Multiple Investment Horizons and Stock Price Dynamics

Alexander Subbotin

UNIVERSITY OF PARIS 1
(PANTHÉON - SORBONNE)

Ref. 2010TA010003

Multiple Investment Horizons and Stock Price Dynamics

by

Alexander Subbotin

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy
in Economics

in the

Ecole doctorale Economie Panthéon Sorbonne (EPS)
Centre d'Economie de la Sorbonne/CNRS
Department of Economics (UFR - 02)

PhD Advisor: Professor Thierry Chauveau

Committee in charge:
Professor Cars Hommes
Professor Yakov Mirkin
Professor Richard Topol
Professor Georges Prat
Professor Bernard De Meyer
Professor Thierry Chauveau

February 10, 2010

Declaration of Authorship

I, Alexander Subbotin, declare that this thesis titled, "Multiple Investment Horizons and Stock Price Dynamics" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

"We all live under the same sky, but we don't all have the same horizon"

Konrad Adenauer

UNIVERSITY OF PARIS 1 (PANTHÉON - SORBONNE)

Abstract

Ecole doctorale Economie Panthéon Sorbonne (EPS)
Centre d'Economie de la Sorbonne/CNRS
Department of Economics (UFR - 02)

Doctor of Philosophy

by Alexander Subbotin

Market prices of risky assets are driven by the actions of economic agents that have different investment horizons: they adjust their portfolios at various frequencies and observe returns at different scales. In this thesis we examine the problem of multiple investment scales from three different angles. In the first place, we study the theoretical implications of the heterogeneity of investors' decision horizons for price dynamics. This analysis is carried in the frameworks of complete and bounded rationality. Second, different time series models of price volatility are examined in view of their capacity to represent the properties of stock returns simultaneously at various time scales. Finally, a method of measuring volatility at multiple scales with wavelet filters is developed, with application to the detection of financial crises.

Acknowledgements

In the first place I would like to thank my advisor Thierry Chauveau from University of Paris-1 for his supervision, advice, corrections and guidance that gave me extraordinary experiences throughout this work. I am specially grateful for his collaboration on the articles that formed are part of this thesis.

I gratefully acknowledge Kateryna Shapovalova for her useful comments, unfailing support and specially for her contribution to the essay that constitutes the third chapter of this thesis.

My thanks go to Thierry Michel and Bertrand Maillet, discussions with whom were helpful for the preparation of the fourth chapter of this thesis.

I am very grateful to Elena Buyanova and Nicolay Berzon from Higher School of Economics for their support and advice. I want to thank Yakov Mirkin for his useful remarks and suggestions.

I would also like to thank the members of the thesis jury who kindly accepted to report on my work.

Thanks to Sunil Patel, who helped with the latex layout of this thesis.

It is a pleasure to pay tribute also to all my present and past colleagues from the Higher School of Economics, Ecole doctorale Economie Panthéon Sorbonne at University of Paris - 1, AORA, Variances, Altran CIS, Renaissance Finance and Natixis, with whom I had opportunity to collaborate during the last five years and whose help enabled the preparation of this thesis.

Last but not the last, I thank my wife and my parents. Without their understanding and endless patience doing all this work would be impossible.

Contents

List of Figures

List of Tables

Abbreviations

ACF	autocorrelation function
AR	autoregression
ARCH	autoregressive conditional heteroscedasticity
ARSV	autoregressive stochastic volatility
ARMA	autoregressive moving average
CCC	constant conditional correlations
DCC	dynamic conditional correlations
DWT	discrete wavelet transform
FIGARCH	fractionally integrated generalized autoregressive conditional heteroscedasticity
GARCH	generalized autoregressive conditional heteroscedasticity
iid	independent and identically-distributed
IGARCH	integrated autoregressive conditional heteroscedasticity
IMS	index of market shocks
LA	least asymmetric
MIS	multiple investment scales
MODWT	maximum overlap discrete wavelet transform
MCM	multiplicative cascade model
MPT	modern portfolio theory
MRW	multi-fractal random walk
MSE	mean square error
MVS	market volatility scale
OU	Ornstein-Uhlenbeck process
RV	realized volatility
SMS	scale of market shocks
SV	stochastic volatility

Chapter 1

Introduction

The central topic of this thesis is modelling and analysis of the stock price fluctuations on a market, where multiple investment scales co-exist. The intuition behind our approach is that different types of investors (institutionals, portfolio managers in investment funds, intraday speculators, etc.) have characteristic frequencies of operations on the market and various horizons of decision taking. Their actions can produce price fluctuations in different bands of spectrum. Also, all of these investors look at the same series of prices at different scales, with different resolution.

The concept of multiple scales, though appealing, lacks theoretical study in the existing literature. It is important to explore if, in presence of multiple scales, the equilibrium price and return of the risky asset can be defined and how. It is interesting whether this approach can explain the stylized patterns in returns, such as volatility clustering. Besides, econometric models of price volatility need to be examined for their compatibility with the multi-scale modelling. Volatility measurement methods, taking into account the spectral properties of fluctuations, could also be helpful. These issues are subjects of the three essays, composing this thesis.

The implications of the multiple investment scales hypothesis for the properties of the time series of risky returns are first studied from the theoretical viewpoint in a simplified heterogeneous agents framework. Then we address practical and empirical issues: choice of volatility time series models, accounting for multiple time scales, and volatility measurement at multiple scales.

The first essay studies the effect of multiple investment horizons and investors' bounded rationality on the price dynamics. We adopt the heterogeneous agents framework, established in Anufriev et al. (2006), and introduce multiple scales in it. We consider a pure exchange economy with one risky asset, populated with agents maximizing

CRRA-type expected utility of wealth over discrete investment periods. An investors' demand for the risky asset may depend on the historical returns, so that the model can encompass a wide range of behaviourist patterns. The necessary conditions, under which the risky return can be a stationary *iid* process, are established. The compatibility of these conditions with different types of demand functions in the heterogeneous agents framework are explored.

We find that conditional volatility of returns cannot be constant in many generic situations, especially if agents with different investment horizons operate on the market. In the latter case the return process can display conditional heteroscedasticity, even if all investors are so-called "fundamentalists" whose demand for the risky asset is subject to exogenous *iid* shocks. We show that the heterogeneity of investment horizons can be a possible explanation of different stylized patterns in stock returns, in particular, mean-reversion and volatility clustering.

In the second essay, we overview different methods of stock prices and exchange rates volatility modelling, focusing on their ability to reproduce the empirical properties in the corresponding time series. The properties of price fluctuations vary across the time scales of observation. The adequacy of different models for describing price dynamics at several time horizons simultaneously is the central topic of this study. We propose a detailed survey of recent volatility models, accounting for multiple horizons. These models are based on different and sometimes competing theoretical concepts. They belong either to GARCH or stochastic volatility model families and often borrow methodological tools from statistical physics. We compare their properties and comment on their practical usefulness and perspectives.

In the third essay, the volatility of a stock market index is decomposed both in time and scale using wavelet filters. We design a probabilistic indicator for volatilities, analogous to the Richter scale in geophysics. The peak-over-threshold method is used to fit the generalized Pareto probability distribution for the extreme values in the realized variances of wavelet coefficients. The indicator is computed for the daily Dow Jones Industrial Average index data from 1896 to 2007 and for the intraday CAC40 data from 1995 to 2006. The results are used for comparison and structural multiresolution analysis of extreme events on the stock market and for the detection of financial crises.

Chapter 2

Price Dynamics in a Market with Heterogeneous Investment Horizons and Boundedly Rational Traders

2.1 Introduction

Up to now, the heterogeneous markets literature almost exclusively focuses on the expectations of market agents, according to which investors are classified into "fundamentalists", "chartists" and "noise traders". It is shown that the interaction, herding behavior and strategy switching of heterogeneous agents transform noise process and create persistent trading volume, excess volatility, fat tails, clustered volatility, scaling laws (see Hommes (2006) and LeBaron (2006) for surveys on interacting agents models). Andersen (1996) interprets the aggregated volatility as the manifestation of numerous heterogeneous information arrivals. Limits to arbitrage, market psychology, heuristics and biases, which are subject of behavioral finance, can also be helpful to explain empirical evidence (see Barberis and Shleifer, 2003).

A number of analytically solvable models were proposed to explore the dynamics of financial market with heterogeneity coming from boundedly rational beliefs of investors about future returns. Brock and Hommes (1998) proposed a model, where investors switch between a number of strategies according to expected or realized excess profits. Stylized simple strategies describe patterns in investors' behavior that are commonly observed empirically - chartism and trend-following. Chiarella and He (2001), Anufriev

et al. (2006) and Anufriev (2008) studied an artificial market populated with investors, following heterogeneous strategies and maximizing the expected CRRA utility. Compared to earlier studies that use CARA utilities, they make investment decisions depend on wealth, which is undoubtedly more realistic but technically more difficult. Vanden (2005) introduces a more sophisticated step-wise dependence of the risk aversion on wealth and finds that this can have important consequences for return dynamics. Recently Weinbaum (2009) showed that heterogeneous risk preferences and risk sharing can be the source of volatility clustering.

To our knowledge, all the above-mentioned models of heterogeneity ignore one of its important sources, which is different investment scales. By investment scales we mean typical periods between two consecutive adjustments of investment portfolio, peculiar to a certain type of investors. The heterogeneity of the market with respect to agents' operations frequencies is further referred to as the Multiple Investment Scales (MIS) hypothesis. We suppose that investors maximize expected utility of wealth at the end of some investment period. We call the typical length of this period as investment horizon (or scale).

Earlier the effect of heterogeneity in investment horizons was studied in Anufriev and Bottazzi (2004). They derive a fixed point for the price of the risky asset dynamics under the assumption that agents maximize expected CARA utility over different periods in future. But their model disregards the effect of various frequencies of portfolio adjustments and, due to the constraints of the CARA assumption, does not realistically account for the dynamics of wealth. They conclude that heterogeneity of investment horizons alone is not enough to guarantee the instability of the fundamental price and the emergence of the non-trivial price dynamics, such as volatility clustering or serial correlations. In this paper we derive the opposite conclusion, which is close to that obtained in Chauveau and Topol (2002). Working in a different framework, they explained volatility clustering of OTC exchange rates by market microstructure effects, unifying intraday and interday dynamics.

Though not examining the MIS hypothesis analytically, several earlier studies evoke the heterogeneity of investment horizons as a possible explanation of the stylized facts in stock price volatility. The assumption that price dynamics is driven by actions of investors at different horizons serves as a micro-economic foundation of the volatility models in Müller et al. (1997). They suppose that there exist volatility components, corresponding to particular ranges of stock price fluctuation frequencies, that are of unequal importance to different market participants. These participants include intraday speculators, daily traders, portfolio managers and institutional investors, each having a characteristic time of reaction to news and frequency of operations on the market. So

frequencies of price fluctuations depend on the periods between asset allocation decisions, and/or the frequencies of portfolio readjustments by investors.

An important question, answered in this paper, is whether the presence of *(i)* contrarian and trend-following investors and *(ii)* heterogeneous information arrivals on the market are necessary properties for an interacting agents model to reproduce the stylized facts of the return volatility dynamics. We show that, under some conditions, volatility clustering can arise even in an economy populated with fundamentalist traders only, given that they adjust their portfolios with different frequencies. We also propose a study of the joint effect of the MIS hypothesis and of the bounded rationality in investment strategies.

The rest of the paper is organized as follows. In the next section we introduce the general setting of the model. Section 2.3 describes the equilibria in the one-scale model with boundedly rational investors, re-examining the conclusions of Anufriev et al. (2006) and preparing the ground for the study of the multi-scale case. In section 2.4 we derive the equilibrium in the MIS case and establish the properties of the return dynamics. In section 2.5 we illustrate our findings with simulation examples. In conclusion the main results are summarized and possible model extensions are discussed.

2.2 A Model for Joint Dynamics of Stock Price and Wealth with Multiple Investment Scales

In this section we formulate the model and then discuss its various possible specifications and assumptions. The general setup follows the lines of Chiarella and He (2001) and Anufriev et al. (2006), to which we add the MIS hypothesis and some constraints on investors' behavior, discussed later. Where possible, we keep the same notation as in Anufriev et al. (2006), to enable the easy comparison of results.

Consider a two-assets market where N agents operate at discrete dates. The risk-free asset yields a constant positive interest R_f over each period and the risky asset pays dividend D_t at the beginning of each period. The price of the risk-free asset is normalized to one and its supply is absolutely elastic. The quantity of the risky asset is constant and normalized to one, while its price is determined by market clearing by a Walrasian mechanism. The Walrasian assumption means that all agents determine their demand for the risky asset taking the price of the risky asset P_t as parameter. In other, though this price is unobserved at the moment when investors form their demand, they calculate the demand for the risky asset at every possible price and submits this

to a hypothetical Walrasian auctioneer. The price is then set so that the total demand across all agents equals to one.

The demand of the risky asset is formulated in terms of the shares of wealth of agents, so that $x_{t,i}$ stands for the share of wealth that investor i with wealth $W_{t,i}$ wishes to invest in the risky asset. The corresponding number of units of the asset is $\frac{W_{t,i}x_{t,i}}{P_t}$. The market clearing condition imposes:

$$\sum_{i=1}^{N} x_{t,i}W_{t,i} = 1$$

The wealth of each investor evolves according to the below equation:

$$W_{t,i} = (1 - x_{t-1,i})W_{t-1,i}(1 + R_f) + \frac{x_{t-1,i}W_{t-1,i}}{P_{t-1}}(P_t + D_t) =$$
$$(1 - x_{t-1,i})W_{t-1,i}(1 + R_f) + x_{t-1,i}W_{t-1,i}(1 + R_t + \varepsilon_t),$$

where D_t is a dividend payment, whose ration to price is supposed to be an *iid* random variable ε_t, and R_t is the return on the risky asset. We define the total return by

$$Y_t = \frac{P_t + D_t}{P_{t-1}}.$$

Following Anufriev et al. (2006), we rewrite the model in rescaled terms which allows to eliminate the exogenous expansion due to the risk-free asset growth from the model:

$$w_{t,i} = \frac{W_{t,i}}{(1 + R_f)^t},\ p_t = \frac{P_t}{(1 + R_f)^t},\ e_t = \frac{\varepsilon_t}{1 + R_f},\ y_t = \frac{Y_t}{1 + R_f}.$$

By consequence, the rescaled return on the risky asset is defined by:

$$r_t = \frac{p_t}{p_{t-1}} - 1 = \frac{1 + R_t}{1 + R_f} - 1 = \frac{R_t - R_f}{1 + R_f}. \tag{2.1}$$

In these terms the whole system dynamics simplifies to:

$$\begin{aligned} p_t =& \sum_i x_{t,i}w_{t,i}, \\ w_{t,i} =& w_{t-1,i}\left[1 + x_{t-1,i}\left(r_t + e_t\right)\right]. \end{aligned} \tag{2.2}$$

Proposition 2.1. *The rescaled price dynamics, solving the dynamic system (2.2), verifies:*

$$p_t = p_{t-1}\frac{\sum_i w_{t-1,i}\left(x_{t,i} - x_{t-1,i}x_{t,i}\right) + e_t\sum_i x_{t,i}x_{t-1,i}w_{t-1,i}}{\sum_i w_{t-1,i}\left(x_{t-1,i} - x_{t,i}x_{t-1,i}\right)},$$

Proof. See Anufriev et al. (2006). □

Proposition 2.1 describes the equilibrium price dynamics in the sense that at each period t Walrasian equilibrium is achieved on our two-asset market. It is straightforward to see that the equilibrium return must satisfy:

$$r_t = \frac{\sum_i w_{t-1,i} \left(x_{t,i} - x_{t-1,i} + x_{t,i} x_{t-1,i} e_t \right)}{\sum_i w_{t-1,i} \left(x_{t-1,i} - x_{t,i} x_{t-1,i} \right)}, \tag{2.3}$$

if the rescaled return is defined by (2.1). Note that equation (2.3) explicitly specifies the return r_t conditionally to the information set at period $t-1$, if and only if we impose additional assumptions: both the demand $x_{t,i}$ and the dividend yield e_t must be independent of the current price level p_t.

