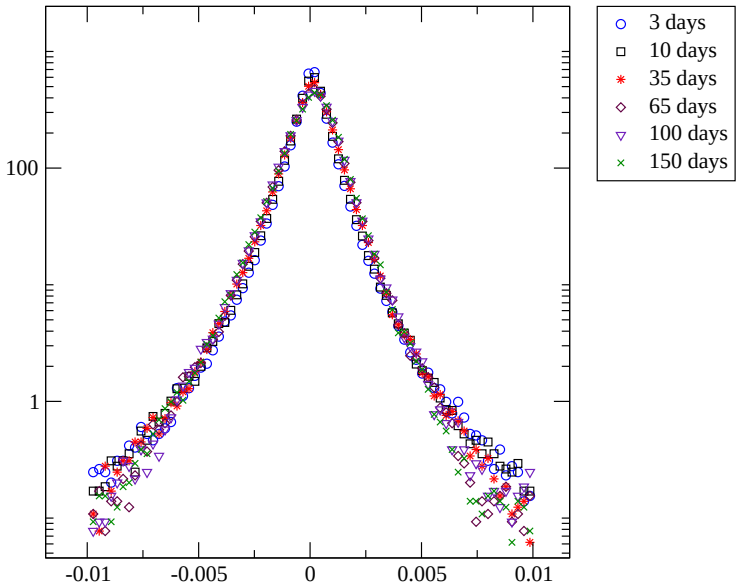


Figure 2.1: Scaling collapse of the PDF of the returns over different time intervals for the S&P500 index (1985-2010). The collapse has been obtained by using Eq. 1.8, with D=0.5. The details of the derivation of this picture are presented in Sec. 3.1

can both be used to equivalently sample the PDF p_τ. In the Fourier space, the above relations read:

$$\tilde{p}_\tau^{(2)}(q,q) = \tilde{g}\left((2\tau)^D q\right), \tag{2.9}$$

$$\tilde{p}_\tau^{(2)}(q,0) = \tilde{g}\left(\tau^D q\right), \tag{2.10}$$

$$\tilde{p}_\tau^{(2)}(0,q) = \tilde{g}\left(\tau^D q\right). \tag{2.11}$$

Now we take advantage of the market behaviour, where successive returns (for τ larger than few minutes) are uncorrelated:

$$\langle x_1 x_2 \rangle_{p_\tau^{(2)}} \equiv \int dx_1 dx_2 \; x_1 x_2 p_\tau^{(2)}(x_1, x_2) = 0, \tag{2.12}$$

that is one evidence of the *efficient market hypothesis* (23) which inhibits from taking profit from the sign values of asset returns. Eq. 2.12 implies $\langle x^2 \rangle_{p_{2\tau}} = \langle (x_1 + x_2)^2 \rangle_{p_\tau^{(2)}} = 2\langle x^2 \rangle_{p_\tau}$. Using the self-similarity property Eq. 1.8, one easily gets: $\langle x^2 \rangle_{p_\tau} = \int x^2 p_\tau(x) dx = \int x^2 \frac{1}{\tau^D} g\left(\frac{x}{\tau^D}\right) dx \propto \tau^{2D}$. And so $2\tau^{2D} = (2\tau)^{2D}$, whence the result: $D = 1/2$.

So self similarity, stationarity and absence of linear correlations imply $D = 1/2$, a value confirmed by repeated analysis of data for all developed markets (21). One example of this scaling is shown in Fig. 2.1.

One important step is then to generalize the content of equation Eq. 2.9 in order to write a general expression for $p_\tau^{(2)}$ in terms of p_τ alone. If the increments x_0 and x_1 were independent, the result would be straightforward:

$$\begin{aligned}\tilde{p}_\tau^{(2)}(q_1,q_2) &= \int dx_1 dx_2 e^{iq_1x_1} e^{iq_2x_2} p_\tau^{(2)}(x_1,x_2) \\ &= \int dx_1 dx_2 e^{iq_1x_1} e^{iq_2x_2} p_\tau(x_1) p_\tau(x_2), \\ &= \tilde{g}\left(\tau^{1/2}q_1\right)\tilde{g}\left(\tau^{1/2}q_2\right). \end{aligned} \tag{2.13}$$

But to take into account the strong non linear correlations of the financial data (see Sec. 1.1.2) one may introduce a *generalized product*, consistent with the set of equations Eq. 2.9 - Eq. 2.11, to write an expression like $\tilde{p}_\tau^{(2)}(q_1,q_2) = \tilde{g}\left(\tau^{1/2}q_1\right) \otimes \tilde{g}\left(\tau^{1/2}q_2\right)$. As can easily be checked, one possible formal solution is obtained by defining: $\tilde{g}(q_1) \otimes \tilde{g}(q_2) \equiv \tilde{g}\left(\sqrt{q_1^2+q_2^2}\right)$. And so:

$$\begin{aligned}\tilde{p}_\tau^{(2)}(q_1,q_2) &= \tilde{g}\left(\tau^{1/2}q_1\right) \otimes \tilde{g}\left(\tau^{1/2}q_2\right) \\ &= \tilde{g}\left(\sqrt{\tau q_1^2 + \tau q_2^2}\right). \end{aligned} \tag{2.14}$$

The generalization to more than two consecutive returns is performed following the same reasoning. The characteristic function of n consecutive increments obeys the following relation:

$$\tilde{p}_\tau^{(n)}(\overbrace{0,\ldots,0}^{j},\overbrace{\tilde{q},\ldots,\tilde{q}}^{m},\overbrace{0,\ldots,0}^{n-j-m}) = \tilde{g}\left((m\tau)^D\tilde{q}\right) \tag{2.15}$$

$$= \tilde{g}\left(\sqrt{m\tau}\tilde{q}\right), \tag{2.16}$$

where in Eq. 2.16 the result $D = \frac{1}{2}$ coming from self-similarity and uncorrelation of returns has been introduced. Following the steps leading to Eq. 2.14, a possible formal solution for the general case is analogously found:

$$\tilde{p}_\tau^{(n)}(q_1,\ldots,q_n) = \tilde{g}\left(\sqrt{\tau q_1^2 + \ldots + \tau q_n^2}\right). \tag{2.17}$$

Provided that this $\tilde{g}$ function is a well-defined characteristic function[1], *i.e.* its inverse-transform satisfies all the defining conditions of a PDF, Eq. 2.14 determines the construction of the joint PDF for the successive returns from the knowledge of a single

[1] Schoenberg's theorem states that this condition is satisfied, for every $n \in \mathbb{N}_1$, only if g is a Gaussian mixture (2, 43, 49).

univariate function $g(x)$, whose choice describes the entire process. A really remarkable point!

In the next section we will relax the stationarity assumption to include in the model stylized facts related to the non-stationarity of the returns.

2.2.3 The non-stationary model

The rich and powerful framework developed in the previous section is based on self-similarity, market efficiency and stationarity of increments. Both in an intra- and an inter-day context, this last assumption on market data crashes against the presence of *exogenous* dynamical effects that, in specific time windows, may generate non-stationary effects and so a reasonable bit of non-stationarity must be included in the model, in some way. Fortunately the model is flexible enough to be enriched with such a feature.

The introduction of a time-inhomogeneous dynamics should in principle give the same results as those of the forementioned model, when long time windows are considered. In fact, when looking at the returns over large time lags or averaging over large time windows, one suppresses the peculiar time-inomogeneities through the averaging process. For instance, the scaling exponent D that, as will be discussed in a moment, can be empirically evaluated below or over $1/2$ in specific parts of the series, reverts back to the value $1/2$, expected for the stationary case, when the analysis extends over larger times, when the singular, exogenous events undergo the averaging process.

In the stationary case of Eq. 2.2 the unconditional PDF is identical for every increment. Now we want to develop a time-inhomogeneous scheme, to be applied after the occurrence of strong exogenous events, that are supposed to affect the dynamics. A simple assumption of time-inhomogeneity due to exogenous effects is given by this alternative property:

$$\tilde{p}_{X_\tau(1+(i-1)\tau)}(q) = \tilde{p}_{X_{a_i\tau}(1)}(q), \quad \text{for } \tau = 1, 2, \dots \quad \text{and } i = 1, 2, \dots \ . \tag{2.18}$$

In this non-stationarity property the unconditional PDF is still identical for every increment, but only up to a rescaling factor a_i depending on the chosen time interval $(i-1)\,\tau$. The coefficients a_i characterize the inhomogeneity.

The ensemble PDF's become time-dependent and, in general, $p_{X_\tau(t)}(x) \neq p_{X_\tau(1)}(x)$. Instead, from Eq. 2.18 the following relation holds:

$$p_{X_\tau(1+(i-1)\tau)}(x) = p_{X_{a_i\tau}(1)}(x). \tag{2.19}$$

The set of equations Eq. 2.6 - Eq. 2.8 becomes:

$$\int dx_1 dx_2 p_{X_\tau(t),X_\tau(t+\tau)}(x_1,x_2)\delta(x-x_1-x_2) = p_{X_{2\tau}(t)}(x), \tag{2.20}$$

$$\int dx_2 p_{X_\tau(t),X_\tau(t+\tau)}(x_1,x_2) = p_{X_\tau(t)}(x_1), \tag{2.21}$$

$$\int dx_1 p_{X_\tau(t),X_\tau(t+\tau)}(x_1,x_2) = p_{X_\tau(t+\tau)}(x_2). \tag{2.22}$$

Because of inhomogeneity, the right hand side of the two last equations no longer coincides. However, we proceed still assuming that a scaling property holds, possibly with different D and g. Here is the self-similarity hypothesis in the special case of time-inhomogeneity:

$$\tilde{p}_{X_\tau(1+(i-1)\tau)}(q) = \tilde{p}_{X_{a_i\tau}(1)}(q) = \tilde{p}_{X_1(1)}\left((a_i\tau)^D q\right) \equiv \tilde{g}\left((a_i\tau)^D q\right), \tag{2.23}$$

while the equivalent property in the inverse-transformed space reads:

$$p_{X_\tau(1+(i-1)\tau)}(x) = \frac{1}{(a_i\tau)^D} g\left(\frac{x}{(a_i\tau)^D}\right). \tag{2.24}$$

We are now ready to determine the coefficients a_i. Clearly, $a_1 = 1$ (see Eq. 2.18). To calculate a_i for $i > 1$, we start setting $t = 1$ and we again impose that the successive increments are linearly uncorrelated

$$\langle x_\tau(1)x_\tau(1+\tau)\rangle_{p_{X_\tau(1)X_\tau(1+\tau)}} \equiv \int dx_1 dx_2 x_1 x_2 p_{X_\tau(1)X_\tau(1+\tau)}(x_1,x_2) = 0, \tag{2.25}$$

and we match the following conditions:

$$\langle x_{2\tau}(1)^2\rangle \equiv \int dx x^2 p^{(2)}_{X_{2\tau}(1)}(x) = (2\tau)^{2D}\int dy y^2 g(y) \tag{2.26}$$

$$\begin{aligned}\langle (x_\tau(1)+x_\tau(1+\tau))^2\rangle &\equiv \int dx_1 dx_2 (x_1^2+x_2^2) p_{X_\tau(1),X_\tau(1+\tau)}(x_1,x_2)\\ &= (\tau)^{2D}\int dy y^2 g(y) + (a_2\tau)^{2D}\int dy y^2 g(y).\end{aligned} \tag{2.27}$$

Thus we obtain $a_2^{2D} = 2^{2D} - 1$.

By iteratively matching relations as $\langle x_{i\tau}(1)^2\rangle = \langle (x_\tau(1) + \ldots + x_\tau(1+(i-1)\tau))^2\rangle$ we are able to derive all the a_i's:

$$\begin{aligned}a_i^{2D} &= i^{2D} - \sum_{l=1}^{i-1} a_l\\ &= i^{2D} - (i-1)^{2D}.\end{aligned} \tag{2.28}$$

We are now able to write a solution for the set of equations

$$\tilde{p}_{X_\tau(1),X_\tau(1+\tau)}(q,q) = \tilde{p}_{X_{2\tau}(1)}(q) = \tilde{g}\left((2\tau)^D q\right), \tag{2.29}$$

$$\tilde{p}_{X_\tau(1),X_\tau(1+\tau)}(q,0) = \tilde{p}_{X_\tau(1)}(q) = \tilde{g}\left((\tau)^D q\right), \tag{2.30}$$

$$\tilde{p}_{X_\tau(1),X_\tau(1+\tau)}(0,q) = \tilde{p}_{X_\tau(1+\tau)}(q) = \tilde{g}\left((a_2\tau)^D q\right). \tag{2.31}$$

derived, using Eq. 2.23, from Eq. 2.20 - Eq. 2.22 (for $t = 1$).

Using again the aforementioned generalized product, it's easy to check that one possible formal solution is given by

$$\begin{aligned} \tilde{p}_{X_\tau(1),X_\tau(1+\tau)}(q_1,q_2) &= \tilde{g}\left(\tau^D q_1\right) \otimes \tilde{g}\left(a_2\tau^D q_2\right) \\ &= \tilde{g}\left(\sqrt{\tau^{2D}q_1^2 + (a_2\tau)^{2D}q_2^2}\right), \end{aligned} \tag{2.32}$$

and, extending the result for k consecutive intervals, the more general relation can be verified:

$$\begin{aligned} \tilde{p}_{X_\tau(1),\ldots,X_\tau(1+(k-1)\tau)}(q_1,\ldots,q_k) &= \tilde{g}\left((a_1\tau)^D q_1\right) \otimes \ldots \otimes \tilde{g}\left((a_k\tau)^D q_k\right) \\ &= \tilde{g}\left(\sqrt{(a_1\tau)^{2D}q_1^2 + \ldots + (a_k\tau)^{2D}q_k^2}\right). \end{aligned} \tag{2.33}$$

Now that Eq. 2.33 has been established, we can obtain an important scaling relation for the PDF of the return $X_\tau(1+t)$, for $\tau = 1, 2, \ldots$ and $t = 0, 1, \ldots$.

In the Fourier space the following relations hold:

$$\begin{aligned} \tilde{p}_{X_\tau(1+t)}(q) &= \tilde{p}_{X_1(1+t),\ldots,X_1(t+\tau)}(\underbrace{q,\ldots,q}_{\tau}) \\ &= \tilde{p}_{\underbrace{X_1(1),\ldots,X_1(t)}_{t},\underbrace{X_1(1+t),\ldots,X_1(t+\tau)}_{\tau}}(\underbrace{0,\ldots,0}_{t},\underbrace{q,\ldots,q}_{\tau}). \end{aligned} \tag{2.34}$$

Using the result of Eq. 2.33, one finally gets:

$$\begin{aligned} \tilde{p}_{X_\tau(1+t)}(q) &= \tilde{p}_{\underbrace{X_1(1),\ldots,X_1(t)}_{t},\underbrace{X_1(1+t),\ldots,X_1(t+\tau)}_{\tau}}(\underbrace{0,\ldots,0}_{t},\underbrace{q,\ldots,q}_{\tau}) \\ &= \tilde{g}\left(\sqrt{a_{t+\tau}^{2D} + \ldots + a_{t+1}^{2D}}\; q\right) \qquad (2.35) \\ &= \tilde{g}\left(\sqrt{(t+\tau)^{2D} - t^{2D}}\; q\right). \qquad (2.36) \end{aligned}$$

We may now revert to the ordinary probability space to obtain the generalized time-inhomogeneous scaling property:

$$p_{X_\tau(1+t)}(x) = \frac{1}{\sqrt{(t+\tau)^{2D} - t^{2D}}} g\left(\frac{x}{\sqrt{(t+\tau)^{2D} - t^{2D}}}\right). \tag{2.37}$$

It is evident that if $D = \frac{1}{2}$ the process is still stationary (the t-dependence vanishes), while the choice $D \neq \frac{1}{2}$ may account for non-stationarity features. When scaling relations with $D \neq \frac{1}{2}$ and/or g non-Gaussian are in force, the associated process is said to scale *anomalously*. One consequence of the *anomalous scaling* is the presence of strong non-linear correlations in the process. As shown by Baldovin and Stella (9, 49),

this generalized product can be the basis of a CLT for the probability distribution of the sum of many critically correlated random variables.

Summarizing the so far depicted results, we may say that the introduction of a generalized product can be used to express the joint characteristic function of returns in term of a single function g. This function can also be used to take into account the time-inhomogeneous properties of the return series and to write an anomalous scaling relation, valid across large time intervals (t) and across different time lags (τ).

2.2.4 Specializing the model

The results of the previous section need to be specialized and clarified. In this section we want to specialize the model for an application on a long series of daily returns. To do so, we need to answer some important questions.

The model assumes that there is a time ($t = 1$) where the process *starts* and the subsequent return PDF are shaped and derived from the PDF of this first return. Can this special time be identified? Are there some or many *restarts* in the series, that reset the value of the coefficients a_i? Is it important to identify them or not?

After these general remarks, another important question is about the value of the scaling parameter D and the shape of the scaling function g. How can these be determined?

One first key observation is that the time-inhomogeneity must vanish, for t exceeding a value $\tilde{t}$, where the Gaussian crossover occurs in force of CLT (scaling breakdown). So, if the available history is sufficiently long (of the order of the year), we assume the validity of the following approximation:

$$p_\tau(x) = \frac{1}{\tilde{t}} \sum_{t=1}^{\tilde{t}} p_{X_\tau(t)}(x). \tag{2.38}$$

A second remark is about the volatility. The model predicts that the volatility, *i.e.* the second moment of the returns, should have the following behaviour, as a function of t:

$$\begin{aligned} \langle x^2 \rangle_{p_{X_\tau(1+t)}} &= \int x^2 p_{X_\tau(1+t)}(x) dx \\ &= \int x^2 \frac{1}{\sqrt{(t+\tau)^{2D} - t^{2D}}} g\left(\frac{x}{\sqrt{(t+\tau)^{2D} - t^{2D}}} \right) dx \\ &\propto \left((t+\tau)^{2D} - t^{2D} \right). \end{aligned} \tag{2.39}$$

For $D = 0.5$ the volatility is constant in time t. For $D < 0.5$ the volatility decreases, with a maximum for $t = 1$; the opposite when $D > 0.5$. This remark suggests that (see, *e.g.*, 18), as far as D is estimated to be below 0.5, the times when local bursts in volatility occur may be considered as times where the process has a restart, so that the a_i are reset to 1. We will go deep along this line of thought in Sec. 3.3.

Baldovin and Stella have also shown (8) that for the DJI index the value of the empirically evaluated scaling exponent D that accounts for the time-inhomogeneity is approximately 0.24. The result is derived through a multiscaling analisys and in their work they show that the value is confirmed by the behaviour of the volatility autocorrelation. Nevertheless, when the analysis extends across the restarts (the exogenous events), the value $D = 0.5$ for the scaling exponent is confirmed.

In general, the value of D is established through an analysis of the non-linear moments, but also through a χ^2 minimization of the fitting of the scaling function g. We'll see this later in some applications.

The most delicate point of the theory is perhaps the identification of the restarts. The idea of identifying them with the local bursts in the volatility is not always satisfactory, and the more recent developments suggest that they can be also introduced by means of a stochastic Poisson process (working paper by Stella, Baldovin and Zamparo).

From all the above discussion, it seems obvious that the contact with real data is made by choosing a specific scaling function g, which best fits the rescaled and aggregated data over different τ's. Some efforts have been made to identify the more suitable function; in the following we report four possible candidates.

- The first proposal by Baldovin and Stella was the 3-parameter CF

$$\tilde{g}(k) = \exp\left(\frac{-Bk^2}{1 + C_\alpha |k|^{2-\alpha}}\right). \tag{2.40}$$

  In (46) it has been shown that the PDF associated to this CF has power law tails with exponent $5 - \alpha$. However this choice is not satisfying in that the explicit formula for the PDF is not computable, thus preventing from an explicit unfolding of the theory.

- Challet and Peirano (18) proposed instead a 2-parameter function: the multivariate Student distribution, defined as

$$s^{(n)}_{\nu,\lambda}(x_1, \ldots, x_n) = \frac{\Gamma\left(\frac{\nu}{2} + \frac{n}{2}\right)}{\pi^{\frac{n}{2}} \Gamma\left(\frac{\nu}{2}\right) \lambda^n} \frac{1}{\left(1 + \frac{x_1^2}{\lambda^2} + \ldots + \frac{x_n^2}{\lambda^2}\right)^{\frac{\nu}{2} + \frac{n}{2}}} \quad (n \geq 1). \tag{2.41}$$

The choice of these authors has been that of setting

$$p_{X_\tau(1),\ldots,X_\tau(1+(n-1)\tau)}(x_1,\ldots,x_n) = \frac{1}{\sqrt{\tau^n}} s^{(n)}_{\nu,\lambda}(\frac{x_1}{\sqrt{\tau}},\ldots,\frac{x_n}{\sqrt{\tau}}). \tag{2.42}$$

This Student distribution provides both analytic tractability and good fit to real market data (see (18) for the details).

- Convex combinations of Gaussian PDF's with varying widths (see, *e.g.*, 52, and references therein) have been used for example in (7). The idea has been developed in an intraday context, but can be naturally extended to an interday landscape, too. The g function is the following:

$$g(x) = \int_0^{+\infty} d\sigma\rho(\sigma)\frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}. \tag{2.43}$$

where $\rho(\sigma)$ is a normalized, positive measure in $]0,+\infty[$. Incidentally, the normalization of ρ implies the correct normalization of g:

$$\int_{-\infty}^{+\infty} g(x)dx = \int_0^{+\infty} d\sigma\rho(\sigma)\int_{-\infty}^{+\infty} dx\frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} = \int_0^{+\infty} d\sigma\rho(\sigma).$$

As already mentioned, the Schoenberg's Theorem (2, 43, 49) shows that the more general g's that satisfy the model conditions are in the form of Gaussian mixtures. For this reason, in the following, we will exclusively refer to this choice.

2.2.5 The Gaussian mixture formulation

Given the form of Eq. 2.43 for the scaling function g, one naturally faces the task of detecting the most suitable form for the $\rho(\sigma)$. This can be easily identified, *e.g.* by matching its moments with those of g, and by relating the large behavior of $\rho(\sigma)$ with the large $|x|$ behavior of $g(x)$. For instance, ρ may decay as a power law at large σ's if the moments of g are expected to be infinite above a given order.

The choice for the shape of the ρ function can be guided, although not directly, by the distribution of the volatility. Anyway, the evaluation of the parameters of these ρ's is done by fitting the scaling function g.

We will use, *e.g.*, the following alternative parametrizations:

i) A form that provides power law decay for large σ is:

$$\rho(\sigma) = A\,\frac{b^{\nu-\lambda-1}\sigma^\lambda}{b^\nu+\sigma^\nu}. \tag{2.44}$$

This parametrization, used for instance in (49), has three free parameters (ν, λ and b), while $A = A(\nu, \lambda, b)$ is a normalization factor derived by imposing:

$$1 \equiv \int_0^{+\infty} d\sigma \rho(\sigma) = \int_0^{+\infty} d\sigma A \, \frac{b^{\nu-\lambda-1}\sigma^{\lambda}}{b^{\nu} + \sigma^{\nu}}. \tag{2.45}$$

ii) A common choice is to use the *inverse gamma function*:

$$\rho(\sigma) = \frac{\sigma_0^{\alpha}}{\Gamma(\alpha)\sigma^{1+\alpha}} \exp\left(-\frac{\sigma_0}{\sigma}\right); \tag{2.46}$$

as will be clear in a moment, this function probably best fits the data both for low and large values of σ. It has two parameters.

iii) Finally, another standard reference function is the *log-normal* distribution:

$$\rho(\sigma) = \frac{1}{\sigma\sqrt{2\pi\sigma_0^2}} e^{-\frac{(\ln\sigma-\mu)^2}{2\sigma_0^2}}, \tag{2.47}$$

which again has only two parameters.

To evaluate the difference between these three, we show how they best approximate the daily volatility of the S&P500 in the period 1983-2010. To do so, according to (17), we use the following definition of the daily *volatility proxy*:

$$\sigma_p = \frac{1}{N_d} \sum_{i=1}^{N_d} |x_{5'}(i)|, \tag{2.48}$$

where $x_{5'}(i)$ is the i-th return of the day, on a 5 minutes scale, and N_d is the total number of these intraday returns.

The distribution of the so defined σ's is shown in Fig. 2.2, as long with the three different fitting functions.

Looking at Fig. 2.2, the inverse gamma seems to be the best fitting function for the volatility proxy. This evidence is not directly related to the optimal choice of the ρ function and, as will be seen hereafter, the other choices give a comparably good agreement for the scaling collapse, too.

2.2.6 The model for daily series

In this section we summarize the results so far achieved and we describe the model for reproducing daily series. The subject is still developing in these days (10) and particular efforts are devoted to the problem of the calibration, which is essential in an interday context.

Anyway, the results so far established are:

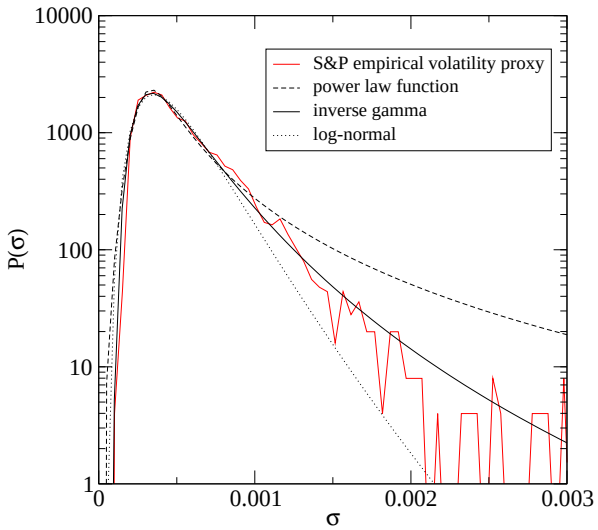


Figure 2.2: Empirical distribution of the volatility proxy of the S&P500 dataset (from 1983 to 2010, 4969 days, N_d=75). The best fits with the candidates for the ρ functions are given, too.

- The time-inhomogeneous scaling property of Eq. 2.37 of the returns of a financial time series suggests a general form for the ensemble PDF of the returns themselves.

- Starting from particular days, which can be identified with peculiar events or shocks in the market history (or possibly also with a realization of a stochastic Poisson process), the PDF of the daily returns evolves in time as follows (see Eq. 2.37, specialized for $\tau = 1$, and for $t = 1, 2, \ldots$):

$$p_{X_1(t)}(x) = \frac{1}{\sqrt{t^{2D} - (t-1)^{2D}}} g\left(\frac{x}{\sqrt{t^{2D} - (t-1)^{2D}}}\right). \tag{2.49}$$

- The scaling function g can be written as a Gaussian mixture (Eq. 2.43), weighted with a ρ function which assumes, *e.g.*, one of the forms of Eq. 2.44, Eq. 2.46 or Eq. 2.47. The resulting form reads:

$$g(x) = p_{X_1(1)}(x) = \int_0^{+\infty} d\sigma \rho(\sigma) \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}. \tag{2.50}$$

- According to Eq. 2.24, the same scaling function can be used to build the ensemble PDF at different times, distant $(i-1)$ unit time intervals from the origin time

$t = 1$. The PDF is simply rescaled via the a_i parameters.

$$\begin{aligned} p_{X_1(1+(i-1))}(x) = p_{X_1(i)}(x) &= \frac{1}{a_i^D} g\left(\frac{x}{a_i^D}\right) \\ &= \int_0^{+\infty} d\sigma \rho(\sigma) \frac{e^{-\frac{x^2}{2a_i^{2D}\sigma^2}}}{\sqrt{2\pi a_i^{2D}\sigma^2}}. \end{aligned} \tag{2.51}$$

At the initial, special time $i = 1$ one has $a_i = a_1 = 1$. Analogous results can be derived for different τ's:

$$\begin{aligned} p_{X_\tau(1)}(x) &= \frac{1}{\tau^D} g\left(\frac{x}{\tau^D}\right) \\ &= \int_0^{+\infty} d\sigma \rho(\sigma) \frac{e^{-\frac{x^2}{2\tau^{2D}\sigma^2}}}{\sqrt{2\pi\tau^{2D}\sigma^2}}. \end{aligned} \tag{2.52}$$

- The joint PDF function for k successive returns can be constructed, thanks to the introduction of a generalized product in Fourier space, using Eq. 2.33, which again is based on the same scaling function g. As an example, one can write

$$p_{X_\tau(1),\ldots,X_\tau(1+(k-1)\tau)}(x_1,\ldots,x_k) = \int_0^{+\infty} d\sigma \rho(\sigma) \prod_{i=1}^{k} \frac{e^{-\frac{x_i^2}{2(a_i\tau)^{2D}\sigma^2}}}{\sqrt{2\pi(a_i\tau)^{2D}\sigma^2}}. \tag{2.53}$$

Eq. 2.53 can easily be checked to be in agreement with Eq. 2.33. Noticing that $\tilde{g}(q) = \int_{-\infty}^{+\infty} e^{iqx} g(x) = \int_0^{+\infty} d\sigma \rho(\sigma) e^{-\frac{\sigma^2 q^2}{2}}$, one easily gets:

$$\begin{aligned} &\tilde{p}_{X_\tau(1),\ldots,X_\tau(1+(k-1)\tau)}(q_1,\ldots,q_k) = \\ &= \int_0^{+\infty} d\sigma \rho(\sigma) \prod_{i=1}^{k} e^{-\frac{(a_i\tau)^{2D}\sigma^2 q_i^2}{2}} \\ &= \int_0^{+\infty} d\sigma \rho(\sigma) e^{-\frac{\sigma^2}{2}\left(\sum_{i=1}^{k}(a_i\tau)^{2D} q_i^2\right)} \\ &= \tilde{g}\left(\sqrt{\sum_{i=1}^{k}(a_i\tau)^{2D} q_i^2}\right). \end{aligned} \tag{2.54}$$

- The main practical issues associated with the model are:
  - Identifying the scaling parameter D and the scaling function g.
  - Identifying (whenever possible) the days when restarts occur (and wherein $a_i = a_1 = 1$).

2.3 The intraday formulation

The powerful scheme for the modeling of financial data in an *interday* scale can be applied in an *intraday* scale, too. If the discussed model is able to reproduce the financial time series of daily returns, this is essentially due to the scaling hypothesis. Looking back at the equations derived in Section 2.2, and particularly at Eq. 2.18, one immediately realizes that the model is constructed so as to be applied to different time-scales τ. The parameters a_i ruling the rescaling of the PDF of the returns are τ-independent (see Eq. 2.28) and so the rescaling of a PDF at time t depends only on the number of time intervals separating the time t from the initial, special time where $i = 1$.

So it is possible to extend the results of Sec. 2.2 to an intraday scale, as long as the linear correlations still vanish (we will confirm that, for the more recent high frequency databases, Eq. 2.25 is valid down to a 5-10 minutes scale).

To do so, it is important to identify the counterpart of the market events (or shocks) that act as a restart for the model. In other words we should define a procedure to highlight the particular times along the series where the condition $a_i = a_1 = 1$ can be imposed. Since every day is composed of many intervals, the restarting event should be identified in a higher detail. In principle we need to identify the restart time with a precision that matches the time width τ.

Within the intraday range, another procedure is available. It takes its origin from the results of Bassler *et al.* (11) who analyzed the EUR/$ exchange rate in an intraday scale, identifying a general behaviour of the volatility, which can be at the basis of a restarting procedure for this intraday model, too. Following the evidence that the volatility shows a clear intraday pattern, the cited authors assume that the stochastic part of the intraday variations of the prices follow the same underlying stochastic process every day. Even if these realizations, corresponding to the prices of every single day, are not independent (and higher order correlations over different days are effectively detected), these realizations constitute an ensemble of histories that can be used, *e.g.*, to calculate ensemble averages.

But what is more relevant, for our present ideas, is the fact that the beginning of the day can be considered, for all practical purposes, as a restart point, a time where we can reset the a_i's.

The generalization works as follows: be

$$x_1^{(1)}(1), x_1^{(1)}(2), \ldots, x_1^{(1)}(N_d), \ \ldots, \ x_1^{(M)}(1), x_1^{(M)}(2), \ldots, x_1^{(M)}(N_d)$$

a given interday series of returns, being N_d the number of daily successive intervals of given size $\tau = 1$ and M the number of days in the series. So, from Eq. 2.49 the following scaling relation holds, for $j = 1, \ldots, M$ and $i = 1, \ldots, N_d$:

$$p_{X_1^{(j)}(i)}(x) = \frac{1}{\sqrt{i^{2D} - (i-1)^{2D}}} g\left(\frac{x}{\sqrt{i^{2D} - (i-1)^{2D}}}\right). \tag{2.55}$$

This scaling property is valid for every day in the series, given that a scaling function g and a scaling parameter D may be uniquely identified.

As in the interday model, the scaling function defines a PDF for the first return of the day:

$$p_{X_1^{(j)}(1)}(x) = g(x) = \int_0^{+\infty} d\sigma \rho(\sigma) \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}. \tag{2.56}$$

By employing Eq. 2.55, one may analogously derive the PDF for the i-th return of the day ($i = 1, \ldots, N_d$). Remembering that $a_i^D = \sqrt{i^{2D} - (i-1)^{2D}}$ (Eq. 2.28), one finally gets:

$$p_{X_1^{(j)}(i)}(x) = \int_0^{+\infty} d\sigma \rho(\sigma) \frac{e^{-\frac{x^2}{2a_i^{2D}\sigma^2}}}{\sqrt{2\pi a_i^{2D}\sigma^2}}. \tag{2.57}$$

Along the same reasoning as before, one can define a joint PDF for k successive returns. In the present case, Eq. 2.53 takes the following special form:

$$p_{X_1^{(j)}(1),\ldots,X_1^{(j)}(k)}(x_1, \ldots, x_k) = \int_0^{+\infty} d\sigma \rho(\sigma) \prod_{i=1}^{k} \frac{e^{-\frac{x_i^2}{2a_i^{2D}\sigma^2}}}{\sqrt{2\pi a_i^{2D}\sigma^2}}. \tag{2.58}$$

Note that the j-dependence is suppressed, in the previous formulas, according to the hypothesis that every single day is a realization of the same stochastic process. We will use all these results in Chap. 4.

2.4 Introducing the skewness: the leverage effect

As remarked by some authors (18), one of the main deficiencies of the theory so far explained is the lack of skewness for the PDF's and the failure to reproduce one important stylized fact as the leverage effect. In this section we investigate the matter and we try to show a possible solution.

The leverage effect can be defined in various ways, but essentially it is a negative correlation between past returns and future volatilities (*i.e.* large returns are associated to future periods of high volatility). An economic interpretation of this effect is still

controversial (see (12) for a review) and as long with a satisfying quantitative description of it. This is also reflected by the lower number of investigation papers devoted to it with respect to other stylized facts as, *e.g.*, the volatility clustering.

A suggested indicator for the quantitative analysis of the leverage effect is the leverage correlation function, defined in (16) as:

$$L_\tau(T) = \frac{1}{K}\langle |x_\tau(t+T)|^2 x_\tau(t)\rangle \tag{2.59}$$

where the normalization factor K takes the following suitable form:

$$K = \langle x_\tau(t)^2\rangle^2. \tag{2.60}$$

The empirical average is thus realized by averaging over t.

We want to investigate if the model described in the previous sections may approximate the empirical results. The model generates symmetric PDF's. Looking at the general form of the model PDF's (Eq. 2.43) one immediately realizes that $g(x) = g(-x)$ and this symmetry extends, due to scaling, to all the PDF's, for every t and τ (Eq. 2.37). There is a link between the leverage effect and the skewness of the price returns distribution: if we look at the theoretical counterpart of Eq. 2.59, taking *e.g.* $\tau = T = 1$, we have (for $i = j+1$)

$$\begin{aligned} L_1^{(th)_{i,j}}(1) &= \langle x_i^2 x_j\rangle_p \\ &= \int_0^{+\infty} d\sigma\rho(\sigma) \int\int dx_i dx_j\ x_i^2 x_j\ \frac{e^{-\frac{x_i^2}{2a_i^{2D}\sigma^2}}}{\sqrt{2\pi a_i^{2D}\sigma^2}} \frac{e^{-\frac{x_j^2}{2a_j^{2D}\sigma^2}}}{\sqrt{2\pi a_j^{2D}\sigma^2}}. \end{aligned} \tag{2.61}$$

Clearly the integral vanishes, because of the symmetry properties of the x_j integrand. The same result is obtained for different T's and τ's. Thus, the absence of skewness of the PDF's in this symmetric model forbids any correlation of odd powers of x, leverage effect included.

One possible way to introduce a certain degree of skewness, without altering the general scheme of the so far exposed model, is that of defining the scaling function g as a Gaussian mixture not only with respect to the variance σ, but also with respect to the mean μ, as follows:

$$g(x) = p_{X_1(1)}(x) = \int_0^{+\infty}\int_0^{+\infty} d\sigma d\mu\ \psi(\sigma,\mu)\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}. \tag{2.62}$$

As a simplest example, one may write

$$\psi(\sigma,\mu) = \frac{1}{2}\rho_1(\sigma)\delta(\mu-\mu_1) + \frac{1}{2}\rho_2(\sigma)\delta(\mu-\mu_2), \tag{2.63}$$

with ρ_1 and ρ_2 two normalized measures. The ψ function is itself a normalized measure, as

$$\int_0^{+\infty}\int_0^{+\infty} d\sigma d\mu\ \psi(\sigma,\mu) = \frac{1}{2}\int_0^{+\infty} d\sigma\rho_1(\sigma)\cdot 1 + \frac{1}{2}\int_0^{+\infty} d\sigma\rho_2(\sigma)\cdot 1 = 1. \tag{2.64}$$

The introduction of a mixture of non-zero centered Gaussians provide a good scheme for obtaining non-zero odd-power moments. In fact, with the definition of Eq. 2.63, Eq. 2.62 becomes:

$$g(x) = p_{X_1(1)}(x) = \int_0^{+\infty} d\sigma\ \frac{\rho_1(\sigma)}{2}\frac{e^{-\frac{(x-\mu_1)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} + \int_0^{+\infty} d\sigma\ \frac{\rho_2(\sigma)}{2}\frac{e^{-\frac{(x-\mu_2)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}. \tag{2.65}$$

and so, for instance, the mean value of the distribution of the first return is:

$$\begin{aligned}\langle x\rangle_p &= \int dx\ xP_{X_1(1)}(x)\\ &= \int_0^{+\infty} d\sigma\ \frac{\rho_1(\sigma)}{2}\int dx\ x\frac{e^{-\frac{(x-\mu_1)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} + \int_0^{+\infty} d\sigma\ \frac{\rho_2(\sigma)}{2}\int dx\ x\frac{e^{-\frac{(x-\mu_2)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\\ &= \frac{\mu_1}{2}+\frac{\mu_2}{2}.\end{aligned} \tag{2.66}$$

As long as $\mu_1 \neq -\mu_2$ the PDF of the first return is non-zero centered. One may keep the non-arbitrage condition by imposing $\mu_1 = -\mu_2 = \overline{\mu}$. Even with this choice the model can have an asymmetry in the PDF, thanks to the arbitrary form of the ρ_i functions. In passing, we notice that one can always impose $\overline{\mu} \geq 0$, upon possibly exchanging the indexes of the ρ's.

Baldovin and Stella have shown (49) that a generic form for the joint PDF of N successive returns p_N that preserves the scaling hypothesis is possible even with these non-zero centered mixtures:

$$p_{X_1(1),\ldots X_1(N)}(x_1,\ldots,x_N) = \int_0^{+\infty}\int_0^{+\infty} d\sigma d\mu\ \psi(\sigma,\mu)\prod_{i=1}^{N}\frac{e^{-\frac{(x-b_i\mu)^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}}, \tag{2.67}$$

with[1] $a_i = i^{2D} - (i-1)^{2D}$ and $b_i = i^D - (i-1)^D$. This formula is compatible with Eq. 2.62, as it should.

We can use Eq. 2.67 and Eq. 2.63 (with $\mu_1 = -\mu_2 = \overline{\mu}$) to calculate the expected values of the generic i-th return ($i = 1, 2, \ldots$):

$$\langle x_i\rangle_p = \int dx\ x\ p_{X_1(i)}(x)$$

[1] Notice the slight change in the notations: the exponent D is now included in the definition of a_i (compare with Eq. 2.28).

$$
\begin{aligned}
&= \int_0^{+\infty} d\sigma \ \frac{\rho_1(\sigma)}{2} \int dx \ x \frac{e^{-\frac{(x-b_i\overline{\mu})^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} + \int_0^{+\infty} d\sigma \ \frac{\rho_2(\sigma)}{2} \int dx \ x \frac{e^{-\frac{(x+b_i\overline{\mu})^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \\
&= \frac{b_i\overline{\mu}}{2} - \frac{b_i\overline{\mu}}{2} = 0. \qquad (2.68)
\end{aligned}
$$

Analogously, upon defining $\langle\sigma^\alpha\rangle_\rho \equiv \int d\sigma\sigma^\alpha\rho(\sigma)$, we derive the second and the third moment:

$$
\begin{aligned}
\langle x_i^2\rangle_p &= \int dx \ x^2 \ p_{X_1(i)}(x) \\
&= \int_0^{+\infty} d\sigma \ \frac{\rho_1(\sigma)}{2} \int dx \ x^2 \frac{e^{-\frac{(x-b_i\overline{\mu})^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} + \int_0^{+\infty} d\sigma \ \frac{\rho_2(\sigma)}{2} \int dx \ x^2 \frac{e^{-\frac{(x+b_i\overline{\mu})^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \\
&= \int_0^{+\infty} d\sigma \ \frac{\rho_1(\sigma)}{2} \int dx \ (x+b_i\overline{\mu})^2 \frac{e^{-\frac{x^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \\
&\quad + \int_0^{+\infty} d\sigma \ \frac{\rho_2(\sigma)}{2} \int dx \ (x-b_i\overline{\mu})^2 \frac{e^{-\frac{x^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \\
&= b_i^2\overline{\mu}^2 + \frac{a_i^2\left(\langle\sigma^2\rangle_{\rho_1} + \langle\sigma^2\rangle_{\rho_2}\right)}{2} \qquad (2.69)
\end{aligned}
$$

and

$$
\begin{aligned}
\langle x_i^3\rangle_p &= \int dx \ x^3 \ p_{X_1(i)}(x) \\
&= \int_0^{+\infty} d\sigma \ \frac{\rho_1(\sigma)}{2} \int dx \ (x+b_i\overline{\mu})^3 \frac{e^{-\frac{x^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \\
&\quad + \int_0^{+\infty} d\sigma \ \frac{\rho_2(\sigma)}{2} \int dx \ (x-b_i\overline{\mu})^3 \frac{e^{-\frac{x^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \\
&= \int_0^{+\infty} d\sigma \ \frac{\rho_1(\sigma)}{2} (b_i^3\overline{\mu}^3 + 3b_i\overline{\mu}a_i^2\sigma^2) \\
&\quad + \int_0^{+\infty} d\sigma \ \frac{\rho_2(\sigma)}{2} (-b_i^3\overline{\mu}^3 - 3b_i\overline{\mu}a_i^2\sigma^2) \\
&= \frac{3}{2} b_i a_i^2 \overline{\mu} \left(\langle\sigma^2\rangle_{\rho_1} - \langle\sigma^2\rangle_{\rho_2}\right) \qquad (2.70)
\end{aligned}
$$

The *skewness* $\gamma_1 \equiv \frac{\langle x^3\rangle}{\langle x^2\rangle^{\frac{3}{2}}}$ of the distribution, defined as the third standardized moment, is then larger than zero if $\langle\sigma^2\rangle_{\rho_1} > \langle\sigma^2\rangle_{\rho_2}$ and negative if $\langle\sigma^2\rangle_{\rho_1} < \langle\sigma^2\rangle_{\rho_2}$ (we had set $\overline{\mu} \geq 0$). Referring to the financial data we may impose the second inequality, since we want the fatter tail on the right (negative) side of the PDF's.

We are now ready to verify the existence of a *leverage effect* in this model: the new form of Eq. 2.61 is:

$$
\begin{aligned}
L_1^{(th)_{i,j}}(1) &= \langle x_i^2 x_j \rangle_p \\
&= \iint d\sigma d\mu \; \psi(\sigma,\mu) \iint dx_i dx_j \; x_i^2 x_j \; \frac{e^{-\frac{(x_i - b_i\mu)^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \frac{e^{-\frac{(x_j - b_j\mu)^2}{2a_j^2\sigma^2}}}{\sqrt{2\pi a_j^2\sigma^2}}, \\
&= \int d\sigma \; \frac{\rho_1(\sigma)}{2} \int dx_i \; x_i^2 \; \frac{e^{-\frac{(x_i - b_i\overline{\mu})^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \int dx_j \; x_j \; \frac{e^{-\frac{(x_j - b_j\overline{\mu})^2}{2a_j^2\sigma^2}}}{\sqrt{2\pi a_j^2\sigma^2}} \\
&\quad + \int d\sigma \; \frac{\rho_2(\sigma)}{2} \int dx_i \; x_i^2 \; \frac{e^{-\frac{(x_i + b_i\overline{\mu})^2}{2a_i^2\sigma^2}}}{\sqrt{2\pi a_i^2\sigma^2}} \int dx_j \; x_j \; \frac{e^{-\frac{(x_j + b_j\overline{\mu})^2}{2a_j^2\sigma^2}}}{\sqrt{2\pi a_j^2\sigma^2}} \\
&= \int d\sigma \; \frac{\rho_1(\sigma)}{2} \left[\left(b_i^2\overline{\mu}^2 + a_i^2\sigma^2 \right) \cdot b_j\overline{\mu} \right] \\
&\quad - \int d\sigma \; \frac{\rho_2(\sigma)}{2} \left[\left(b_i^2\overline{\mu}^2 + a_i^2\sigma^2 \right) \cdot b_j\overline{\mu} \right] \\
&= \frac{a_i^2 b_j \overline{\mu}}{2} \left(\langle \sigma^2 \rangle_{\rho_1} - \langle \sigma^2 \rangle_{\rho_2} \right). \qquad (2.71)
\end{aligned}
$$

We will check later (Sec. 4.1.3.1) the agreement of this formula with the data. Here we just stress what follows:

- the asymmetric features of the $g(x)$ make a *leverage effect* emerge in the model;
- the sign of the effect is in agreement with the choice $\langle \sigma^2 \rangle_{\rho_1} < \langle \sigma^2 \rangle_{\rho_2}$ that, as seen above, gives a negative skewness;
- the effect is not symmetric under $i \leftrightarrow j$ exchange, as expected. As a drawback, the effect is not zero for $j > i$, even if smaller than for $i > j$.

3

The interday formulation

An approximate answer to the right problem
is worth a good deal more than
an exact answer to an approximate problem.
J.W. Tukey

For in a minute there are many days.
W. Shakespeare

3.1 Detecting scaling symmetries

In the present and in the following chapter we will separately study interday (Chap. 3) and intraday returns statistics (Chap. 4). We will verify, over different time scales and time lags, the scaling assumption of Eq. 1.8 on which the theory of Chap. 2 is based. Moreover, we will explore some consequences and effects deriving from this scaling approach, that also focus on the predictive power of the model. However, before approaching this task, some general remarks about the way of detection of scaling symmetries are in order.

Since the idea of scaling was brought into the financial market analysis, many studies have been devoted to the verification of the scaling hypothesis and to its consequences (see, *e.g.*, 11, 21, 34, 37, 48). Given the importance of the topic, we decided to explore and to test further the scaling hypothesis, principally aiming at the validation of a method that may drive with sufficient reliability to the identification of the scaling function and of the scaling exponent. In addition, in case of returns calculated over different trading days in non continuous markets, we explored the effect of the inclusion or of the exclusion of the overnight return. We will refer in particular to the following high frequency dataset: the collection of S&P500 index values from 9:35 to 16:00 NY

time, from September 30th 1985 to October 20th 2010. We deal altogether with 6253 days, with 78 returns per day (and one additional overnight return).

3.1.1 The scaling over different time scales: dealing with overnight returns

The range of validity of Eq. 1.8 is not obviously infinite. As already noted in Sec. 2.2.4, the CLT generates a crossover to normal scaling for sufficiently large time lags τ. On the other hand, the financial series are finite and actually not so long. Therefore it is difficult to build rich statistics of returns over long times: with our 25 years of data, *e.g.*, only 6252 non overlapping daily returns can be obtained, a really unsatisfactory value to investigate the shape of the histogram of the returns, particularly in the tails! One solution may be that of taking also overlapping returns[1], but at the costly price of having strongly correlated returns.

The construction of the returns preliminarily proceeds as follows: first of all we construct the series of 5-min returns[2] $\{x_5(t)\}$ (neglecting the overnight). Therefore, having obtained 78 returns per day (487734 returns altogether), we profit of the additive property of log returns, to build the returns over different time lags: to generate the series of (overlapping) returns at a $(5\cdot n)$-min scale ($n \in \mathbb{N}_{\geq 2}$), we simply add successive returns:

$$x_{(5\cdot n)}(t) = x_5(t) + \ldots + x_5(t+(n-1)), \quad t = 1, 2, \ldots, 487734 - (n-1). \tag{3.1}$$

The returns series obtained in this way are rescaled according to Eq. 1.8. When dealing with a long time series, the time-inhomogeneous effects, that are evident only in limited time windows, disappear: the inhomogeneities are also leveled off by the sliding window sampling, and then one has $D = 1/2$. The returns on a time scale τ are then divided by $\tau^{1/2}$. The returns are aggregated on an histogram and normalized, again according to Eq. 1.8. The result is presented in Fig. 3.1 for $\tau = 10$ min, $\ldots, 5$ days[3]. In any case, for τ's larger than 10 days, some problems arise with these returns, as can be seen in Fig. 3.2.

The reason for this scaling breakdown must be found in the way the returns are built in Eq. 3.1. When times larger than one day are considered, the overnight return starts

[1] With the expression *overlapping returns* we refer to the returns generated across overlapping time intervals. For instance, given the historical series $S(1), S(2), S(3), S(4)$... the returns $x_2(1) = S(3) - S(1)$ and $x_2(2) = S(4) - S(2)$ are overlapped in the time interval $2 - 3$.

[2] We remember that with *return* we actually decided to refer to the *log-returns* as defined in Sec. 1.1.2

[3] For clarity, we highlight that the returns over a time lag of 1 day (5 days) are built by summing 78 (390) consecutive 5-min returns. We stress again that the overnight return is left out.

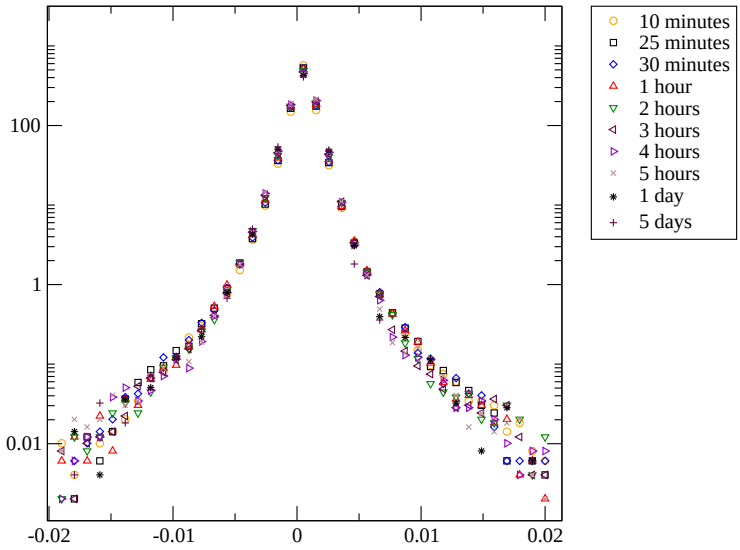


Figure 3.1: Scaling collapse of S&P500 returns (1985-2010), without overnight returns. The returns over a time lag of 1 day (5 days) are built by summing 78 (390) consecutive 5-min returns, again without the overnight return. The scaling hypothesis of Eq. 1.8 is reasonably well verified on the reported time scales.

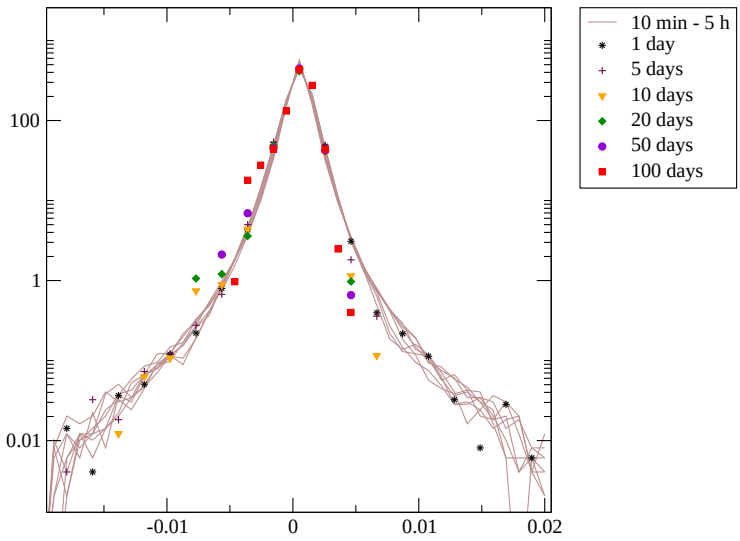


Figure 3.2: Scaling collapse of S&P500 returns (1985-2010), without overnight returns. The returns over a time lag of n days are built by summing $78 \cdot n$ consecutive 5-min returns, without the overnight return. The scaling breaks for τ's larger than 10 days.

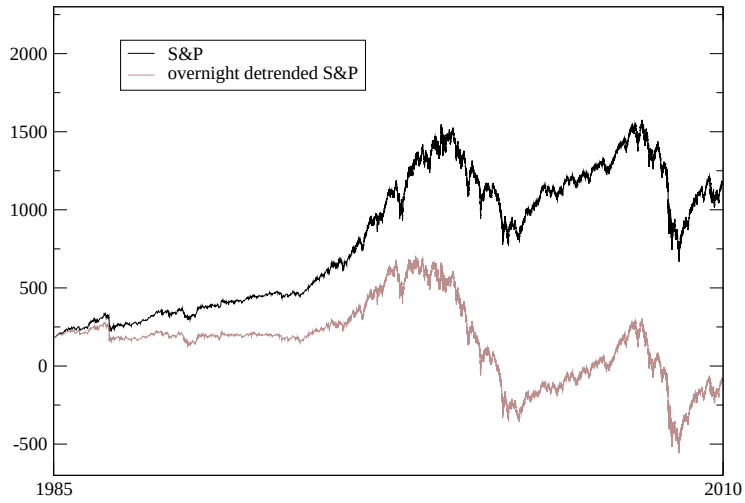


Figure 3.3: Difference between the S&P500 index value and the virtual S&P500 value reconstructed from the returns calculated via Eq. 3.1, without the overnight contribution. The overnight returns significantly contribute to the trend of the index.

being important because it is essential to the trend of the index itself, as evident from Fig. 3.3 where the S&P500 index value is compared to the virtual S&P500 value reconstructed from the returns calculated via Eq. 3.1, without the overnight contribution.

Therefore another procedure has been used to investigate the scaling over large times: directly calculate the return from the index value, neglecting the presence of a *peculiar* return (or better, of a peculiar time window where the market is closed and no trading occurs). The result of this scaling collapse, obtained for $\tau = 3$ days 150 days, is displayed in Fig. 3.4. The inclusion of the overnight returns appears to be essential to the validation of the scaling assumption for time intervals of more than some days.

A direct comparison between the results of Fig. 3.1 and Fig. 3.4 is not possible, essentially because of the different way in which the returns are calculated and of the different treatment of the overnight returns. Moreover in Fig. 3.1 the collapse is referred to a 5 minutes scale, while in Fig. 3.4 to a 1 day scale. Nevertheless, it's surprising to see in Fig. 3.5 the good overlap of the two different sets of returns: those calculated up to a time lag of 5 days via Eq. 3.1 (without the overnight) and the interday sets with the overnight included.

In conclusion, the scaling is verified for both sets of returns. The mixing of the two should be done only after a deep analysis of the effect of the overnight return has been made. In particular, the inclusion of the overnight return is essential to reproduce the

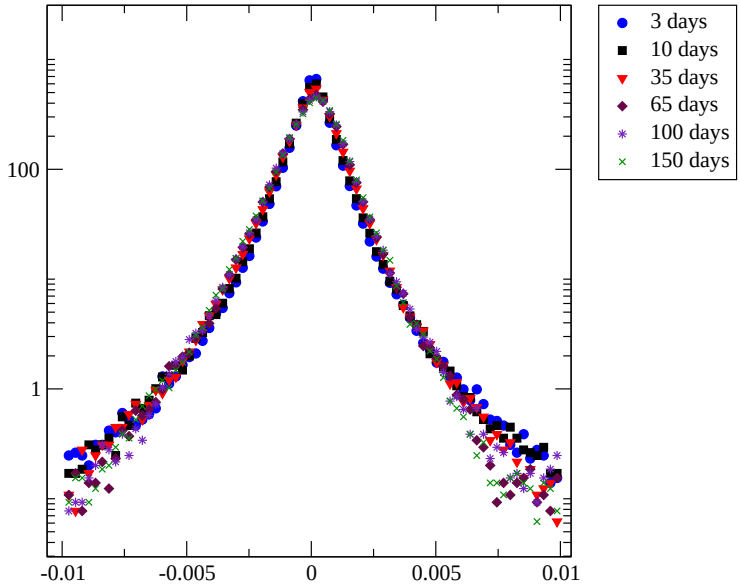


Figure 3.4: Scaling collapse of S&P500 returns (1985-2010) with overnight returns included. The scaling lasts for τ's larger than 10 days, up to about 100 days and more. Compare with Fig. 3.2

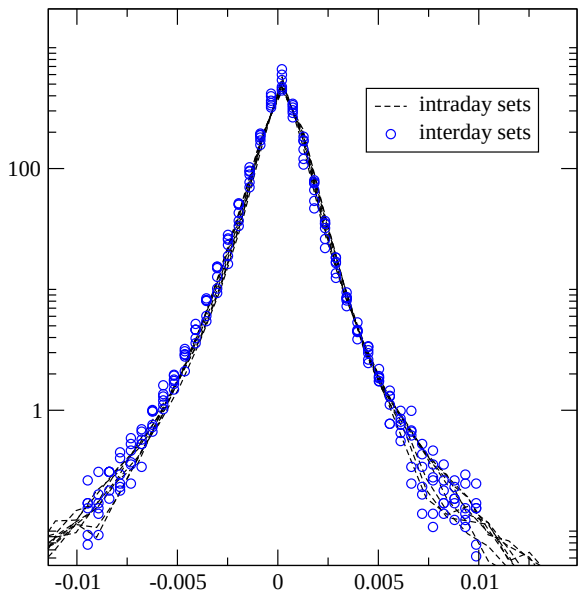


Figure 3.5: Comparison between the scaling collapse of S&P500 *intraday* and interday (overnight included) returns (1985-2010). The expression *intraday sets* here refers to returns calculated via Eq. 3.1, up to a time lag of $390 \cdot 5$ min $= 5$ days. The overlap of the two sets of histograms is not fully expected (see text).

scaling collapse for series of interday returns (Fig. 3.4).

In the following, however, we will keep these two worlds separate, dealing separately with interday (in this chapter) and intraday (in Chap. 4) returns statistics.

3.1.2 Overnight vs. over-weekend returns

In the previous section we left out one potentially relevant issue: during some days (Sundays, holidays, ...) the market is closed. We discarded this problem by linking every market day with the following one, irrespective of the presence of holidays, and then neglecting to discriminate between the different time lag of the returns. We generally referred to *overnight* returns, but actually the difference in time can be as large as *one overnight and one day* (the return after one single day of holiday) or *one overnight and two days* (the return after the weekend). Substantially, we assumed that every weekend and every holiday lasts exactly one night.

In the following, we want to verify whether this assumption is justified. A first key observation is that the scaling collapse of Fig. 3.4 seems not to be affected by problems of this kind: actually, a careful check of the legend suggests that the scaling collapse is obtained right with time lags that do not include the days of market closure. This is a very different result, with respect to the result of Fig. 3.2, obtained neglecting the overnight return, where the scaling was not verified.

Furthermore, a detailed work of selection permits us to separate the *overnight* from the *over-weekend* returns. So, leaving out the poor statistics for returns over single days of holiday and for those of more than two days of market inactivity, we can build an histogram for both sets of selected returns: the result is displayed in Fig. 3.6. The graphs highlight a good overlapping of the two statistics and clearly suggest that the assumption of treating overnight and over-weekend returns in the same way is valid. The result can be extended by induction to the other, more rare, closing intervals.

3.2 Simulating histories

The computer simulation of stochastic series obeying given statistics is a complex task and, unavoidably, we will only face the problem with methods and tools commensurated with the limits and the aims of this thesis. Nevertheless a glimpse of the adopted procedures for the generation of intraday and interday series of returns will be given, for completeness.

The central issue, in this specific context, is the necessity of generating returns series according to PDF's defined as in Eq. 2.52 or in Eq. 2.57 for intraday and interday

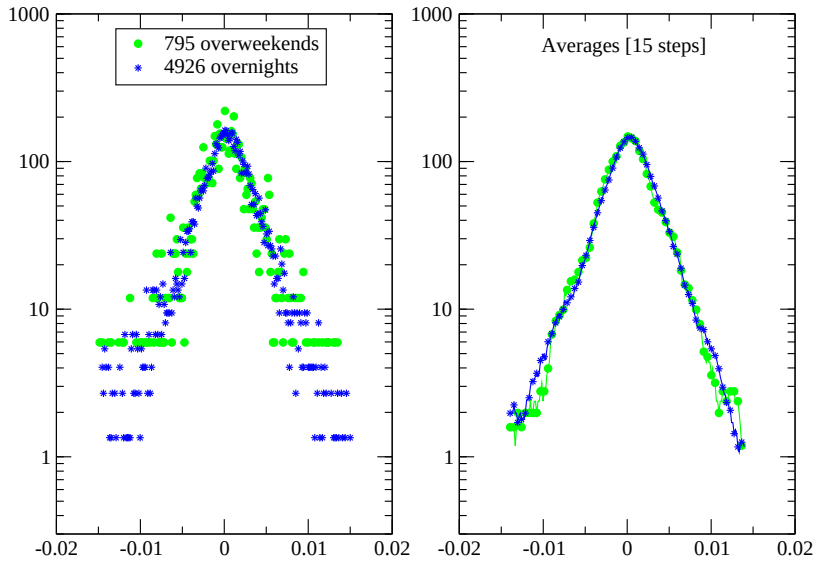


Figure 3.6: Comparison between the returns (S&P500 1985-2010) over one single night and over one weekend (*left*). To smooth the graph, the average (over 15 bins) has been run (*right*).

returns, respectively. As a preliminary to this, one must also decide how to spot the days (in the interday context) or the minutes (in an intraday context) when the reset of the a_i parameters occurs. One adopted solution for the interday scheme is to randomize the reset of the a_i's according to a Poisson distribution of a given variance (typically ranging between $30 - 150$ days). For the intraday scheme, instead, following the reasoning already exposed in Sec. 2.3, the generation of the returns is done according to the choice $a_i = a_1 = 1$ for the first return of the day.

The goal of generating returns with probability given by the desired PDF's is ultimately achieved by application of the so called *rejection sampling*, a method first proposed by von Neumann (50) and also called *accept-reject algorithm* (41). It unfolds in two steps:

I a $\tilde{\sigma}$ is chosen, according to the PDF expressed by the $\rho(\sigma)$ function;

II for every $t = 1, 2, \ldots, t_f$, with t_f the time of the next restart, a Gaussian stochastic variable is generated, with 0 mean and $a_t^2\tilde{\sigma}^2$ variance.

Let us give some more details for I. The procedure necessarily passes through the discretization of the σ variable and of the ρ function: the σ's range in the interval $\sigma_{min} \to \sigma_{Max}$, where the limits are chosen according to the given statistics (for instance,

σ_{min} may be half the lowest and σ_{Max} twice the highest absolute return in the history of real market data to be replicated by the model).

The ρ function is discretized as follows: $\rho(\sigma) \to \rho(\sigma_l)$ with $l = 1, 2, \ldots, l_{Max}$ and $\sigma_1 = \sigma_{min}$, $\sigma_2 = \sigma_{min} + \Delta$, $\ldots$ $\sigma_{l_{Max}} = \sigma_{min} + (l_{Max} - 1)\Delta = \sigma_{Max}$.

The interval Δ is chosen so that $\Delta = (\sigma_{Max} - \sigma_{min})/(l_{Max} - 1)$ and the normalization condition for the ρ is $\sum_{l=1}^{l_{Max}} \Delta \cdot \rho(\sigma_l) = 1$.

The choice of the $\tilde{\sigma}$ is made in the discrete range $\sigma_1, \sigma_2, \ldots \sigma_{l_{Max}}$. One random, uniformly chosen r is picked in the interval $[0, 1[$; then if $\Delta \cdot \rho(\sigma_1) > r$ the choice is $\tilde{\sigma} = \sigma_1$ otherwise if $\sum_{l=1}^{\tilde{l}-1} \Delta \cdot \rho(\sigma_l) \leq r$ and $\sum_{l=1}^{\tilde{l}} \Delta \cdot \rho(\sigma_l) > r$ the choice is $\tilde{\sigma} = \sigma_{\tilde{l}}$. It is easy to convince oneself that, drawing more and more $\tilde{\sigma}$'s, one finally gets a distribution that approximates the ρ function, in its discretized form.

After a $\tilde{\sigma}$ is picked, with the right probability, a Gaussian variable is easily generated, with time-dependent variance $a_t^2 \tilde{\sigma}^2$.

3.3 Omori law for financial aftershocks

3.3.1 Introduction

One critical observation suddenly emerges as one approaches the study of the statistical properties of the returns of the financial time series: intermittent volatility bursts occur, usually followed by periods of relatively high volatility (if compared to the average value of the previous periods) where possibly other bursts occur (see, *e.g.*, 47). These events are referred as "shocks" of the market, and are often anticipating periods of financial crisis (see, *e.g.*, 28). This stylized fact shows clearly the statistical non-independence of successive returns, in an appropriate time scale.

The analysis of shocks and rare events in complex systems is not uncommon in different fields of natural and social sciences. For example we may cite earthquakes, landslides, forest fires and scale free networks: the power-law distributions characterizing these different areas imply that rare, extreme events may occur with a non-negligible probability (see, *e.g.*, 27, and references therein).

In particular, a question may arise, concerning the behaviour of a complex system after a shock or an extreme event. Clearly this is crucial in fields, like earthquakes, where a correct evaluation of the risks related to aftershocks can be useful both from a social and a financial point of view. Omori (38), more than a century ago, described with his famous Omori-law the behaviour of the statistics of aftershocks in the non-stationary period following a big earthquake.

According to his formulation, the number of aftershocks per unit time can be described by a power law, implying also that a characteristic time-scale for the relaxation of the system does not exist.

The idea of Omori (38) was applied in finance by Lillo and Mantegna (30): the authors analyzed the period following two extreme events in the S&P500 index and found good indications of power law relaxation regimes, leading ultimately to the result that, in a given horizon after a financial shock, index returns above a large threshold obey a power-law function which is the analogue of the Omori-law in geophysics.