The simplest assumption about dividends one can suggest to make the model in (2.2) tractable, is that the dividend yield is an *iid* non-negative stochastic process. Following Chiarella and He (2001) and Anufriev et al. (2006), we stick to this assumption, though we are aware of the constraints it imposes. Dividends in our economy are deprived of their own dynamics, but follow the risky asset price. Roughly speaking, the amount of dividends available is supposed to automatically adapt to the fluctuations of the price level, so that the mean dividend yield remains unchanged. In real life dividends are paid by stock issuers and so depend on companies' profits and decided payout ratios. If the supply of the risky asset is fixed, one can hardly expect a perfectly linear dependence between average dividends and prices, though a positive relationship between them does exist. However, for the purposes of our paper, the *iid* assumption for the dividend yield is sufficient.

So far, nothing has been said about the way agents determine the desired proportions of investment in the risky asset. The MIS hypothesis, studied in this paper, implies that some investors do not trade at all time periods and remain passive. During the period, when some investor is out of the market, his share of investment in the risky asset is no longer a result of his decisions but a consequence of price and wealth movements, independent of his will. The following proposition specifies the way investment shares evolve.

Proposition 2.2. *Let $x_{t,i}^{-k}$ be the share of investment in the risky asset of investor i, who actually participated in the trade k periods ago, $k = 1, \ldots, h$ with h being his investment horizon. The investment share verifies the following recurrent relationship:*

$$x_{t,i}^{-k} = \frac{x_{t-1,i}^{-k+1}(1 + r_t)}{1 + x_{t-1,i}^{-k+1}(r_t + e_t)} \tag{2.4}$$

Proof. See Appendix. □

At the period when investor i readjusts his portfolio, his demand for the risky asset $x^0_{t,i}$ is determined according to some investment function. In this paper, we suppose that investment functions are given as the dependence of the share of wealth, invested in the risky asset, on the beliefs about future gains. We also suppose that investment functions are deterministic and do not change over time for the same investor[1]. The beliefs are based on the past observations of prices and dividends, without any private information that could be used to forecast future returns. Moreover, each investment function is supposed to be independent of the current wealth, which is a natural assumption in the CRRA framework. So investor i's function reads:

$$x^0_{t,i} = f_i(r_{t-1}, \ldots, r_{t-L_i}, e_{t-1}, \ldots, e_{t-L_i}) \tag{2.5}$$

where L_i is the maximum lag for historical observations used by the agent i, which can be finite or infinite.

In this paper we will in particular focus on the case of preferences that corresponds to the maximization of the mean-variance CRRA expected utility of wealth[2]. Let us suppose that investors, possibly operating over different time scales, maximize a mean-variance expected utility:

$$\max_{x^0_{t,i}} \left\{ \mathrm{E}_{t-1,i}(W_{t+h,i}) - \frac{\gamma_i}{2W_{t,i}} \mathrm{Var}_{t-1,i}(W_{t+h}) \right\} \tag{2.6}$$

with operators $\mathrm{E}_{t-1,i}(\cdot)$ and $\mathrm{Var}_{t-1,i}(\cdot)$ standing for the beliefs of agent i about the mean and variance given the information at time $t-1$. The information set of period $t-1$ includes the prices of the risky asset and the dividends at time $t-1$ and earlier. The coefficient γ_i is a positive constant that measures the risk aversion of investor i. The time horizon of decision taking, denoted h, corresponds to the period of time when investor i does not readjust his portfolio. The number of units of risky asset in investor's possession remains constant over $[t; t+h]$, while the share of investment in the risky asset may evolve. We assume that dividends, paid by the risky asset during this period, are accumulated on the bank account, yielding the risk-free rate.

Proposition 2.3. *The solution $x^{0*}_{t,i}$ of the maximization problem (2.6) is approximately given by:*

$$x^{0*}_t \approx \frac{\mathrm{E}_{t-1,i}\left[\sum_{k=1}^{h}(e_{t+k} + r_{t+k})\right]}{\gamma_i \mathrm{Var}_{t-1,i}\left[\sum_{k=1}^{h}(e_{t+k} + r_{t+k})\right]} \tag{2.7}$$

[1] Note that this does not exclude functions, corresponding to investment strategies that evolve according to predefined rules.

[2] Most of our results are valid also in the case of the general investment function (2.5), not necessarily representing the beliefs about mean and variance. This will be specially indicated further in the paper.

Proof. See Appendix. □

Chiarella and He (2001) show that the expression (2.7) with $h = 1$ also emerges as an approximative solution in the maximization problem with the power utility function. This approximation, however, consists in a discretization of a continuous-time process with Gaussian increments and thus it can be far from the real solution for non-infinitesimal time units. So we prefer to work with mean-variance maximization directly. Alternatively, an investment function of the form (2.7) could be set on an *a priori* basis since it describes the behavior of a mean-variance investor with constant relative aversion to risk.

Notice that if the return process is *iid*, $\mathrm{E}_{t-1}[y_{t,t+h}] = h\,\mathrm{E}_{t-1}[r_{t+1} + e_{t+1}]$ and $\mathrm{Var}_{t-1}[y_{t,t+h}] = h\,\mathrm{Var}_{t-1}[r_{t+1} + e_{t+1}]$. This ensures that if, in addition, risk aversion is homogeneous for investors at all scales ($\gamma_i = \gamma$), the demand for the risky asset does not depend on the investment horizon. We maintain the assumption of homogeneous risk aversion throughout this paper.

In equation (2.7) the share of wealth to be invested in the risky asset depends exclusively on the beliefs of agents about future yields. In the heterogeneous agents literature these beliefs are based on historical prices of the risky asset up to a certain lag. Following Chiarella and He (2001), we do not include p_t in the information set for beliefs formations in order to avoid unnecessary complexity. Nevertheless, in the MIS case the aggregate demand on the risky asset naturally depends on the current price level. Indeed, suppose that the previous date, when investor i participated in the trade, was $t-k$ and that at this date the share of wealth $x^0_{t-k,i}$ he invested in the risky asset was determined according to (2.5). Then it follows from (2.4) that his current investment share $x^{-k}_{t,i}$ depends on the historical returns and dividend yields up to the lag $L_i + k - 1$, but also on the current return and the dividend yield, which are unknown before the trade at date t. So equation (2.3) does not explicitly specify the dynamics of the risky return.

In the following section we study the dynamics of the price and wealth in the model with one scale, which is a particular case of the model, introduced in the previous section. We further refer to it as the benchmark model. We extend the analysis of Anufriev et al. (2006) in several aspects, also important in the MIS case, studied later.

2.3 Equilibria in the One-Scale Model with Bounded Rationality

As we have mentioned before, in the one-scale case, equation (2.3) completely and explicitly describes the dynamics of the return on the risky asset under the market clearing condition. By specifying the demand function, one can determine the equilibrium price and return. This equilibrium dynamics was earlier studied in Anufriev et al. (2006), who replace the actual dividend yield by its mean and work with the so-called "determenistic skeleton" of the system. In the deterministic case the (rescaled) return is constant: $r_t = r$. The authors prove that two types of equilibria are possible: either a single agent survives [3], or many agents survive, but in both cases the equilibrium share of investment in the risky asset and the steady growth rate of its price are determined in a similar way. They must satisfy the relationship, which is easily obtained from (2.2) for a single-agent case, when we set $x_t = x_{t-1}$ for all t. This relationship is called the Equilibrium Market Line (EML) and reads:

$$x = \frac{r}{r + \bar{e}} \tag{2.8}$$

where $\bar{e}$ is the mean dividend yield.

The demand functions of investors depend on a single variable and are of the form:

$$x = f(r_{t-1}, \ldots, r_{t-L}) = f(r, \ldots, r) = \tilde{f}(r) \tag{2.9}$$

The equilibrium points are determined as the intersections of the demand curve $\tilde{f}(r)$ and the EML. It is shown that, if multiple agents survive, their demand functions must all intersect the EML at the same equilibrium point. Stability conditions, depending on the properties of derivatives of $f_i(\cdot)$ with respect to returns at different lags, are established. We refer the reader to the original paper of Anufriev et al. (2006) for further details.

In our approach, the main difference is that we are interested in the stochastic properties of the return process. In particular, we establish analytically, under what conditions the dynamics of returns is "simple" (*iid*) and when it displays "interesting" dynamic patterns (conditional heteroscedasticity and/or serial correlations). In our view, this type of approach is appropriate for the study of boundedly rational behavior of agents, whose investment functions are based on beliefs about mean and variance of the return process. This point is explained further.

For the case of multiple agents with heterogeneous investment functions, Anufriev et al. (2006) determine which form of the demand function "dominates" the others. For

[3] *i.e.* his share in the total wealth does not decrease to zero in infinite time

example, if a trend-follower (investor who strongly extrapolates past returns) meets a fundamentalist (investor, whose demand function is independent of the price history), we can predict which of them survives, depending on the respective form of their investment functions. Under some conditions, the trend-follower outperforms the fundamentalist and survives. A striking feature of the model is that equilibria are possible for almost any, and even completely senseless, demand functions and can even be stable.

The problem here is with bounded rationality. More precisely, it is important to what extent the rationality is bounded. In Anufriev et al. (2006) and Chiarella and He (2001), investment functions are given *a priori*, and though they formally depend on the beliefs of agents about the mean and variance of future returns, there are no constraints on how these beliefs should be related to the true quantities.

Bounded rationality means that agents may not know the true model. But in equilibrium, when the return on the risky asset is supposed to be constant, it is hard to admit that the beliefs have nothing to do with reality. Besides, the stability of such equilibria hardly makes sense from the economic point of view, since agents would have incentives to change their strategies, if they were allowed to.

In Brock and Hommes (1998) agents are allowed to switch between strategies, according to the profits they yield in the past. The agents can thus be claimed to be *procedurally rational*, because they try to rationally choose strategies according to some criteria. In our case, a more exact definition of procedural rationality can be helpful to study the model analytically. We restrain the class of admissible investment functions, considerably reducing the possibilities for non-rationality of economic agents, without necessarily imposing rational expectations.

Investment functions basically define how agents' beliefs about future returns are formed, i.e. they are concise descriptions of the outputs of the beliefs-making procedure. The rationality of such procedure can be tested in some simple reference case, where the outcome of the procedure is expected to correspond to the rational behavior. In our case of mean-variance investor, we require that for the *iid* returns beliefs about mean and variance of the process should be unbiased. This is formalized in the following definition.

Definition 2.4. An investment function of the form:

$$x^0_{t,i} = \tilde{f}_i\left(\mathrm{E}_{t-1,i}[y_{t,t+h}], \mathrm{Var}_{t-1,i}[y_{t,t+h}]\right) \tag{2.10}$$

is called procedurally rational if the beliefs $\mathrm{E}_{t-1,i}(y_{t-1,t+h})$ and $\mathrm{Var}_{t-1,i}(y_{t,t+h})$ about the mean and variance of the future total returns are unbiased estimates of these quantities with finite error, if the true process $y_{t,t+1}$ is *iid*.

This definition is an adaptation of Simon's procedural rationality [4] to our context. It basically states that, if previous observations of returns display no non-trivial dynamic patterns, the beliefs about mean and variance of investors should have no systematic error. Note that in no way we state that returns should actually follow an *iid* process. We only describe the behavior of the investment function in this hypothetical case, in order to impose some constraints on the "reasonability" of the decision taking procedure used by investors.

Note that our definition does not contradict to the concept of bounded rationality, but it requires some moderate degree of consistency in investors' beliefs. Procedurally rational investors can actually be trend-followers or contrarians. Consider, for example, the following specifications for the beliefs about the mean of future returns:

$$\begin{aligned} \mathrm{E}_{t,i}(y_{t+1}) &= c_i + \frac{d_i}{l}\sum_{k=1}^{l} y_{t-k} \qquad (A) \\ \mathrm{E}_{t,i}(y_{t+1}) &= \frac{1-d_i}{L}\sum_{k=1}^{L} y_{t-k} + \frac{d_i}{l}\sum_{k=1}^{l} y_{t-k} \qquad (B) \end{aligned} \tag{2.11}$$

The function of the type, analogous to (2.11A), is used in Chiarella and He (2001) to represent the behavior of heterogeneous investors. Here c_i is some constant that represents the risk premium, required by the investor, and d_i is a behaviorist parameter, which specifies, how investor i extrapolates the performance of the risky asset over l recent periods. If $d_i = 0$, the investor is fundamentalist, if $d_i > 0$ he is a trend-follower, otherwise contrarian (chartist). It is easy to see that this specification does not correspond to our definition of the procedural rationality, unless simultaneously $c_i = 0$ and $d_i = 1$. The function (2.11B) also allows for the extrapolation of the recent returns via the parameter d. If $l < L$, positive d_i corresponds to the trend-following. But this function verifies our condition for procedural rationality: in the *iid* case the expectation of the difference in the short-term and the long-term mean is null.

Suppose that investors' preferences are described by the mean-variance function of the form (2.7), satisfying definition 2.4. Having restrained the set of admissible investment functions, we turn to the study of the price dynamics in the benchmark model. In the following theorem we establish the conditions that must be verified by the investment function to ensure "simple" dynamics of the returns, which can be associated with some steady growth trajectory. It states that the assumption of investors' rational expectations is equivalent to the *iid* dynamics of returns.

[4] "Behavior is procedurally rational when it is the outcome of appropriate deliberations. Its procedural rationality depends on the process that generated it" (Simon, 1976, p.131).