We want to generalize the idea of Lillo and Mantegna (30), trying to investigate the possibility of an ensemble analysis of different shocks, in order to compare the empirical results with those of the model presented in Chap. 2 and in (8, 48).

3.3.2 From the non-stationary behaviour of financial series to the Omori law for finance

The empirical investigation of the behaviour of a financial time series shows the non-stationarity of the underlying stochastic process, in particular time intervals. The volatility of an asset or of an index which is here defined as the absolute value of the asset return, is itself a stochastic process, fluctuating in time. This evidence is clear, *e.g.*, for one of the set we analyze in the following: the S&P500 index daily returns[1] from 1950 to 2010 (see Fig. 3.7). The non-stationarity is revealed in periods of increased and of persistence of high volatility after the occurrence of main shocks. In fact, if we zoom the graph in Fig. 3.7 around one of the peaks of volatility, we can see, for example as in Fig. 3.8, that even before and particularly after the large return X_1, the volatility has a mean value that largely exceeds the average value of the, let say 100, previous days and that a certain evolution in time of the distribution of the volatility is also at first sight detectable.

To be more precise and to try to define a strategy for an application of the Omori law in an interday scale, we wish to adhere as much as possible to the model of Omori (38) in the formulation of Lillo and Mantegna (30). The Omori law $n(t) \propto t^{-p}$ asserts that the number of expected aftershock earthquakes per unit time at time t after the main shock decays as a power law with an exponent p that for earthquakes ranges between 0.9 and 1.5. To avoid divergence at $t = 0$, the Omori law is often rewritten as

$$n(t) = K(t + \tau)^{-p}, \tag{3.2}$$

[1] More precisely, the difference of the log of the S&P500 index values at close time.

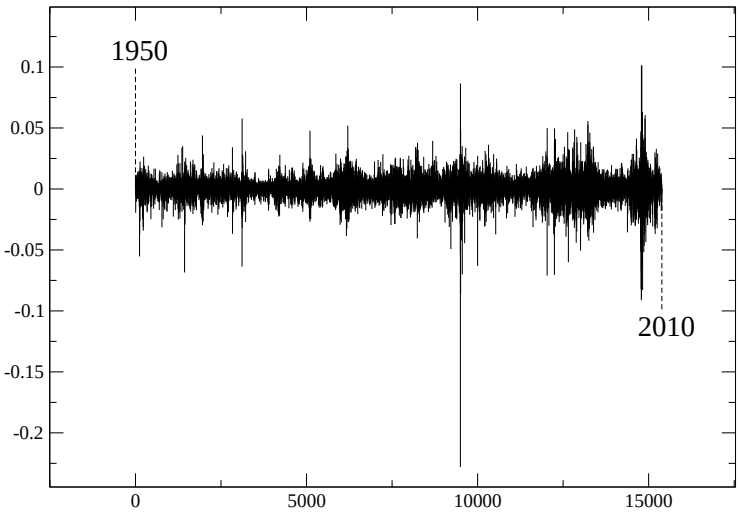


Figure 3.7: S&P500 index log-returns from 1950 to 2010 (15386 records).

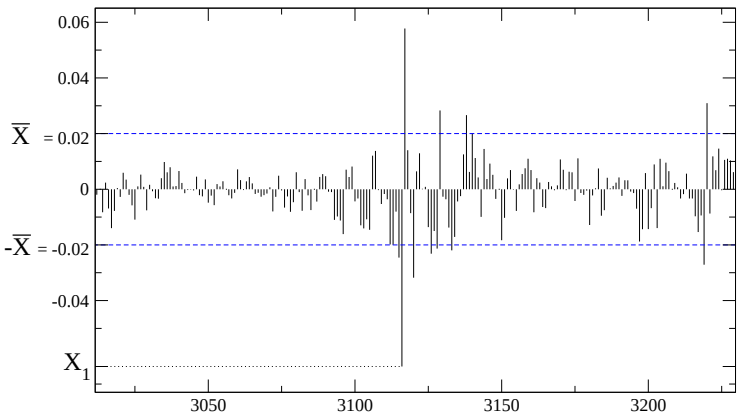


Figure 3.8: S&P500 index log-returns (zoom)

with K and τ two positive constants. For our present purposes it's convenient to write it in an integrated form: the cumulative number $N(t)$ of expected aftershocks is

$$N(t) = K[(t+\tau)^{1-p} - \tau^{1-p}]/(1-p), \tag{3.3}$$

when $p \neq 1$ and $N(t) = K\ln(t/\tau + 1)$ for $p = 1$. This formulation of the Omori law is more convenient because, integrating over time, it reduces the fluctuations due to the limited size of the sample histories. Since the validity of the law is limited to aftershock periods, when applying this law it's important to state what do we mean with "shock". In our financial context, we will refer to a "shock" as to an event (a single return x) whose absolute value exceeds a given threshold X_{Th} which is supposed to be "large" if compared to the average value of the absolute returns. A different number of shocks can be highlighted from the time series, depending on the size of the chosen threshold X_{Th}.

Another question arises as one considers the time distance between successive shocks. Following Lillo and Mantegna (30), we may assume that the time over which the Omori effect can be spotted and analyzed is in the maximum range of 50-70 days, values that come from maximizing the time period investigated whilst assuring that the relaxation is still in progress. What if another shock occurs during that period? How is the process affected? Several different choices have been made during our analysis, both including and disregarding such overlapping events. We will come later to that point: by now we just wanted to focus on the, possibly relevant, freedom of choice for the selection rules for the shocks, affecting the total number of shock events in the time series.

The goal of our study is to spot the Omori-like pattern of the aftershocks statistics by averaging over different events, which are, for the sake of simplicity, supposed to be independent and different realizations of the same stochastic process. We will find a substantial agreement with the Omori law and we will test the empirical results with the related predictions which can be derived from the model presented by Baldovin and Stella (8, 48), that will be hereafter explained.

3.3.3 Aftershocks analysis via Omori law

The datasets used in this analysis are:

(i) S&P500 index daily log-returns (close to close) from 1950 to 2010 (15386 records)

(ii) DJI index daily log-returns (close to close) from 1900 to 2010 (26588 records)

The analysis of the series starts from the definition of the value X_{Th} and $\overline{X}$, the threshold for the main shocks and for the aftershocks, respectively. Working in day time units (1 = 1 day), we scroll the time series of the returns x_t top down until we detect a return $x_{\overline{t}}$ for which $|x_{\overline{t}}| > X_{Th}$. From that point on (we will set $x_1 \equiv x_{\overline{t}}$)[1] we focus on the $t_{max} - 1$ successive returns $x_2, \dots x_{t_{max}}$[2]. We calculate, for that particular time window, the cumulative number of aftershocks (*i.e.* the cumulative number of returns exceeding the threshold $\overline{X}$ from $t = 2$ to $t = t_{max}$, as a function of t). We will denote with $N(1,t)$, for $t = 2, \dots t_{max}$, the cumulative number of detected aftershocks relative to this 1st shock, up to time t.

If $|x_k| > X_{Th}$, for some t_k in the interval $2 \leq t \leq t_{max}$, we set $x_k \equiv x_1$ and we start again selecting $t_{max} - 1$ successive returns, on which we apply the aforementioned procedure of aftershock detection, deleting the results for $N(1,t)$ calculated up to that point. In this way we assume that only the last large shock is important and that the presence of main shocks occurred just before the last one doesn't affect the aftershocks' statistics. The efficiency of this assumption has been checked.

After evaluating $N(1, t_{max})$, we start scrolling down again the series until we reach another day $\tilde{t}$ for which $|x_{\tilde{t}}| > X_{Th}$. Again we set $x_1 \equiv x_{\tilde{t}}$ for this new large shock and, with the same rules as before, we evaluate $N(2,t)$, for $t = 2, \dots t_{max}$, and so on and so forth until we reach the end of the series.

We end up with a set of N_S different shocks, for each one of which we have a detailed record of the cumulative aftershocks occurrence: $N(n,t)$ for $n = 1 \dots N_S$ and $t = 2 \dots t_{max}$. Taken one by one these N_S series of data are very noisy[3] and don't show a clear adherence with the Omori law, but, assuming that the same Omori like behaviour for all the shocks[4] holds, we can tentatively average them over the different n's. We will call $\overline{N}(t)$ this average taken over the different n's. The results are presented in Fig. 3.9 for the S&P500 dataset. The fluctuations are larger in Fig. 3.9(b) than in Fig. 3.9(a), because of the smaller number of shocks spotted for the choice of a larger threshold X_{Th}.

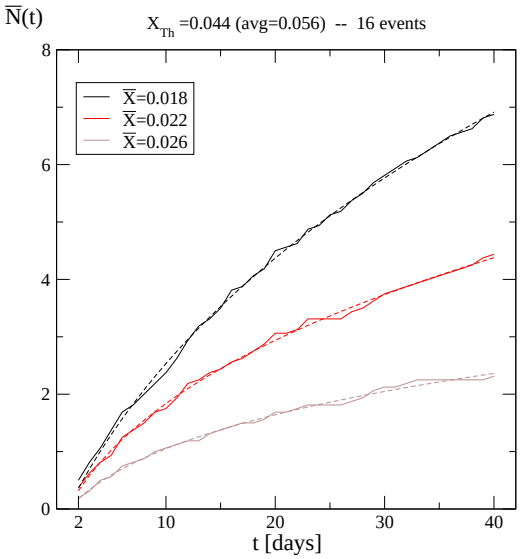
The fit with the Omori law gives the results of Tab. 3.1. However a simple debugging of the fitting process reveals that these parameters do not represent an efficient

[1] The reason for setting $t = 1$ at the shock and not $t = 0$ as prescribed by Eq. 3.3 will be clear below, in section 3.3.4. However, this is related to our choice of numbering the returns of the model series stating from 1.

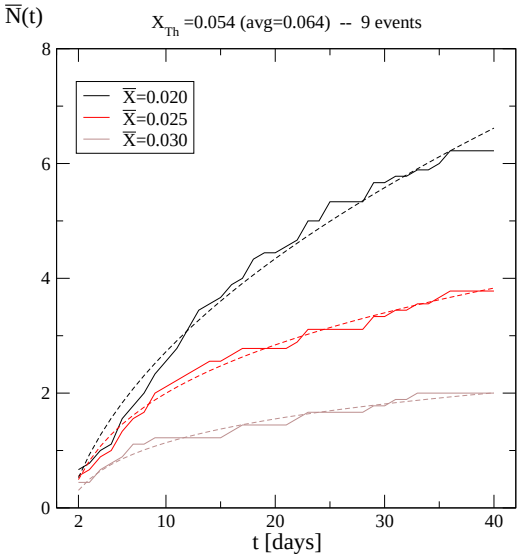
[2] t_{max} is chosen in the range of $30 - 40$ days, to maximize the number of valid shocks.

[3] After all this is not surprising, since the performed analysis highlights aftershock events, that are rare events, statistically distributed in the tails of the returns distribution.

[4] This is an arbitrary assumption, to be checked *a posteriori* after the Omori law parameters have been estimated.



(a)



(b)

Figure 3.9: Cumulative number of aftershocks till time t for the S&P500 index shocks (two different choices for X_{Th} are shown). The Omori-law fit is dashed. Notice the larger fluctuations in (b)

Table 3.1: Omori law parameters for the shock of Fig. 3.9

$\overline{X} =$	0.018	0.022	0.026
	$x_{Th} = 0.44$		
K	1.62	0.74	0.63
τ	7.73	2.99	4.00
$\boldsymbol{p}$	**0.71**	**0.68**	**0.83**
	$x_{Th} = 0.54$		
K	0.63	0.66	0.42
τ	1	1	1
$\boldsymbol{p}$	**0.51**	**0.77**	**0.86**

parametrization of the process. Moving or fixing one of the parameters gives indistinguishable dashed curves in Fig. 3.9 (but significantly different values for the other two). This suggests that a lower number of parameters may suffice to describe the process. This was particularly evident for the records for $X_{Th} = 0.054$, where we decided to set $\tau = 1$ to find reasonable values for K and p.

The choice of the thresholds is crucial, as can be argued comparing the graphs in Fig. 3.10. If X_{Th} is too small, as in Fig. 3.10(a), the growth of $\overline{N}(t)$ is almost linear, indicating that we're not in an Omori regime (the shocks are not sufficiently large). On the other side, if X_{Th} is too large, as in Fig. 3.10(c), the number of shocks is low and the aftershocks' statistics is very poor.

In any case it's remarkable the verification of an Omori-like behaviour, obtained by this averaging process over different shock events. This is another stylized fact which should be predicted by a good model for financial returns. That's what we check in the next section.

3.3.4 The model predictions for the Omori regimes

We want now to show that the model presented in Chap. 2 predicts the Omori-like features highlighted in the previous section, and we check to what extent the Omori-like behaviour of financial aftershocks can be used to calibrate the model itself.

One of the key aspect of the model is the capability of writing explicit expressions for the joint probability of multiple successive returns. Working again in a one-day time scale, indicating with x_i the daily *close to close* return of the i-th day, we can write down

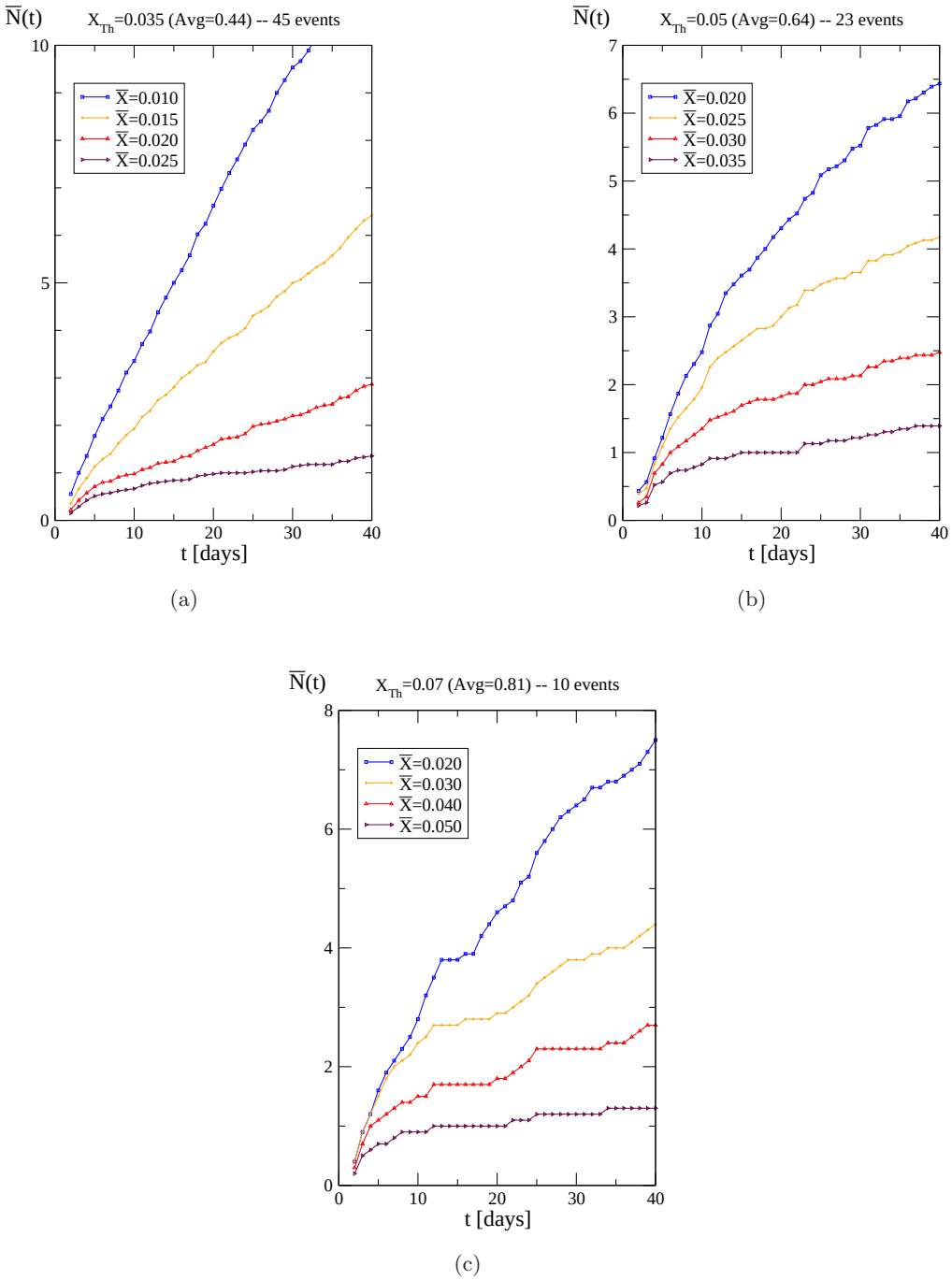


Figure 3.10: Cumulative number of aftershocks till time t for the DJI index shocks (three different choices for X_{Th} are shown).

in the following form the expression for the joint probability density $p(x_1, x_2, \ldots, x_t)$ of getting the returns $x_1, x_2, \ldots, x_t$ at day $1, 2, \ldots, t$, respectively:

$$p(x_1, x_2, \ldots, x_t) = \int_0^\infty d\sigma \rho(\sigma) \prod_{i=1}^{t} \frac{\exp\left(-\frac{x_i^2}{2\sigma^2 a_i^2}\right)}{\sqrt{2\pi\sigma^2 a_i^2}}, \tag{3.4}$$

where

$$a_i \equiv \left[i^{2D} - (i-1)^{2D}\right]^{1/2}, \quad i \in \mathbb{N}_1 \tag{3.5}$$

and where $\rho(\sigma)$ is a normalized function, obtained by fitting the scaling function $g(x)$ of the single returns (see Chap. 2 for the details).

The model can be applied as follows: the basic formulation of the intraday model is that the average stationarity detected across the long time series is caused by the, more or less randomly occurring, resets of the a_i parameters. Here we assume that the occurrence of an extreme event, like the shock, may reset the process, leading to the setting $a_i = a_1$ for the day of the shock. From that day on, unless other shocks occur, the a_i's evolve daily according to Eq. 3.5, at least in the range of interest of the Omori-like decay. This assumption is quite strong and, ultimately, not strictly verifiable. However, we take it as an *ansatz*, and in a following section we will discuss the problem from an alternative perspective, leading to similar results.

To investigate the probability $P(|x_t| > \overline{X} \mid \overline{x}_1)$ of getting in the day t a return $|x_t| > \overline{X}$, given a shock at day 1 ($|\overline{x}_1| > X_{Th}$), we can make use of the definition of conditioned probability density

$$p(x_2, \ldots, x_t | x_1) = \frac{p(x_1, x_2, \ldots, x_t)}{p(x_1)}. \tag{3.6}$$

We can explicitly write that probability by integrating Eq. 3.6 over the variables $x_1, x_2, \ldots, x_t$, with the conditions:

(i) $x_1 = \overline{x}_1$,

(ii) $|x_2|, \ldots |x_{t-1}| < X_{Th} \simeq \infty$,

(iii) $|x_t| > \overline{X}$.

In (i) we condition the probability to the size of the shock at $t = 1$ while in (iii) we restrict the range of the return at time t to values above the aftershock's threshold. With the approximation (ii), for simplicity, we just neglect the conditional constraint from the $t-2$ returns $x_2, \ldots, x_{t-1}$[1]. However, since the threshold X_{Th} is chosen to

[1]The inclusion of the more correct condition $|x_2|, \ldots |x_{t-1}| < X_{Th}$ may be tested, too. Further developments may include the verification of the influence of this choice, also in combination with the different possible methods for the selection of the shocks.

be large, compared to the square root of the second moment of the distribution of the returns, the assumption (ii) is guaranteed in practice. The result is straightforward:

$$P(|x_t| > \overline{X} \mid \overline{x}_1) = \frac{2\int_0^\infty d\sigma\rho(\sigma)\frac{\exp\left(-\frac{\overline{x}_1^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}}\int_{\overline{X}}^\infty \frac{\exp\left(-\frac{x_t^2}{2\sigma^2 a_t^2}\right)}{\sqrt{2\pi\sigma^2 a_t^2}}dx_t}{\int_0^\infty d\sigma\rho(\sigma)\frac{\exp\left(-\frac{\overline{x}_1^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}}}, \tag{3.7}$$

where, as can be seen from the factor 2, the parity properties of $p(x_1, x_2, \ldots, x_t)$ have been taken into account.

To compare the results of the previous section with the theoretical ones, we also write the average total number of expected aftershocks from day 2 to day t in the following way:

$$N(t) = \frac{1}{K}\sum_{k=1}^{K}\sum_{i=2}^{t} P(|x_{i_k}| > \overline{X} \mid \overline{x}_{1_k}). \tag{3.8}$$

where K is the total number of shocks in the series, and $\overline{x}_{1_1}, \overline{x}_{1_2} \ldots \overline{x}_{1_K}$ the amplitudes of the shocks themselves.

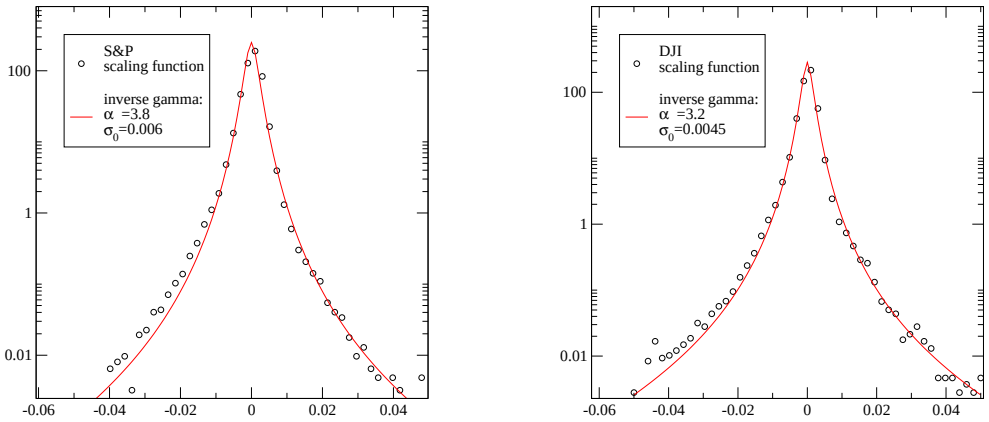
Given the shape of $\rho(\sigma)$, the model has only one parameter, the scaling exponent D, which turns out to be crucial. To see it, we first define the explicit shape of the $\rho(\sigma)$, by taking an *inverse-gamma* density:

$$\rho(\sigma) = \frac{\sigma_0^\alpha}{\Gamma(\alpha)\sigma^{1+\alpha}}\exp\left(-\frac{\sigma_0}{\sigma}\right), \tag{3.9}$$

where $\sigma_0 > 0$ is the scale factor, and $\alpha > 0$ is the form factor. This function is a convenient and common choice in finance when one wants to fit the realized volatilities of an asset (see, *e.g.*, 17). The scaling collapse is done according to the time-inhomogeneous scaling formula Eq. 2.37 with $t = 1$ and $\tau = 1, 2, \ldots, 20$. The fit of the factors has been obtained via χ^2 minimization and the results are presented, both for S&P500 and DJI, in Fig. 3.11.

The results of the fitting show that this model cannot grasp the asymmetry of the scaling function, for which an extended development of the theory is in progress (see, *e.g.*, the results anticipated in Sec. 2.4).

Now the key parameter is the scaling exponent D, which has been estimated during the scaling collapse procedure, by optimizing the collapse itself. The theoretical behaviour of the aftershocks' statistics (calculated via Eq. 3.7 and Eq. 3.8) is deeply affected by the value of D. The results are plotted in Fig. 3.12 and Fig. 3.13.



(a) (b)

Figure 3.11: Scaling functions for the S&P500 (a) and DJI (b) datasets. The best fit with an inverse-gamma ρ function is also shown. The collapse has been realized with $D = 0.25$ for (a) and $D = 0.2$ for (b).

3.3.5 Further detailed analysis

Remarkably the model is able to reproduce the general feature of the Omori-like aftershocks' statistics, but a detailed check of Fig. 3.12 and of Fig. 3.13 shows that the values of D appear somewhat larger than those derived via scaling analysis (Fig. 3.11). However, given the variability over different choices of $\overline{X}$ (compare, *e.g.* Fig. 3.12(a) and Fig. 3.12(e) or Fig. 3.12(b) and Fig. 3.12(f)), we can accept the result of $D \approx 0.25 \div 0.30$ for the S&P500 dataset and $D \approx 0.22 \div 0.27$ for the DJI dataset.

We must remember, in any case, that the empirical results also have a statistical error, coming from the averaging process over the different shocks. This error is quite large, as can be seen for example in Fig. 3.14.

In addition to this, the model is really sensible to the fitted parameters of the scaling function, too. In fact, as can be seen by parallel comparison of the Fig. 3.15(a) and Fig. 3.15(b), even if, with different tuned values of σ_0 and α, a substantial agreement on the shape of the scaling function can be still detected in Fig. 3.15(a), the results for the aftershock statistics depicted *e.g.* in Fig. 3.15(b) are significantly different, eventually leading, in our scheme, to a different estimation on the value of D.

This result is not surprising indeed, and takes us back to the problem of the poor statistics of our datasets: a good fitting of the scaling function, especially on the tails of the distribution, is not possible, because of the limited size of the datasets themselves. Even in the largest of the two, the DJI index, the number of data is just 26588, a

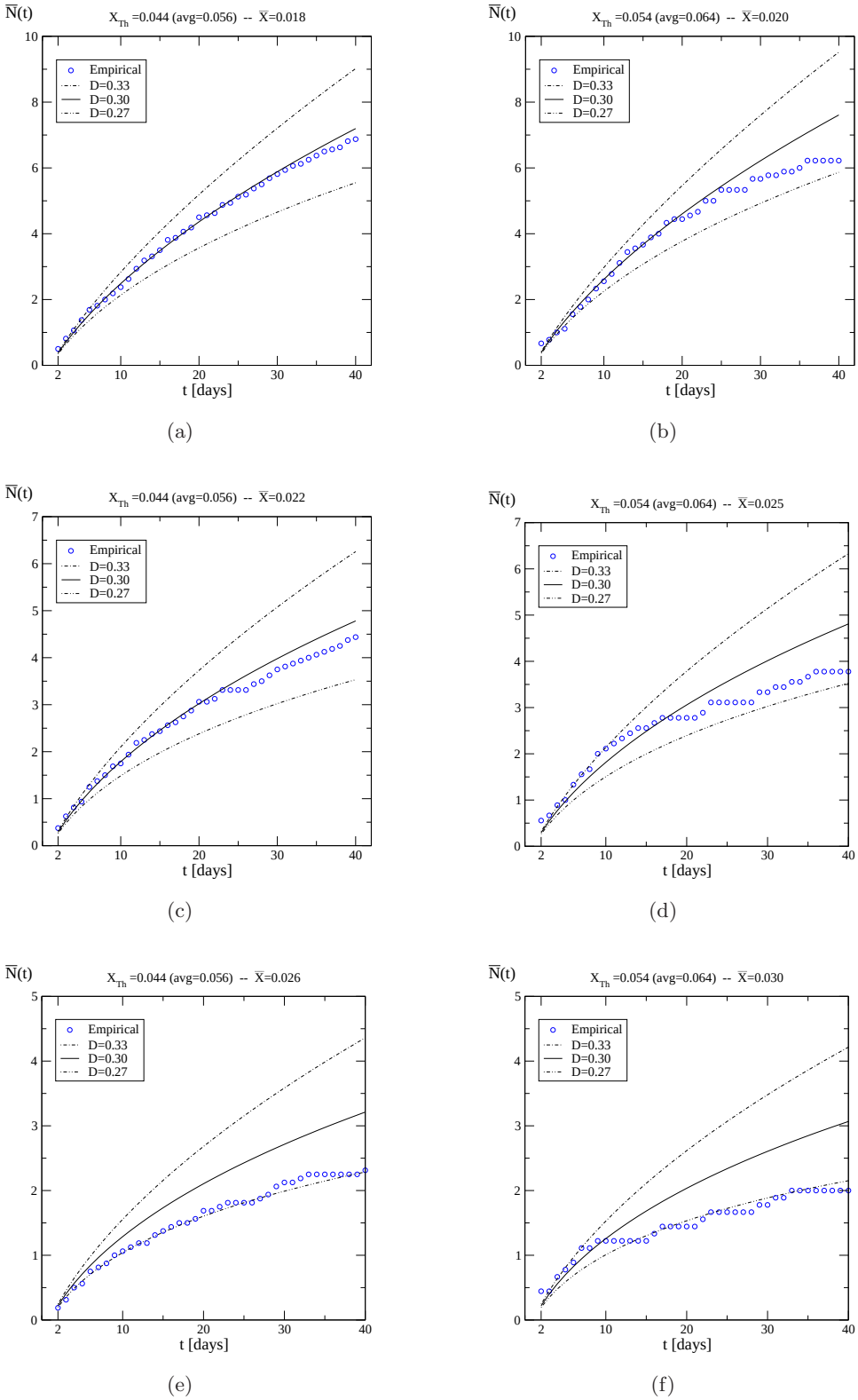


Figure 3.12: Empirical vs. theoretical results for the S&P500 dataset, for two different values of X_{Th} (left and right columns) and three different values of $\overline{X}$ (rows). These values of the thresholds are the same as those of Fig. 3.9. The best value for the scaling exponent is around $D = 0.3$.

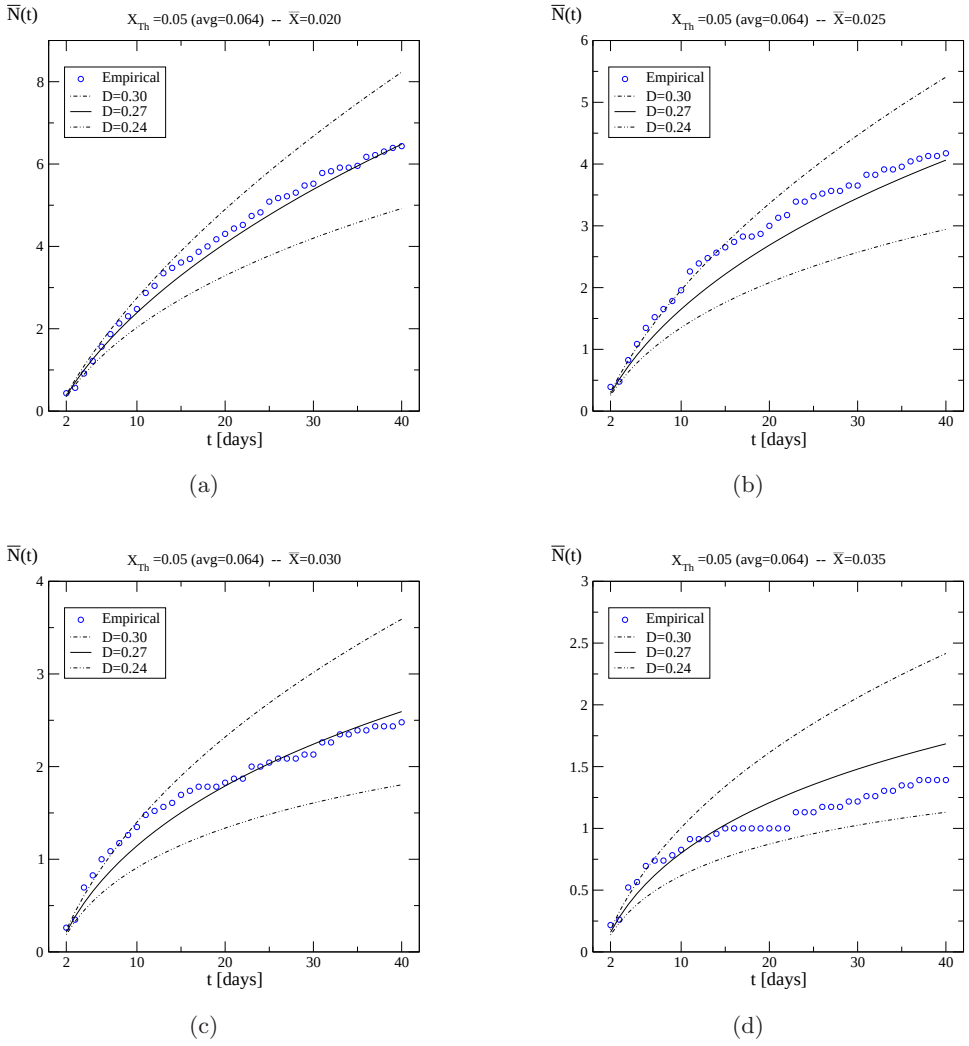


Figure 3.13: Empirical vs. theoretical results for the DJI dataset, for the choice $X_{Th} = 0.050$ (see Fig. 3.10(b)). The best value for the scaling exponent is around $D = 0.27$.

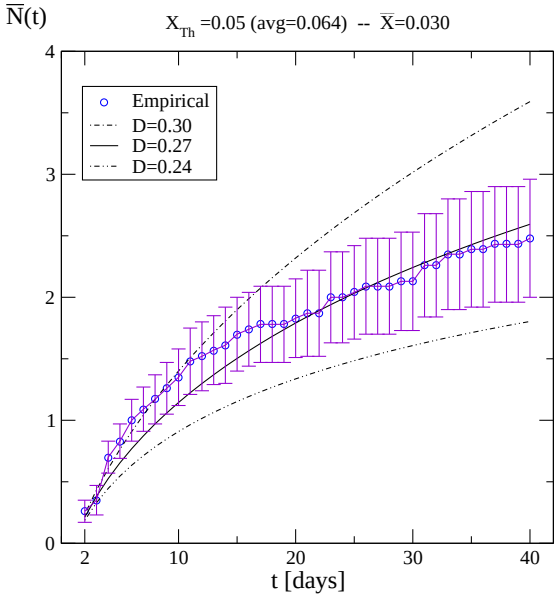
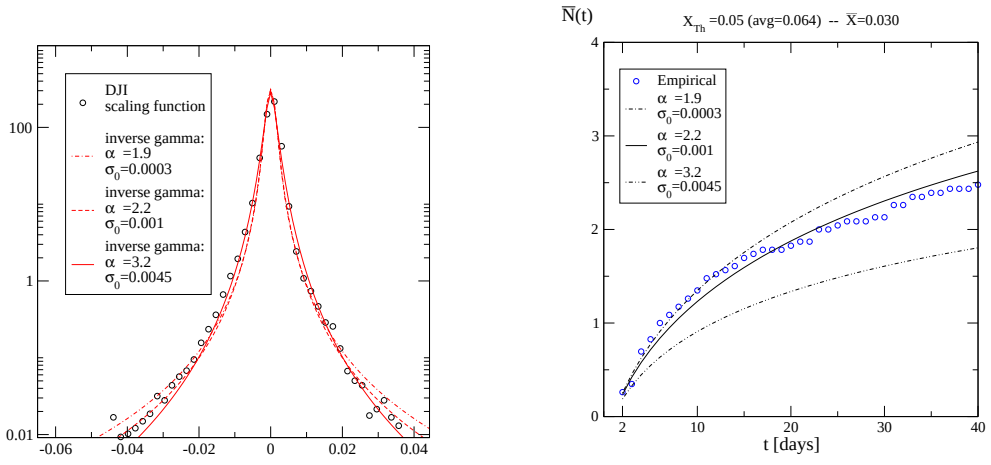


Figure 3.14: Error bars for the empirical averages over different shocks in one of the studied cases (DJI dataset, $X_{Th} = 0.050$, $\overline{X} = 0.030$, 23 shocks). Notice that this graph is that for which we detected the Omori effect with the largest number of aftershocks; in all the other cases, *e.g.* in the S&P500 dataset, the number of events was smaller, ultimately leading to even larger error bars.



(a) (b)

Figure 3.15: DJI dataset: a slight difference in the fits of the scaling function $g(x)$ (a) leads to significantly different theoretical predictions for the aftershock statistics (b). In (a) and (b), to highlight the influence of the parameters σ_0 and α, the same value $D = 0.24$ has been used.

number far not sufficient for good statistical stability, especially considering the fat-tailed behaviour of the scaling function.

3.3.5.1 Beyond Omori law: an ensemble description of the shocks

Given our *ansatz* of setting $a_i = a_1$ for the day of the shock, we can see that, at least formally, all the shocks and the subsequent 39-*days long* periods are treated by the model in the same way, as they were realizations of the same stochastic process (with the only difference that the conditioning size of the shock varies).

In this sense, we can assume that the collection of 40-*returns long* periods that is extracted for every choice of X_{Th} constitutes a small ensemble of histories obeying the same statistics. Tentatively, on this dataset of "shock periods" we can try to investigate the model predictions, in comparison to the empirical evidences coming from averaging procedures.

We will focus on the DJI index dataset, with $X_{Th} = 0.05$. This choice gives 23 shock events and consequently an *ensemble* of $L = 23$ periods, each with 40 returns $x_{t,l},\ t = 1, 2, \dots 40,\ l = 1, \dots L$. Even within the choice of this largest dataset, the expected agreement will be surely affected by large statistical errors.

We want to check, in particular, the behaviour of the empirical moments of the distribution of the returns, after the shock, as a function of time:

$$\langle |x_t|^\alpha \rangle_e = \frac{1}{L} \sum_{l=1}^{L} |x_{t,l}|^\alpha \quad t = 1, 2 \dots 40. \tag{3.10}$$

and compare it with the theoretical counterpart:

$$\langle |x_t|^\alpha \rangle_p = \int_0^\infty d\sigma \tilde{\rho}(\sigma) \int_{-X_{Th}}^{X_{Th}} |x|^\alpha \frac{\exp\left(-\frac{x^2}{2\sigma^2 a_t^2}\right)}{\sqrt{2\pi\sigma^2 a_t^2}} dx, \quad t = 1, 2, \dots 40. \tag{3.11}$$

where $\tilde{\rho}(\sigma)$ is the scaling function obtained by scaling collapse of the considered returns only (see the aforementioned procedure). It's no surprise that, having only $23 \times 40 = 920$ returns, the scaling collapse will have a poor statistical stability, especially on the tails. Anyway, a scaling function can be found, via χ^2 minimization. This gives, for the series in object, the following results: $D = 0.25$, $\alpha = 3.68$, $\sigma_0 = 0.012$. The scaling function and the comparison between the empirical and the theoretical results for the moments with $\alpha = 1,\ 0.5,\ 0.25$ are given in Fig. 3.16.

Despite the statistical uncertainties and the relevant fluctuations emerging from the analysis, the result proves that the aggregation of the different shocks can be useful to validate the model forecasts, especially for what concerns the time-inhomogeneity

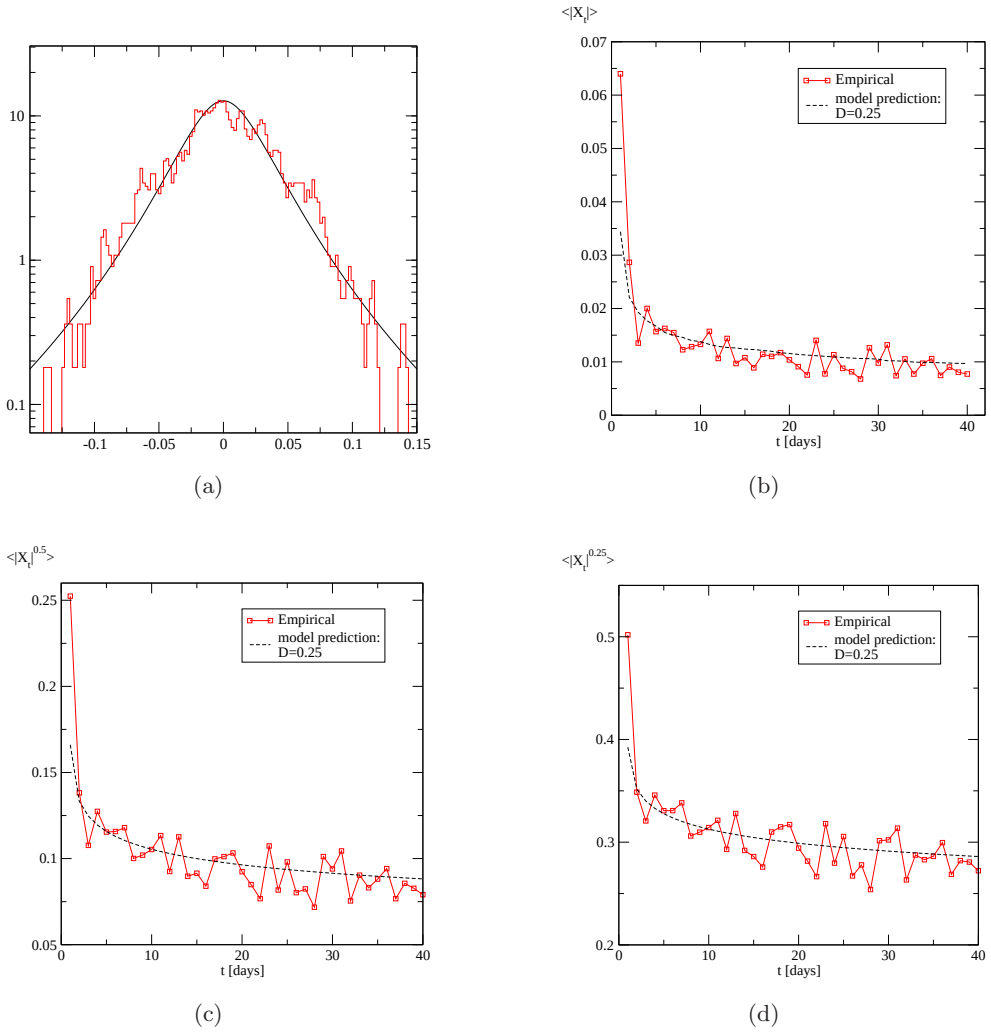


Figure 3.16: DJI dataset: (a) The scaling collapse and the scaling function of the returns of the 23 detected shock periods; (b)(c)(d) Comparison between the empirical absolute moments and the model prediction, as a function of time. For the theoretical prediction, the optimized scaling function calculated in (a) has been used. The agreement is remarkable, despite the statistical uncertainty coming from the limited size of the dataset.

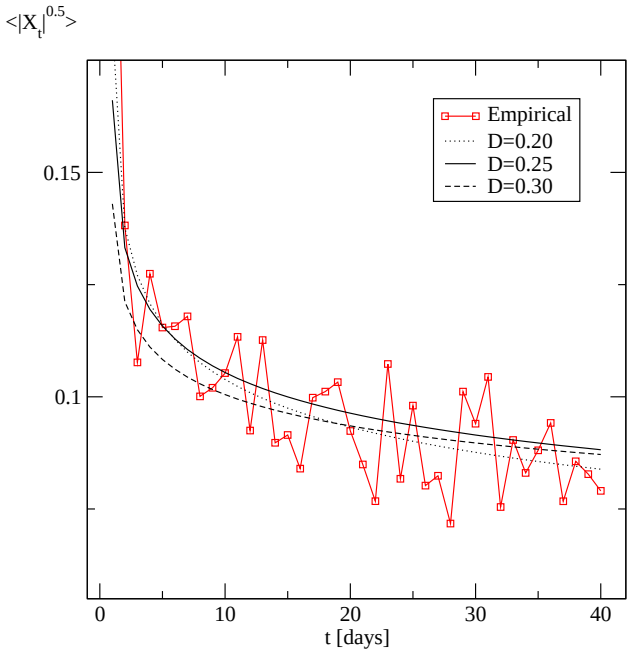


Figure 3.17: DJI dataset. Comparison between the model predictions of the behaviour of the 0.5 moment, over different scaling parameters (D). The fluctuation of the empirical data forbids any preferable selection of D.

properties of the returns. The agreement between the empirical and the theoretical values is verified upon different choices of α, and it's remarkable that the result is obtained just via χ^2 minimization of the scaling collapse.

Nevertheless, the strong uncertainties deriving from the limited size of the dataset forbids us a deeper, more interesting, step. The described analysis, in fact, does not seem to represent a good method for calibrating the model itself, or to define the limits of the scaling parameter D. Every choice of D, in fact, gives a slightly different scaling function (*i.e.* different α and σ_0). The combination of the three new parameters produces not so different theoretical predictions (Fig. 3.17).

At this level of analysis, however, a joint check of the behaviour of the moments and of the scaling should be done, in order to find the optimizing parameter for both indicators. Ultimately this check produces, for the DJI dataset, an optimized value of D around 0.25 (the details are not given).

3.3.6 Stability of the analysis vs. stability of the model

After checking the consistency of the aforementioned results, one of the questions that may naturally arise regards the stability of the system. In a nutshell, given the wide

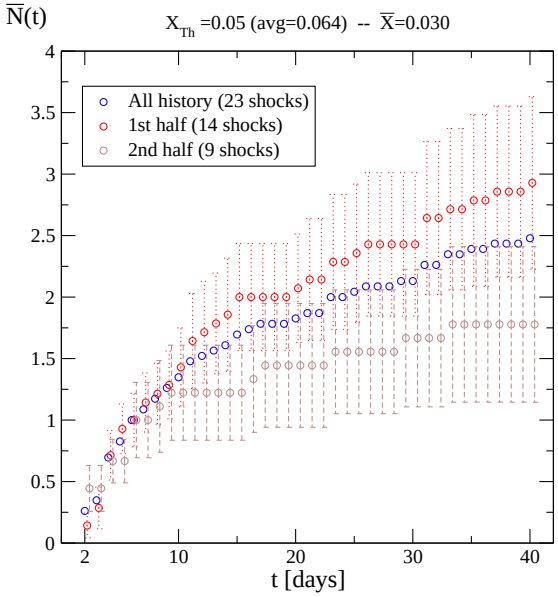


Figure 3.18: Results of the cumulative aftershock analysis for the whole DJI history and for tho subsets (the 1st and the 2nd half of the history, respectively). The error bars for these two last sets are given. [For ease of reading, the sets have been slightly horizontally shifted]

error bars coming from the empirical analysis (see, *e.g.*, Fig. 3.14) and given the large fluctuations of the checked indicators, how does the model deal with (and describe) such a variability? In addition, in Eq. 3.8 and Eq. 3.7 the shock sizes enter in the calculation and might have an important effect.

We want to check, in the present section, if the variability of the sizes of the shocks, their order, and, more generically, the realized history of returns affect the results of the analysis. For example what happens if we split our dataset and we check the Omori-law behavoiur in the two subsets, separately? Do we obtain analogous results?

Again, we focus on the largest of the two datasets at our disposal, the DJI index. We split the history into two parts and we perform the same empirical analysis as in section 3.3.3; we set $X_{Th} = 0.05$ and $\overline{X} = 0.03$ and we present the results in Fig. 3.18, together with the result coming from the whole DJI history.

Some remarks are necessary, here. First of all, we see that the empirical averages significantly differ from each other, though showing the same Omori-like pattern. Secondly we emphasize the relevance of the error bars, coming from the averaging process over the shocks: the regions of uncertainty overlap for all the three sets. Finally, we highlight that the curves, coming from different sets (roughly the entire, the first and the second part of the last century of data), may not be directly comparable, because

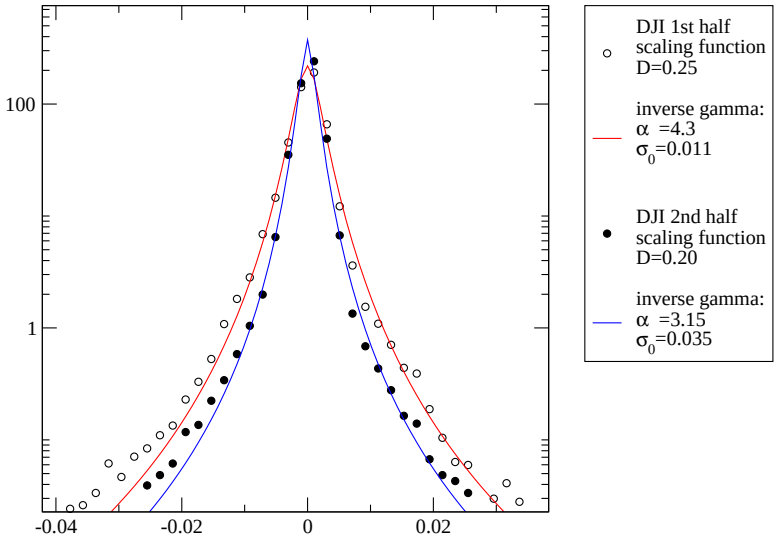


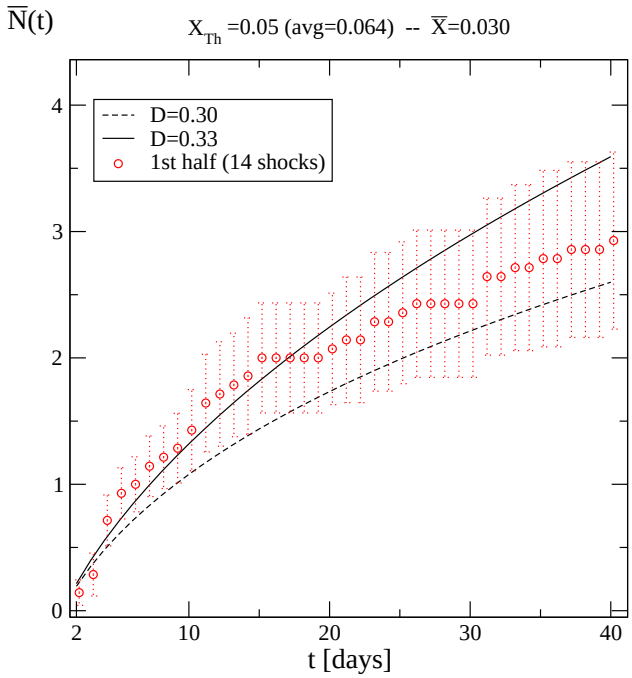
Figure 3.19: Scaling functions for the 1st and the 2nd half of the DJI history, respectively. In the legend, the values of the evaluated fitting parameters are shown.

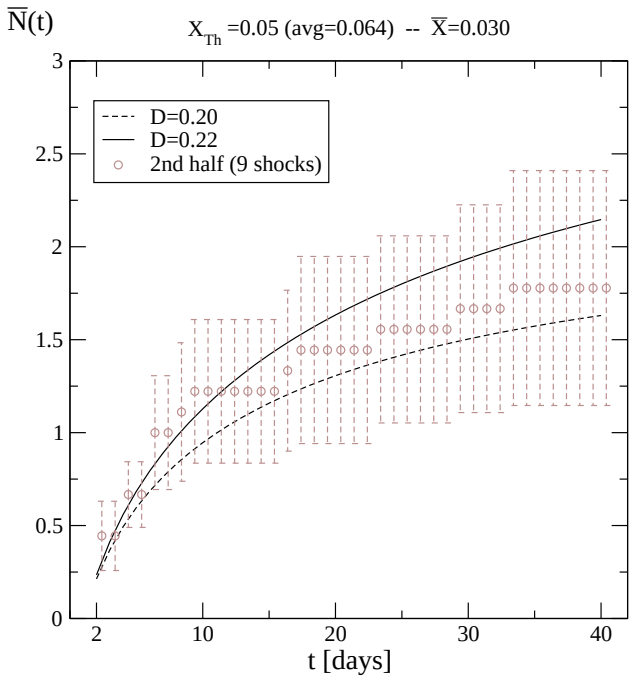
the statistics of the underlying process (or, in our framework, the scaling function) may be different.

To check this idea in detail, we build the scaling function for both sets and we try to evaluate the model prediction on the basis of the two different scaling functions, separately. The resulting scaling functions have been plotted together in Fig. 3.19. The fitting of the scaling collapse gives a different scaling exponent for the tho halves of the dataset. In particular, there seems to be a scaling exponent $D = 0.25$ for the first half of the last century and $D = 0.20$ for the second half.

The comparison between the empirical and the theoretical curves of the Omori aftershocks' statistics is given in Fig. 3.20. The result is good for Fig. 3.20(b), with the value of the scaling parameter D well reproducing the statistics of the expected aftershocks. On the other hand, the result of Fig. 3.20(a) highlights an overestimation of D.

The result and the poor agreement must be read in comparison with all the results presented up to now. We have found altogether the existence of an Omori-like behaviour over different indexes and for different levels of the threshold X_{Th}, but with a quite large statistical error affecting our estimations of the scaling exponent D. Moreover, having one only history for each index, and having such large error bars, we cannot decide upon the goodness of our estimate. We should do statistics over different histories, too. In order to proceed further in this direction, numerical simulation is needed.





(b)

Figure 3.20: Comparison between model prediction and empirical data for the two distinct subsets of the DJI history. The value of D is larger in the first subset than in the second, as expected from the results on the scaling collapse (see Fig. 3.19). In (b) the estimated value of D agrees with the scaling analysis: $0.20 - 0.22$ vs. 0.20. On the other hand, in (a) the estimated value is definitely larger than expected: $0.30 - 0.33$ vs. 0.25.

3.3.7 Omori law verification in computer generated histories

The above quoted analyses and results have shown a general qualitative agreement between the Omori-like features of the aftershock statistics and the model predictions following from the Baldovin and Stella model assumptions (8). On the other hand, a large statistical uncertainty always accompanies these results, sometimes being so large to make the model predictions very poor: certainly the Omori analysis seems unfit for the calibration of the model itself.

One of the problems affecting this study is the presence of one single (and not so long) history, for both indexes. One should in fact be aware of the chance of getting scattered results, far from the expected averages, especially when dealing with statistically non significant events. This is in fact the case here, as can be checked by looking at the error bars in Fig. 3.14: dealing with a low number of shocks ultimately leads to a large standard deviation for the expected number of aftershocks, increasing as a function of t.

There is a double source of uncertainty: from one side, dealing with a low number of shocks ultimately leads to a large, increasing as a function t, standard deviation for the expected number of aftershocks. From another side, being the historical series not so long, we cannot be sure that the selected samples (the spotted aftershock regimes) are good statistical representatives for the process.

First of all, in this section, we want to make this last remark more quantitative. To do so, we generate many histories taking advantage of a calibration system for the model that is actually in course of development (10). Using the calibrated parameters coming from the analysis of the given S&P500 historical series we generate histories of different length and we compare the results of the aftershock statistics analysis as a function of the length of these series. Supposing that the model is ergodic, and the analyses performed until now seem to guarantee it, the results should be more and more stable as the length of the histories increases, ultimately leading to an asymptotic result that can be considered, in the ergodic hypothesis, as the model prediction for the process. This asymptotic convergence is quite easily verified taking very long series.

In Fig. 3.21 the results coming from eight simulated histories of length equal to that of the available real S&P500 history (15386 returns in Fig. 3.21(a)) and eight histories ten times longer (153860 returns in Fig. 3.21(b)) are depicted. The graphs show that we're still too far from the ergodic convergence, especially in the case of Fig. 3.21(a).

The error bars should have been depicted, too, but they have been omitted for sake of clarity. Anyway, these error bars, coming from the averaging process over the different shocks are approximately of the same size of the spread between the different

simulations, both for Fig. 3.21(a) and Fig. 3.21(b). As expected, a longer history has also a larger number of shocks, and so the standard deviation for $\overline{N}(t)$ is lower, too.

The model prediction is achieved taking very long series (more than 6 million returns): the error bars, in this case, are below the size of the circular markers in the graphs.

It should be now clear that with a single history of small length, as that of our datasets, the Omori-like analysis cannot precisely identify the value of the scaling exponent. Due to the fact that our histories are far from the lenght needed to satisfy ergodicity, the stability of the result is not guaranteed, and the value of D which best fits the curves may significantly differ from the correct one.

The agreement between the simulations and the real history is confirmed, in a statistical sense, for different choices of the threshold (Fig. 3.22). The behaviour of the historical S&P500 series lays in the range of those of the simulated ones, for different choices of the thresholds.

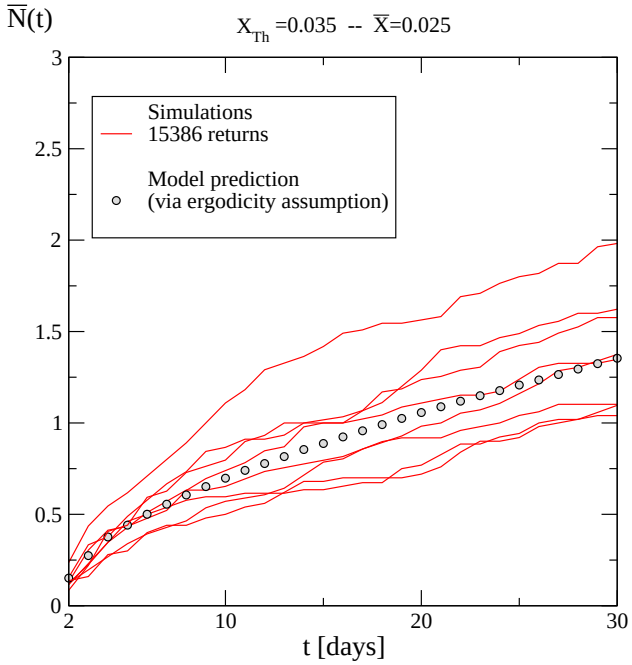
Thus the particular realized history can be considered as one peculiar, non representative, sample in an ensemble of short histories. Even if it has the statistical properties of the ensemble (and actually it generates the ensemble itself), the series cannot be useful for the detection of some parameters, such as the scaling exponent D, just by mere application of an Omori-like analysis.

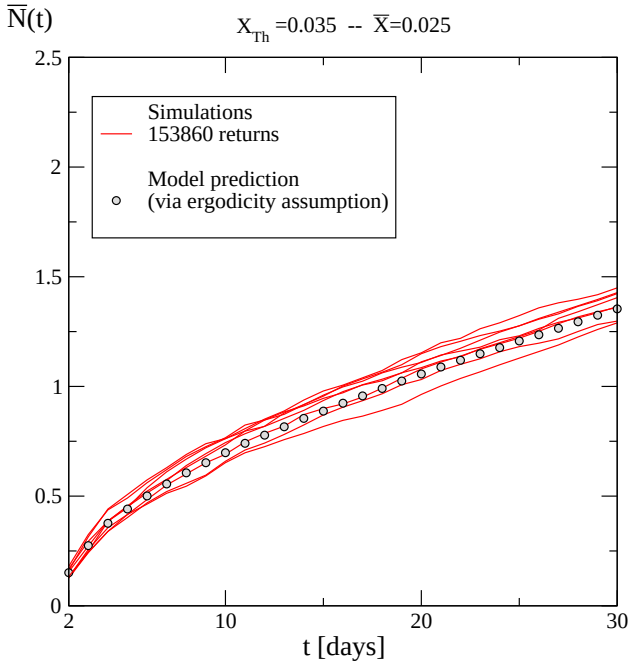
On the other hand, the analysis provides a further test of the capabilities of the model and defines a new indicator for checking the calibration of the model: even if it cannot help as a stand-alone calibrating machine, the Omori analysis can be used in combination with the estimation of other indicators for the determination of the parameters of the model.

3.3.8 Final remarks

In this chapter the Omori law for earthquakes initially presented by Omori (38), and verified in a financial context on the basis of high frequency data by Lillo and Mantegna (30), has been tested in the day to day return series of the DJI and S&P500 indexes. The validation of the law in this field suggested to test the capability of the model of Baldovin and Stella (8) with respect to the prediction of the aftershocks' statistics.

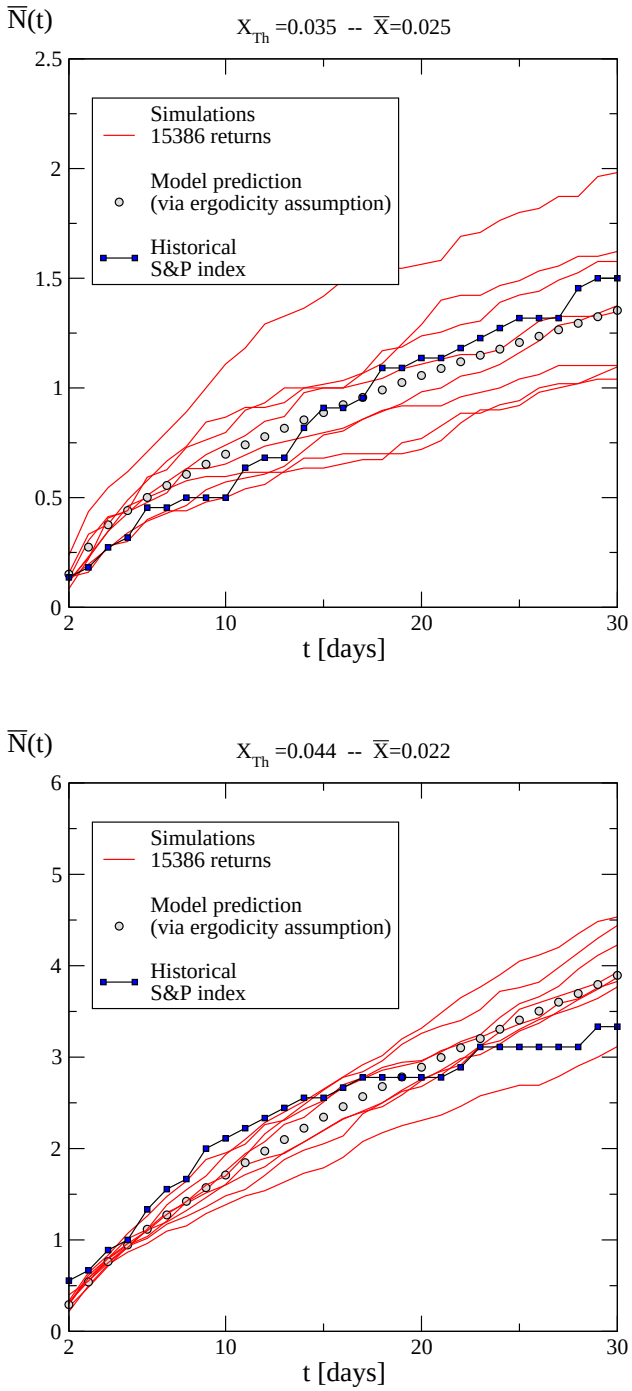
A key role is played by the scaling exponent D, which sensibly affects the predictions. However, this role cannot be exploited to try to determine its correct value, because, due to the poor statistics of the shock events and to the short length of the aftershocks' histories, the results of the prediction are not stable, as verified by splitting in two the series or by running computer simulations.





(b)

Figure 3.21: Each graph presents the Omori-like behaviour detected analyzing eight different simulations of the S&P500 index. In (a) the length of the time-series is equal to that of the historical studied series. In (b) the simulation is ten as longer. The theoretical result coming from the ergodic assumption is also shown.



(b)

Figure 3.22: For two different choices of the thresholds, the comparison between the simulation results, the ergodic theoretical model prediction and the result from the available S&P500 are given. The choice for the scaling parameter is $D = 0.25$. The error bars for $\overline{N}(t)$ in the empirical and the simulation data are not shown for clarity, but they are approximately the same size as the spread between the different simulations.

Eventually we can say that the model suggests that the detection of the scaling parameter D cannot rely on the Omori analysis alone. Even if a value for D can be derived from the aftershocks' analysis, this value is influenced and weakened in its stability by the low numbers of aftershocks that can be isolated in so short series.

However it is remarkable that the Omori-like features can be reproduced by the model and that, if the statistical uncertainties are taken into account, this analysis may represent a further test for the calibration of the model.

4

The intraday model

Winning is optional, trying your best is not.
J. Parshall and A. Tully

4.1 The intraday analysis of returns series

In this chapter we take advantage of the scheme developed in Section 2.3 to reproduce the financial data in an intraday context. The aim of this work is to calibrate the model to reproduce the features of the available histories of high frequency data as well as possible and to try to develop an intraday strategy that can possibly employ the predictive power of the model to generate useful and profitable indications of buy and sell operations. Since the scheme applies to high frequency data (with a time interval between the returns that is typically as low as 5 or 10 minutes), the possible available datasets are only those of the last 20-30 years, back to the time when the first high frequency registrations were made in the markets.

We will concentrate our analysis mainly on these two datasets:

- EUR/$ exchange rate from March 1st 2000 to March 1st 2005,
- S&P500 index from September 30th 1985 to October 20th 2010;

but some compendiary analyses will be performed on the following two assets, too:

- IBM stock price from January 2nd 2003 to March 30th 2010,
- CAT stock price from January 2nd 2003 to December 12th 2008.

The idea we will follow is that of applying the model to the different datasets, to derive the scaling exponent D and the scaling function $g(x)$. We will then compare the empirical results of some particular correlators with the theoretical predictions: we will

verify the agreement between them. Finally, and specifically for the S&P500 dataset only, we will develop a buy/sell strategy to take profit of the predictive power of the intraday scheme. The resulting strategy will give a total gain/loss amount that can be easily compared to analogous results derived from other standard strategies (GARCH models). In this way the accuracy of the model will be verified. And its usefulness, too.

4.1.1 Testing the intraday model: ensemble of histories from EUR/$ exchange rate

In order to detect the possible presence of nonstationarity at an intraday time-scale for the distribution of the increments, one would need to have access to many realizations of the same process, repeated under similar conditions. Quite remarkably, high-frequency financial time series offer an opportunity of this kind, in which it is possible to directly sample an ensemble of histories. The themes exposed in this section follow the scheme of (7).

In (11) it has been proposed that when considering high-frequency EUR/$ exchange rate data as recorded during the first three hours of the New York market activity, different process realizations can tentatively be identified in the daily repetitions of the trading. This idea is motivated by the detection of a daily pattern in the evolution of the intraday volatility, which is particularly evident in specific time windows, that are associated with the opening and closure time of the main stock exchange markets.

This gives the interesting possibility of estimating quantities related to ensemble-, rather than time-averages. In this section we profit of this opportunity by showing that a proper analysis of the statistical properties of this ensemble of histories naturally leads to the identification and validation of the stochastic model of market evolution presented in Chap. 2. As therein already described, the main idea at the basis of this model is that the scaling properties of the return distribution are sufficient to fully characterize the process in the time range within which they hold.

The main feature of this time-inhomogeneous model is the reconstruction of the joint PDF of the returns, $p^{(t)}(x_1, x_2, \ldots, x_t)$, on the basis of the scaling symmetry revealed by the PDF of the total return[1] $x_t(t) \equiv \ln S(t) - \ln S(0) = \sum_{i=1}^{t} x_1(i) \equiv \sum_{i=1}^{t} x_i$. The PDF of the variable $X_t(t)$, $p_{X_t(t)}(x)$, satisfies a scaling symmetry with scaling exponent D and scaling function g, if, for every t,

$$t^D p_{X_t(t)}(t^D x) = g(x). \tag{4.1}$$

[1] In this chapter, we switch to the more intuitive notation for the returns: $x_\tau(t) \equiv \ln S(t) - \ln S(t-\tau)$. The important formulas Eq. 4.1 and Eq. 2.37 are identical as one sets in Eq. 2.37 $t = 0$ and $\tau = t$.

Under the premises of this scaling symmetry, $p^{(t)}(x_1, x_2, \ldots, x_t)$ is then reconstructed as

$$p^{(t)}(x_1, x_2, \ldots, x_t) = \int_0^\infty d\sigma \rho(\sigma) \prod_{i=1}^{t} \frac{\exp\left(-\frac{x_i^2}{2\sigma^2 a_i^2}\right)}{\sqrt{2\pi\sigma^2 a_i^2}}, \tag{4.2}$$

where we used the notation

$$a_i \equiv \left[i^{2D} - (i-1)^{2D}\right]^{1/2}, \tag{4.3}$$

and $\rho(\sigma) \geq 0$ is a PDF for a mixture of Gaussian distributions with different widths σ. See Chap. 2 and (8) for additional details.