Theorem 2.5. *In the benchmark model with homogeneous procedurally rational agents the return process can be iid with finite mean and variance if and only if investors have rational expectations. In this case the mean and variance of the return process are uniquely defined by the mean dividend yield and investors' risk aversion.*

Proof. The homogeneity of agents means that they all have the same investment functions and, in particular, the same risk aversion $\gamma_i = \gamma$. In the benchmark model they also use the same information, so $x_{t,i} = x_{t,j}, \forall t, i, j$ and we can drop the second subscript. Thus this case is analogous to a single-agent model with a representative agent. Simplifying (2.2), it is straightforward to see that the returns do not directly depend on the wealth dynamics, since we have:

$$r_t = \frac{x_t - x_{t-1} + e_t x_t x_{t-1}}{(1 - x_t)x_{t-1}} \tag{2.12}$$

If r_t is an *iid* process, then r_t is independent of the returns' history $r_{t-1}, r_{t-2}, \ldots$, but it is also independent of $x_t, x_{t-1}, \ldots$ since the latter depend only on past returns. Consider the stochastic process $r_{t|t-1}$ of returns, conditional to the information at period $t-1$, which is defined as the set $I_t = \{r_{t-1}, r_{t-2}, \ldots; x_t, x_{t-1}, \ldots\}$. It follows from the above that this process is also *iid*.

The quantities $x_{t|t-1}$ and $x_{t-1|t-1}$ are both deterministic since the investment function at time t depends only on returns at time $t-1$ and earlier. So the conditional mean and variance of returns are:

$$\mathrm{E_{t-1}}(r_t) = \frac{x_t - x_{t-1} + \bar{e} x_t x_{t-1}}{(1 - x_t)x_{t-1}} \tag{2.13}$$

$$\mathrm{Var_{t-1}}(r_t) = \sigma_e^2 \frac{x_t^2}{(1 - x_t)^2} \tag{2.14}$$

with $\bar{e}$ and σ_e^2 the mean and variance of the dividend yield process respectively (both are supposed to be constant). Note that here the operators $\mathrm{E_t}(\cdot)$ and $\mathrm{Var_t}(\cdot)$ no longer refer to the agent's beliefs, but to the mathematical expectation and variance of random variable.

We have shown that the process $r_{t|t-1}$ is *iid*. Then it follows from (2.13) that $x_t = x_{t-1} = x^*$ and equation (2.12) simplifies to:

$$r_t = e_t \frac{x^*}{1 - x^*} \tag{2.15}$$

The investment function $f(r_{t-1}, \ldots, r_{t-L})$ takes the value x^* with probability 1 for all values $r_{t-1}, \ldots, r_{t-L}$ drawn from an *iid* process if and only if it is a constant function in any domain where the vector $r_{t-1}, \ldots, r_{t-L}$ takes values with non-zero probability.

Since the return dynamics, given by (2.15), is *iid*, procedural rationality implies that the beliefs of investors are unbiased:

$$\begin{aligned} \mathrm{E}_{t-1}\left[r_{t+1} + e_{t+1}\right] =& \bar{e}\frac{x^*}{1-x^*} + \bar{e} = \frac{\bar{e}}{1-x^*} \\ \mathrm{Var}_{t-1}\left[r_{t+1} + e_{t+1}\right] =& \mathrm{Var}_{t-1}\left[\frac{e_{t+1}}{1-x^*}\right] = \frac{\sigma_e^2}{(1-x^*)^2} \end{aligned} \tag{2.16}$$

Then, according to (2.7) with $h = 1$, the investment share satisfies:

$$x^* = \frac{\frac{\bar{e}}{1-x^*}}{\gamma \frac{\sigma_e^2}{(1-x^*)^2}} \tag{2.17}$$

From (2.17) we obtain a unique solution for x^*[5]:

$$x^* = \frac{\bar{e}}{\gamma\, \sigma_e^2 + \bar{e}} \tag{2.18}$$

This proves that if returns are *iid*, then the investment share x^*, computed from (2.18), uniquely determines the mean and variance of the process $r_{t|t-1}$ (or, in other words, the necessary conditions for the *iid* return dynamics). It is easy to see that the solution we derived corresponds to the case where investors have rational expectations.

It can be shown straightforwardly, that these conditions are also sufficient. It suffices to plug the constant x^* in the equation (2.12) for returns and then verify that the expectation and variance of returns are constant and given by (2.13) and (2.14) respectively. □

An important consequence of theorem 2.5 is that, in the benchmark model with homogeneous procedurally rational investors, returns on the risky asset never have simple *iid* dynamics, unless the investors have rational expectations. Note that equation (2.14) describes conditional volatility dynamics in the model. It follows from (2.14) that for $0 < x_t < 1$, conditional variance always increases with x_t. If the investment function depends positively on the historical mean of returns and negatively on their historical variance (which is an appropriate assumption in a procedurally rational context), then the conditional variance is a decreasing function of historical variance and increasing

[5] Analoguous computation in terms of not-rescaled variables gives $x^* = \mathrm{E}_{t-1}[\varepsilon_{t+1}]\left(\gamma\, \mathrm{Var}_{t-1}[\varepsilon_{t+1}] + \mathrm{E}_{t-1}[\varepsilon_{t+1}]\right)^{-1}$, which is slightly different from (2.18) because of the first order approximation. This difference is of no incidence in our context.

function of historical returns. At the same time, volatility has the same "memory" as the squared share of investment in the risky asset, which is determined by investors' beliefs. If the latter are adjusted slowly, then volatility also adjusts slowly.

Now consider the return dynamics in a more general case, when homogeneous agents having arbitrary (not necessarily boundedly rational) investment functions of the form

$$x_t = f(r_{t-1}, \dots, r_{t-L}, e_{t-1}, \dots, e_{t-L}).$$

The stochastic process (2.12) for the return dynamics is non-linear. We study the properties of its first-order Taylor linearization in the neighborhood of the expected return, $r_{t-k} = \bar{r}$ for all k, with arbitrary values of e_{t-k}. We denote $\widetilde{e}_t = e_t - \bar{e}$ and $\widetilde{r}_t = r_t - \bar{r}$ the deviations of dividend yield and return from their average values. We also denote f'_k the first derivative of $f(\cdot)$ with respect to r_{t-k} for $k = 1, \dots, L$. The form of the return process is given by the following theorem.

Theorem 2.6. *In the benchmark model with homogeneous agents, if the return process is covariance stationary, it satisfies:*

$$\begin{aligned} r_t &= \frac{\bar{x}}{1-\bar{x}}\bar{e} + \widetilde{r}_t \\ \widetilde{r}_t &= \sum_{k=1}^{L+1} a_k \widetilde{r}_{t-k} + v_t \widetilde{e}_t \\ v_t &= \frac{\bar{x}}{1-\bar{x}} + \sum_{k=1}^{L} b_k \widetilde{r}_{t-k} \end{aligned} \tag{2.19}$$

with:

$$\begin{aligned} a_1 &= \frac{f'_1[1-\bar{x}\,(1-\bar{e}\,)]}{\bar{x}\,(1-\bar{x}\,)^2} \\ a_k &= \frac{f'_k[1-\bar{x}\,(1-\bar{e}\,)] + f'_{k-1}(\bar{x}\,-1)}{\bar{x}\,(1-\bar{x}\,)^2}, \;\; k \in \{2, \dots, L\} \\ a_{L+1} &= \frac{f'_L}{\bar{x}\,(1-\bar{x}\,)} \\ b_k &= \frac{f'_k}{(1-\bar{x}\,)^2}, \;\; k \in \{1, \dots, L\} \\ \bar{x} &= \end{aligned} \tag{2.20}$$

Proof. See Appendix. □

Equations (2.19) can be written in the equivalent form:

$$\begin{aligned} \widetilde{r}_t &= \sum_{k=1}^{L+1} a_k \, \widetilde{r}_{t-k} + \sigma_e \left(u_t + \frac{\bar{x}}{1-\bar{x}} \right) \varepsilon_t \\ u_t^2 &= \sum_{k=1}^{L} b_k^2 \, \widetilde{r}_{t-k}^2 + 2 \sum_{\substack{i,j \in \{1,\dots,L\} \\ i \neq j}} b_i b_j \widetilde{r}_{t-i} \widetilde{r}_{t-j} \end{aligned} \tag{2.21}$$

with ε_t a standardized independent white noise. This stresses the ARCH-like nature of the returns process. Note that its mean is described by an expression, equivalent to the definition of equilibrium on the EML in (Anufriev et al., 2006).

Now we can turn to the case with heterogeneous agents, i.e. the case when $x_{t,i}$ are determined in a different way by each investor. Again we restrict the investment functions to mean-variance and procedurally rational. Theorem 2.7 shows that the simple *iid* dynamics does not appear generically if investors are heterogeneous.

Theorem 2.7. *In the benchmark model with heterogeneous procedurally rational agents the return process can be iid with finite mean and variance only if the aggregate share of wealth invested in the risky asset is constant. In this case the mean and variance of the return process are proportional to the mean and variance of the dividend yield.*

Proof. See Appendix. □

Basically this theorem says that if the aggregate share of investment in the risky asset is subject to stochastic shocks or fluctuations, the return dynamics is almost surely not trivial and displays dynamic patterns. The situation, when the aggregate investment function is constant and returns are *iid*, can arise only when the dependence of the individuals' procedurally rational investment functions on the past returns is not characterized by prevailing patterns. More precisely, individual deviations $\nu_{t,i} = x_{t,i} - \bar{x}$ from some constant investment share $\bar{x}$, weighted by the wealth portions of agents $\xi_{t,i}$, are eliminated by aggregation with probability one:

$$P(\sum_{i=1}^{N} \xi_{t,i} \nu_{t,i} = 0) = 1$$

for all t. For this condition to be true, some form of the law of large numbers must be satisfied and, moreover, the expectation of $\nu_{t,i}$, conditional on past returns, must be constant. This is improbable in the situation, when all investors base their expectations on the same vector of realized past returns and this vector is not constant.

2.4 Equilibria with Multiple Investment Scales

In the previous section we considered the case when investors have the same investment horizons, but possibly different investment functions. Now we come back to th MIS hypothesis and study another source of heterogeneity, related to investment horizons. Now assume that there exist H investment scales with portfolio readjustment periods $h = 1, \dots, H$ time units, so that each agent has a characteristic investment scale that does not change. Suppose that within each investment scale investors are homogeneous, *i.e.* have the same specifications of demand function. Finally, suppose that at each date the wealth of investors, having the same investment scale, is distributed so that a constant part of this wealth, equal to $1/h$ belongs to the investors, rebalancing their portfolios at the current date.

The latter assumption does not necessarily imply that the wealth can be redistributed between different groups of investors in a given period. Rather, it means that there is a large number of investors, going in and going out, and they have random dates of intervention on the market but fixed frequencies of trades. So the composition of each cohort of investors may change, but its average share of wealth remains constant.

Under these simplifying assumptions, we can aggregate all investors, acting at the same scale h, and replace them by a representative agent, whose share of wealth, invested in the risky asset, satisfies:

$$x_{t,h} = \frac{1}{h} \sum_{k=0}^{h-1} x_t^{-k} \tag{2.22}$$

Equations (2.2), describing the dynamics of the system, are still true, but now the subscript i corresponds to the investment scale and the wealth $w_{t,i}$ is the aggregate wealth of a class of investors, having the same investment horizon. In section 2.2 we derived equation (2.4) that describes the evolution of the share of investor's wealth, invested in the risky asset, when he does not trade. Then the complete system of equations, describing the dynamics of risky return, reads:

$$\begin{aligned} r_t &= \frac{\sum_{h=1}^{H} w_{t-1,h} \left(x_{t,h} - x_{t-1,h}\right) + e_t \sum_{h=1}^{H} x_{t,h} x_{t-1,h} w_{t-1,h}}{\sum_{h=1}^{H} w_{t-1,h} x_{t-1,h} \left(1 - x_{t,h}\right)} \\ x_{t,h} &= \frac{1}{h} \sum_{k=0}^{h-1} x_{t,h}^{-k} \\ x_{t,h}^{-k} &= \frac{x_{t-1,h}^{-k+1}(1 + r_t)}{1 + x_{t-1,h}^{-k+1}(r_t + e_t)} \end{aligned} \tag{2.23}$$

As noted above, an important feature of (2.23) is that it describes the return dynamics only implicitly, because the investment share for all but the shortest scales inevitably depends on current return. The relation between the price and the dividend process becomes non-linear and complicated, because it includes previous dividends. For the general equation of price dynamics, we can prove that:

Theorem 2.8. *Whatever the number of scales H, there always exists at least one positive market clearing price for which the return r_t satisfies (2.23).*

Proof. See Appendix. □

It is important to specify conditions, under which the multi-scale dynamics does not degenerate, that is the portions of wealth, held by the agents, investing at each scale, do not tend to zero as time tends to infinity.

More precisely, denote $\xi_{t,h}$ the portion of wealth that belongs to investors of type h.

Definition 2.9. The MIS dynamics, described by equation (2.23), is called non-degenerating, if for any investment scale h such as $\xi_{0,h} > 0$ we have :

$$P(\xi_{t,h} = 0) = 0,$$

when t approaches infinity.

In the following theorem we establish the necessary and sufficient conditions that provide for the non-degenerating dynamics in the MIS system. Denote $g_{t,h}$ the growth rate of wealth of investors of type h at time t:

$$g_{t,h} = \frac{w_{t,h}}{w_{t-1,h}} = 1 + x_{t-1,h}(r_t + e_t).$$

We suppose that the stochastic process $\ln(g_{t,h})$ is covariance-stationary. Furthermore, we suppose that it verifies the following conditions on its memory:

$$N^{-1}||\{\mathrm{Cov}\left(\ln(g_{t+i,h}), \ln(g_{t+j,h})\right)\}_{\{i=1,\dots,N, j=1,\dots,N\}}||_2 \leq C \tag{2.24}$$

for all positive N and some finite C. This technical condition, implying that $\ln(g_{t,h})$ is a stochastic process with bounded spectral density, ensures that the average growth rates of wealth converge almost surely to their expectation as time tends to infinity. This result is proved in Ninness (2000).

Theorem 2.10. *The multiple investment scales dynamics, described by equation (2.23), is non-degenerating if and only if for any h:*

$$\mathrm{E}\left[\ln(g_{t,i})\right] = \mathrm{E}\left[\ln(g_{t,j})\right],\ \forall\ i,j \in \{1,\dots,H\}$$

Proof. See Appendix. □

To interpret the theorem, notice that the log growth rate of the wealth is approximately equal to the product of the total return on the risky asset and the share of wealth, invested in the risky asset at the previous period. Thus, for the model to be non-degenerating, investors should either have the same average share of investment in the risky asset, or lower investment shares should be compensated by positive correlation of the investment share with future return. A particular case of the non-degenerating system is non-predictive equal-in-law investment shares:

$$\begin{aligned} x_{t,i} \overset{L}{=} & x_{t,j}\ \forall\ i,j \in \{1,\dots,H\},\ \forall\ t \\ \mathrm{Cov}\left(x_{t-i,h}, r_t\right) = & 0,\ \forall h \in \{1,\dots,H\},\ \forall i,t. \end{aligned} \tag{2.25}$$

Note that in MIS system the existence of autocorrelations in returns implies correlation of the investment shares with the future returns. Moreover, the latter is higher for investors at longer scales, because at each time period there are more passive investors, whose investment shares depend on past returns, even if elementary investment functions are constant. Thus condition (2.25) is related to the absence of serial correlations in returns.