We are working under the hypothesis of Section 2.3. In particular, we will focus on the historical returns of the EUR/$ exchange rate, on a 10 minutes scale. The period under analysis is from March 1st 2000 to March 1st 2005, with t ranging in almost three hours after 10.00 a.m., NY time.

Following the ideas of (11), we also assume that every single day may be considered as a repetition of a random stochastic process and that the opening time of the NYSE, a particular moment of the day where a lot of information is delivered into the market, represents a *restart* for the process itself. We therefore set $a_i = a_1 = 1$ for the first return after 10.00 a.m.

Working in time units of tens of minutes, we thus obtain an ensemble of $M = 1,282$ realizations $\left\{x_1^l(t)\right\}_{l=1,2,\ldots,M}$ of the discrete-time stochastic process $X_1(t)$, with t ranging in almost three hours after 10.00 a.m., i.e., $1 \leq t \leq 17$. Below, the superscript "e" labels quantities empirically determined on the basis of this ensemble.

A fundamental element for the use of the model is the calibration of the parameters. The scaling exponent D and the scaling function g have here a central role as they appear in Eq. (4.1) and thus influence the joint PDF of a given daily realization of the process.

For the purposes of this study, we follow the practice of adopting a two step calibration procedure. First we evaluate D, and then we use the calibrated D in order to obtain a data-collapse which allows identifying g.

A quantitative way to calibrate D is offered by the analysis of the moments of $X_t(t)$: $\langle |x_t(t)|^q \rangle_e \equiv \frac{1}{M} \sum_{l=1}^{M} |x_t^l(t)|^q$. These indicators have a theoretical counterpart: a simple integration shows that if $p_{X_t(t)}(x)$ satisfies Eq. (4.1) independently of g, one must also have $\langle x^q \rangle_{p_{X_t(t)}} = \int x^q p_{X_t(t)}(x) dx = \int x^q \frac{1}{t^D} g\left(\frac{x}{t^D}\right) dx \propto t^{qD}$. The comparison between theoretical and empirical non-linear moments is displayed in Fig. 4.1(a) and the best fitted value for the scaling exponent is $D = 0.364...$ Another key observation is that the empirical second moment $\langle x_1(t)^2 \rangle_e \equiv \sum_{l=1}^{M} [x_1^l(t)]^2 / M$ systematically decreases as

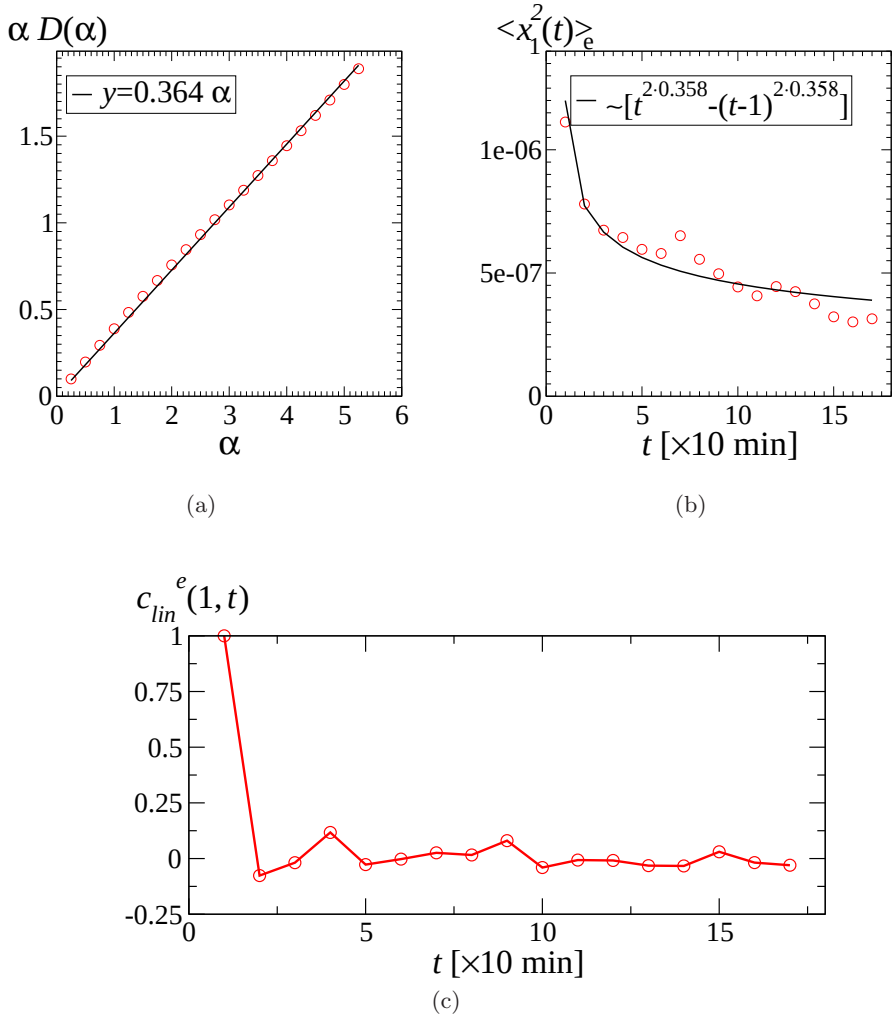


Figure 4.1: EUR/$ dataset: empirical ensemble analysis of the returns. (a) Analysis according to the ansatz in Eq.4.1. The straight line characterizes a simple-scaling behavior with a best fitted $D = 0.364$. (b) The line is given by $\langle\sigma^2\rangle_\rho \left[t^{2D} - (t-1)^{2D}\right]$, with $\langle\sigma^2\rangle_\rho = \langle x_1^2\rangle_p = 2.3 \cdot 10^{-7}$ and the best fitted $D = 0.358$. (c) The linear correlation vanishes for non-overlapping returns.

a function of t in the interval considered (see Fig. 4.1(b)). Its theoretical counterpart is

$$\begin{aligned} \langle x_t^2 \rangle_p \equiv \langle x^2 \rangle_{p_{X_1(t)}} &= \int x^2 p_{X_1(t)}(x) dx = \int x^2 \int_0^\infty d\sigma \rho(\sigma) \frac{\exp\left(-\frac{x^2}{2\sigma^2 a_t^2}\right)}{\sqrt{2\pi\sigma^2 a_t^2}} \\ &\propto a_t^2 = \left[t^{2D} - (t-1)^{2D}\right]. \end{aligned} \tag{4.4}$$

Here we can also see if the form of the a_i coefficients is compatible with the i-dependence already implied by the non-stationarity. Eq. 4.3 appears to be reasonably well compatible with the trend of the empirical mean square elementary returns $\langle |x_1(t)|^2 \rangle_e$. Indeed, given $\langle \sigma^2 \rangle_\rho = \langle x_1^2 \rangle_p = 2.3 \cdot 10^{-7}$, the best fit in Fig. 4.1(b) is obtained with $D = 0.358\ldots$ in the expression for $\langle x_t^2 \rangle_p$. The expectation value of σ^2 is with respect to the ρ entering the integral representation (Eq. 2.56) already chosen for g. Remarkably, the value of D is very close to the estimate of D obtained above through the analysis of the moments of $p_{X_t(t)}$.

In passing, an important empirical fact (Fig. 4.1(c)) is that the linear correlation between returns for non-overlapping intervals

$$c_{lin}^e(1,t) \equiv \frac{\frac{1}{M}\sum_{l=1}^M \left[x_1^l(1)\ x_1^l(t)\right]}{\sqrt{\langle |x_1(1)| \rangle_e \ \langle |x_1(t)| \rangle_e}}, \tag{4.5}$$

with $t = 2, \ldots,$ is negligible in comparison with the correlation of the absolute values of the same returns. At this time scale, also correlators of odd powers of a return with odd or even powers of another return are negligible. Only even powers of the returns are strongly correlated. This element is at the basis of our choice to set at 10 minutes the shortest time scale and justifies the application of the model [see, *e.g.* Eq. 2.12 or Eq. 2.25].

The scaling collapse, obtained via Eq. 4.1, is depicted in Fig. 4.2. The scaling function g identified by that collapse plot is manifestly non-Gaussian. It may also be assumed to be even to a good approximation[1]. In our case the set of data on which we can count to construct histograms of g is relatively poor. So, our determinations of ρ will be rather qualitative.

Once identified ρ, we use it for a weighted representation of the joint PDF's of the successive elementary returns $x_t \equiv x_1(t)$, $t = 1, 2, \ldots, 17$ generated in the process (Eq. 4.2).

[1] We have detrended the data by subtracting from $x_1^l(t)$ the average value $\sum_{l=1}^M x_1^l(t)/M$. Data skewness can be shown to introduce deviations much smaller than the statistical error-bars in the analysis of the correlators.

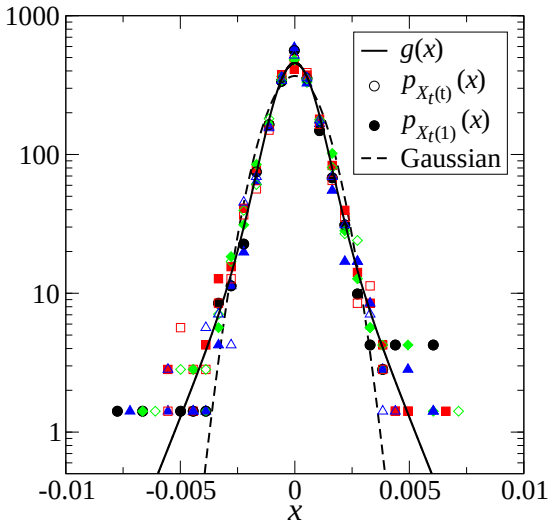


Figure 4.2: EUR/$ dataset. Non-Gaussian scaling function g. Empty [full] symbols are obtained by rescaling $p_{X_t(t)}$ $[p_{X_1(t)}]$ according to Eq. (4.1) [Eq. (2.55)] for $t = 1, 5, 10, 17$.

4.1.1.1 Correlations analysis

As discussed above, the identified ρ function may be used to reconstruct the joint PDF of the returns x_i's as in Eq. 4.2. In this section we elaborate further on this point, by performing a detailed comparison between model predictions (based on an explicit expression for ρ) and empirical determinations of various two-point correlators.

Considering the data collapse of both $p_{X_t(t)}$ and $p_{X_1(t)}$ in Fig. 4.2, we use the power law functional form of Eq. 2.44 for ρ [see also (49)]:

$$\rho(\sigma) = A\,\frac{\sigma^{\lambda}}{d+\sigma^{\nu}}, \quad \sigma \in [\sigma_{min}, +\infty[, \quad 0 < \lambda < \nu, \tag{4.6}$$

where A is a normalization factor, and $d > 0$ is a parameter influencing the width of the distribution g. Notice that $\rho(\sigma) \sim \sigma^{-(\nu-\lambda)}$ for $\sigma \gg 1$. The rational behind this choice for ρ is that one can use the exponents λ, ν to reproduce the large $|x|$ behavior of $g(x)$, and then play with the other parameters to obtain a suitable fit of the scaling function, for instance the one reported in Fig. 4.2. In any case, at this level of analysis, a clear indication about the tail behaviour is not available, and so the distinction between power-law or exponential tails is merely arbitrary.

The first two-point correlator we consider in our analysis is

$$\kappa_{\alpha,\beta}(1,n) \equiv \frac{\langle |x_1(1)|^{\alpha}\ |x_1(n)|^{\beta}\rangle_p}{\langle |x_1(1)|^{\alpha}\rangle_p\ \langle |x_1(n)|^{\beta}\rangle_p} = \frac{\langle |x_1|^{\alpha}\ |x_n|^{\beta}\rangle_p}{\langle |x_1|^{\alpha}\rangle_p\ \langle |x_n|^{\beta}\rangle_p}, \tag{4.7}$$

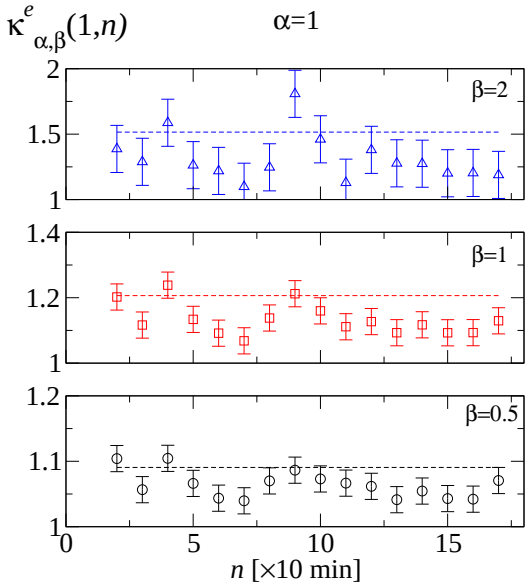


Figure 4.3: Constancy of $\kappa^e_{\alpha,\beta}$ for the EUR/$ dataset. Dashed lines are model predictions.

with $n > 1$, and $\alpha, \beta \in \mathbb{R}_+$. A value $\kappa_{\alpha,\beta} \neq 1$ means that returns on non-overlapping intervals are dependent. Using Eq. 4.2 it is possible to express a general many-return correlator in terms of the moments of ρ. For example, from Eq. 4.2 we have

$$\langle |x_1|^\alpha \ |x_n|^\beta \rangle_p = B_\alpha \, B_\beta \ \ a_1^\alpha \, a_n^\beta \, \langle \sigma^{\alpha+\beta} \rangle_\rho, \tag{4.8}$$

with

$$B_\alpha \equiv \int_{-\infty}^{+\infty} dx \ |x|^\alpha \ \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \tag{4.9}$$

and

$$\langle \sigma^\lambda \rangle_\rho = \int_0^{+\infty} d\sigma \ \rho(\sigma) \sigma^\lambda. \tag{4.10}$$

Thus we obtain

$$\kappa_{\alpha,\beta}(1,n) = \frac{\langle \sigma^{\alpha+\beta} \rangle_\rho}{\langle \sigma^\alpha \rangle_\rho \, \langle \sigma^\beta \rangle_\rho} = \frac{B_\alpha B_\beta}{B_{\alpha+\beta}} \frac{\langle |x_1|^{\alpha+\beta} \rangle_p}{\langle |x_1|^\alpha \rangle_p \, \langle |x_1|^\beta \rangle_p}. \tag{4.11}$$

Two model predictions in Eq. 4.11 are: (i) Despite the non-stationarity of the increments x_i's, $\kappa_{\alpha,\beta}(1,n)$ is independent of n; (ii) The correlators are symmetric, i.e., $\kappa_{\alpha,\beta} - \kappa_{\beta,\alpha} = 0$.

We can now compare the theoretical prediction of the model for $\kappa_{\alpha,\beta}(1,n)$, Eq. 4.11, with the empirical counterpart

$$\kappa^e_{\alpha,\beta}(1,n) \equiv \frac{\sum_{l=1}^M \left[\left| x_1^l \right|^\alpha \ \left| x_n^l \right|^\beta \right]}{\frac{1}{M} \sum_{l=1}^M \left| x_1^l \right|^\alpha \ \sum_{l=1}^M \left| x_n^l \right|^\beta}, \tag{4.12}$$

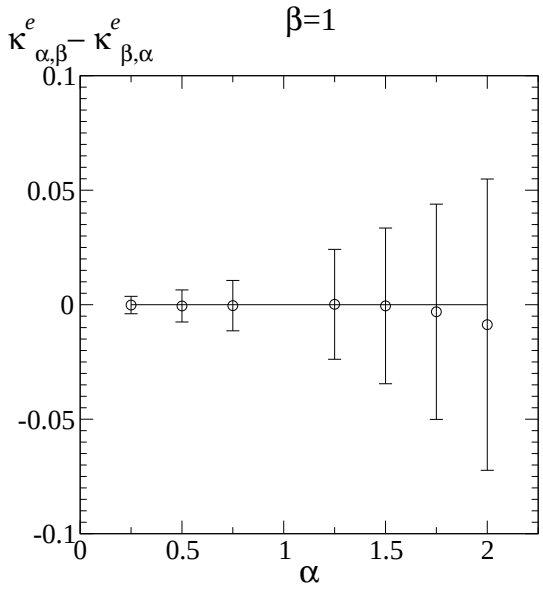


Figure 4.4: Symmetry of $\kappa^e_{\alpha,\beta}$ for the EUR/$ dataset. Error-bars are determined as in Fig. 4.3.

which we can calculate from the EUR/$ dataset. Notice that once ρ is fixed to fit the one-time statistics in Fig. 4.2, in this comparison we do not have any free parameter to adjust. Also, since our ensemble is restricted to $M = 1,282$ realizations only, large fluctuations, especially in two-time statistics, are to be expected.

Fig. 4.3 shows that non-overlapping returns are strongly correlated in the about three hours following the opening of the trading session, since $\kappa^e_{\alpha,\beta} \neq 1$. In addition, the constancy of $\kappa^e_{\alpha,\beta}$ is clearly suggested by the empirical data. In view of this constancy, we can assume as error-bars for $\kappa^e_{\alpha,\beta}$ the standard deviations of the sets $\left\{\kappa^e_{\alpha,\beta}(1,n)\right\}_{n=2,3,\ldots,17}$. The empirical values for $\kappa^e_{\alpha,\beta}$ are also in agreement with the theoretical predictions for $\kappa_{\alpha,\beta}$ based on our choice for ρ. In this and in the following comparisons it should be kept in mind that, although not explicitly reported in the plots, the uncertainty in the identification of ρ of course introduces an uncertainty in the model predictions for the correlators.

In Fig. 4.4 we report that also the symmetry $\kappa_{\alpha,\beta} = \kappa_{\beta,\alpha}$ is empirically verified for the EUR/$ dataset. The validity of this symmetry for a process with non-stationary increments is quite remarkable.

A classical indicator of strong correlations in financial data is the volatility auto-

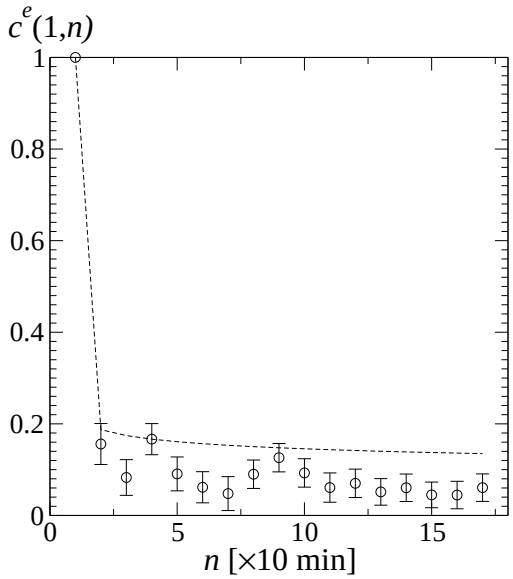


Figure 4.5: Volatility autocorrelation for the EUR/$ dataset. Dashed line is the model prediction.

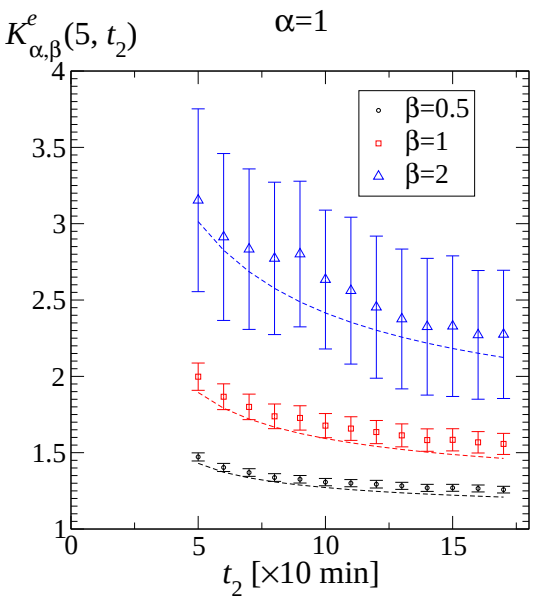


Figure 4.6: Correlators $K^e_{\alpha,\beta}$ for the EUR/$ dataset. Dashed lines are model predictions.

correlation, defined as

$$c(1,n) \equiv \frac{\langle |x_1|\ |x_n| \rangle_p - \langle |x_1| \rangle_p \langle |x_n| \rangle_p}{\langle |x_1|^2 \rangle_p - \langle |x_1| \rangle_p^2}. \tag{4.13}$$

In terms of the moments of ρ, through Eq. 4.8 we have the following expression for c:

$$c(1,n) = \frac{B_1^2\ a_1\ a_n\ \left[\langle \sigma^2 \rangle_\rho - \langle \sigma \rangle_\rho^2\right]}{a_1^2\ \left[B_2\ \langle \sigma^2 \rangle_\rho - B_1^2 \langle \sigma \rangle_\rho^2\right]}. \tag{4.14}$$

Unlike $\kappa_{\alpha,\beta}$, c is not constant in n. The comparison with the empirical volatility autocorrelation,

$$c^e(1,n) \equiv \frac{\sum_{l=1}^M \left[\left|x_1^l\right|\ \left|x_n^l\right|\right] - \frac{1}{M} \sum_{l=1}^M \left|x_1^l\right|\ \sum_{l'=1}^M \left|x_n^{l'}\right|}{\sum_{l=1}^M \left|x_1^l\right|^2 - \frac{1}{M} \sum_{l=1}^M \left|x_1^l\right|\ \sum_{l'=1}^M \left|x_1^{l'}\right|}, \tag{4.15}$$

yields a substantial agreement (see Fig. 4.5). The error-bars in Fig. 5 are obtained by dynamically generating many ensembles of $M = 1,282$ realizations each, according to Eq. 4.2 with our choice for ρ, and taking the standard deviations of the results. Again, the uncertainty associated to the theoretical prediction for c is not reported in the plots. Problems concerning the numerical simulation of processes like the one in Eq. 4.2 are discussed in (49) and in Sec. 3.2.

A further test of our model can be made by analyzing, in place of those of the increments, the non-linear correlators of $x_t(t)$, with varying t. To this purpose, let us define

$$K_{\alpha,\beta}(t_1,t_2) \equiv \frac{\langle |x_{t_1}(t_1)|^\alpha\ |x_{t_2}(t_2)|^\beta \rangle}{\langle |x_{t_1}(t_1)|^\alpha \rangle\ \langle |x_{t_2}(t_2)|^\beta \rangle}, \tag{4.16}$$

with $t_2 \geq t_1$. Model calculations similar to the previous ones give, from Eq. 4.2,

$$K_{\alpha,\beta}(t_1,t_2) = \frac{B^{(2)}_{\alpha,\beta}(t_1,t_2)}{t_1^{\alpha D} t_2^{\beta D} B_{\alpha+\beta}} \frac{\langle |x_1|^{\alpha+\beta} \rangle_p}{\langle |x_1|^\alpha \rangle_p\ \langle |x_1|^\beta \rangle_p}, \tag{4.17}$$

where

$$B^{(2)}_{\alpha,\beta}(t_1,t_2) \equiv \int_{-\infty}^{+\infty} dx_1\ |x_1|^\alpha\ \frac{\exp\left(-x_1^2/\left(2t_1^{2D}\right)\right)}{\sqrt{2\pi t_1^{2D}}}$$
$$\int_{-\infty}^{+\infty} dx_2\ |x_2|^\beta\ \frac{\exp\left[-(x_1-x_2)^2/\left(2t_2^{2D}-2t_1^{2D}\right)\right]}{\sqrt{2\pi\left(t_2^{2D}-t_1^{2D}\right)}}. \tag{4.18}$$

According to Eq. 4.17, $K_{\alpha,\beta}$ is now identified by both ρ and D. Moreover, it explicitly depends on t_1 and t_2. The comparison between Eq. 4.17 and the empirical quantity

$$K^e_{\alpha,\beta}(t_1,t_2) \equiv \frac{\sum_{l=1}^M \left[\left|x^l_{t_1}(t_1)\right|^\alpha\ \left|x^l_{t_2}(t_2)\right|^\beta\right]}{\frac{1}{M} \sum_{l=1}^M \left[\left|x^l_{t_1}(t_1)\right|^\alpha\right]\ \sum_{l=1}^M \left[\left|x^l_{t_2}(t_2)\right|^\beta\right]}, \tag{4.19}$$

reported in Fig. 4.6 (the error-bars are determined as in Fig. 4.5) supplies an additional validation of this intraday model.

4.1.2 A lesson form larger datasets: the S&P500 index

In this section we develop the intraday analysis for the dataset of the S&P500 index ranging from September 30th 1985 to October 20th 2010. For each single day l ($1 \leq l \leq M$), we consider the index values, $\{s^l(t)\}$ ($t = 0, 1, \ldots, 20$), every 10 minutes between 9:40 a.m. (when we set $t = 0$) and 13:00 a.m. ($t = 20$), New York time[1]. With respect of the previous Section, we are working with a larger set: we have $M = 6283$ realizations of the stochastic variable $X_1(t) \equiv \ln S(t) - \ln S(t-1)$ with $1 \leqslant t \leqslant 20$. This larger database reduces to some extent the statistical fluctuations of the empirical quantities. Interestingly, the same features extracted from the EUR/$ exchange rate also characterize the S&P500. The results of this section will be used as a basis for the development of an intraday trading strategy in Section 4.2.

The procedure for the analysis follows the same steps as above, but with a different choice for the form of the ρ function. In fact, we adopt here the inverse gamma function of Eq. 2.46, that we rewrite here for clearness:

$$\rho(\sigma) = \frac{\sigma_0^\alpha}{\Gamma(\alpha)\sigma^{1+\alpha}} \exp\left(-\frac{\sigma_0}{\sigma}\right). \tag{4.20}$$

The parameter $\sigma_0 > 0$ is called *scale factor*, whilst $\alpha > 0$ is named *form factor*. The choice of this function is here motivated by the wish to adhere to the common choice in finance to fit the realized volatilities of an asset (see, *e.g.*, 17) with such a function. Nevertheless, as already remarked, the differences between the results obtained with different forms of the ρ, at this level of analysis, are not discriminable. As a simplifying assumption, we disregard the weak skewness of the return distribution, whose inclusion in the modeling framework is delayed till Section 4.1.3.

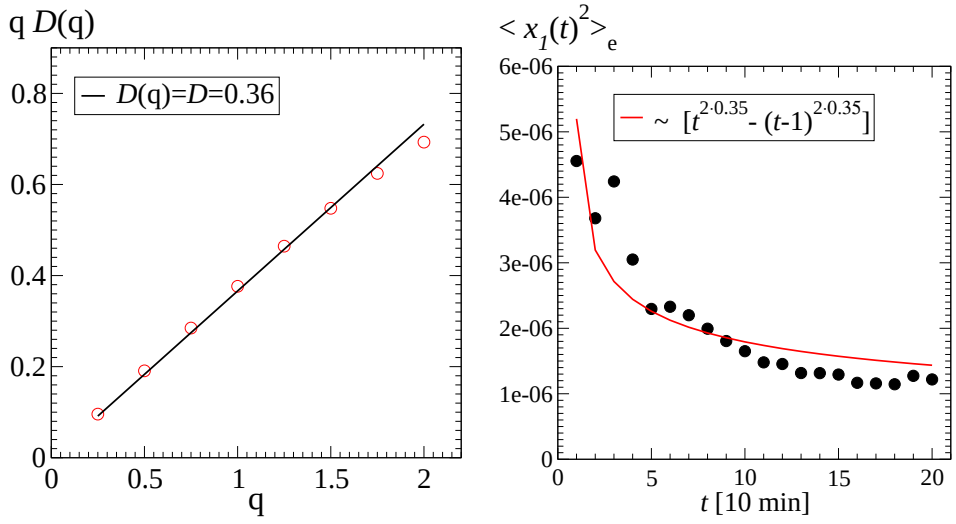
The resulting scaling function

$$g(x) = \int_0^\infty \rho(\sigma) \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} d\sigma \tag{4.21}$$

has then a power-law decay for large $|x|$ with exponent $\alpha + 1$, whereas σ_0 simply sets the scale of its width. Within this parametrization, g is thus identified in terms of the two parameters σ_0 and α.

The scaling parameter D is again selected (see Section 4.1.1) by checking the consistency of the analysis of the non linear moments $\langle |x_t(t)|^q \rangle_e \equiv \frac{1}{M} \sum_{l=1}^M |x_t^l(t)|^q$ (Fig. 4.7(a)) with the trend of the second moment, with respect to the time t, $\langle x_1(t)^2 \rangle_e \equiv \sum_{l=1}^M [x_1^l(t)]^2/M$ (Fig. 4.7(b)). Taking advantage of the knowledge of the scaling exponent $D = 0.36$, a χ^2 minimization algorithm shows that the values $\sigma_0 = 0.0045$

[1] Again, we work in a 10 minutes' time scale



(a) (b)

Figure 4.7: Empirical ensemble analysis of the returns (S&P500 dataset). (a) Analysis according to the ansatz in Eq. (4.1). The straight line characterizes a simple-scaling behavior with a best fitted $D = 0.36$. (b) The line is given by $\langle \sigma^2 \rangle_\rho \ \left[t^{2D} - (t-1)^{2D}\right]$, with the best fitted $D = 0.35$.

and $\alpha = 3.69$ adequately fit (see Fig. 4.8) the data-collapse obtained via Eq. (4.1). These parameters, calibrated upon the entire available history of the S&P500 constitute the basis for the development of the *in sample* testing of the strategy presented in Section 4.2.

But, before approaching the strategy problem, we want to see what are the differences with respect to the EUR/$ analysis. First of all, referring to Fig. 4.7(a), we can see that the analysis is limited to $q = 2$, while we went up to $q = 5$ in Fig. 4.3. That is because in the present case we detect a multiscaling effect at $q \simeq 2$ [for multiscaling in S&P500 index see, *e.g.*, (51)]. So, the linear fitting must be restricted only to the region of low moments. Correspondingly, the analysis of the correlators of Eq. 4.11 and Eq. 4.12 that leads to Fig. 4.9 must rise some remarks:

- The error bars in Fig. 4.9 are reduced with respect to those in Fig. 4.3. This is essentially due to the larger dataset of the S&P500, that reduces the uncertainty in the empirical indicators.
- There is a slight discrepancy for what concerns the two graphs for $\beta = 0.5$ and $\beta = 1$, where the theoretical prediction seems too large.
- Even if the top graph in Fig. 4.9 seems to show correct results, it has to be noticed that the data are in a multiscaling regime, since $\alpha + \beta = 3$.

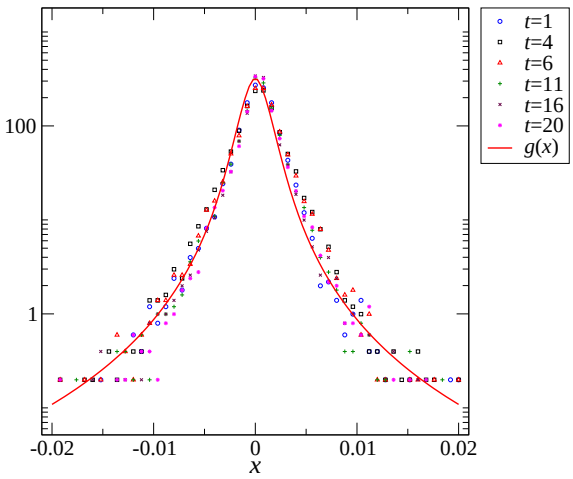


Figure 4.8: The scaling function for the S&P500, after aggregation. The g function is constructed via Eq. 4.21 with a ρ function chosen in the form of an inverse gamma, whose best fitted parameters are $\sigma_0 = 0.0045$ and $\alpha = 3.69$. A certain asymmetry is clearly spotted.

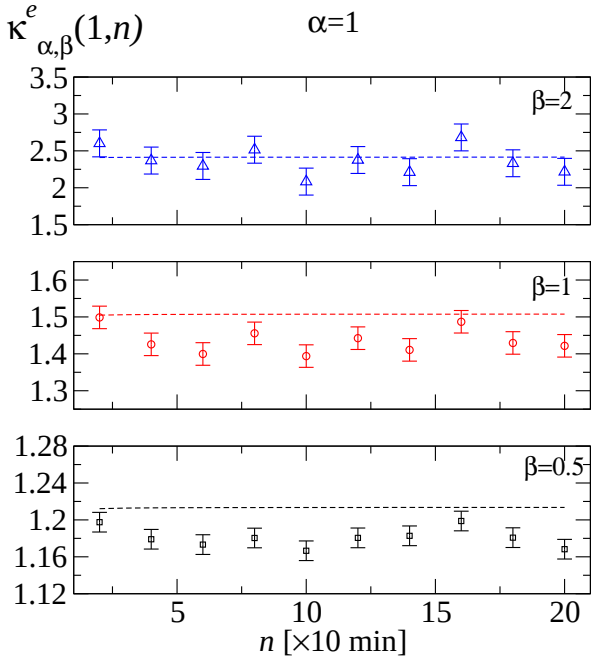


Figure 4.9: Constancy of $\kappa^e_{\alpha,\beta}$ for the S&P500 dataset. Dashed lines are model predictions. Note the slight discrepancy and compare with the later results of Fig. 4.15

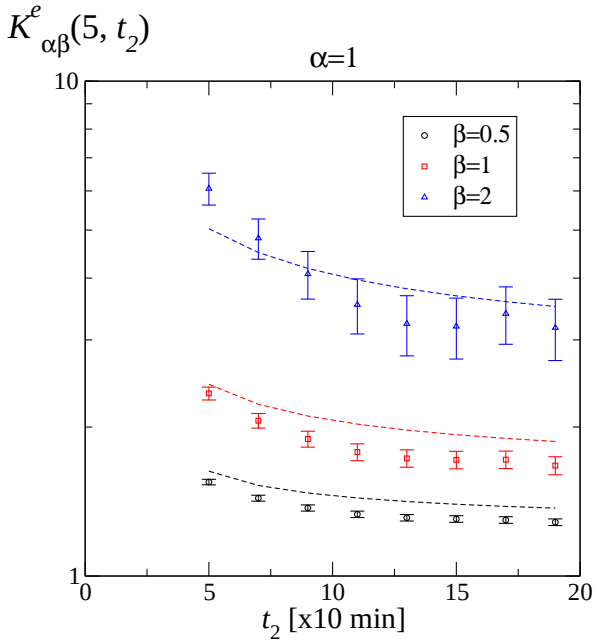


Figure 4.10: Correlators $K^e_{\alpha,\beta}$ for the S&P500 dataset. Dashed lines are model predictions. Note, again a certain discrepancy and compare with the later results of Fig. 4.16

.