By analogy with the one-scale case, we analyze the equilibrium dynamics of the system (2.23). First let us study the "mean" dynamics, supposing $e_t = \bar{e}$. The following theorem shows that there exists an equilibrium path $r_t = \bar{r}$ that solves the deterministic analog of (2.23).

Theorem 2.11. *The dynamic system (2.23) with $e_t = \bar{e}$ has a unique equilibrium solution with constant return:*

$$\begin{aligned} \bar{r} &= \frac{\bar{x}}{1-\bar{x}}\bar{e}\,, \\ f_h(\bar{r}\,,\dots,\bar{r}\,) &= \bar{x} \end{aligned} \tag{2.26}$$

Proof. Suppose that the system has some equilibrium solution $r = \bar{r}$. This implies:

$$\frac{\sum_{h=1}^{H} \bar{w}_{t-1,h}\left[\bar{x}_{t,h} - \bar{x}_{t-1,h} + \bar{e}\bar{x}_{t,h}\bar{x}_{t-1,h}\right]}{\sum_{h=1}^{H} \bar{w}_{t-1,h}\bar{x}_{t-1,h}\left(1-\bar{x}_{t,h}\right)} = \bar{r} < \infty \tag{2.27}$$

Besides, it is easy to notice that, for any h, investment in the risky asset is constant, because the investment functions depend only on the past realizations of returns and dividend yields, equal to $\bar{r}$ and $\bar{e}$ respectively:

$$\bar{x}_{t,h} = f_h(\bar{r}\,, \ldots, \bar{r}\,, \bar{e}\,, \ldots, \bar{e}\,) = \bar{x}_h.$$

At the same time conditions (2.25) implies that average investment shares are equal for all types of investors. Thus the trajectories of wealth satisfy:

$$\bar{w}_{t-1,h} = w_{0,h}\,[1 + \bar{x}(\bar{r} + \bar{e})]^{t-1}\,.$$

Thus equation (2.27) simplifies to:

$$\frac{\bar{e}\bar{x}}{1-\bar{x}} = \bar{r},$$

which is equivalent to:

$$\bar{x} = \frac{\bar{r}}{\bar{r} + \bar{e}}. \tag{2.28}$$

We need to verify that (2.28) is compatible with the multi-horizon dynamics of the investment shares, characterized by passiveness of a part of agents at some time periods. Recall that the investment in risky asset of each type of agents h is the mean of investments of agents that readjusted their portfolios with $0, \ldots, h-1$ periods ago. But whenever the readjustment takes place, the investment share, depending on lagged returns and dividend yields, always takes the same value $\bar{x}^0_h$. At the next period after portfolio readjustment the investment share of the passive investor becomes:

$$\frac{\bar{x}^0_h(1+\bar{r})}{1 + \bar{x}^0_h(\bar{r} + \bar{e})}$$

We define the function $g = \Re \to \Re$ as:

$$g(x) = \frac{x(1+\bar{r})}{1 + x(\bar{r} + \bar{e})} \tag{2.29}$$

and $g^k(x)$ as a k-times composition of function $g(\cdot)$, that is $g \circ g \ldots \circ g(x)$, with $g^0(x)$ defined as $g(x) = x$. Then for any h we have:

$$\bar{x}_{\,h} = \frac{1}{h}\sum_{k=0}^{h} g^k(\bar{x}^{\,0}_h) \tag{2.30}$$

Now notice that

$$g(\frac{\bar{r}}{\bar{r} + \bar{e}}) = \frac{\bar{r}}{\bar{r} + \bar{e}},$$

which implies that $\bar{x}_h^0 = \frac{\bar{r}}{\bar{r}+\bar{e}}$ satisfies equation (2.30). This proves that if the equilibrium return exists, it satisfies:

$$\bar{r} = \frac{\bar{e}\bar{x}}{1-\bar{x}},$$

and thus is uniquely defined. □

We can now study the properties of the stochastic process for the risky returns and compare the results with those, obtained for the one-scale case. As before, we will proceed by the linearization of the dynamic system. Define the following function $F : \Re^{t-1} \times \Re^{t-1} \times \Re \times \Re \to \Re$:

$$\begin{gathered} F(r_1, \ldots, r_{t-1}, e_1, \ldots, e_{t-1}, r_t, e_t) = \\ \frac{\sum_{h=1}^{H} w_{t-1,h} \left[x_{t,h} - x_{t-1,h} + e_t x_{t,h} x_{t-1,h} \right]}{\sum_{h=1}^{H} w_{t-1,h} x_{t-1,h} \left(1 - x_{t,h}\right)} - r_t \end{gathered} \tag{2.31}$$

with $x_{t,h}$ defined as in (2.23). The following theorem describes the equilibrium dynamics in the neighborhood of the solution of the deterministic analog of the system.

Theorem 2.12. *In the model with homogeneous rational agents and multiple investment scales, the return process is approximately described by:*

$$\begin{aligned} r_t &= \bar{r} + \hat{r}_t + \bar{V}\widetilde{e}_t, \\ \hat{r}_t &= \sum_{k=1}^{H-1} A_k \hat{r}_{t-k} + V_t \widetilde{e}_t, \\ V_t &= \sum_{k=1}^{H-1} B_k \hat{r}_{t-k} \end{aligned} \tag{2.32}$$

where:

$$\begin{aligned} A_k &= \frac{a_k - b_k}{(1-c)(1+\bar{r})^k}, \\ B_k &= \frac{\bar{x}\,(1-2\bar{x})(c\,b_k - a_k)}{(1-\bar{x})(1+\bar{r})^{k+1}(1-c)}, \\ \bar{V} &= \frac{\bar{x}}{1-\bar{x}}, \\ a_k &= \sum_{h=k+2}^{H} \frac{h-k-1}{h} \xi_{0,h}\,, \\ b_k &= \sum_{h=k+1}^{H} \frac{h-k}{h} \xi_{0,h}\,, \\ c &= \sum_{h=1}^{H} \frac{h-1}{h} \xi_{0,h}\,. \end{aligned}$$

Proof. See Appendix. □

The result of theorem 2.12 shows that the return dynamics in the multiscale model with rational investors is very close to the one-scale dynamics in the rational expectations case, the only difference being the term $\hat{r}_t$. It represents the deviation from the hypothetical trajectory of returns, that would be realized in a one-scale market, and can be interpreted as error correction term. Note that there is no constant in the volatility of the disturbance term, which means that the correction term either vanishes or explodes, depending on the values of coefficients A_k, B_k and the variance of $\widetilde{e}_t$. We will study its behavior for plausible values of parameters in the following section. Note that the dynamics in the multiscale case is considerably different from the general one-scale case, described by theorem 2.6. There the terms, containing serial correlations and heteroscedasticity, are not vanishing, while in the multiscale rational case their presence is temporal after a shock, in the absence of which they completely disappear from the return dynamics.

Theorem 2.12 refers to the case, when investors' demand functions at the times of portfolio readjustment are trivial: investment shares are constant at the level, corresponding to the rational equilibrium, which coincides with the one-scale equilibrium. In practice, investment decisions may depend on the historical returns, so the framework of procedural rationality would be more adequate for modeling. One can establish a general analytical representation of the return dynamics in this case. Indeed, equations (2.41), (2.43) and (2.44) in the proof of theorem 2.12 (see Appendix) still hold, but instead of (2.42) we need to adopt a general form for the investment functions, as in theorem 2.6. However, in our view, such general representation would be of little practical value. Instead, using simulation, we explore the return dynamics, corresponding to concrete stylized examples of investment functions. This issue is addressed in the following section.

2.5 Simulation Study

We determined the equations of the risky return dynamics in the case of the market, populated with rational participants, acting at one and several investment horizons. We also established the framework for the study of the procedurally rational investment, that can incorporate behavioral patterns, such as trend extrapolation and contrarian strategies. Our goal in this section is to explore the empirical properties of the return series, generated by different versions of our model, and to associate the properties of

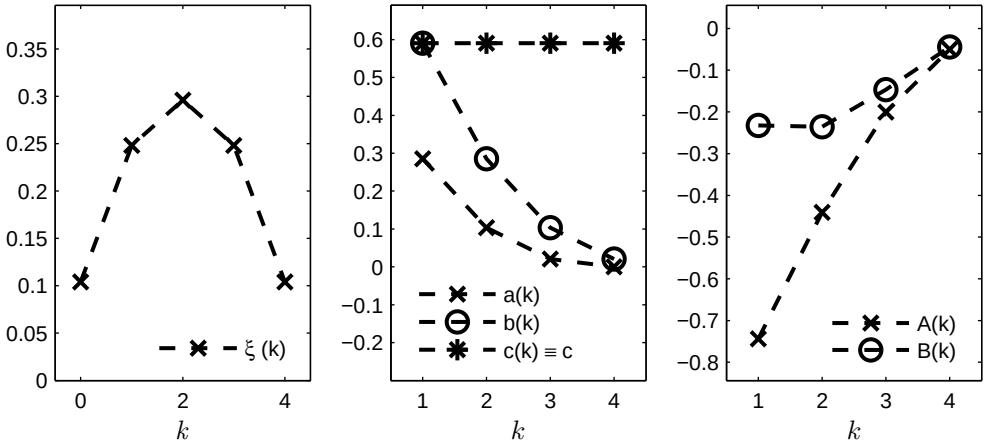
the model with the stylized patterns, observed on real market data: contrarian returns, trend formation and conditional heteroscedasticity.

From observation of (2.32) it is clear that introducing multiple scales changes the way, in which the dynamic system for the risky return reacts to shocks. These shocks could be of completely exogenous or of behavioral nature. We will first study the case when, along with the "normal" disturbance term, interpreted as dividend yield, the model is occasionally perturbated by exogenous shock on returns, unrelated to the investment functions. Such abnormal returns can reflect deviations from market clearing equilibrium at some time periods.

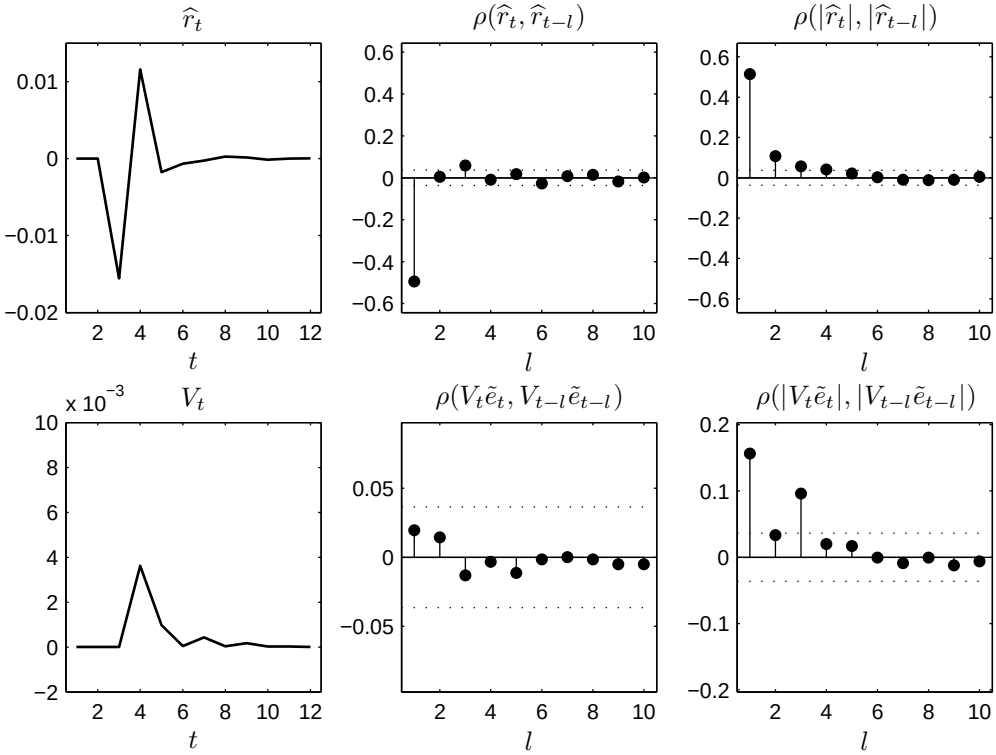
Returns trajectories are simulated for a market with five horizons, where abnormal returns occur at random periods, on average once per 50 trades. We are interested in the values of coefficients A_k and B_k, that determine the way the shock at period t is reverberated at future dates. Note that in a one-scale model such shocks have absolutely no incidence on future returns. The above-mentioned coefficient depend on a_k, b_k and c, that characterize how initial wealth is distributed among investors.

FIGURE 2.1: Shocks to returns, ∩-shape wealth distribution

(a) Model Parameters



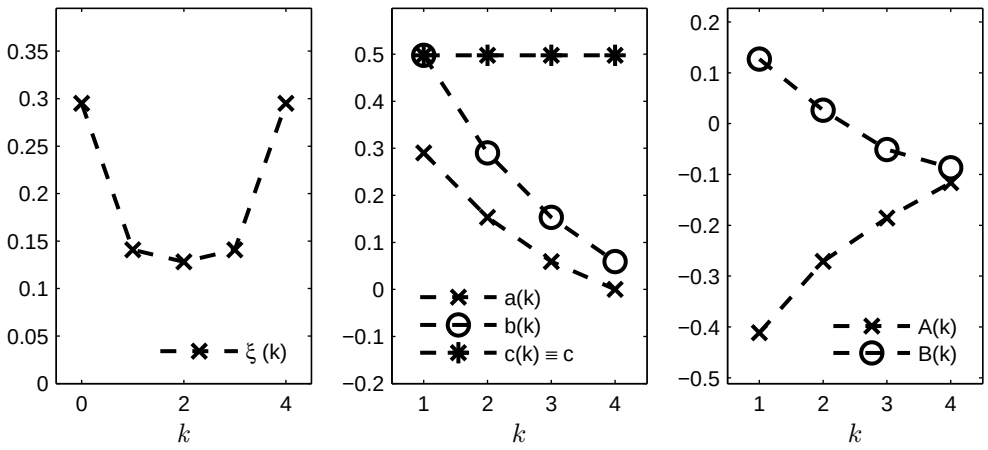
(b) Reaction to Exogenous Shocks



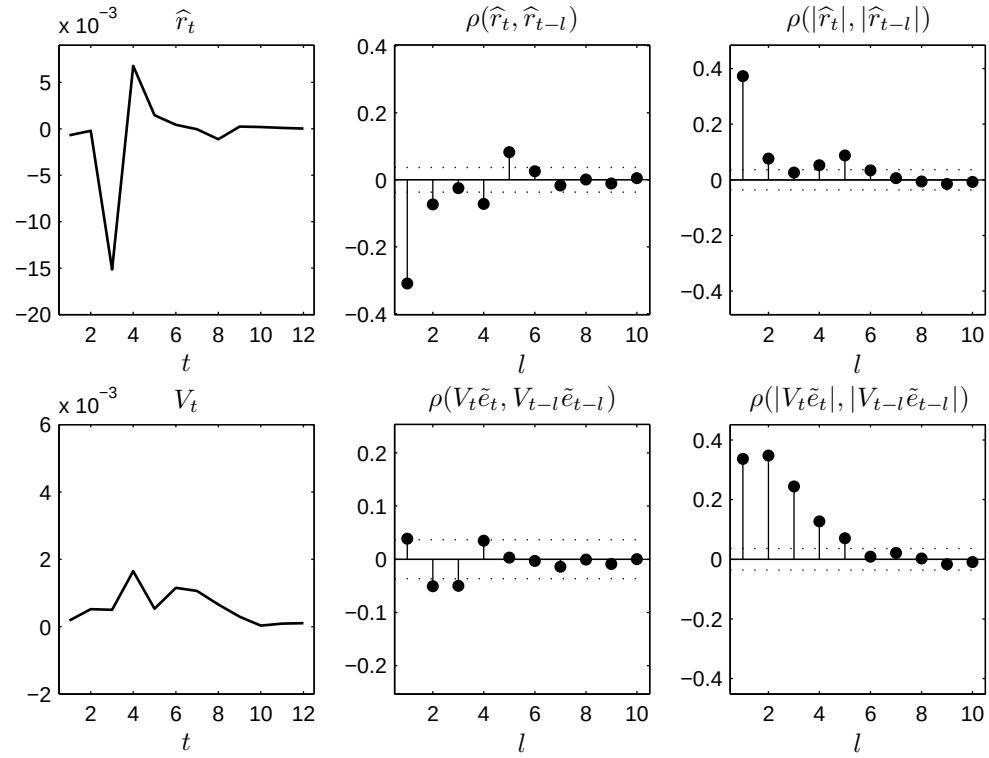
Model with $H = 5$, where initial wealth is distributed according to discretized $\beta(2,2)$ distribution, $e \sim N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Gaussian shocks with variance $\frac{\bar{x}^2}{(1-\bar{x})^2}\sigma_\epsilon^2$ are applied to the $\widetilde{r}_t$ series at random dates with frequency $\frac{1}{10\,H}$, i.e. on average every 50 points. Autocorrelations are estimated on a 10 000 - periods simulation path.

FIGURE 2.2: Shocks to returns, ∪-shape wealth distribution

(a) Model Parameters



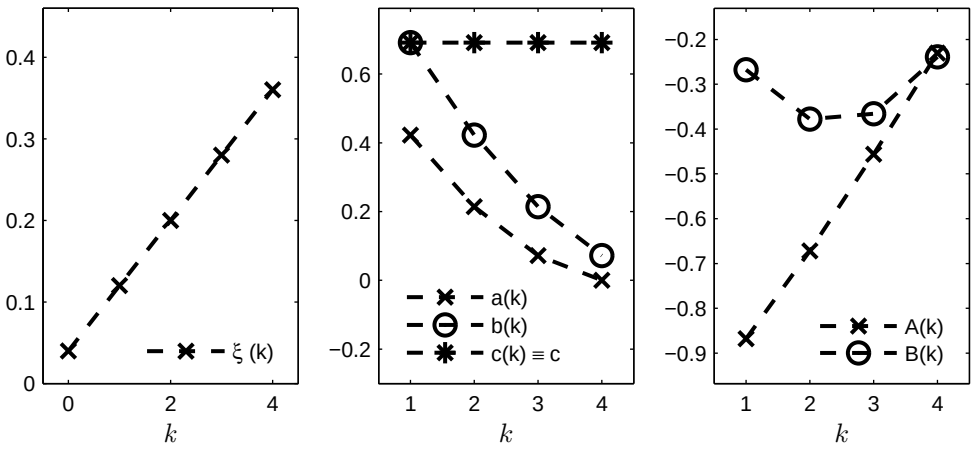
(b) Reaction to Exogenous Shocks



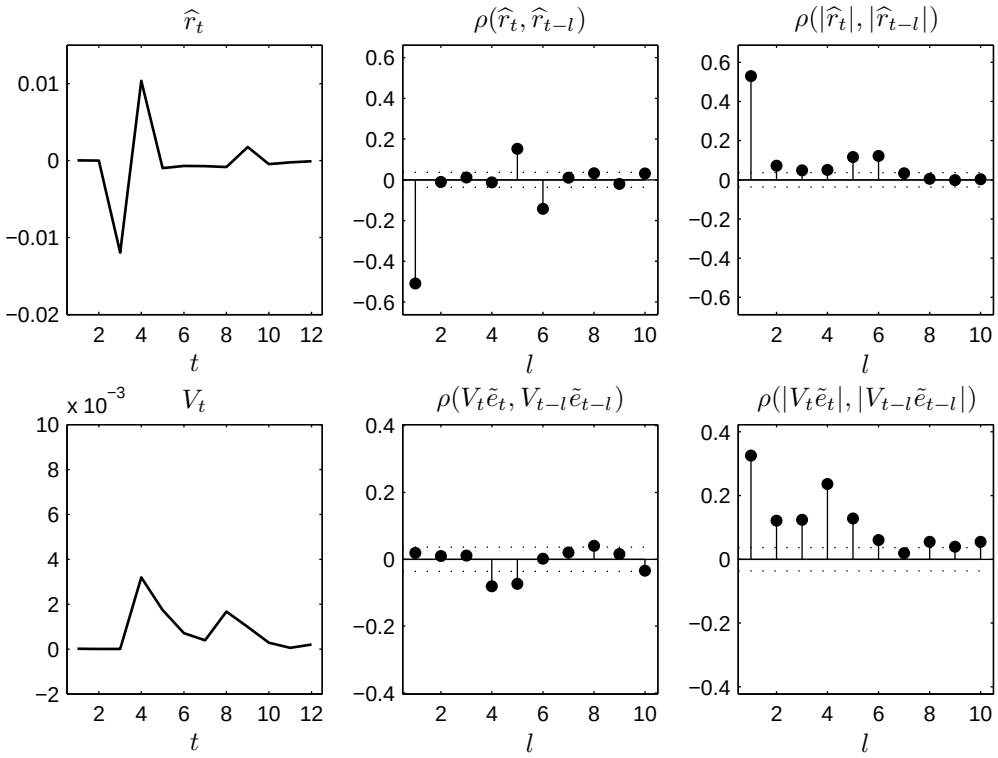
Model with $H = 5$, where initial wealth is distributed according to discretized $\beta(2,2)$ distribution, $e \sim N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Gaussian shocks with variance $\frac{\bar{x}^2}{(1-\bar{x})^2}\sigma_\epsilon^2$ are applied to the $\widetilde{r}_t$ series at random dates with frequency $\frac{1}{10\,H}$, i.e. on average every 50 points. Autocorrelations are estimated on a 10 000 - periods simulation path.

FIGURE 2.3: Shocks to returns, wealth shares increase with investment horizon

(a) Model Parameters



(b) Reaction to Exogenous Shocks



Model with $H = 5$, where initial wealth is distributed according to discretized $\beta(2, 2)$ distribution, e $N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Gaussian shocks with variance $\frac{\bar{x}^2}{(1-\bar{x})^2}\sigma_\epsilon^2$ are applied to the $\widetilde{r}_t$ series at random dates with frequency $\frac{1}{10\,H}$, i.e. on average every 50 points. Autocorrelations are estimated on a 10 000 - periods simulation path.

On Figures 2.1 - 2.3 we report the results for three characteristic cases, respectively: ∩-shape density of wealth distribution around the maximum at horizon 3, ∪-shape with symmetric distribution of wealth with maxima at the shortest and longest horizons and the case, when the wealth share linearly increases with investment horizon, so that more wealth is allocated to long-term investors. The first row of plots on each figure represents the wealth distribution and coefficients a_k, b_k, c, A_k and B_k. The second row represents examples of the trajectory of the error correction term $\hat{r}_t$ for 10 periods after a shock occurs and sample autocorrelation function for $\hat{r}_t$ and $|\hat{r}_t|$, estimated over a 10 000 trading periods simulation path. The third row zooms on the volatility component of the error term, V_t.

Coefficients A_k are always negative at their magnitude decreases with the lag at the rate, depending on the form of the initial wealth distribution. In all cases, this leads to significant anticorrelation with lag one. This result is spectacular because it shows that even without any behavioral or other hypotheses about investment functions our multiscale model generates "contrarian returns".

The presence of small, but significant serial anticorrelations in the series of stock returns is one of the stylized facts about stock price dynamics, known since Fama (1965). In practice, the presence of statistical arbitrage can reduce these autocorrelations, but these possibilities are limited by various transaction costs, so anticorrelations at daily frequencies often remain noticeable (see Jegadeesh and Titman, 1995). In more recent studies, anticorrelations in returns are evidenced for many "other than US" markets (see Lee et al., 2003). Our findings add one more possible explanation.

Non-technically, the error-correction effect in our model can be described as follows. At period t abnormally high (or low) return drives upwards (or downwards) passive investors' share of wealth, invested in the risky asset. At period $t+1$ those of them, who participate in the market, readjust their portfolios to achieve the target allocation. This triggers risky asset return and wealth of passive investors in the direction, opposite to the initial shock. At the next period, investors, who were passive in the previous two periods and currently participate in the trade, readjust their wealth shares with regards to the composite effect of the two previous fluctuations, an so on.

The conditional heteroscedasticity effect is also present and its importance depends on the wealth distribution across scales, with more wealth allocated to longer scales meaning more memory in volatility. However, this effect is relatively small in magnitude and is only slightly reflected in the autocorrelogram of the $|\hat{r}_t|$ series, dominated by the anticorrelation effect. It is almost unnoticeable in the r_t series, to which white noise with variance $\frac{\bar{x}}{1-\bar{x}}\sigma_e^2$ is added (in our example, the noise standard deviation is 3×10^{-3}).

Now consider another type of shock - deviation of the investment function from the equilibrium level. We simulate models with 1-5 horizons, perturbated by random small and non-persistent fluctuations of $x^0_{t,h}$, that occur on average once per $10\,h$ periods. In Table 2.1 we report the mean and standard deviation levels for the models with different number of scales. The average return remaining constant, we observe that the global volatility level drops down as the number of horizons increases. In a multi-scale model exogenous shocks are somewhat diluted, because many of the investors do not participate in the trades when the shock occurs. They are, however, affected by the abnormal return, generated at this period.

TABLE 2.1: Volatility of returns in the multi-scale model with shocks to the investment function

	1 scale	2 scales	3 scales	4 scales	5 scales
Mean return	0.009	0.009	0.009	0.009	0.009
Volatility	0.025	0.021	0.017	0.015	0.014

Models with $H = 1, \ldots, 5$, where initial wealth is distributed according to discretized $\beta(2,2)$ distribution, $e \sim N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Gaussian shocks with variance 0.01^2 are applied to the demand functions series at random dates with frequency $\frac{1}{10\,H}$, i.e. on average every $10-50$ points, depending on investment scale. Parameters are estimated on a 10 000 - periods simulation path.

On Figure 2.4 we represent the dynamic properties of returns in the models with different number of scales. For the one-scale case we observe large abnormal returns at the periods of shocks and large returns at the following period, explained by the reversion to normality of the risky asset weights in investors' portfolios. On the ACF for returns we find strong negative autocorrelation at the first lag, characteristic to the MA(1) process. Note that this is not the case for the shocks in returns, unrelated to demand, that were discussed previously. In the one-scale model, the latter do not trigger subsequent correction.

In the multi-horizon model with h scales, the shocks to $x^0_{t,h}$ have impact on the aggregated investment function $x_{t,h}$ during $h-1$ periods, the time necessary to rebalance all portfolios after a shock. This causes a lasting impact on the volatility term $\frac{x_t}{1-x_t}$ and creates deviations from the equilibrium trajectory, defined by theorem 2.12 with $\hat{r}_t = 0$. The deviation, in its turn, "activates" the error correction mechanism, described above in details. We find that, in the case of shocks to the demand function, the conditional heteroscedasticity effect is no longer negligible. It manifests itself by the emergence of the significant autocorrelations in absolute returns up to $h-1$ lags. For illustration, we calibrate an MA(1) - GARCH(1,1) model of the form $r_t = \mu + \varepsilon_t + \alpha\varepsilon_{t-1}$, $\varepsilon_t \sim T(0, \sigma_t, k)$, $\sigma_t = c + \phi\, r_{t-1} + \theta\, \sigma_{t-1}$ (see Table 2.2) and find significant moving average and autoregressive components in the conditional volatility process.

For the moment we assumed that disturbances to the system were purely exogenous. We did not suppose any particular type of behavior, such as trend-following or contrarian, and we did not make the shocks depend on the past history of returns. It is important that even in this simple case we find that the MIS model can generate interesting dynamic patterns in returns. Certainly, deviations from rational equilibrium in investment functions can be associated to investors' behavior and can be present at every period, unlike our stylized example, designed for illustration purposes.