We will settle these last two questions in Section 4.1.3. By now we content ourselves with having lower statistical errors, as can also be seen from the analysis of the theoretical (Eq. 4.17) and empirical (Eq. 4.19) process correlators, reported in Fig. 4.10.

Despite the slight discrepancies, the general trend of the curves is replicated by the model, together with the symmetry of the empirical correlators $\kappa_{\alpha,\beta} = \kappa_{\beta,\alpha}$, as shown in Fig. 4.11.

For what concerns the volatility autocorrelation (see Eq. 4.14 and Eq. 4.15) the agreement is remarkable for this dataset: the results are presented in Fig. 4.12.

The results of this section have highlighted some discrepancies between the empirical and the theoretical results. The presence of multiscaling effects and the asymmetry of the scaling function make the above related analysis too approximate. In the next section the asymmetric model will be employed, with significantly improved results.

4.1.3 The skewed intraday model: an improved analysis for the S&P500 index

In this section we profit from the scheme of Sec. 2.4 to try to improve the results of the above reported analysis on the S&P500 dataset.

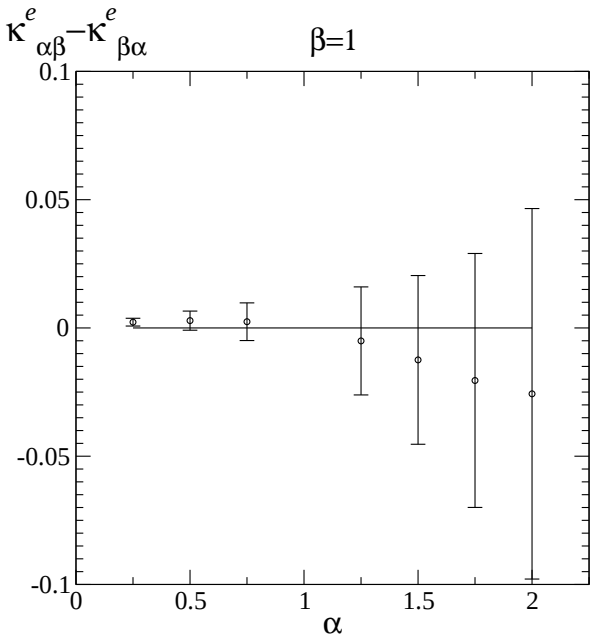


Figure 4.11: Symmetry of $\kappa^e_{\alpha,\beta}$ for the S&P500 dataset. Error-bars are determined as in Fig. 4.3.

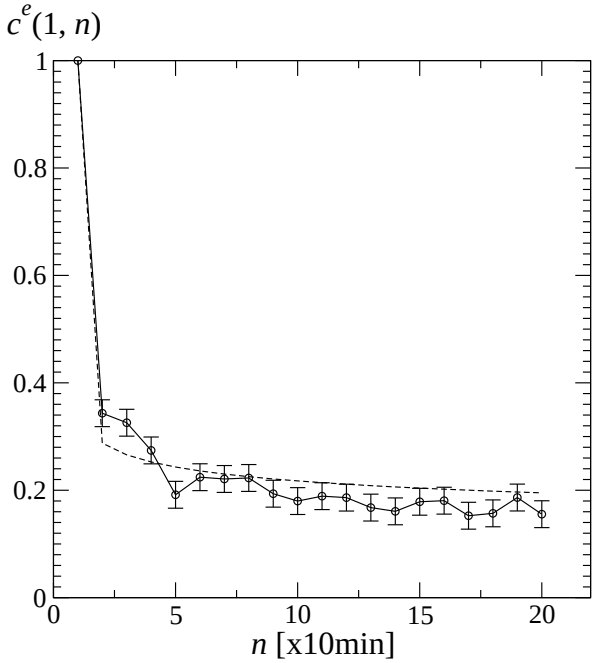


Figure 4.12: Volatility autocorrelation for the S&P500 dataset. Dashed line is the model prediction.

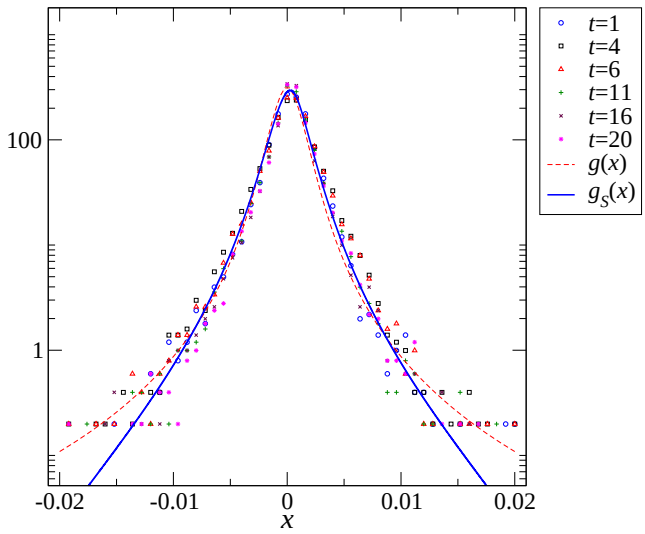


Figure 4.13: The skewed scaling function for the S&P500. The g_S function is constructed via Eq. 2.62 with a ψ function chosen in the form of $\psi(\sigma,\mu) = k\rho_1(\sigma)\delta(\mu-\overline{\mu}) + (1-k)\rho_2(\sigma)\delta(\mu+\overline{\mu})$ with $k = 0.7$, $\overline{\mu} = 4\cdot 10^{-4}$ and $\rho_i(\sigma)$ two inverse gamma functions of the form Eq. 2.46 with parameters $\alpha_1 = 3.7$, $\sigma_{0_1} = 0.0045$, $\alpha_2 = 3.4$ and $\sigma_{0_2} = 0.0049$. The asymmetry is grasped. The old symmetric g function of Fig. 4.8 is given, for reference.

The first goal will be that of correcting the fit of the scaling function of Fig. 4.8. Therein, some fitting problems can be detected, both in the center of the distribution and in the first part of the tails, closer to the maximum (and in particular to the right). The introduction of two non zero-centered inverse gamma functions provides a certain evident improvement: we write a *skewed* scaling function $g_S(x)$ using Eq. 2.62 with a ψ form that is a simple modification of Eq. 2.63:

$$\psi(\sigma,\mu) = k\rho_1(\sigma)\delta(\mu-\overline{\mu}) + (1-k)\rho_2(\sigma)\delta(\mu+\overline{\mu}). \tag{4.22}$$

As before, the fitting of the parameters has been produced by means of a χ^2-minimization algorithm, for different choices of the weight k and of the shift parameter $\overline{\mu}$.

The qualitative agreement that is recognizable in Fig. 4.13 can find a strong quantitative support through the analysis of the above considered correlators (Sec. 4.1.1.1).

To start with, let us check the all important volatility autocorrelation $c(1,n)$ (Eqs. 4.14, 4.15 and Fig. 4.14). The introduction of a skewed model generates a little correction, whose effect is to approximate better the empirical data.

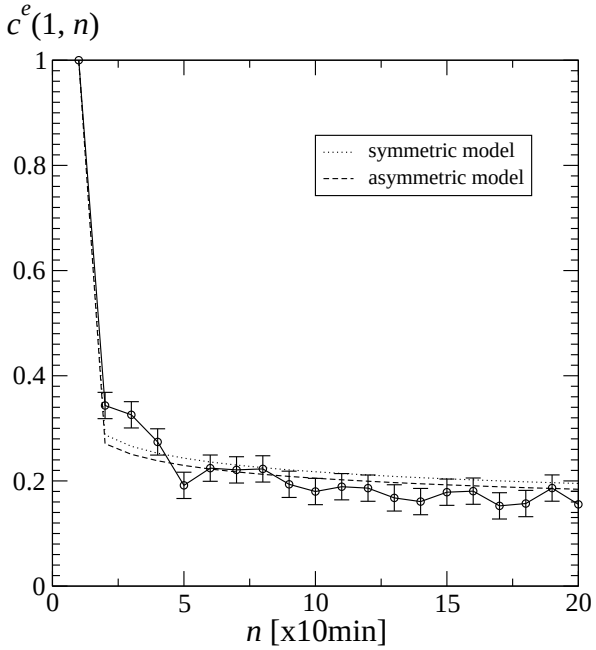


Figure 4.14: Volatility autocorrelation for the S&P500 dataset. The asymmetric model gives a theoretical prediction that is more centered with the empirical data.

But the most significant improvement can be found when analyzing the correlators $\kappa^e_{\alpha,\beta}$ of Eqs. 4.11 and 4.12. The agreement, especially for the two lower moments depicted in Fig. 4.15, is remarkable. The same degree of improvement is to be found by comparing the graphs of Fig. 4.16: there, in the asymmetric model, the empirical (Eq. 4.19) and the theoretical (Eq. 4.17) behaviour for the $K^e_{\alpha,\beta}$ correlators is found in good agreement, especially for the lowest moments.

We've already noticed that, for problems of multiscaling, the analysis of the correlators should be limited to overall values of the moments below, let say, 2. Beyond this limit the scaling exponent D takes other, more unstable, values. For this reason, for lower values of α and β, the analysis of Fig. 4.15 has been be repeated. The results are displayed in Fig. 4.17, where the agreement between the theoretical and empirical values is relevant.

In this section we significantly improved the results of Sec. 4.1.2 for the S&P500 intraday dataset by introducing a skewed scaling function g_S in the form of Eq. 4.22. The asymmetric features of the scaling function of the returns forces the employment of the model detailed in Sec. 2.4. The model grasps the features of the scaling function as well as all the so far studied low-order correlators.

Differently from the EUR/$ dataset, where the symmetric model already produced

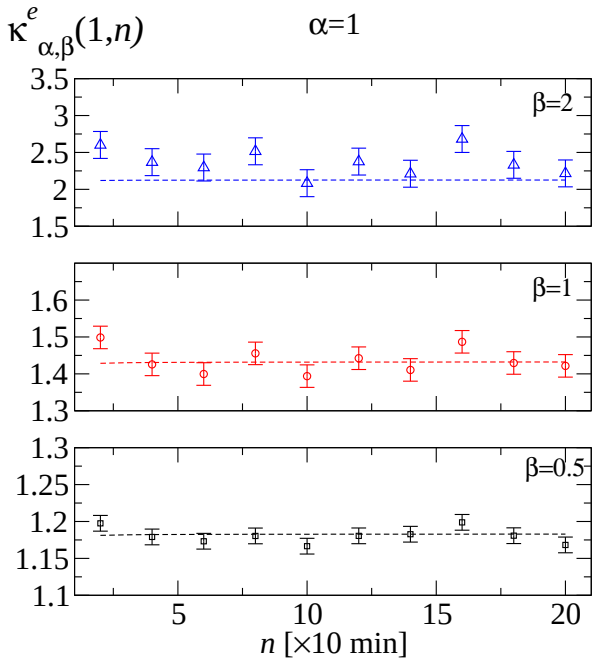


Figure 4.15: Constancy of $\kappa^e_{\alpha,\beta}$ for the S&P500 dataset. Dashed lines are model predictions for the skewed model. Note that the theoretical lines are much more centered, compared to Fig. 4.9.

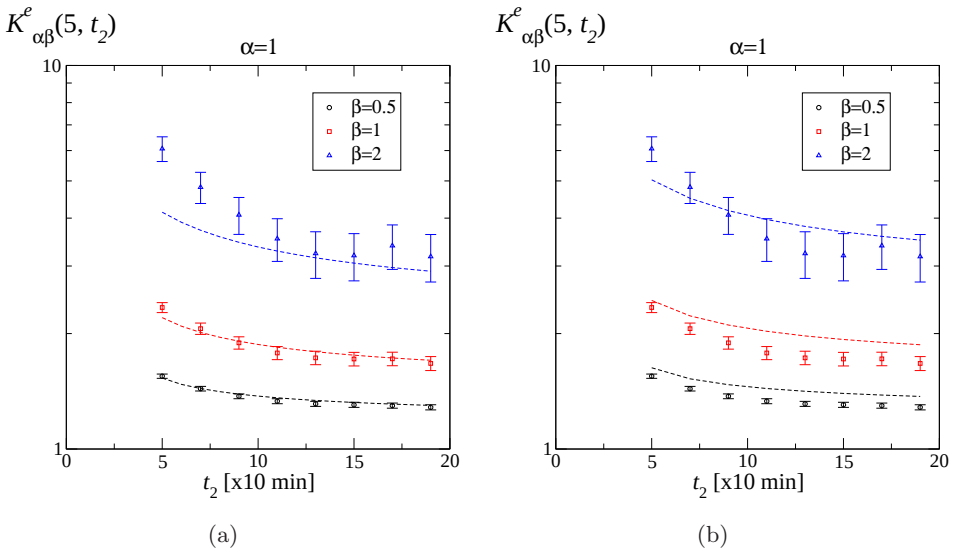


Figure 4.16: Correlators $K^e_{\alpha,\beta}$ for the S&P500 dataset. Dashed lines are model predictions. (a) The theoretical results for the skewed model; (b) the same graph of Fig. 4.10 (symmetric model) is given for reference.

.

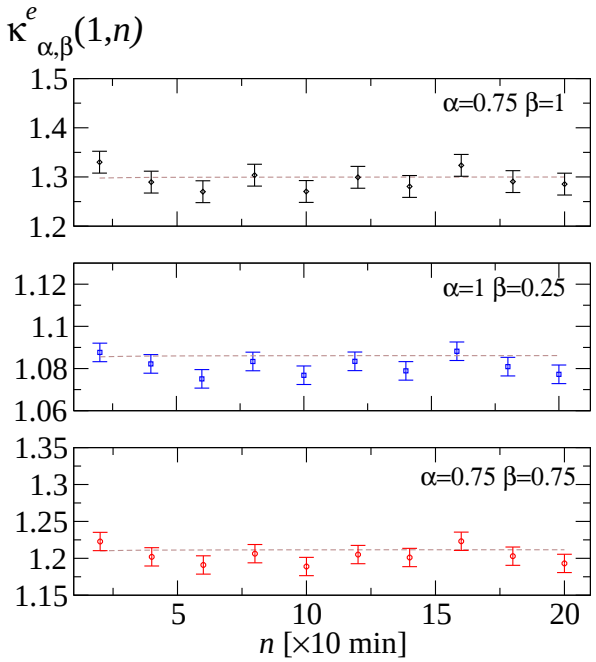


Figure 4.17: S&P500 dataset. Correlators $\kappa^e_{\alpha,\beta}$ for lower moments. Dashed lines are model predictions for the skewed model.

satisfactory predictions for the correlators and for the scaling collapse (7), the S&P500 intraday dataset requested the application of a skewed scheme, whose introduction produced so relevant improvements to be regarded as a good test for the applicability of the general theory of the model of Baldovin and Stella (8, 49) to non-symmetric datasets.

4.1.3.1 Sneaking a look at the intraday leverage effect

As a concluding task, thanks to the relevant presence of asymmetry, the leverage effect for this S&P500 intraday dataset can tentatively be investigated.

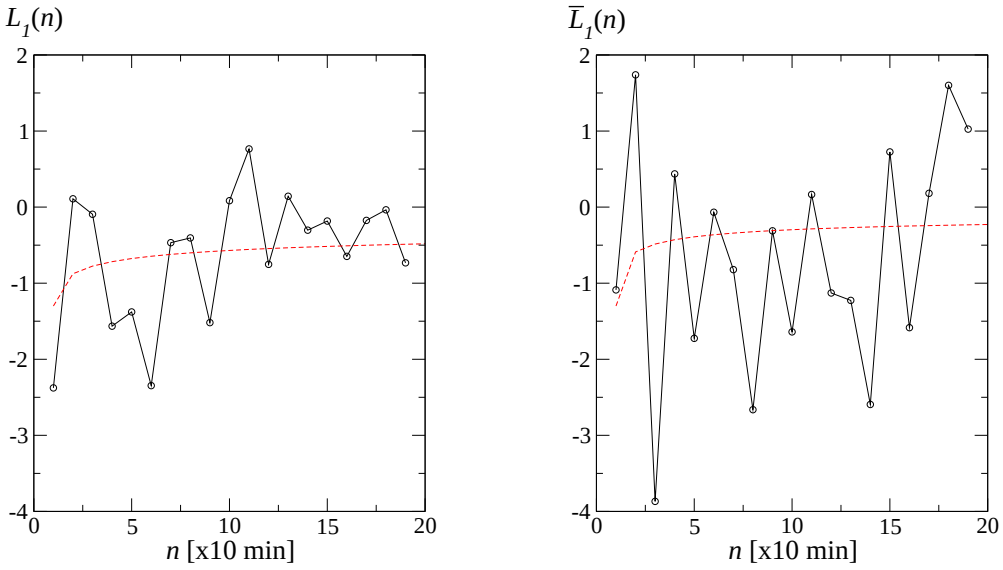
In particular, we want to analyze the return-volatility correlation during the considered three hours of market. To do so, we might look at the natural intraday specifications of Eq. 2.59 :

$$\mathcal{L}_1(n) = \frac{\langle x_1(1)|x_1(n)|^2\rangle_e}{\langle x_1(1)^2\rangle_e{}^2} \tag{4.23}$$

and

$$\overline{\mathcal{L}}_1(n) = \frac{\langle |x_1(1)|^2 x_1(n)\rangle_e}{\langle x_1(1)^2\rangle_e{}^2}. \tag{4.24}$$

where, as before, the subscript e indicates the average taken over the ensemble of histories and n is the intraday time, in 10-min units. The idea is here to refer to the first



(a) (b)

Figure 4.18: Return-volatility correlators. (a) $L_1(n)$ correlator between the first return and the following absolute returns of the day (leverage effect); (b) $\overline{L}_1(n)$ correlator between the first absolute return and the following returns of the day. Notice the large fluctuations of the indicators (larger in (b)). Dashed lines are model predictions [as an approximation we depict the trend $\sim (n^{2D} - (n-1)^{2D})^{1/2}$ in (a) and $\sim (n^{D} - (n-1)^{D})^{1/2}$ in (b)].

return of the day and to watch at the n-dependence of $\mathcal{L}_1(n)$ and $\overline{\mathcal{L}}_1(n)$. Unfortunately, the empirical results are affected by large statistical noise, so it's convenient to verify the following two indicators instead:

$$L_1(n) = \frac{\langle x_1(1)|x_1(n)|\rangle_e}{\langle x_1(1)|x_1(1)|\rangle_e} \tag{4.25}$$

and

$$\overline{L}_1(n) = \frac{\langle |x_1(1)|x_1(n)\rangle_e}{\langle |x_1(1)|x_1(1)\rangle_e}. \tag{4.26}$$

Looking at the results in Fig. 4.18, some remarks are to be highlighted:

- the indicators of Eq. 4.25 and Eq. 4.26 are still very noisy. Some more definite trend can be spotted for $L_1(n)$ in Fig. 4.18(a) , but the overall result is not significative;

- given these fluctuations, the theoretical results are not comparable with the empirical evidence. Moreover, given the level of uncertainty, the approximation depicted in Fig. 4.18 is justified;

- the lower level of noise in Fig. 4.18(a) with respect to Fig. 4.18(b) may indicate that a leverage effect is effectively present, in the sense that there is a certain definite negative correlation between the first return of the day and the subsequent absolute returns. The opposite, *i.e.* the correlation between the first absolute return and the following signed returns, is less evident, as expected from considering the efficient market hypothesis. Nevertheless, we remind that the model in analysis prescribes a negative correlation for these returns, too (see Eq. 2.71).

We need to remind that in Eq. 4.25 and Eq. 4.26 the total exponent of the correlators is 2: near to the upper boundary of the scaling validity range. To get rid of possible multiscaling effects, one may consider *e.g.* the following correlators, instead:

$$L_1^{(0.25)}(n) = \frac{\langle x_1(1)|x_1(n)|^{0.25}\rangle_e}{\langle x_1(1)|x_1(1)|^{0.25}\rangle_e} \tag{4.27}$$

and

$$\overline{L}_1^{(0.25)}(n) = \frac{\langle |x_1(1)|^{0.25}x_1(n)\rangle_e}{\langle |x_1(1)|^{0.25}x_1(1)\rangle_e}. \tag{4.28}$$

However, the behaviour of these two indicators is still affected by large statistical error.

As a conclusion, we may say that the model prediction for the leverage effect cannot be directly verified on the data, because the statistical fluctuations derived from the relatively small size of the ensemble do not allow a clear identification of the trends of the leverage effect indicators. The testing of the model predictions for the leverage effect in an intraday scale certainly needs other techniques and indicators to be adequately investigated.

4.2 Intraday trading rules

4.2.1 Building *in sample* and *out of sample* quantiles sets

The analysis performed in Sec. 4.1.2 on the full history of the S&P500 dataset (from 1985 to 2010) will be here used to analyse the $in-sample$ performances of a trading strategy based on the model outcomes, as discussed in the following sections. However, $out\ of\ sample$ analyses are also needed. Starting from our 25 years' database, we decided to use 15 years of data (from 1985 to 1999) to build a scaling function to be used to initialize the trading strategy for the year 2000. We then repeatedly shifted year by year, always using the 15 previous years to calibrate the model, until 2009 (the last complete year). So, *e.g.*, we used the data from 1994 to 2008 to fit the scaling function for the strategy to be used for year 2009. As a simplifying assumption, we employ a symmetric model, with a scaling function g built with a $\rho(\sigma)$ shaped as an inverse gamma

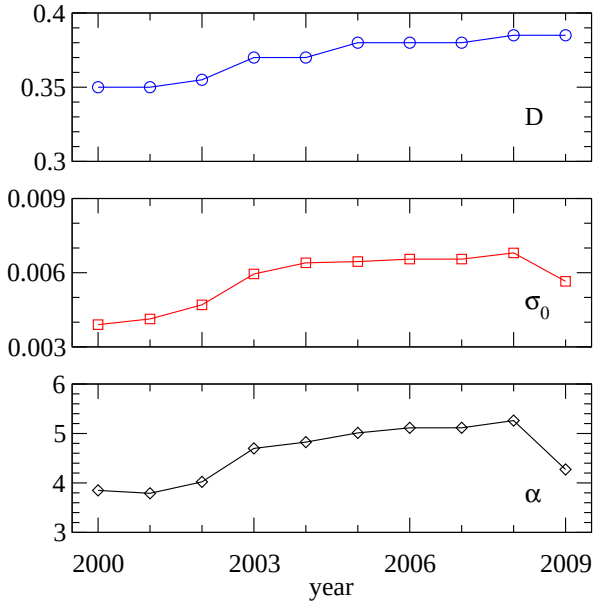


Figure 4.19: Fitting parameters for the $\rho(\sigma)$ in the *out of sample* analysis (history = 15 previous years) of the S&P500 dataset.

function (Eqs. 4.20 and 4.21). Fig. 4.19 shows the results of the χ^2-minimization, and the optimal values of the parameters used for running the *out of sample* strategies from 2000 and 2009. Notably, the scaling parameter D is increasing over time and stabilize from 2005 to 2009. Differently, the parameters characterizing the inverse-gamma density show evidence of a clear drop once the crisis period of 2008 is introduced in the calibration sample. See, *e.g.*, the comparison between the inverse-gamma's of year 2005 and 2009 presented in Fig. 4.20.

4.2.2 Density forecasts and trading signals

The model described in Chap. 2 has implicit density forecasting abilities. In fact, assuming correct model specification, given the parameter values and the first daily return, we know the PDF of the daily evolution of returns. Furthermore, the knowledge is not limited to each single return, but extends to any aggregation of returns. Such understanding allows the construction of a trading strategy based on the density forecasts. Furthermore, at each point in time within the day, the forecasts for the remaining part of the day might be updated using the additional intra-daily information. For this reason, we consider in the following two different trading approaches: the first just makes use of the index value, the first available for the day, and then considers

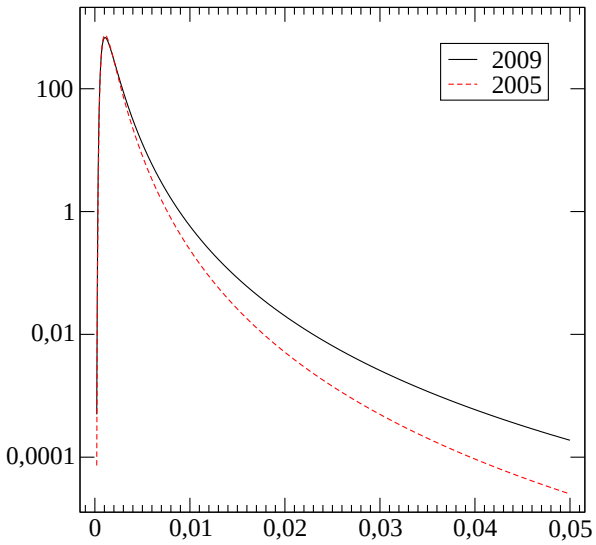


Figure 4.20: Comparison between the inverse-gamma functions for the S&P500, calibrated before and after the 2008 crisis.

the density forecasts made for the remaining part of the day (we call it *Unconditional trading*); the second approach allows for an update of the density forecasts using more recent returns, too (we denote it *Conditional trading*, since it conditions the forecasts to the new information). Both trading approaches extract trading signals from density forecasts: if at a given point within the intra-daily range under study the observed market price is above (below) the $1-q$ (q) quantile of the predicted price density we have a buy (sell) signal.

4.2.2.1 Unconditional trading

Taking advantage of the predictive features of the model, we calculate for every day the quantiles for the expected price at time t, starting from the opening value S_0 which is the only conditioning information. Note that the model, in computing the density forecasts, does not include the set of information of the prices observed in the past. It is assumed that all the information contained in the past observations has been compressed in the model parameter, namely the scaling exponent D. At this stage we consider an unconditioned situation. We assume that the probability density for the return is shaped according to the formula for the historical $g(r)$, with the time

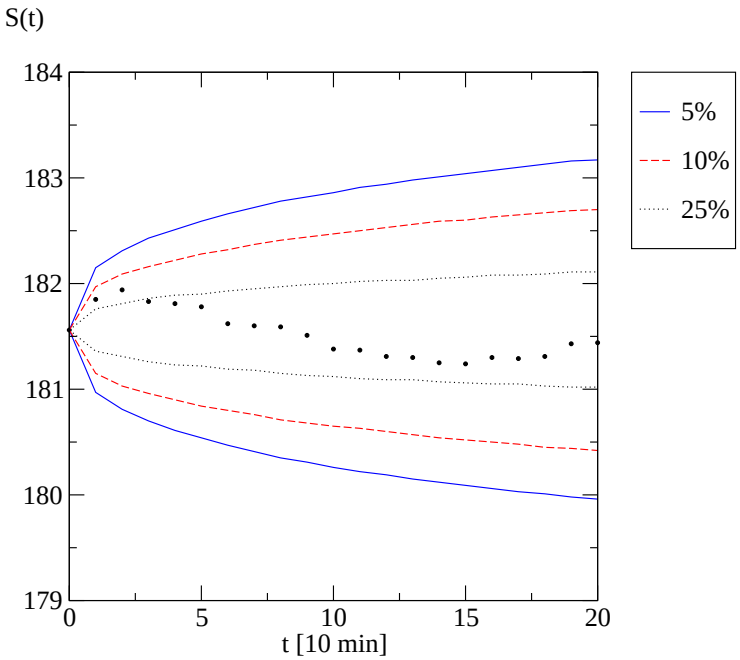


Figure 4.21: Upper and lower expected index values for the first day of the S&P500 dataset (September 30th, 1985), compared to real prices (in circles)

inhomogeneity enclosed in the a_t's:

$$P_t(x) \equiv P(x_t) = \int_0^\infty \rho(\sigma) \frac{e^{-\frac{x^2}{2\sigma^2 a_t^2}}}{\sqrt{2\pi\sigma^2 a_t^2}} d\sigma \quad , \quad a_t \equiv \left[t^{2D} - (t-1)^{2D}\right]^{1/2}.$$

The values of the expected limit returns x_t^{min}, $t = 1, \ldots, 20$ for the lower quantiles q (for $q = 5\%$ (0.05), 10% (0.1), 25% (0.25)) are then easily obtained by numerically solving the following equation with respect to x_t^{min}

$$q = \int_{-\infty}^{x_t^{min}} dx \int_0^\infty \rho(\sigma) \frac{e^{-\frac{x^2}{2\sigma^2 a_t^2}}}{\sqrt{2\pi\sigma^2 a_t^2}} d\sigma. \tag{4.29}$$

Due to the symmetry of the employed scaling function, the corresponding values x_t^{max} are simply obtained via sign flip: $x_t^{max} = -x_t^{min}$.

The upper ($S_t^{q,max}$) and lower($S_t^{q,min}$) expected values of $S(t)$ can be easily calculated, because $S(t)$ is an increasing function of x_t and so the quantiles of S are directly related to the quantiles of x. Summarizing, for every choice of q, for every time t from 1 to 20, two price values are obtained. With probability $1 - 2q$ the price at time t is placed between these values $S_t^{q,min}$ and $S_t^{q,max}$. As an illustration, in Fig. 4.21 the results of the *in sample* analysis for the first day of the dataset (September 30th, 1985) are shown.

The comparison between these extreme values and the actual real market price can lead us to the definition of a buy or sell strategy.

The empirical analysis considers both *in sample* and *out of sample* cases. The difference between the two is that for the former a unique PDF is used, calibrated within the 25 years' dataset, while in the latter case the scaling function is calibrated every year taking into account the previous 15 years' history. For both approaches we follow the aforementioned calibration and forecasting procedure.

4.2.2.2 Conditional trading

More interestingly, exploiting the non-Markovian character of our model, we can use the value of, say, the first return of the day to condition the subsequent expected evolution of the index. Using our model prediction for the joint PDF of the returns [Eq. (4.2)] we thus can build a conditioned scheme where the first return is used to condition the successive probability densities:

$$P^{(t|1)}(x_2,\ldots,x_t|x_1) \equiv \frac{P^{(t)}(x_1,x_2,\ldots,x_t)}{P^{(1)}(x_1)} = \frac{\int_0^\infty d\sigma\rho(\sigma)\prod_{i=2}^{t}\frac{\exp\left(-\frac{x_i^2}{2\sigma^2 a_i^2}\right)}{\sqrt{2\pi\sigma^2 a_i^2}}}{\int_0^\infty d\sigma\rho(\sigma)\frac{\exp\left(-\frac{x_1^2}{2\sigma^2 a_1^2}\right)}{\sqrt{2\pi\sigma^2 a_1^2}}} \quad , \quad t\leq 20. \tag{4.30}$$

More generally, we can write the following expression, conditioning to the first s returns the PDF of the successive $t-s$ returns, with $t\leq 20$ and $s\leq t$:

$$P^{(t|s)}(x_{s+1},\ldots,x_t|x_1,\ldots x_s) \equiv \frac{P^{(t)}(x_1,x_2,\ldots,x_t)}{P^{(s)}(x_1,\ldots,x_s)} = \frac{\int_0^\infty d\sigma\rho(\sigma)\prod_{i=s+1}^{t}\frac{\exp\left(-\frac{x_i^2}{2\sigma^2 a_i^2}\right)}{\sqrt{2\pi\sigma^2 a_i^2}}}{\int_0^\infty d\sigma\rho(\sigma)\prod_{k=1}^{s}\frac{\exp\left(-\frac{x_k^2}{2\sigma^2 a_k^2}\right)}{\sqrt{2\pi\sigma^2 a_k^2}}}. \tag{4.31}$$

In Fig. 4.22 some interesting features of the applied model can be detected. In particular, it is clear the influence of the nearest past returns to determine the expected amplitude of the next ones. Conditional trading will thus be based on bounds which are not fixed within the day but may vary according to the new information included in the most recent market prices.