TABLE 2.2: Estimation of GARCH(1,1) in a 5-scale model

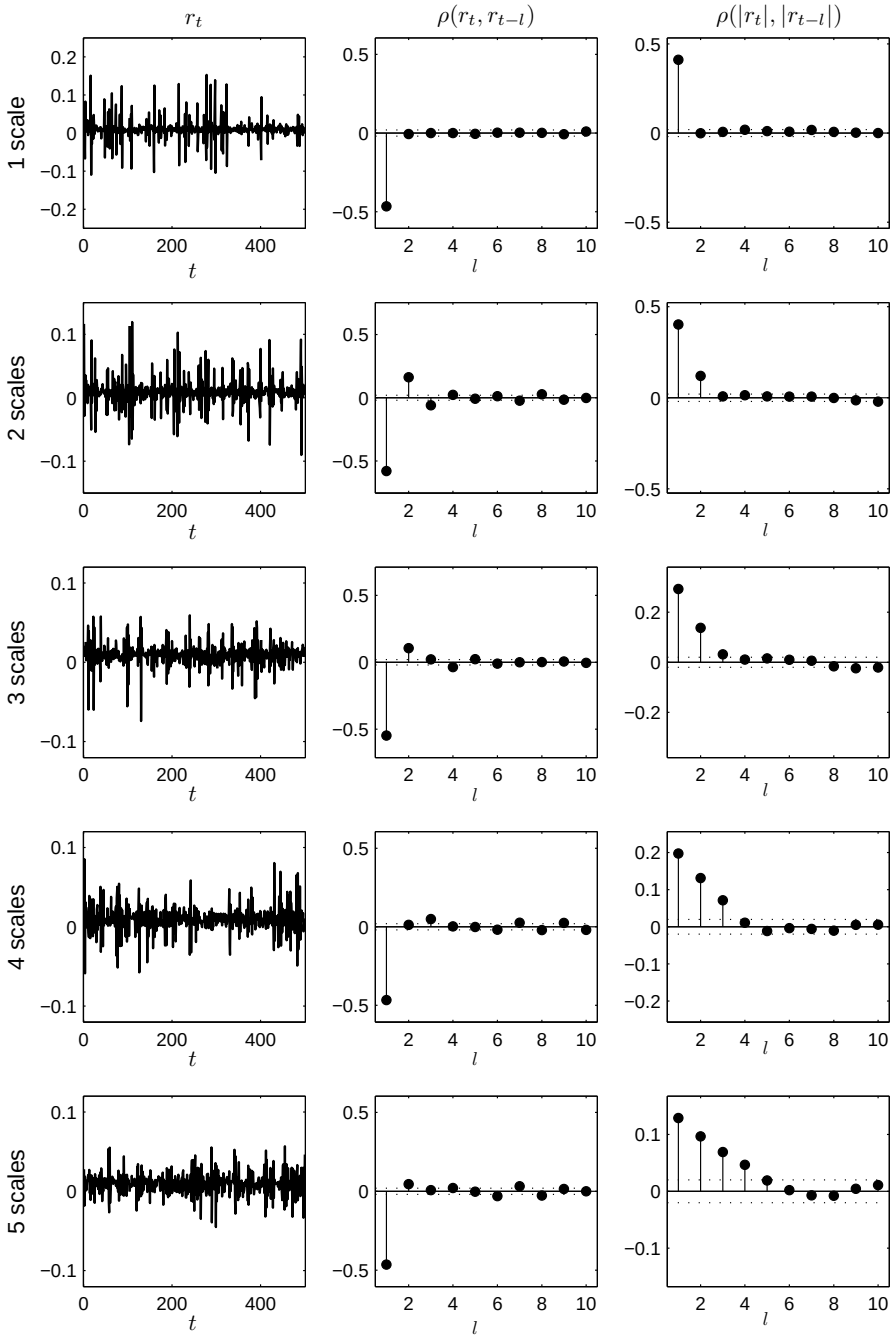
	Value	Standard Error	T-Statistic
$\widehat{\mu}$	0.0091	6×10^{-5}	165.0157
$\widehat{\alpha}$	-0.4103	0.0102	-40.1992
$\widehat{c}$	6×10^{-5}	4×10^{-6}	14.7496
$\widehat{\theta}$	0.2684	0.0314	8.5379
$\widehat{\phi}$	0.3702	0.0290	12.7593
$\widehat{k}$	4.2249	0.2228	18.9672
Log-likelihood:	31 025		

Data generating process: models with 5 horizons, where initial wealth is distributed according to discretized $\beta(2,2)$ distribution, $e \quad N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Gaussian shocks with variance 0.01^2 are applied to the demand functions series at random dates with frequency $\frac{1}{10\,H}$, i.e. on average every $10-50$ points, depending on the investment scale. 10 000 simulations. Estimated model specification: $r_t = \varepsilon_t + \alpha\varepsilon_{t-1}$, $\varepsilon_t \quad T(0, \sigma_t, k)$, $\sigma_t = c + \phi\, r_{t-1} + \theta\, \sigma_{t-1}$.

On Figure 2.5 we present the results of simulation of a five-scale model with procedurally rational investors that have investment functions, corresponding to definition 2.4 and the equation of the mean expected returns is of the form (2.11B). For simplicity, we suppose that $L = \infty$, so that the "long term" estimate of average return is the rational equilibrium mean return. We choose l to be equal to h, so that the short-term estimate is in fact the last return, observed on each scale. The parameter d, if the behavior of investor is contrarian ($d < 0$) or trend-following ($d > 0$). Note that $d > 0$ can also correspond to the case, when investors are fundamentalists, but estimate expected returns on a short historical sample.

Also for simplicity, we assume that agents' strategies are switching, so that the coefficient d can take three values: $d_{TF} = 3 \times 10^{-3}\sqrt{h}$, $d_C = 1.5 \times 10^{-3}\sqrt{h}$ and $d_R = 0$. The value d_{TF} is chosen higher than d_C because the trend-following behavior seems to be more common, but also because this specification ensures insignificant autocorrellations of returns, making the simulation results more realistic. The normalization factor $\sqrt{h}$ corresponds to the speed of convergence of the empirical mean estimate to the true

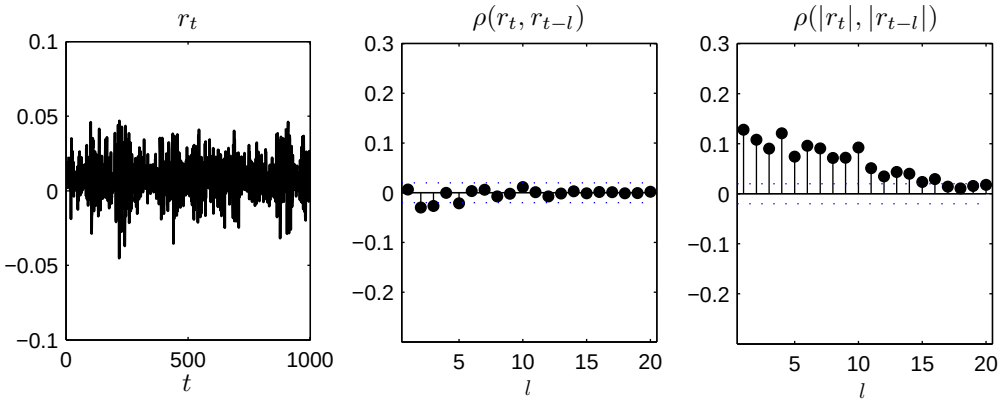
FIGURE 2.4: Shocks to the investment functions



Models with $H = 1, \ldots, 5$, where initial wealth is distributed according to discretized $\beta(2, 2)$ distribution, $e \approx N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Gaussian shocks with variance 0.01^2 are applied to the demand functions series at random dates with frequency $\frac{1}{10\,H}$, i.e. on average every $10 - 50$ points, depending on the investment scale. Autocorrelations are estimated on a 10 000 - periods simulation path.

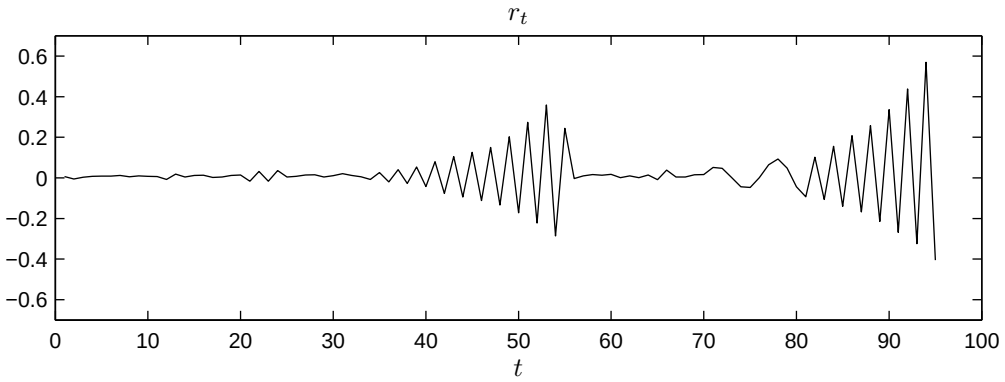
value and ensures that the magnitude of noise in investment function is the same at all scales. The strategies are chosen independently by investors at each scale according to a Markovian transition matrix. At each period the probability to continue using the same

FIGURE 2.5: Multiscale model with switching between contrarian and trend-following investment



Model with $H = 5$, where initial wealth is distributed according to $\beta(2,2)$-distribution, $e \sim N(0.03, 0.02^2), \bar{x} = 0.75$. Switching rules are described in the text. 10 000 simulations.

FIGURE 2.6: One-scale model with switching between contrarian and trend-following investment



Model with $H = 1$, $e \sim N(0.03, 0.02^2), \bar{x} = 0.75$. Switching rules are described in the text. Simulation until the market "explodes" (no market clearing possible) - 96 periods.

strategy, as in the previous periods, is $1 - \frac{1}{10\,h}$ and the probability to switch to one of the other two states is $\frac{1}{20\,h}$[6].

The resulting dynamics is characterized by insignificant autocorrelations in returns and low, but lasting and significant autocorrelations in absolute returns, which is very characteristic of real stock markets (see, for example Subbotin et al. (2010b) or Cont (2001)).

[6]We could make the switching endogenous and depending on the success of the corresponding strategy in the past, as in Brock and Hommes (1998). This would be theoretically more justified but more complicated, while the effect on the return dynamics is barely the same. Our focus being on the multiple scales and not on the behaviorist patterns, we prefer the simpler option.

Recall that the ARCH-effect can also be obtained in a one-scale model with the same type of investment strategies switching. The important feature of the multi-scale case is that is smooths and mutually mitigates the effects that are present on different scales, so that important deviations from the rational behavior do not lead to explosions of the risky asset price. For illustration, consider the one-scale model with the same investment function, as in our five-scales example. The system is unstable and explodes, as shown on Figure 2.6. Thus, introducing multiple scales we extend the set of available strategies and add stability to the dynamic system.

2.6 Conclusion

We have shown that the risky return process in the rational equilibrium for investors with constant relative risk aversion is similar for the case of one and multiple investment scales. However, this result does not hold if the system is subject to shocks, exogenous or related to deviations from rational behavior. In fact, the main difference between the multi-scale and the one-scale models is the way the resulting dynamic system reacts to shocks.

We first demonstrated that the multi-scale model with exogenous shocks to returns displays serial anticorrelations, which is in line with empirical evidence on the so-called contrarian profits. Popular explanations of reversion in the stock returns include overreaction to firm specific information (Jegadeesh and Titman, 1995), measurement errors, related to the bid-ask spread (Boudoukh et al., 1994; Conrad et al., 1997), lead-lag effect, supposing that some stocks react to news faster than others (Chou et al., 2007; Lo and MacKinlay, 1990), time varying systematic risk (Chan, 1988) and some others. Our model contains no overreaction, but includes error correction that consists in temporarily passive market participants' tendency to adjust their portfolios back to the target weights, after the latter have deviated because of price fluctuations. Our theoretical findings contribute to the mean reversion literature, offering one more plausible explanation for the anticorrelation in returns.

More importantly, we demonstrate that the multi-scale model with shocks to investment functions generates conditional heteroscedasticity. Up to now, explanations of conditional heteroscedasticity in the theoretical literature were almost exclusively based on the switching between contrarian and trend-following strategies (Anufriev et al., 2006; Brock and Hommes, 1998; Chiarella and He, 2001; Weinbaum, 2009) or on wealth-dependent relative risk aversion (Vanden, 2005) . Unlike the first class of models, we do not need to stipulate any special patterns in investment behavior. Even exogenous *iid* disturbances to demand functions generate the GARCH effects. The mechanics of

this effect is methodologically close to the relative risk aversion step function model in Vanden (2005). In both models volatility depends on the demand for the risky asset, expressed as portion of wealth. In Vanden (2005) volatility changes due to the step-wise dependence of the relative risk aversion on wealth. Thus the relation between the price of the risky asset and the value of investment in it cannot remain constant. In our model the demand does not perfectly follow price fluctuations, because some of the agents remain passive at each trading date.

Our multi-scale framework is perfectly compatible with the analysis of the contrarian and the trend-following behavior. Our analysis is restricted to the so-called procedurally rational investment strategies, ensuring that, in the case of the *iid* returns, market participants do not make systematic errors in estimating mean and variance. We prove that in the one-scale case, which was earlier studied in Anufriev et al. (2006), equilibrium returns are never *iid*, unless procedurally rational investment functions degenerate to the truly rational (constant) investment functions. Naturally, this result also holds for the multi-scale case. One of the consequences of introducing the multiple scales is that the behavioral shocks to investment functions are smoothed in time and lessened in magnitude, which adds stability to the system.

We notify the reader that, though our results are rather general, they are nevertheless based on a series of strong assumptions. The most important of them include *(i)* the absence of information signals, related to future returns, other than contained in historical prices; *(ii)* constant and exogenous frequencies of market participants' interventions in the market; *(iii)* the assumption that the dividend yield is and *iid* random variable. These assumptions, whose implications on our model are briefly discussed below, can be subjects of further model extensions.

Introducing signals in our model would lead to further heterogeneity of market agents, that will have different times of reaction to news. Obviously, investors acting at large scales would react to news with some delay, which can reinforce the effects, already observed in the multi-scale model. Technically, the model with signals can be close to the one studied in this paper. Our dynamic system is innovated with one variable interpreted as dividend yield. It could be replaced by agents' expectation of expected futures dividends, with one noteworthy difference that this expectation would not be immediately accounted in wealth.

We supposed that frequencies of trading are fixed and exogenous, but they could be made dependent on market conditions. It would be natural to assume that, in a high volatility market, investors tend to readjust their portfolios more often. Trading frequency could be made completely endogenous by introducing implicit transaction

costs. In particular, this approach could be helpful in explaining the patterns in trading volume.

Finally, our model, similarly to Chiarella and He (2001) and Anufriev et al. (2006), suggests that the average dividend yield is proportional to price. In economic terms, this means that corporate profits exhibit constant returns to scale, which implies that the equilibrium return can be determined endogenously by the risk-return preferences of the market agents. It would be interesting to study the impact of heterogeneity in the investment behavior in a more consistent macroeconomic setting, where dividend yields are limited by the real economies' productivity.

All the three extensions can be based on the dynamic model with multiple investment scales, studied in this paper. The importance of the results, presented here, is that they establish a general framework, which can be used for further research on more specific problems.