4.2.3 Developing and applying intraday strategies

Denoting as t the time in days, and as t_k the period k within day t, the analysis described above permits the detection of all those times t_k when the index value $S(t_k)$

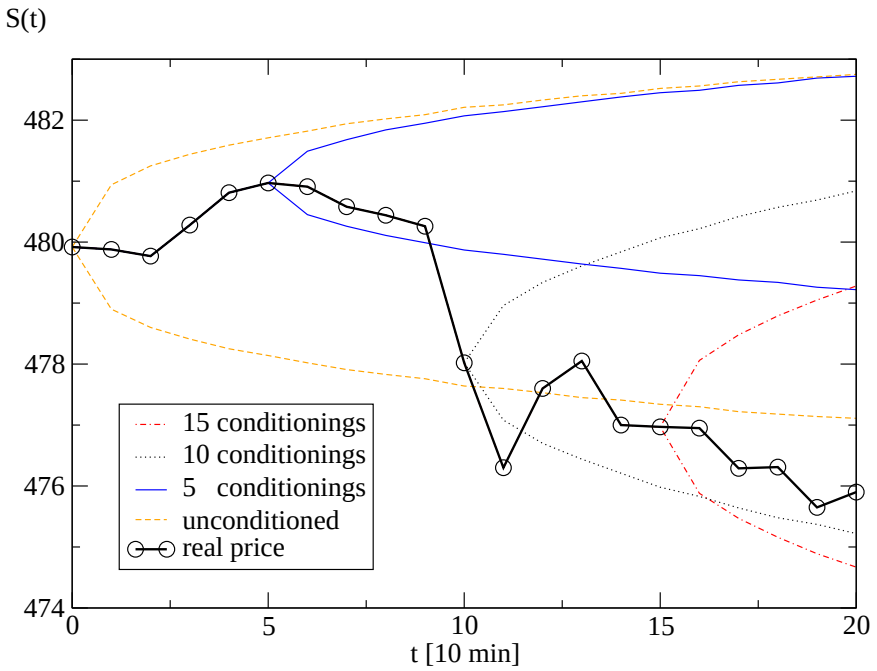


Figure 4.22: Upper and lower expected index values [q=10%] for one random day of the dataset (February 4th, 1994). Different number of conditioning returns are considered.

brakes through the expected upper ($S_{t_k}^{q,max}$) or lower ($S_{t_k}^{q,min}$) quantile[1]. This quantile violation can be used to define a trading strategy. Buy and sell signals will be identified comparing the observed stock price at time t_k with the quantiles at level q. Note that the ensemble property we observe lasts from 9:30 a.m. to 1 p.m. As a consequence, the trading strategy we consider operates within this time lapse. Furthermore, to avoid the impact of news arrivals between 1 p.m. and the opening of the next day, the trading strategy opens and closes positions within the day. As we mentioned when describing the dataset we are using, the time interval we consider is 10 minutes and the model we propose provides a sort of price density forecasts given (conditioning on) the opening price of a given day. In addition, as the model works using returns, the density forecasts (and therefore the quantiles at level q) will be available from 9:40 a.m.

Within a given day, given the quantiles at level q, the trading signals and the trading activity are defined as follows:

A: If there are no open positions

[1] We recall that quantiles for time t_k have been simulated using the proposed model and conditioning on some information set within the day (see the previous section; a discussion on this aspect will follow in this empirical analysis).

A.i: Buy if $S(t_k) > S_{t_k}^{q,max}$

A.ii: Sell if $S(t_k) < S_{t_k}^{q,min}$

B: If there are open positions

B.i: Close a long position if $S(t_k) < S_{t_k}^{q,max}$

B.ii: Close a short position if $S(t_k) > S_{t_k}^{q,min}$

B.iii: Close long or short positions if they are still in place at 1 p.m.

By construction, multiple trades are possible within volatile days. Differently, in trending days, single operations will take place, while during stable days no positions will be taken.

Our purpose is to monitor the performances of the trading strategy defined above. Therefore, we simulate the evolution of a trader using our strategy and having an initial cash amount equal to 1 million. Given the previous comments on the ensemble property and its link with the trading strategy, at the beginning of the day and at 1 p.m. the simulated portfolio is entirely composed by cash. Furthermore, in order to avoid losses larger than the portfolio value when implementing short positions, we limit the investment value to 90% of the overall portfolio value[1]. For symmetry, we apply the same rule also for long positions. As a result, when trades are created, 10% of the portfolio remains in cash. Once a signal is observed, the trade is executed at the price $S(t_k)$.

Three different quantile levels will be considered: 5%, 10%, and 25%. Furthermore, given the discussion in the previous section, we simulated both unconditional and conditional quantiles for all the three levels previously defined. When conditioning will be considered, we produce quantiles using the opening price (no conditioning), or conditioning on the first 3, 6 or 9 returns. Higher numbers of conditioning points are not used given that, within the day, the model proposed works only from the opening up to 1 p.m at a 10-minute frequency, thus with a total of 21 observations per day.

We first analyse the trading strategy in-sample, in order to evaluate its abilities in terms of yearly profits and average return by trade (in basis points). At this stage, we also verify the impact of the different number of conditioning returns.

[1]We do not take into account the margins generally required when creating short position. We motivate this choice by the need of evaluating the strategy abilities on both long and short trades without penalizing short positions, as would be the case when margins larger than 10% would be required. Furthermore, given that the trades will last at maximum for 4 hours and 20 minutes (from 9:40 to 1 p.m.), we believe that an implicit margin of 10% will be sufficient.

Table 4.1: Average profit per trade in basis points (1985-2010)

	25%	10%	5%
All trades			
0	2.619	1.602	0.590
3	3.788	2.842	1.862
6	3.449	3.027	2.313
9	3.243	3.070	2.184
Long trades			
0	1.599	1.180	-0.483
3	3.409	2.748	1.840
6	3.094	2.942	2.564
9	2.529	2.652	1.886
Short trades			
0	3.577	1.977	1.470
3	4.173	2.932	1.882
6	3.824	3.110	2.087
9	3.996	3.469	2.455

The first column reports the conditioning elements (0 stands for no conditioning). The first row reports the quantile level.

Table (4.1) reports the average return of the trades generated by the strategy during the range from October 1985 to October 2010. Profits are indicated as average basis points per trade, and distinguishing also between long and short trades. This allows verifying if the strategy better identifies signals in a specific direction. Three elements clearly appear. At first, the profit decreases with decreasing quantile values, as if stronger signals provide smaller performances (such an evidence is not related to the trade sign). We explain such an unexpected result by the fact that trades are valued at the price $S(t_k)$ and not at the price at which a trade could have been executed once the quantile violation had been observed. This inconsistency (which cannot be easily solved) reduces the potential profit coming from relevant (and unexpected) price movements. Those, in fact, are the price changes detected by the model-based quantiles at the 5% level.

The second relevant comment we make refers to the relation between average profit and conditioning information. We observe that a conditioning on the first returns of the day increases the profit. Such a result holds independently of the trade sign. This is somewhat expected as during the first part of the day the model adapts its behaviour to the most recent data. Therefore, it provides a better fit to the ensemble property within the remaining part of the day. Considering the number of conditioning, we note that the relation between conditioning points and profits is not linear, and has a maximum between one and six conditioning points. Considering this result and the previous comment on the quantile/profit relation, we decided to use in the out-of-sample analysis just three conditioning points (value with maximum profit for 25% quantiles).

Table (4.2) reports the same quantities as Table (4.1), but by year (we drop 1985 and 2010 where only part of the year was available). In general, the use of a 25% quantile provides the higher average profit. Furthermore, we observe that the largest part of negative values are identified in the range 2004-2007, where the market was clearly upward trending and in a low volatility phase. This is a further expected outcome, since the model detects violations which are associated with large movements. The largest profits are located during 1986 and 1987, in a high volatility period.

Overall, the average profit per trade is relatively small, if not negative. A first not completely satisfactory result. Nevertheless, we observe that during some specific market phases, the trading strategy provides large average profits per trade, see for instance the 90's.

Beside the impact of conditioning and quantile level, a further element is of relevant interest: the number of trades created by the strategy in a given time interval. Table (4.3) reports several elements for the most recent years. The first column just repeats

Table 4.2: Average profit per trade in basis points (yearly values)

Year	Unconditional			Conditional 3 points			Conditional 6 points		
Quantile	25%	10%	5%	25%	10%	5%	25%	10%	5%
1986	8.31	6.06	6.02	11.39	9.95	9.96	11.84	12.21	12.13
1987	12.32	11.87	12.79	17.14	15.50	13.60	15.04	11.31	12.11
1988	3.26	1.98	1.24	7.34	6.00	4.02	8.83	8.37	7.43
1989	5.11	2.69	-1.33	6.88	5.87	7.83	7.26	5.95	4.02
1990	9.00	8.20	5.58	11.11	10.19	5.64	10.90	7.84	3.62
1991	4.46	1.13	1.58	6.35	5.65	2.55	7.54	7.48	4.12
1992	3.16	-0.25	-3.76	4.45	5.29	7.71	3.62	3.78	3.71
1993	1.54	-0.30	-2.39	3.38	4.51	2.87	3.15	4.04	3.80
1994	1.90	-0.13	-3.79	2.76	1.48	-0.89	3.77	2.10	1.18
1995	0.08	-2.36	-11.08	2.31	1.46	-0.26	2.14	1.23	0.13
1996	0.84	-1.63	-2.36	2.16	0.89	-1.28	3.19	2.37	0.28
1997	2.62	1.33	1.48	4.85	3.90	2.90	3.81	3.49	1.51
1998	3.37	3.49	0.37	4.73	0.35	-0.76	3.91	3.96	4.24
1999	1.48	0.84	-2.02	2.82	2.57	0.79	3.18	6.09	4.98
2000	0.02	-2.81	-0.88	2.78	0.81	1.72	0.32	1.14	3.74
2001	0.60	0.29	0.29	2.69	1.64	2.23	0.03	-2.14	-2.33
2002	2.65	0.28	0.35	3.49	0.17	-0.62	-0.55	-0.77	-2.64
2003	1.81	0.38	2.29	0.40	-0.30	-1.96	0.99	-0.51	-0.52
2004	-1.26	-1.66	-1.74	-0.38	-0.37	0.16	0.02	0.90	0.47
2005	-1.91	-2.06	0.52	-1.23	-0.25	0.98	-0.58	-1.36	-1.22
2006	0.38	-0.34	-1.20	-0.68	-1.63	-1.70	0.21	-0.52	-2.29
2007	-1.68	-4.88	-9.35	-1.55	-3.42	-4.99	-0.91	-2.49	-3.36
2008	4.06	2.37	-1.73	4.37	3.43	-0.50	3.03	-1.08	-2.16
2009	2.20	4.23	2.88	3.48	2.19	2.63	2.49	3.88	2.85

the content of Table (4.2), while the following columns separately consider Long and Short trades, and for those distinguish between true and false signals, where the former are the signals really providing a positive trade profit. Note that in the trading simulations we execute trades at the 10 minute frequency; the distinction between true and false signals might not be optimal. In fact, executing trades few seconds after the violation of quantiles would improve the model performances, thus potentially increasing the number of true signals. In light of this comment, the relative small number of true signals leading to positive trade profits [see the last two columns of Table (4.3)] would increase and should not be considered as a negative outcome. A further element supporting the strategy is the average profit of trades where the signals are considered as true. In fact, both for long and short trades, average profits are sensibly larger, peaking around 80 basis points (on average). Furthermore, false signals lead to trades with small losses (compared to the gains). We also note that the losses due to false signals are larger during volatile market phases, as in 2008 and 2009 (and this holds for both long and short trades).

With respect to the number of trades, this is much larger when using 25% quantiles compared to 5% quantiles, while there are small differences between the use of Conditional and Unconditional quantiles. Long and short trades are almost numerically equivalent and there are no differences between trending and volatile market phases. Overall, the number of trades, irrespectively of the sign, increases during volatile periods, an expected result. Finally, as we previously anticipated, the number of true signals is relatively small, and there are again not many differences with respect to the trade sign.

Overall, Table (4.3) show evidence of some potential interest in the strategy, since the average profit for true signals is quite elevate (in particular compared to the overall average profit).

While the previous tables where focusing on the return over single trades (on average), Table (4.4) focuses on the overall profit of the strategy over single years (assuming a starting cash amount of 1 million). The returns are reported in percentages, and show evidence of positive performances in most periods. Comparing first the Conditional versus Unconditional quantiles, we observe that conditional modeling is clearly better: its returns are higher apart in few cases and its the standard deviation is smaller[1]. Contrasting the 25% and 5% quantiles, the use of narrower bands for the identification of

[1]The standard deviation is computed over the daily returns of the simulated portfolio and then annualized. Note that days when trades are not created will provide a zero return since we did not assume any remuneration for the bank account.

the signals provide larger returns over the years. This potentially exposes the portfolio to a number of trades generated by false signals, but the profits coming from true signals will balance them. Such a result holds irrespectively of the conditioning type.

Finally, if we compare the performances of the trading strategy (25% quantiles) to that of the underlying equity index [see the last two colums of Table (4.4)], we note a relevant positive result: when the market is experiencing high volatility, our strategy provides positive returns with a volatility smaller than that of the market, and this is particularly evident when the market has yearly negative returns; on the contrary, when the market is in a low volatility, our strategy has small or negative returns; finally, our strategy has always a volatility smaller than that of the market. This finding suggests that our strategy could be used to hedge the market volatility, since it provides positive returns in case of high market volatility, and with smaller risk. This is further confirmed by the correlation between market returns and standard deviation with the corresponding values of our strategy [see the last row of Table (4.4)]: positive and very high correlation between standard deviations, and low negative correlation between returns.

The promising in-sample performances of our strategy are confirmed in the out-of-sample results. In this further evaluation, we compare our model to a more traditional approach based on GARCH models [see (22) and (14)]. Moving from the point of view of ensembles toward that of financial time series, several elements characterizing high frequency data have to be considered. First of all, and most relevant, the entire daily set of high frequency returns will be used, and not just the range 9:30 a.m. to 1 p.m. Second, the periodic behaviour of the intra-daily volatility is considered [see (4), among others]. To capture this element, together with variance asymmetry, we consider as a competing model an asymmetric GARCH, the GJR [see (25)] with a periodic deterministic variance component. Our choice is motivated by the relative simplicity of the competitor, a kind of benchmark, and by the possibility of easily generating from that model density forecasts at a given quantile under a distributional assumption for the model innovations. The competing model is given as follow:

- the empirical returns on a 10-minute time scale are represented as: $x_t^l = m_t^l \, \epsilon_t^l$, where t identifies the 10-minute period within day l with a range which is now $t = 1, 2, ...T$ ($T = 39$ for our dataset), m_t^l is a deterministic periodic function, and ϵ_t^l is the stochastic component; this model implies that returns are generated as $X_t^l = \mathbb{N}\left(0, m_t^l \, \mathbb{V}_\epsilon\right)$, where $\mathbb{N}(\mu, \sigma)$ indicates a Gaussian random variable of mean μ and variance σ, and $\mathbb{V}_\epsilon$ is the (stochastic) variance of the random component;

Table 4.3: Average profit per trade and number of trades: long/short trades, false/true signals

	Average profit (bp)							Number of trades								
	All	Long			Short			All	Long			Short			% True	
		All	True	False	All	True	False		All	True	False	All	True	False	Long	Short
25% - Conditional 3 points																
2005	-1.23	-0.10	21.43	-6.76	-2.24	15.63	-7.15	422	199	47	152	223	48	175	23.6%	21.5%
2006	-0.68	-0.58	15.16	-6.28	-0.76	15.29	-5.65	438	207	55	152	231	54	177	26.6%	23.4%
2007	-1.55	-4.01	13.05	-9.36	1.73	29.41	-9.19	462	264	63	201	198	56	142	23.9%	28.3%
2008	4.37	-0.67	51.60	-20.27	9.90	66.53	-18.77	505	264	72	192	241	81	160	27.3%	33.6%
2009	3.48	2.57	36.17	-12.87	4.42	53.20	-16.31	463	235	74	161	228	68	160	31.5%	29.8%
5% - Conditional 3 points																
2005	0.98	2.36	25.96	-3.26	-0.63	9.69	-1.92	97	52	10	42	45	5	40	19.2%	11.1%
2006	-1.70	-0.91	15.26	-4.69	-2.25	27.18	-6.01	90	37	7	30	53	6	47	18.9%	11.3%
2007	-4.99	-6.31	17.56	-6.71	-3.86	11.24	-8.34	130	60	1	59	70	16	54	1.7%	22.9%
2008	-0.50	-3.59	31.85	-16.55	2.61	61.09	-19.88	253	127	34	93	126	35	91	26.8%	27.8%
2009	2.63	2.68	53.40	-10.21	2.59	53.19	-13.78	169	79	16	63	90	22	68	20.3%	24.4%
25% - Unconditional																
2005	-1.91	0.36	21.13	-7.61	-3.48	15.40	-8.67	407	166	46	120	241	52	172	27.7%	21.6%
2006	0.38	0.47	19.63	-7.50	0.30	26.70	-7.78	369	177	52	125	192	45	131	29.4%	23.4%
2007	-1.68	-3.21	16.58	-9.91	-0.13	35.74	-11.42	479	241	61	180	238	57	169	25.3%	23.9%
2008	4.06	-4.02	57.18	-28.56	13.19	71.86	-21.66	560	297	85	212	263	98	158	28.6%	37.3%
2009	2.20	0.01	50.94	-23.19	4.41	55.57	-19.72	521	262	82	180	259	83	167	31.3%	32.0%
5% - Unconditional																
2005	0.52	-0.75	14.39	-6.25	2.64	26.78	-0.38	24	15	4	11	9	1	8	26.7%	11.1%
2006	-1.20	-1.08	22.61	-4.82	-1.33	22.42	-9.81	41	22	3	19	19	5	14	13.6%	26.3%
2007	-9.35	-9.69	38.70	-12.62	-9.17	18.42	-16.98	103	35	2	33	68	15	53	5.7%	22.1%
2008	-1.73	-8.57	58.81	-31.19	5.59	83.91	-28.44	362	187	47	140	175	53	122	25.1%	30.3%
2009	2.88	5.61	62.38	-14.47	0.71	53.57	-18.05	302	134	35	99	168	44	124	26.1%	26.2%

Table 4.4: In sample yearly return and standard deviation compared with the S&P500 Index

	25% - C. 3 p.		5% - C. 3 p.		25% - Unc.		5% - Unc.		S&P500	
	Return	Dev.st	Return	Dev.st	Return	Dev.st	Return	Dev.st	Return	Dev.st
1986	40.47	4.88	11.36	2.98	33.53	5.23	5.06	2.99	14.620	14.635
1987	79.71	7.78	21.58	6.31	61.65	12.98	24.05	11.28	2.028	32.013
1988	26.55	4.41	5.33	2.47	12.14	4.99	0.97	2.68	12.401	17.019
1989	23.91	3.81	7.67	2.55	16.45	4.09	-0.88	1.82	27.250	13.006
1990	47.37	4.98	9.65	3.19	40.56	6.21	7.07	3.83	-6.559	15.886
1991	23.62	4.20	3.43	2.18	17.52	4.86	1.33	2.37	26.307	14.242
1992	14.84	3.26	5.99	1.68	9.97	3.29	-1.82	1.11	4.464	9.644
1993	10.39	2.76	2.27	1.65	4.06	2.96	-0.76	1.27	7.055	8.567
1994	9.55	3.04	-0.67	1.14	6.03	3.61	-1.98	1.40	-1.539	9.805
1995	7.54	2.59	-0.22	1.36	0.28	2.93	-2.80	0.99	34.111	7.776
1996	7.87	4.15	-1.38	2.14	2.97	4.68	-1.61	1.87	20.264	11.734
1997	20.57	6.02	4.86	4.38	12.49	6.30	2.06	3.45	31.008	18.059
1998	20.03	5.58	-1.40	3.26	15.79	8.29	-0.07	5.95	26.669	20.207
1999	11.98	5.77	1.53	3.13	6.65	7.40	-3.95	3.90	19.526	17.999
2000	11.47	8.30	2.48	5.05	-2.08	10.07	-2.99	6.39	-10.139	22.134
2001	11.63	7.07	4.26	4.41	1.80	9.57	0.80	6.18	-13.043	21.471
2002	15.22	8.26	-1.74	5.40	11.67	10.65	0.43	7.68	-23.366	25.926
2003	1.90	5.38	-2.91	2.80	9.45	7.79	3.99	4.80	26.380	17.000
2004	-1.61	3.56	0.20	1.74	-4.92	3.77	-0.75	1.05	8.993	11.049
2005	-4.68	3.21	0.97	1.19	-6.81	3.39	0.10	0.57	3.001	10.243
2006	-2.82	2.80	-1.42	1.15	1.22	3.70	-0.48	0.92	13.619	9.985
2007	-7.40	3.90	-6.06	1.49	-7.96	4.83	-8.67	2.39	3.530	15.928
2008	19.95	10.70	-2.23	6.99	18.47	15.45	-8.63	12.55	-38.486	40.810
2009	15.43	7.97	3.99	5.43	9.63	10.29	8.29	7.68	24.707	27.177
Corr.	-0.08	0.95	0.04	0.94	-0.08	0.98	0.15	0.98		

- m_t^l is a periodic deterministic variance modeled similarly to (5), but using dummy variables instead of harmonics; we might represent returns as

$$\ln[(X_t^l)^2] = \ln[(m_t^l)^2] + \ln[(\varepsilon_t^l)^2], \tag{4.32}$$

with

$$\ln[(m_t^l)^2] = a_1 + \sum_{j=1}^{T} a_j d_{t,j}^l, \tag{4.33}$$

where $d_{t,j}^l$, $j = 2, ...T$ is a dummy variable assuming value 1 when $j = l$ and zero otherwise, while $a_1, a_2 \dots a_T$ are parameters to be estimated;

- furthermore, the stochastic term ε_t^l follows a GJR model (Glosten et al. 1993) allowing thus for the decomposition

$$\varepsilon_t^l = \sigma_t^l \, Z_t^l, \tag{4.34}$$

where $Z_t^l = \mathbb{N}(0,1)$ and the conditional variance is given by

$$(\sigma_t^l)^2 = \omega + \Big(\alpha_0 + \alpha_1 \, I(\overline{\epsilon}_t^l < 0)\Big) \, (\overline{\epsilon}_t^l)^2 + \beta \, (\overline{\sigma}_t^l)^2, \tag{4.35}$$

where

  - $(\overline{\epsilon}_t^l)^2 \equiv (\epsilon_{t-1}^l)^2$ if $t > 1$ and $(\overline{\epsilon}_t^l)^2 \equiv (\epsilon_T^{l-1})^2$ if $t = 1$,
  - $(\overline{\sigma}_t^l)^2 \equiv (\sigma_{t-1}^l)^2$ if $t > 1$ and $(\overline{\sigma}_t^l)^2 \equiv (\sigma_T^{l-1})^2$ if $t = 1$,
  - $I(\overline{\epsilon}_t^l < 0)$ is equal to one when $\overline{\epsilon}_t^l$ is negative and zero otherwise,
  - ω, α_0, α_1 and β are parameters to be estimated. These parameters must satisfy the constraints for variance positivity and covariance stationarity $\omega > 0$, $\alpha_0 > 0$, $\alpha_1 > 0$, $\beta > 0$ and $\alpha_0 + 0.5\alpha_1 + \beta < 1$ (under an assumption of symmetry for the density characterizing Z_t^l.

The estimation of the model proceeds by steps. At first the periodic component is estimated by linear regression using equations (4.32) and (4.33). The fitted periodic component is used to recover the estimated values of ϵ_t^l. Over those, the GJR parameters are estimated by Quasi Maximum Likelihood approaches using a Gaussian likelihood. Given the estimated parameters, and under a Gaussian density for the innovations z_t^l, we generate possible paths for the future evolution of the conditional variance $(\sigma_t^l)^2$, of the innovations ϵ_t^l, and of the returns x_t^l (the periodic component m_t^l is purely deterministic and is thus simply replicated in the forecasting exercise). Under the distributional hypothesis, the needed quantiles are then determined and used as an alternative input for the identification of the trading signals.

Table 4.5: Out of sample average profit per trade in basis points (2000-2010)

	25%	10%	5%
	All trades		
Unc.	0.772	-0.058	-0.267
Cond. (3)	1.565	0.465	-0.098
GARCH	-1.488	-3.207	-0.628
	Long trades		
Unc.	-1.131	-0.854	-1.260
Cond. (3)	0.962	0.250	-0.513
GARCH	-2.812	-3.950	-0.546
	Short trades		
Unc.	2.597	0.661	0.613
Cond. (3)	2.145	0.655	0.284
GARCH	-0.489	-2.594	-0.696

The first column reports the model: Unc. stands for our model with no conditioning; Cond. (3) refers to the model with a conditioning to the first three returns of the day; finally, GARCH identifies the conditional variance specification with deterministic periodic component. The first row reports the quantile level.

In Table 4.5 we report the out-of-sample average profit per trade considering our Unconditional and Conditional quantiles as well as the GJR model. As we mentioned in the previous section, the out-of-sample evaluation focuses only of the range 2000 to 2010, since the period 1985-1999 is used to calibrate the ensemble-based model. Results for our model are similar to the in-sample outcomes, with conditional modeling providing better results. Furthermore, both Conditional and Unconditional model specifications have performances largely better than the GRJ model. The difference between the model-based strategies appear more clearly in Table 4.6 that contains annual returns of the simulated portfolios. We note here that using the 25% quantiles together with a conditioning on the first three returns of the day provides the best results in high volatility market phases (larger than 20% annualized daily market volatility). On the contrary, when the volatility is lower, there is not a clear preference across the models (GJR included). The yearly volatility of the GJR is lower compared to our models as well as to the market, but the yearly profits are clearly unsatisfactory. In addition, the last row reports the correlations between the strategies return and standard deviations with the corresponding market quantities: the results confirm the previous findings for our strategy, while for the GJR we have lower correlation in the standard deviations.

Table 4.6: Out of sample yearly return and standard deviation compared with the S&P500 Index

	25% - C. 3 p.		5% - C. 3 p.		25% - Unc.		5% - Unc.		25% - GJR		5% - GJR		S&P500	
	Return	Dev.st	Return	Dev.st	Return	Dev.st	Return	Dev.st	Return	Dev.st	Return	Dev.st	Return	Dev.st
2000	12.96	8.36	-1.29	5.31	-2.15	10.09	-3.46	6.56	-10.88	4.88	-1.95	2.46	-10.14	22.13
2001	12.61	7.19	5.28	4.35	2.15	9.49	0.49	6.28	1.75	5.62	0.52	3.00	-13.04	21.47
2002	16.21	8.49	-0.93	5.53	11.03	10.76	0.69	7.68	-7.83	5.54	2.29	3.39	-23.37	25.93
2003	2.93	5.35	-0.61	2.95	8.43	7.63	4.13	4.69	9.39	4.94	-1.27	2.17	26.38	17.00
2004	-1.08	3.52	-0.69	1.60	-6.60	3.66	-0.66	0.99	0.70	1.69	-0.44	0.44	8.99	11.05
2005	-5.68	3.00	0.88	1.13	-6.80	3.27	0.06	0.52	-1.86	1.38	0.14	0.25	3.00	10.24
2006	-3.08	2.61	-0.99	1.04	0.79	3.56	-0.84	0.85	3.11	1.90	0.09	0.61	13.62	9.99
2007	-9.25	3.91	-4.96	1.49	-10.39	4.79	-8.71	2.22	-4.27	1.61	-0.98	0.45	3.53	15.93
2008	22.73	10.66	-1.90	7.70	18.30	15.46	-8.43	12.70	-7.02	4.22	-1.49	1.38	-38.49	40.81
2009	18.36	8.17	1.74	5.41	12.31	10.15	8.05	7.50	3.99	4.91	-0.78	1.33	24.71	27.18
Corr.	-0.52	0.95	0.06	0.96	-0.30	0.97	0.62	0.98	0.78	0.64	-0.19	0.45		

Summarizing our findings, we can state that the proposed model has some potential for the development of trading strategies aimed at hedging the volatility risk, since their performances are positive during high market volatility, and characterized by a lower risk compared to the market index. Signals extracted from the model can be directly used, or could be considered as confirmatory signals for other strategies working with high frequency data, or could be used to detect relevant market movements.

In this empirical example we do not consider several elements that could have an impact on the trading strategy profits. We motivate this by the need of evaluating the model in comparison to a simple benchmark. Across the elements we did not include, we have the trading costs. Once those are introduced, the profits reported in the previous tables would be sensibly reduced. However, the trading strategy we implement was based on a fixed frequency database, using a 10-minute interval. This has a relevant impact on the trading outcomes. In fact, if a quantile violation is observed at time t_k, we execute the trade at time t_k with the price observe at that point in time. However, the violation could have taken place in each instant between time t_{k-1} and t_k. A trader using our approach would produce quantiles to be used for each period of 10 minutes, but will immediately detect the violation, and operate in the market soon after the violation (assuming she fully trusts the signal). On the contrary, working with a fixed time span of 10 minutes, we lose part of the potentially relevant content of the signal, since the price at time t_k might be significantly different from the unknown price observed at the trade execution just after the violation occurred between time t_{k-1} and t_k. A further element not included in our trading example is the remuneration of the bank account. In addition, overnight liquidity operations could be introduced given that the portfolio is entirely into cash from 1 p.m. of day t up to 9:39 a.m. of day $t+1$. Finally, we note that even the trading strategy could be improved, for instance introducing stop-loss and take-profit bounds on the implemented orders.

4.3 Testing the intraday model with stock quotes returns

In this section we employ the intraday model to analyze the behaviour of two stock quotes: International Business Machines (NYSE: IBM) and Caterpillar, Inc. Common Stock (NYSE: CAT). We will note that the model reproduces, within certain limits, both return statistics and we will particularly highlight some differences between the two quotes, that are interesting for matters of scaling and multiscaling.

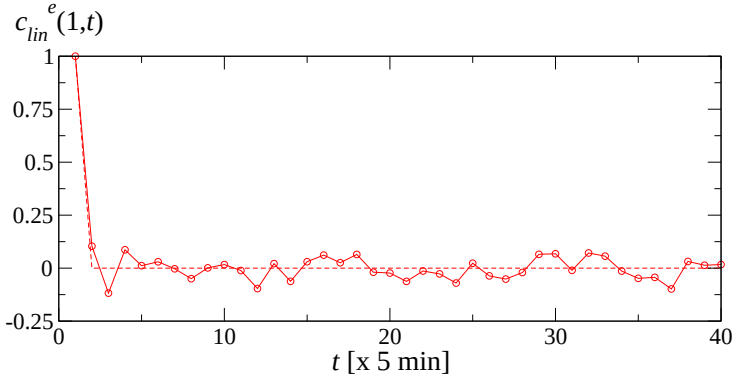


Figure 4.23: Linear correlator $c^e_{lin}(1,t)$ for the IBM dataset. Dashed line is the model prediction. Note that the quick decay to zero permits to employ the model in a 5-minutes time scale.

4.3.1 The IBM stock quote

Looking at the rapid vanishing of the linear correlator $c^e_{lin}(1,t)$ (Eq. 4.5) of the returns of the IBM stock quote in Fig. 4.23, we immediately understand that the rapid decay to zero permits to assert that, for the period in analysis, in a 5-minutes time scale, the correlations are absent, and then the model can be applied to such a time based return scale. The dataset of returns for the IBM quote consists of 1809 days, with 40 returns per day from 9:40 a.m. to 00:55 p.m. NY time.

The decay of the second moment in the considered time window [see Fig. 4.24(b)] follows the same decaying trend as in the S&P500 dataset, with an estimated scaling exponent D around 0.32. On the other hand, the analysis of the trend of the moments [Fig. 4.24(a)], for low values of the moment q, shows that the scaling ansatz of Eq. 4.1 is valid for the IBM datset, with a best fitted $D = 0.31$. The two values are in quite good agreement, and this permit us to deepen the analysis following the above mentioned procedure.

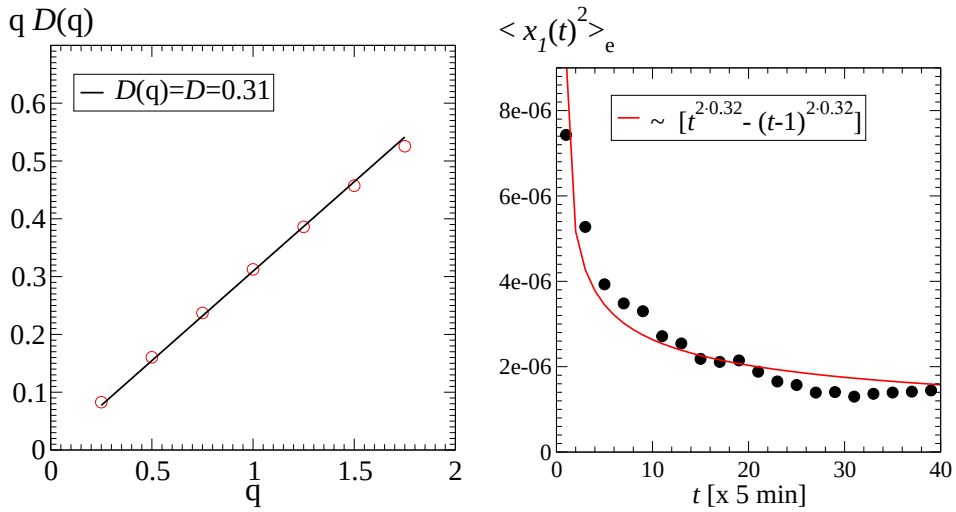
First of all we build the scaling function, by collapsing the aggregated returns via Eq. 4.1. The presence of a certain skewness suggests to use the asymmetric model presented in Sec. 2.4 with a ψ function chosen *e.g.* as follows:

$$\psi(\sigma,\mu) = \frac{1}{2}\rho_1(\sigma)\delta(\mu-\overline{\mu}) + \frac{1}{2}\rho_2(\sigma)\delta(\mu+\overline{\mu}), \tag{4.36}$$

with

$$\rho_i(\sigma) = A_i \, \frac{b_i^3\sigma^4}{b_i^7+\sigma^7} \tag{4.37}$$

and $\overline{\mu} = 6\cdot 10^{-4}$, $b_1 = 0.0012$, $b_2 = 0.00125$. A_i's are two proper normalization factors. The resulting scaling function, together with the collapsed distributions for different



(a) (b)

Figure 4.24: Empirical ensemble analysis of the returns for the IBM dataset. There is a certain agreement between the analysis in (a) performed according to the ansatz in Eq. 4.1, where the straight line characterizes a simple-scaling behavior with a best fitted $D = 0.31$ and the trend of the second moment in (b) where the line is given by $\sim \left[t^{2D} - (t-1)^{2D}\right]$, with $D = 0.32$.

time lags, are displayed in Fig. 4.25. The fitting of the function g to the real data is remarkable, both in the center and in the tails of the distributions. The verification of the scaling hypothesis can lead us to study the behaviour of some of the correlators introduced in Sec. 4.1.1. For example, the volatility autocorrelation (Eqs. 4.14 and 4.15) displays a very good agreement between empirical and theoretical trends (Fig. 4.26). Other agreements concern the $\kappa^e_{\alpha,\beta}$ correlators (Fig. 4.27), their $\alpha-\beta$ symmetry and the $K^e_{\alpha,\beta}$ process correlators (these two last are not displayed for conciseness). Referring for example to Fig. 4.27 we must still remember that the results are significative only up to an overall moment of about 2. A clear multiscaling deviation for $q \geq 2$ is detected, and so the model validity is restricted to this value of the moments. This may explain the large oscillations in the upper graph of Fig. 4.27.