2.7 Appendix. Proof of Theorems and Propositions.

Proof of proposition 2.2

Proof. Consider an investor, who does not operate on the market in current period. He has a portion x_t of his wealth invested in the risky asset. In the next period the wealth he detains in the risky shares becomes (we omit the subscript i for simplicity):

$$\frac{x_t W_t P_{t+1}}{P_t} = \frac{x_t w_t p_{t+1}(1+R_f)^{t+1}}{p_t} = x_t w_t (1+r_{t+1})(1+R_f)^{t+1}$$

His wealth invested in the risk-free asset reads:

$$W_t(1-x_t)(1+R_f) + \frac{D_{t+1} W_t x_t}{P_t} = w_t(1-x_t)(1+R_f)^{t+1} +$$
$$+ e_{t+1} w_t x_t (1+R_f)^{t+1} = (1+R_f)^{t+1} w_t [1 - x_t(1-e_{t+1})]$$

Therefore, the share of investment in the risky asset satisfies:

$$x_{t+1}^{-1} = \frac{x_t w_t (1+r_{t+1})(1+R_f)^{t+1}}{x_t w_t (1+r_{t+1})(1+R_f)^{t+1} + (1+R_f)^{t+1} w_t [1 - x_t(1-e_{t+1})]} =$$
$$= \frac{x_t(1+r_{t+1})}{1 + x_t(r_{t+1} + e_{t+1})}$$

Equation (2.4) is obtained if instead of x_t the same argument is applied to x_{t-1}^{-k+1}. □

Proof of Proposition 2.3

Proof. The evolution of an investor's wealth between t and $t+h$ reads (we drop the subscript i to ease the notation):

$$W_{t+h} = x_t \frac{W_t}{P_t} \left(P_{t+h} + \sum_{k=1}^{h} D_{t+k} (1+R_f)^{h-k} \right) + (1-x_t) W_t (1+R_f)^h.$$

Let $Y_{t,t+h}$ be the total return on the risky asset for the passive investor between t and $t+h$. Then we have:

$$Y_{t,t+h} = \frac{P_{t+h}}{P_t} + \sum_{k=1}^{h} \frac{D_{t+k}}{P_t} (1+R_f)^{h-k} - 1 =$$
$$\prod_{k=1}^{h} (1+R_{t+k}) + \sum_{k=1}^{h} \varepsilon_{t+k} \prod_{i=1}^{k-1} (1+R_{t+i})(1+R_f)^{h-k} - 1.$$

The solution of the maximization problem (2.6) reads:

$$x_t^* = \frac{1 + \mathrm{E}_{t-1}[Y_{t,t+h}] - (1+R_f)^h}{\gamma \mathrm{Var}_{t-1}[Y_{t,t+h}]}. \tag{2.33}$$

For small returns, (2.33) can be approximated by:

$$x_{t,i}^* = \frac{1 + \mathrm{E}_{t-1}[Y_{t,t+h}] - h\,R_f}{\gamma_i \mathrm{Var}_{t-1}[Y_{t,t+h}]},$$

so that in terms of rescaled variables the total riksy return can be written:

$$1 + Y_{t,t+h} \stackrel{def}{=} (1+R_f)^h (1 + y_{t,t+h}) =$$
$$(1+R_f)^h \left[\prod_{k=1}^{h} (1 + r_{t+k}) + \sum_{k=1}^{h} e_{t+k} \prod_{i=1}^{k-1} (1 + r_{t+i}) \right],$$

or with the first order precision:

$$1 + Y_{t,t+h} \approx 1 + y_{t,t+h} \approx h\,R_f + \sum_{k=1}^{h} r_{t+k} + e_{t+k}.$$

For the optimal investment share, this gives approximative solution:

$$x_{t,i}^* \approx \frac{\mathrm{E}_{t-1}[y_{t,t+h}]}{\gamma_i \mathrm{Var}_{t-1}[y_{t,t+h}]}$$

□

Proof of Theorem 2.6

Proof. Equation (2.12) can be written in the form:

$$r_t = \frac{x_t - x_{t-1} + \bar{e}\,x_t x_{t-1}}{x_{t-1} - x_t x_{t-1}} + \frac{x_t x_{t-1}}{x_{t-1} - x_t x_{t-1}} \widetilde{e}_t.$$

We define the function:

$$F(r_{t-1}, \ldots, r_{t-L-1}) = \frac{x_t - x_{t-1} + \bar{e}\,x_t x_{t-1}}{x_{t-1} - x_t x_{t-1}}.$$

The function F depends on the history of returns upto the lag $L+1$ because it the term x_{t-1}, determined by investors from L observations of returns prior to $t-1$. Denote F'_k the derivative of F with respect to its argument number k, evaluated in the point

$(\bar{r}\,,\ldots,\bar{r}\,)$. Then we have:

$$F_1' = \frac{(f_1' + f_1'\bar{x}\,\bar{e}\,)[\bar{x}\,(1-\bar{x}\,)] + \bar{x}\,f_1'(\bar{x}\,-\bar{x}\,+\bar{x}\,^2\bar{e}\,)}{\bar{x}\,^2(1-\bar{x}\,)^2} = \frac{f_1'[1-\bar{x}\,(1-\bar{e}\,)]}{\bar{x}\,(1-\bar{x}\,)^2}$$

For $k = 2,\ldots,L$ we obtain:

$$F_k' = \frac{\bar{x}\,(1-\bar{x}\,)[f_k' - f_{k-1}' + \bar{x}\,\bar{e}\,(f_k' + f_{k-1}')] - \bar{x}\,^2\bar{e}\,[f_{k-1}'(1-\bar{x}\,) - \bar{x}\,f_k']}{\bar{x}\,^2(1-\bar{x}\,)^2} = \frac{f_k'[1-\bar{x}\,(1-\bar{e}\,)] + f_{k-1}'(\bar{x}\,-1)}{\bar{x}\,(1-\bar{x}\,)^2}$$

Finally, for the last term:

$$F_{L+1}' = \frac{\bar{x}\,(1-\bar{x}\,)(-f_L' + \bar{x}\,\bar{e}\,f_L') - \bar{x}\,^2\bar{e}\,(1-\bar{x}\,)f_L'}{\bar{x}\,^2(1-\bar{x}\,)^2} = -\frac{f_L'}{\bar{x}\,(1-\bar{x}\,)}$$

In a similar way, define

$$G(r_{t-1},\ldots,r_{t-k-1}) = \frac{x_t x_{t-1}}{x_{t-1} - x_t x_{t-1}}.$$

and compute its derivatives:

$$G_1' = \frac{\bar{x}\,(1-\bar{x}\,)\bar{x}\,f_1' + \bar{x}\,f_1'\bar{x}\,^2}{\bar{x}\,^2(1-\bar{x}\,)^2} = \frac{f_1'}{(1-\bar{x}\,)^2}$$

For $k = 2,\ldots,L$:

$$G_k' = \frac{[\bar{x}\,f_k' + \bar{x}\,f_{k-1}']\bar{x}\,(1-\bar{x}\,) - \bar{x}\,^2[f_{k-1}'(1-\bar{x}\,) - \bar{x}\,f_k']}{\bar{x}\,^2(1-\bar{x}\,)^2} = \frac{f_k'}{(1-\bar{x}\,)^2}$$

The last term reads:

$$G_{L+1}' = \frac{\bar{x}\,f_L'\bar{x}\,(1-\bar{x}\,) - \bar{x}\,^2 f_L'(1-bx)}{\bar{x}\,^2(1-\bar{x}\,)^2} = 0$$

Using these results, we can rewrite the stochastic process (2.12) for returns in a linearized form:

$$\widetilde{r}_t = -\bar{r}\, + F(\bar{r}\,,\ldots,\bar{r}\,) + \sum_{k=1}^{L+1} F_i'\widetilde{r}_{t-k} + v_t\widetilde{e}_t$$

$$v_t = G(\bar{r}\,,\ldots,\bar{r}\,) + \sum_{k=1}^{L} G_i'\widetilde{r}_{t-k}$$

Since $\widetilde{r}_t$ is a zero-mean process, we impose:

$$\bar{r} = F(\bar{r}, \ldots, \bar{r}) = \frac{\bar{x}\,\bar{e}}{1-\bar{x}}$$

which immediately gives (2.19). □

Proof of Theorem 2.7

Proof. The equilibrium dynamics of returns can be derived from (2.2) in a way, similar to the homogeneous case:

$$r_t = \frac{\sum_i w_{t-1,i}\,(x_{t,i} - x_{t-1,i}) + e_t \sum_i x_{t,i}x_{t-1,i}w_{t-1,i}}{\sum_i w_{t-1,i}\,(x_{t-1,i} - x_{t,i}x_{t-1,i})} \tag{2.34}$$

If r_t is *iid*, the same is true for $r_{t|t-1}$. For the mean of $r_{t|t-1}$ to be constant, it is necessary that:

$$\begin{gathered}\sum_i \phi_{t-1,i}\,(x_{t,i} - x_{t-1,i}) = c_1 \sum_i \phi_{t-1,i}\,(x_{t-1,i} - x_{t,i}x_{t-1,i}) \\ \sum_i \phi_{t-1,i}x_{t,i}x_{t-1,i} = \tilde{c}_2\,(x_{t-1,i} - x_{t,i}x_{t-1,i})\end{gathered} \tag{2.35}$$

with c_1 and $\tilde{c}_2$ two constants and $\phi_{t,i} = \frac{w_{t,i}}{\sum_i w_{t,i}}$. Then $\mathrm{E}_{t-1}(r_t) = c_1 + \tilde{c}_2\bar{e}$ and $\mathrm{Var}_{t-1}(r_t) = \tilde{c}_2^2\sigma_e^2$. Simplifying (2.35) and setting $c_2 = \frac{\tilde{c}_2}{1+\tilde{c}_2}$, we get:

$$\begin{gathered}\sum_i \phi_{t-1,i}x_{t,i} = [1 - c_1(1-c_2)] \sum_i \phi_{t-1,i}x_{t-1,i} \\ \sum_i \phi_{t-1,i}x_{t,i}x_{t-1,i} = c_2 \sum_i \phi_{t-1,i}x_{t-1,i}\end{gathered} \tag{2.36}$$

with c_1 and c_2 are two constants. The case of constant average investment function is a particular case of (2.36) when we chose $c_1 = 0$. Note that the quantity $\sum_i \phi_{t-1,i}x_{t,i}x_{t-1,i}$ must also be constant in this case.

On the other hand, if returns are *iid*, procedurally rational investors must have investment functions, based on unbiased beliefs. Thus, for any t the latter can be written in the form:

$$x_{t,i} = \frac{\mathrm{E}\,(r_t) + \upsilon_{t,i}}{\gamma_i \mathrm{Var}\,(r_t) + \zeta_{t,i}} \tag{2.37}$$

with $\upsilon_{t,i}$ and $\zeta_{t,i}$ are centered random variables. But (2.37) implies that in (2.36) the growth rate of the aggregate investment share $1 - c_1(1-c_2)$ cnanot be deterministic and different from one. Thus the case of the constant average investment function, where $c_1 = 0$, is the only one compatible with the *iid* dynamics of returns in the heterogeneous case. The mean and variance of returns in this case are proportional to the mean and

variance of the dividend yield, but are not defined uniquely and depend on the parameter c_2. □

Proof of Theorem 2.8

Proof. Let us denote

$$\begin{aligned}
\Xi(r) =& r - \frac{\sum_{h=1}^{H} w_{t-1,h}\left(x_{t,h} - x_{t-1,h}\right) + e_t \sum_{h=1}^{H} x_{t,h}x_{t-1,h}w_{t-1,h}}{\sum_{h=1}^{H} w_{t-1,h}x_{t-1,h}\left(1 - x_{t,h}\right)} \\
x_{t,h} =& \frac{1}{h}\sum_{k=0}^{h-1} x_{t,h}^{-k} \\
x_{t,h}^{-k} =& \frac{x_{t-1,h}^{-k+1}(1+r)}{1 + x_{t-1,h}^{-k+1}(r+e_t)}
\end{aligned}$$

We prove that continuous function $\Xi(r)$ at least once takes the value zero in the interval $]-1;\infty[$ (the lower bound for r ensures that stock price stays positive). It is easy to show that:

$$\lim_{r\to -1} \Xi(r) < 0$$

First, notice that $\lim_{r\to -1} x_{t,h}^{-k} = 0,\ ,\forall k$, and so the right side of the expression for $\Xi(r)$ does not depend on r. Then, using that $0 \le x_{t,h} \le 1$ and by consequence

$$\sum_{h=1}^{H} w_{t-1,h}x_{t-1,h}\left(1 - x_{t,h}\right) > 0,$$

we can show that the statement $\lim_{r\to -1} \Xi(r) > 0$ is equivalent to:

$$-\sum_{h=1}^{H} w_{t-1,h}x_{t-1,h}\left(1 - x_{t,h}\right) - \sum_{h=1}^{H} w_{t-1,h}\left(x_{t,h} - x_{t-1,h}\right) - e_t\sum_{h=1}^{H} x_{t,h}x_{t-1,h}w_{t-1,h} > 0.$$

The latter inequality simplifies to:

$$-\sum_{h=1}^{H} w_{t-1,h}x_{t-1,h}x_{t,h} - \sum_{h=1}^{H} w_{t-1,h}x_{t,h} - \sum_{h=1}^{H} w_{t-1,h}x_{t,h}, x_{t-1,h}e_t > 0$$

which is evidently absurd.