However, in the range of validity of the model, the empirical results for all the considered indicators on the IBM dataset are well predicted and reproduced. The leverage effect has been tested, too, but with non-significative results, in that the statistical error related to the finite size of the dataset does not permit an evaluation of the model estimations (the error bars are about twice as large as those of the S&P500 dataset).

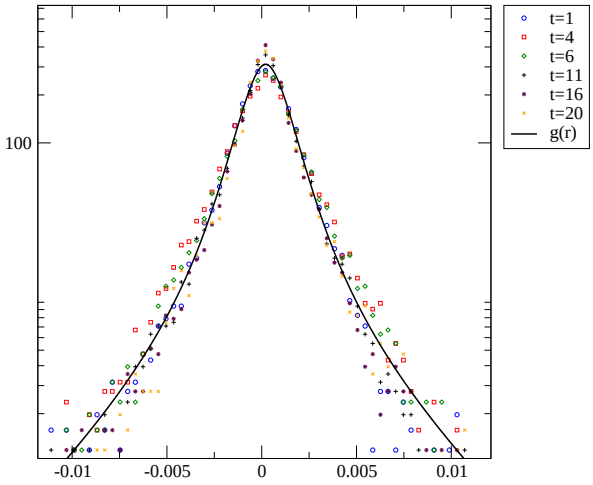


Figure 4.25: The scaling function for the IBM, after aggregation. The g function is constructed via Eq. 2.62 with a ψ function chosen in the form of Eq. 4.36.

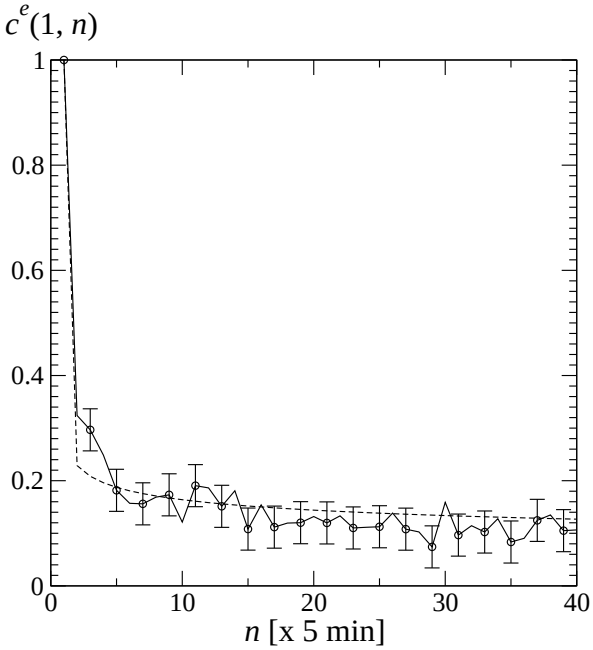


Figure 4.26: The volatility autocorrelation for the IBM. The dashed line is the model prediction.

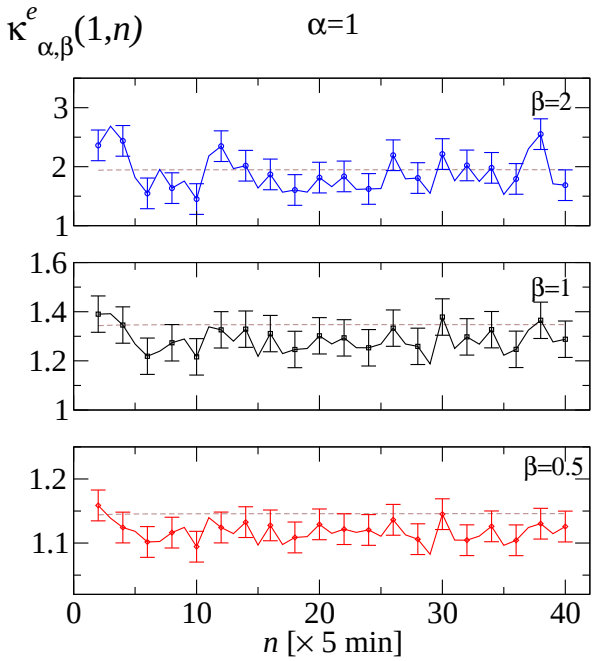


Figure 4.27: Constancy of $\kappa^e_{\alpha,\beta}$ (Eq. 4.12) for the IBM dataset. Dashed lines are model predictions (Eq. 4.11).

4.3.2 The CAT stock quote

The analysis of the CAT quote presents some relevant peculiarities, that specify the level of applicability of the intraday model. It is therefore of some interest to go into the details of the analysis for this asset, too: even if the results will be somehow poor, they all the same give a good insight into the limits of the described procedure. In Fig. 4.28 the trend of the CAT stock price for the period is shown, together with the log-returns. The graph of the trend presents two distinct regimes: approximately in the first two thirds of the graph the price is uniformly increasing and, beside some isolated spikes in the volatility, the returns are quite uniformly distributed. In the remaining part of the graph (approximately the last two years) the trend is no more clear: the price oscillates with a period of some months and, toward the end of 2008, there is a large downfall in the quotation. Correspondingly, the volatility is more clustered and, around the end of the period, significantly increased. As a consequence, we are led to ask ourselves if this pattern may influence the analysis. The first two steps of our work are, as before, the calculation of the empirical second moment $\langle x_1(t)^2\rangle_e$ and the analysis of the scaling of the moments. We work in a 10-minutes' time scale, and we have at our disposal $M = 1497$ complete days. After obtaining some preliminary, non satisfying results, we

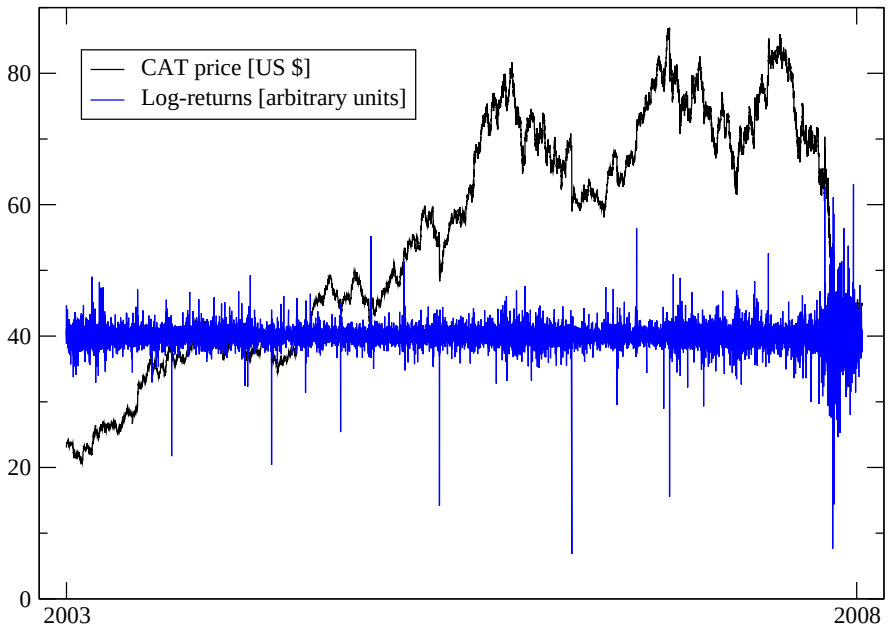


Figure 4.28: Overview of the quotations of the CAT stock, from 2003 and 2008. See the increased volatility in the more recent years (2008 crisis).

decided to split our database into three parts relative to the years 2003-4, 2005-6 and 2007-8, respectively. The trend of the empirical second moment for the three subsets is displayed, for the whole time window from 9:30 a.m. to 4 p.m., in Fig. 4.29. While the first two subsets present an analogous behaviour, the last one is completely out of that range. Even if a similar decaying trend can be identified in the first three hours from opening, the scale of the moment is different. This difference persists all day long and, around the end of the trading time, the last dataset shows a clear increase of the volatility that is not so evident for the first two. Considering instead the first three hours from opening, the scaling of the moments presents a peculiar theme (Fig. 4.30): the subsets for the first four years generate a quasi-linear, although slightly divergent, trend; on the contrary, the last subset displays a definite multiscaling effect above $q \sim 2$. Moreover, the multiscaling effect in the last two years, due to the larger values for the returns in that period, dominates the analysis jointly performed on all the available years. In fact the trend of the last subset is almost superposed to that of the complete dataset. Analogous different results are spotted considering the distribution of the returns (Fig. 4.31): again the PDF's for the first four years are similar, while they differ from that of 2007-8.

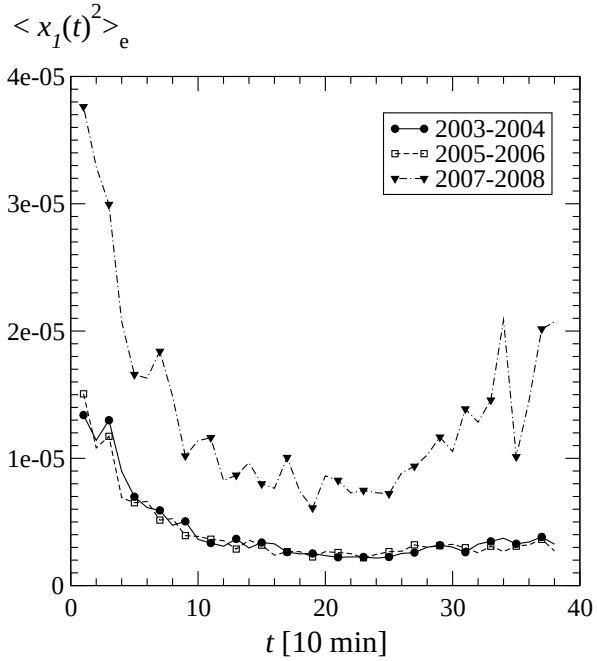


Figure 4.29: The average empirical second moment of the CAT dataset during the full time of trading. The difference between the two earlier subsets and the latest is remarkable.

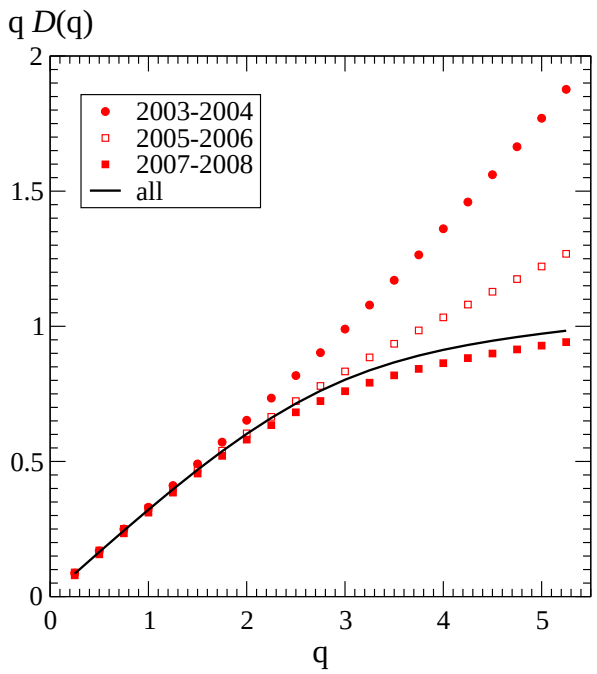


Figure 4.30: The scaling analysis of the moments according to Eq. 4.1 (CAT dataset). The black full line is relative to the whole period 2003-2008. Notice the more definite multiscaling effect in the last subset, that drives the trend for the complete dataset.

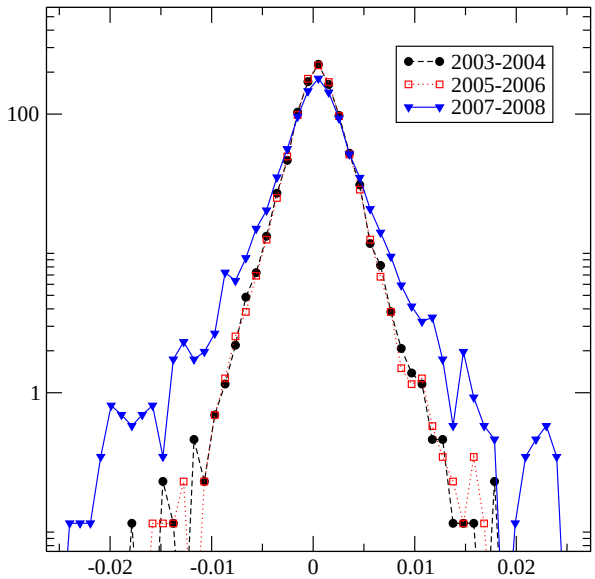
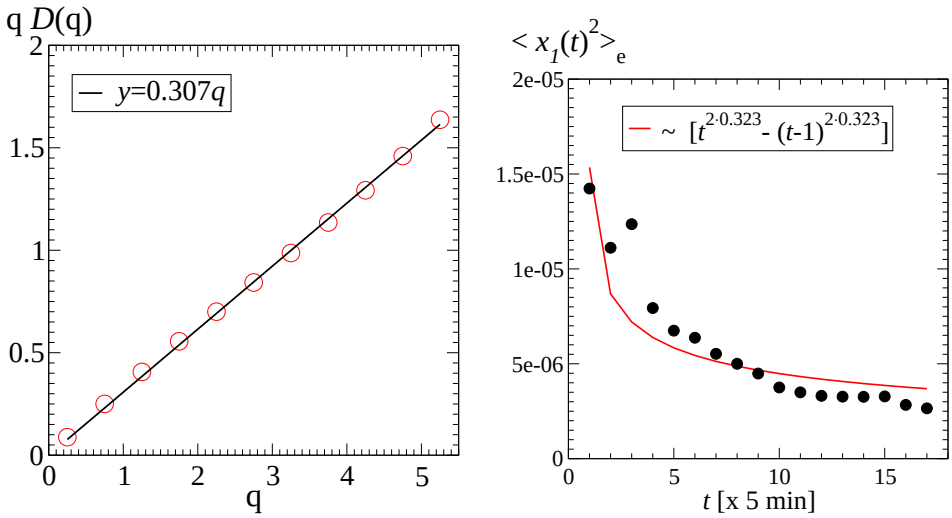


Figure 4.31: The distribution of the returns for the three CAT subsets.



(a) (b)

Figure 4.32: Empirical ensemble analysis of the returns for the CAT dataset (years 2003-2006). The analysis of the scaling of the non-linear moments (a) gives a best fitted $D = 0.307$. The trend of the second moment in (b), where the line is given by $\sim \left[t^{2D} - (t-1)^{2D}\right]$, gives a best fitted scaling exponent $D = 0.323$.

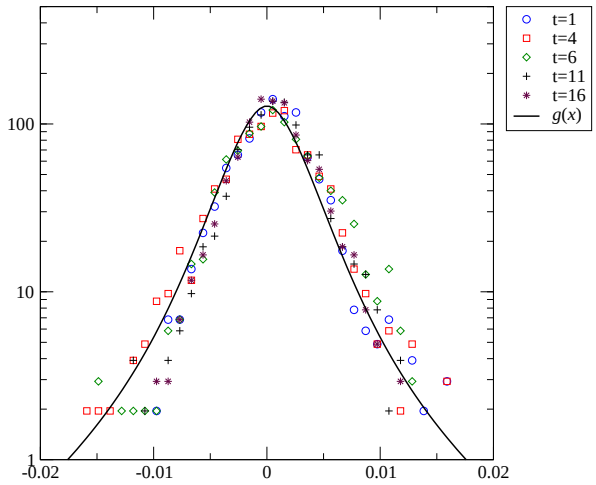


Figure 4.33: The scaling function for the CAT dataset (2003-2006) The collapse is obtained with a scaling exponent $D = 0.315$.

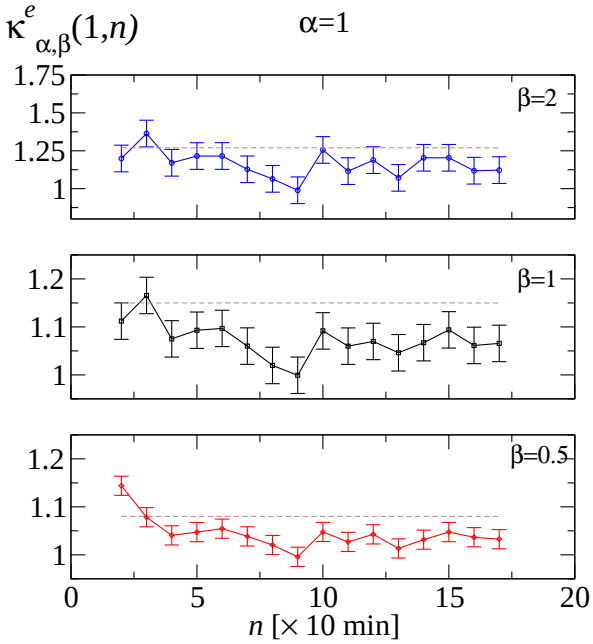


Figure 4.34: Constancy of $\kappa^e_{\alpha,\beta}$ (Eq. 4.12) for the CAT dataset. Dashed lines are model predictions (Eq. 4.11). The error bars are the standard deviation of the empirical κ's.

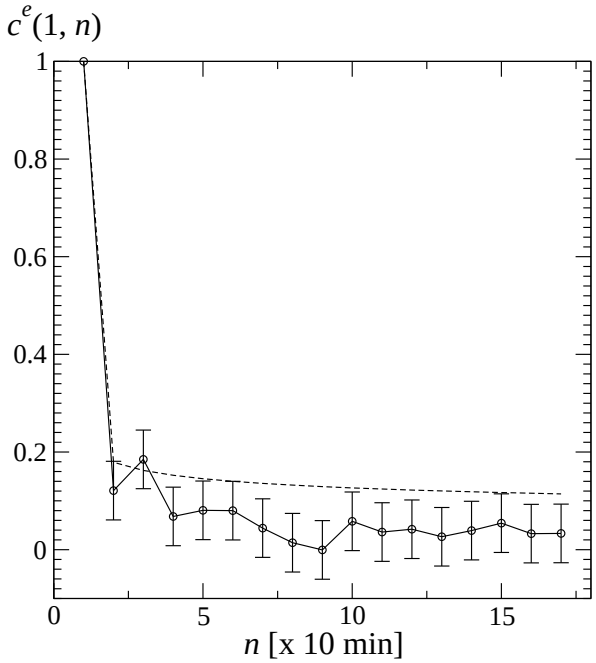


Figure 4.35: The volatility autocorrelation for the CAT. The dashed line is the model prediction. A large, not shown uncertainty comes from the possible forms of the ρ function.

Given the empirical evidence that the years 2003-2006 are to be treated under a different framework respect to the years 2007-2008, we decided to limit our analysis to the first four years, thus working with a more limited ($M = 999$ days) but more homogeneous dataset. In these years of quasi-linear increasing trend of the asset value the multiscaling effect is almost negligible [Fig. 4.32(a)]. For this reduced 4-years dataset we developed the same analysis as in Sec. 4.1.1, starting from the estimation of the scaling exponent: the two estimates derived from the graphs of Fig. 4.32 are not so distant and give, altogether, $D = 0.315 \pm 0.008$. This parameter is used to collapse the different returns histograms via Eq. 4.1: the result is presented in Fig. 4.33. Unfortunately the very reduced size of the dataset (below 1000 histories) generates a very poor scaling collapse, and consequently the scaling function $g(x)$ is not well determined. The ρ function used to build the g in this symmetric version of the model is the following

$$\rho(\sigma) = A \, \frac{\sigma^4}{0.0025^7 + \sigma^7}, \tag{4.38}$$

where the constant A is a suitable normalization factor.

As a natural consequence of the smallness of the dataset and of the uncertainty in the definition of the scaling function, the predictions about the correlators are less

significative, too. For what concerns the constancy of the $\kappa^e_{\alpha,\beta}$ correlators (Eqs. 4.11 and 4.12) the result is shown in Fig. 4.34. With respect to, *e.g.*, Fig. 4.27 the correlators are reduced and definitely more unstable.

Regarding the volatility autocorrelation (Fig. 4.35) the result is even more unsatisfactory. The model prediction is almost everywhere far from the empirical evidence, even if the error bars are larger than in the previous tests (compare, *e.g.*, with Fig. 4.26). Here it is important to stress that the error coming from the uncertainty about the form of the scaling function is large and that, if included, the overlap between the empirical and the theoretical curves would be achieved.

We limit our analysis to these correlators only; the results are already clear in that they suggest that the CAT database is too small to give good agreements. Even if also in this dataset some of the patterns already highlighted for the EUR/$ or the S&P500 or the IBM datasets are emerging, these patterns are not significative for the validation of the theory. As an important result, however, there is the confirmation that large, and possibly uniform, datasets are essentials to verify the model predictions with a good statistical significance.

5

Conclusions

Here we summarize the main achievements of our work.

The fundamental aspect is the employment of a non-Markovian model for the description of the evolution of market indexes and asset prices, both in an intra- and in an inter-day setting.

Two different formulations for the model have been described, for these two different settings. The first refers to long series of interday returns; the second is related to high frequency intraday datasets. The former, although containing elements of time-inhomogeneity that account for exogenous inputs on the market, is also a stationary model. The latter is built within the assumption of considering ensembles of histories extracted from high frequency daily data: this model is time-inhomogeneous in its range of application and this inhomogeneity is the result of the systematic exogenous effect of the market opening.

Both formulations are based on the anomalous scaling, that is thoroughly verified across different time scales for all the analyzed datasets. In the former case, the scaling property refers to the empirical PDF of the inter-day returns; in the latter, to the PDF of the returns, evaluated from the ensemble of daily histories. The model, is ultimately capable of grasping many of the most common financial stylized facts, but has also a strong predictive power.

Relevant memory effects and correlations are present in both schemes; they have been described and empirically verified.

In practice, by analyzing the interday returns we could detect, both for S&P500 and DJI, an Omori-like behaviour of the aftershocks' statistics. Nevertheless, even if the small total number of shock events in the historical series forbids stable statistics, as verified via computer simulation, the general trend of the financial counterpart of the Omori law for seismic aftershocks is well reproduced within the context of this model.

In the intraday scheme, instead, taking advantage of the predictive properties of the model, we could build a buy/sell strategy that has finally revealed some small arbitrage opportunities. The comparison with another, GARCH-based, strategy has shown the validity of he employed scheme, given that the obtained profit was overall larger. The result shows how an accurate study of the strong correlations in the process permits the foundation of a successful trading strategy, as an alternative to ill-defined *chartist* approaches.

The problem of the calibration is discussed, too. For the interday model a satisfactory procedure is now under development and the analysis of the Omori regimes has revealed itself a good calibration test, if not a good calibration tool. The intraday model, instead, already finds a good support for the calibration from the correlators' analysis, as described in detail in the text.

Further developments will necessarily regard the extension of the intraday time window where the model can be applied and, as soon as a satisfactory calibration scheme will be available also for the interday model, the tasks of defining interday strategies and derivative pricing methods.

References

[1] Admati, A. and Pfleiderer, P., A theory of intraday patterns: volume and price variability. *Review of Financia Studies,* 1988, **1**, (1), 3–40. 5

[2] Aldous, D.J., Exchangeability and related topics. *Lecture Notes in Mathematics,* 1985, **1117**, 1–198. 17, 23

[3] Allez, R. and Bouchaud, J-P., Individual and collective stock dynamics: intra-day seasonalities. *New Journal of Physics,* 2011, **13**, 25010. 5

[4] Andersen, T. and Bollerslev, T., Intraday periodicity and volatility persistence in financial markets. *Journal of empirical finance,* 1997, **4**, 115–158. 5, 96

[5] Andersen, T.G. and Bollerslev, T., Heterogeneous information arrivals and return volatility dynamics: uncovering the long run in high volatility returns. *Journal of Finance,* 1997, **LII 3**, 975–1005. 99

[6] Bachelier, L., Thorie de la spculation. *Annales Scientifiques de l'Ecole Normale Suprieure,* 1900, **3**, (17), 21–86. 1, 2, 7

[7] Baldovin, F., Bovina, D., Camana, F. and Stella, A.L., Modeling the Non-Markovian, Non-stationary Scaling Dynamics of Financial Markets. In: *Econophysics of order-driven markets,* Springer, 2011, Part III, 239–252. 23, 66, 83

[8] Baldovin, F. and Stella, A.L., Scaling and efficiency determine the irreversible evolution of a market. *PNAS,* 2007, **104**, (50), 19741–19744. 14, 22, 41, 43, 60, 61, 67, 83

[9] Baldovin, F. and Stella, A.L., Central limit theorem for anomalous scaling due to correlations. *Phys. Rev. E,* 2007, **75**, (2), 020101. 5, 20

[10] Baldovin, F., Stella, A.L. and Zamparo, M., *To be submitted,* 2012. 24, 60

[11] Bassler, K.E., McCauley, J.L. and Gunaratne, G.H., Nonstationary increments, scaling distributions, and variable diffusion processes in financial markets. *PNAS,* 2007, **104**, (44), 17287–17290. 5, 27, 33, 66, 67

[12] Bekaert, G. and Wu, G., Asymmetric volatility and risk in equity markets. *Rev. Fin. Stud.,* 2000, **13**, 1. 29

[13] Black, F. and Scholes, M., The pricing of options and corporate liabilities. *Journal of Political Economics,* 1973, **81**, 637–654. 6

[14] Bollerslev, T., Generalized autoregressive conditional heteroskedasticity. *J. Econometrics,* 1986, **31**, (3), 307–327. 12, 96

[15] Borland, L., Bouchaud, J-P., Muzy, J-F. and Zumbach, G., The Dynamics of Financial Markets – Mandelbrot's multifractal cascades, and beyond. *Working paper,* 2008, arXiv:cond-mat/0501292 1

[16] Bouchaud, J-P., Matacz, A. and Potters, M., Leverage Effect in Financial Markets: The Retarded Volatility Model. *Physical Review Letters,* 2001, **87**, (22), 228701. 5, 29

[17] Bouchaud, J-P. and Potters, M., Theory of Financial Risk and Derivative Pricing (2nd edn), Cambridge University Press, 2009. 1, 4, 6, 9, 12, 24, 49, 75

[18] Challet, D. and Peirano, P.P., The Ups and Downs of Modeling Financial Time Series with Wiener Process Mixtures. *MPRA Paper,* No. 16358, 2008, http://mpra.ub.uni-muenchen.de/16358/. 22, 23, 28

[19] Clark, P.K., A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica,* 1973, **41**, (1), 135–155. 10

[20] Cont, R., Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance,* 2001, **1**, 223–236. 1

[21] Di Matteo, T., Aste, T., and Dacorogna, M.M., Long-term memories of developed and emerging markets: using the scaling analysis to characterize their stage of development. *Journal of Banking & Finance,* 2005, **29**, (4), 827–851. 16, 33

[22] Engle, R.F., Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica,* 1982, **50**, (4), 987–1007. 11, 96

[23] Fama, E.F., Efficient capital markets: a review of theory and empirical work. *Journal of finance,* 1970, **25**, (2), 383–417. 3, 16

[24] Ghashghaie, S., Breymann, W., Peinke, J., Talkner, P. and Dodge, Y., Turbulent Cascades in Foreign Exchange Markets. *Nature,* 1998, **381**, 767–770. 6

[25] Glosten, L., Jagannathan, R. and Runkle, D., Relationship Between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks. *Journal of Finance,* 1993, **48**, 1779–1801. 12, 96

[26] Gopikrishnan, P., Meyer, M., AMaral, L.A.N. and Stanley, H.E., Inverse cubic law for the distribution of stock price variations. *Eur. Phys. J. B,* 1998, **3**, 139–140. 6

[27] Helbing, D., Ammoser, H., and Khnert, C., Disasters as Extreme Events and the Importance of Network Interactions for Disaster Response Management. In: *Extreme Events in Nature and Society,* Springer, 2006, Part III, 319–348. 40

[28] Johansen, A. and Sornette, D., Critical crashes. *Risk,* 1999, **12**, (1), 91–94. 40

[29] Lévy, P., Calcul des Probabilits. *Gauthier-Villars,* Paris, 1925. 9

[30] Lillo, F. and Mantegna, R.N., Power-law relaxation in a complex system: Omori law after a financial market crash. *Physical Review E,* 2003, **68**, 016119 1–5. 4, 41, 43, 61

[31] Mandelbrot, B.B., The variation of certain speculative prices. *Journal of Business,* 1963, **36**, 394–419. 2, 6, 8

[32] Mandelbrot, B.B. and van Ness, J.W., Fractional Brownian Motions, Fractional Noises and Applications. *The SIAM Review,* 1968, **10**, (4), 422–437. 10

[33] Mantegna, R.N. and Stanley, H.E., An introduction to econophysics, Cambridge University Press, 2000. 1, 3, 4, 6, 10, 12

[34] Mantegna, R.N. and Stanley, H.E., Scaling behaviour in the dynamics of an economic index. *Nature,* 1995, **376**, 46–49. 5, 13, 33

[35] Mantegna, R.N. and Stanley, H.E., Turbulence and financial markets. *Nature,* 1996, **383**, 587–588. 6

[36] Mantegna, R.N. and Stanley, H.E., Stochastic Process with Ultraslow Convergence to a Gaussian: The Truncated Lévy Flight. *Phys. Rev. Lett.,* 1994, **73**, 2946–2949. 10

[37] Müller, U. A., Dacorogna, M., Olsen, R.B., Pictet, O.V., Schwarz, M. and Morgenegg, C., Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking & Finance,* 1990, **14**, (6), 1189–1208. 6, 13, 33

[38] Omori, F., On the aftershocks of earthquakes. *Journal of the College of Science, Imperial University of Tokyo,* 1894, **7**, 111–200. 40, 41, 61

[39] Osborne, M.F.M., Brownian motion in the stock market. *Operations Research,* 1959, **7**, 145–173. 2

[40] Potters, M., Cont, R., and Bouchaud, J-P., Financial markets as adaptative systems. *Europhysics Letters,* 1998,**41** 239. 6

[41] Robert, C.P. and Casella, G., Monte Carlo Statistical Methods. New York: Springer-Verlag, 2004. 39

[42] Seemann, L., McCauley, J.L. and Gunaratne, G.H., Intraday volatility and scaling in high frequency foreign exchange markets. *International Review of Financial Analysis,* 2011, **20**, (3), 121–126.

[43] Schoenberg, I.J., Metric Spaces and Completely Monotone Functions. *Ann. Math.,* 1938, **39**, 811–841. 17, 23

[44] Simkowitz, M.A. and Beedles, W.L., Asymmetric stable distributed security returns. *Journal of the American Statistical Association,* 1980, **75**, 306–312. 4

[45] Sinha, S., Chatterjee, A., Chakraborti, A. and Chakrabarti, B.K., Econophysics: An Introduction, Wiley-VCH, 2010. 13

[46] Sokolov, I.M., Chechkin, A.V. and Klafter, J., Fractional diffusion equation for a power-law-truncated Lévy process. *Physica A,* 2004, **336**, 245–251. 22

[47] Sornette, D., Malevergne, Y., and Muzy, J.-F., Volatility fingerprints of large shocks: endogeneous versus exogeneous. In: *Application of Econophysics, Proceedings of the second Nikkei symposium on econophysics,* H. Takayasu, Springer Verlag, 2004. 40

[48] Stella, A.L. and Baldovin, F., Role of scaling in the statistical modelling of finance. *PRAMANA - Journal of Physics,* Indian academy of Sciences, 2008, **71**, (2), 341–352. 14, 33, 41, 43

[49] Stella, A.L. and Baldovin, F., Anomalous scaling due to correlations: limit theorems and self-similar processes. *J. Stat. Mech.,* 2010, (2), P02018. 17, 20, 23, 24, 30, 70, 74, 83

[50] Von Neumann, J., Various techniques used in connection with random digits. Monte Carlo methods. *Nat. Bureau Standards,* 1951, **12**, 36–38. 39

[51] Wang, F., Yamasaki, K., Havlin, S. and Stanley, H.E., Indication of multiscaling in the volatility return intervals of stock markets. *Phys. Rev. E,* 2007, **77**, (1), 016109. 76

[52] Xu, D. and Wirjanto, T.S., An Empirical Characteristic Function Approach to VaR Under a Mixture-of-Normal Distribution with Time-Varying Volatility. *The Journal of Derivatives,* 2010, **18**, (1), 39–58. 23

Acknowledgements

At the end of these three years of work I would like to acknowledge Prof. Attilio Stella for trusting me and for putting at my disposal his deep knowledge and his trained competence.

I also acknowledge Dr. Fulvio Baldovin for easing the task of my research with his numerous explanations and his precise hints and especially for that touch of sincere spirit of comradeship through which he helped me in most of my difficult days; all this was received by me as a friendly gift, and will remain as one of my dearest memories.

I thank my wife Marika, on whose love rests my true happiness. I thank her for helping me to find my way, and for being always a wise and comprehensive partner.

Finally, I thank the little Martino who, with his pure tenderness, shows me the light of the real Love.