On the other hand, one can show that:

$$\lim_{r\to \infty} \Xi(r) = \infty$$

This follows from $\lim_{r\to\infty} x_{t,h}^{-k} = 1,\ \forall k$, which implies that:

$$\lim_{r\to\infty} = \frac{\sum_{h=1}^{H} w_{t-1,h}\left(x_{t,h} - x_{t-1,h}\right) + e_t \sum_{h=1}^{H} x_{t,h} x_{t-1,h} w_{t-1,h}}{\sum_{h=1}^{H} w_{t-1,h} x_{t-1,h}\left(1 - x_{t,h}\right)} < \infty.$$

The result of the theorem is obtained by continuity. □

Proof of Theorem 2.10

Proof. Denote $\xi_{t,i,j}$ the ratio between the wealth of investors at scale i and at scale j, taken at time t, that is:

$$\xi_{t,i,j} = \frac{w_{t,i}}{w_{t,j}}$$

It is evident that definition (2.24) is equivalent to

$$\xi_{0,i,j} > 0 \implies P(\xi_{t,i,j} = 0) = 0,\ \forall\ i, j \in \{1, \ldots, H\},\ \text{when } t \to \infty \tag{2.38}$$

The ratio of wealth can be written as:

$$\xi_{t,i,j} = \xi_{0,i,j} \frac{\prod_{k=1}^{t} g_{k,i}}{\prod_{k=1}^{t} g_{k,j}}$$

or equivalently in logarithmic terms:

$$\ln(\xi_{t,i,j}) = \ln(\xi_{0,i,j}) + \sum_{k=1}^{t} \left[\ln(g_{k,i}) - \ln(g_{k,j})\right]. \tag{2.39}$$

It follows from (2.24) that, as $t \to \infty$, the sum in (2.39) converges almost surely to:

$$\sum_{k=1}^{\infty} (\mathrm{E}\left[\ln(g_{k,i})\right] - \mathrm{E}\left[\ln(g_{k,i})\right]),$$

which is finite if and only if

$$\mathrm{E}\left[\ln(g_{k,i})\right] = \mathrm{E}\left[\ln(g_{k,i})\right].$$

This condition is evidently equivalent to:

$$\ln(\xi_{t,i,j}) \xrightarrow{a.s.} \ln(\xi_{0,i,j}),$$

which proves the theorem. □

Proof of Theorem 2.12

Proof. Consider equation $F(r_1, \ldots, r_{t-1}, e_1, \ldots, e_{t-1}, r_t, e_t) = 0$ that implicitly defines the equilibrium return dynamics. In order to study conditional volatility of returns, we will linearize the function F in the neighborhood of the point $M_\epsilon \stackrel{def}{=} (\bar{r}, \ldots, \bar{r}, \bar{e}, \ldots, \bar{e}, \bar{r}, e_t)$ rather than the point M (recall that e_t is the source of innovation in the system). A first-order series expansion yields:

$$F(M_\epsilon) + \sum_{i=1}^{t} \frac{\partial F}{\partial r_{t-k}}(M_\epsilon)\widetilde{r}_{t-k} + \sum_{k=1}^{t} \frac{\partial F}{\partial e_{t-k}}(M_\epsilon)\widetilde{e}_{t-k} + \frac{\partial F}{\partial r_t}(M_\epsilon)\widetilde{r}_t = 0. \tag{2.40}$$

This implies that the return satisfies:

$$\widetilde{r}_t = -\frac{F(M_\epsilon)}{\frac{\partial F}{\partial r_t}(M_\epsilon)} - \sum_{k=1}^{t} \frac{\frac{\partial F}{\partial r_{t-k}}(M_\epsilon)}{\frac{\partial F}{\partial r_t}(M_\epsilon)}\widetilde{r}_{t-k} - \sum_{k=1}^{t} \frac{\frac{\partial F}{\partial e_{t-k}}(M_\epsilon)}{\frac{\partial F}{\partial r_t}(M_\epsilon)}\widetilde{e}_{t-k}. \tag{2.41}$$

The share of investment in the risky asset is a function of returns and dividend yields, given by the last two equations of (2.23). We denote:

$$\begin{aligned} \bar{x} &= x_{t-1,h}(M_\epsilon), \\ x_{h,\epsilon} &= x_{t,h}(M_\epsilon) = \frac{1}{h}\left[\bar{x} + (h-1)\frac{\bar{x}\,(1+\bar{r})}{1+\bar{x}\,(\bar{r}+e_t)}\right], \\ x_\epsilon &= \sum_{h=1}^{H} \xi_{0,h}\, x_{h,\epsilon}\,. \end{aligned} \tag{2.42}$$

Now let us obtain the explicit expressions for all elements of equation (2.41):

$$F(M_\epsilon) = \frac{x_\epsilon - \bar{x} + \bar{x}\, x_\epsilon\, e_t}{\bar{x}\,(1-x_\epsilon)} - \bar{r}\,.$$

The partial derivative of F with respect to the past return r_{t-k}, $k = 1, \ldots, t$ reads:

$$\begin{aligned}
&\frac{\partial F}{\partial r_{t-k}}(M_\epsilon) = \Bigg\{ \Bigg[\sum_{h=1}^{H} \frac{\partial w_{t-1,h}}{\partial r_{t-k}}(M_\epsilon)(x_{h,\epsilon} - \bar{x} + \bar{x}\, x_{h,\epsilon}\, e_t) + \\
&\sum_{h=1}^{H} \bar{w}_{t-1,h} \left(\frac{\partial x_{t,h}}{\partial r_{t-k}}(M_\epsilon) - \frac{\partial x_{t-1,h}}{\partial r_{t-k}}(M_\epsilon) + \frac{\partial (x_{t-1,h} x_{t,h})}{\partial r_{t-k}}(M_\epsilon) e_t \right) \Bigg] \times \\
&\left[\sum_{h=1}^{H} \bar{w}_{t-1,h} \bar{x}\,(1 - x_{h,\epsilon}) \right] - \left[\sum_{h=1}^{H} \frac{\partial w_{t-1,h}}{\partial r_{t-k}}(M_\epsilon) \bar{x}\,(1 - x_{h,\epsilon}) + \right. \\
&\left. \sum_{h=1}^{H} \bar{w}_{t-1,h} \left(\frac{\partial x_{t-1,h}}{\partial r_{t-k}}(M_\epsilon) - \frac{\partial (x_{t-1,h} x_{t,h})}{\partial r_{t-k}}(M_\epsilon) \right) \right] \times \\
&\left[\sum_{h=1}^{H} \bar{w}_{t-1,h} (x_{h,\epsilon} - \bar{x} + \bar{x}\, x_{h,\epsilon}\, e_t) \right] \Bigg\} \times \left[\sum_{h=1}^{H} \bar{w}_{t-1,h} \bar{x}\,(1 - x_{h,\epsilon}) \right]^{-2} = \\
&\Bigg\{ \sum_{h=1}^{H} \xi_{0,h} \left(\frac{\partial x_{t,h}}{\partial r_{t-k}}(M_\epsilon) - \frac{\partial x_{t-1,h}}{\partial r_{t-k}}(M_\epsilon) + \frac{\partial (x_{t-1,h} x_{t,h})}{\partial r_{t-k}}(M_\epsilon)\, e_t \right) \times \\
&\sum_{h=1}^{H} \xi_{0,h}\, \bar{x}\,(1 - x_{h,\epsilon}) - \sum_{h=1}^{H} \xi_{0,h} \left(\frac{\partial x_{t-1,h}}{\partial r_{t-k}}(M_\epsilon) - \frac{\partial (x_{t-1,h} x_{t,h})}{\partial r_{t-k}}(M_\epsilon) \right) \times \\
&\left. \sum_{h=1}^{H} \xi_{0,h}\, (x_{h,\epsilon} - \bar{x} + \bar{x}\, x_{h,\epsilon}\, e_t) \right\} \times \left[\sum_{h=1}^{H} \xi_{0,h}\, \bar{x}\,(1 - x_{h,\epsilon}) \right]^{-2} = \\
&\frac{1}{\bar{x}^{\,2}(1 - x_\epsilon)^2} \Bigg\{ \sum_{h=1}^{H} \xi_{0,h} \left[\frac{\partial x_{t,h}}{\partial r_{t-k}}(M_\epsilon)\, (\bar{x}\,(1 - x_\epsilon) + \bar{x}\,(\bar{x}\, e_t + x_\epsilon - \bar{x})) + \right. \\
&\left. \left. \frac{\partial x_{t-1,h}}{\partial r_{t-k}}(M_\epsilon)\, (x_\epsilon\, (\bar{x}\, e_t + x_\epsilon - \bar{x}) - (-\bar{x}\, x_\epsilon + x_\epsilon + \bar{x}\, x_\epsilon\, e_t)) \right] \right\} = \\
&\frac{\sum_{h=1}^{H} \xi_{0,h} \left[\frac{\partial x_{t,h}}{\partial r_{t-k}}(M_\epsilon) \left(\bar{x} + \bar{x}^{\,2} e_t - \bar{x}^{\,2} \right) + \frac{\partial x_{t-1,h}}{\partial r_{t-k}}(M_\epsilon) \left({x_\epsilon}^2 - x_\epsilon \right) \right]}{\bar{x}^{\,2}(1 - x_\epsilon)^2}.
\end{aligned} \tag{2.43}$$

The derivative of F with respect to the current return reads:

$$\frac{\partial F}{\partial r_t}(M_\epsilon) = \frac{\sum_{h=1}^{H} \xi_{0,h} \frac{\partial x_{t,h}}{\partial r_t}(M_\epsilon) \left(\bar{x} + \bar{x}^{\,2} e_t - \bar{x}^{\,2} \right)}{\bar{x}^{\,2}(1 - x_\epsilon)^2} - 1.$$

In a similar way, we find the derivative of F with respect to the past dividend yield e_{t-k}, $k = 1, \ldots, t$:

$$\begin{aligned}
&\frac{\partial F}{\partial e_{t-k}}(M_\epsilon) = \\
&\frac{\sum_{h=1}^{H} \xi_{0,h} \left[\frac{\partial x_{t,h}}{\partial e_{t-k}}(M_\epsilon) \left(\bar{x} + \bar{x}^{\,2} e_t - \bar{x}^{\,2} \right) + \frac{\partial x_{t-1,h}}{\partial e_{t-k}}(M_\epsilon) \left({x_\epsilon}^2 - x_\epsilon \right) \right]}{\bar{x}^{\,2}(1 - x_\epsilon)^2}.
\end{aligned} \tag{2.44}$$

Now let us compute $\frac{\partial x_{t,h}}{\partial r_{t-k}}(M_\epsilon)$. The derivative of the multi-horizon demand function ($h > 1$) with respect to the contemporaneous return reads:

$$\begin{gathered}\frac{\partial x_{t,h}}{\partial r_t}(M_\epsilon) = \frac{1}{h}\frac{\partial}{\partial r_t}\left[\bar{x} + \frac{\bar{x}(1+r_t)}{1+\bar{x}(r_t+e_t)} + \frac{g(\bar{x})(1+r_t)}{1+g(\bar{x})(r_t+e_t)} + \right. \\ \left. \frac{g^2(\bar{x})(1+r_t)}{1+g^2(\bar{x})(r_t+e_t)} + \ldots + \frac{g^{h-2}(\bar{x})(1+r_t)}{1+g^{h-2}(\bar{x})(r_t+e_t)}\right] = \\ \frac{1}{h}\frac{\partial}{\partial r_t}\left[\bar{x} + (h-1)\frac{\bar{x}(1+r_t)}{1+\bar{x}(r_t+e_t)}\right] = \\ \frac{h-1}{h}\frac{\bar{x}-\bar{x}^2+\bar{x}^2 e_t}{[1+\bar{x}(\bar{r}+e_t)]^2}.\end{gathered} \tag{2.45}$$

The derivative with respect to the lagged return is 0 if the lag k is larger or equal to $h-1$. To compute it for the case $0 < k < h-1$ let us denote:

$$\tilde{x} = \frac{\bar{x}(1+r_{t-k})}{1+\bar{x}(r_{t-k}+\bar{e})} = \frac{g^m(\bar{x})(1+r_{t-k})}{1+g^m(\bar{x})(r_{t-k}+\bar{e})},\ \forall m > 0$$

and

$$g_\epsilon(x) = \frac{x(1+\bar{r})}{1+x(\bar{r}+e_t)}.$$

Then using the property that $g(\tilde{x}(M_\epsilon)) = g(\tilde{x}(\bar{x})) = \bar{x}$, we obtain:

$$\begin{gathered}\frac{\partial x_{t,h}}{\partial r_{t-k}}(M_\epsilon) = \frac{1}{h}\frac{\partial}{\partial r_{t-k}}\left[\bar{x} + g_\epsilon(\bar{x}) + g_\epsilon(g(\bar{x})) + g_\epsilon(g^2(\bar{x})) + \ldots + \right. \\ \left. g_\epsilon(g^{k-1}(\bar{x})) + (h-k-1)\ g_\epsilon\left(g^{k-1}(\tilde{x})\right)\right] = \\ \frac{h-k-1}{h}\ \frac{\partial}{\partial x}\ g_\epsilon(M_\epsilon)\ \frac{\partial}{\partial x}g^{k-1}(M_\epsilon)\ \frac{\partial}{\partial r_{t-k}}\tilde{x} = \\ \frac{h-k-1}{h}\ g'_\epsilon(\bar{x})\ [g'(\bar{x})]^{k-1}\ \frac{\partial}{\partial r_{t-k}}\tilde{x}.\end{gathered} \tag{2.46}$$

Computing derivatives in (2.46) yields:

$$\begin{gathered}\frac{\partial}{\partial x}g(M_\epsilon) = \frac{(1+\bar{r})\,[1+\bar{x}(\bar{r}+\bar{e})] - (\bar{r}+\bar{e})(1+\bar{r})\bar{x}}{[1+\bar{x}(\bar{r}+\bar{e})]^2} = \\ \frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+\bar{e})]^2},\end{gathered}$$

$$\frac{\partial}{\partial x}g_\epsilon(M_\epsilon) = \frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+e_t)]^2}$$

and

$$\begin{gathered}\frac{\partial}{\partial r_{t-k}}\tilde{x}(M_\epsilon) = \frac{\bar{x}[1+\bar{x}(\bar{r}+\bar{e})] - \bar{x}^2(1+\bar{r})}{[1+\bar{x}(\bar{r}+\bar{e})]^2} = \\ \frac{\bar{x}-\bar{x}^2+\bar{x}^2\bar{e}}{[1+\bar{x}(\bar{r}+\bar{e})]^2}.\end{gathered}$$

So finally for $k = 1, \ldots, h-2$ we have:

$$\frac{\partial x_{t,h}}{\partial r_{t-k}}(M_\epsilon) = \frac{h-k-1}{h} \frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+e_t)]^2} \left[\frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+\bar{e})]^2}\right]^{k-1} \frac{\bar{x}-\bar{x}^2+\bar{x}^2\bar{e}}{[1+\bar{x}(\bar{r}+\bar{e})]^2}. \quad (2.47)$$

Then, using the fact that:

$$\frac{\partial}{\partial e_{t-k}}\tilde{x}(M_\epsilon) = -\frac{\bar{x}^{\,2}(1+\bar{r}\,)}{[1+\bar{x}(\bar{r}+\bar{e})]^2},$$

we obtain:

$$\frac{\partial x_{t,h}}{\partial e_{t-k}}(M_\epsilon) = -\frac{h-k-1}{h} \frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+e_t)]^2} \left[\frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+\bar{e})]^2}\right]^{k-1} \frac{\bar{x}^{\,2}(1+\bar{r}\,)}{[1+\bar{x}(\bar{r}+\bar{e})]^2}. \quad (2.48)$$

Using previous results, it is easy to notice that for $k = 1, \ldots, h-1$:

$$\begin{aligned}
\frac{\partial x_{t-1,h}}{\partial r_{t-k}}(M_\epsilon) &= \frac{h-k}{h} \left[\frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+\bar{e})]^2}\right]^{k-1} \frac{\bar{x}-\bar{x}^2+\bar{x}^2\bar{e}}{[1+\bar{x}(\bar{r}+\bar{e})]^2}, \\
\frac{\partial x_{t-1,h}}{\partial e_{t-k}}(M_\epsilon) &= -\frac{h-k}{h} \left[\frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+\bar{e})]^2}\right]^{k-1} \frac{\bar{x}^{\,2}(1+\bar{r}\,)}{[1+\bar{x}(\bar{r}+\bar{e})]^2}.
\end{aligned} \quad (2.49)$$

Expressions (2.45),(2.47) - (2.49) can be simplified using the relation:

$$\bar{x} = \frac{\bar{r}}{\bar{r}+\bar{e}}$$

This gives:

$$\begin{aligned}
\frac{\partial x_{t,h}}{\partial r_t}(M_\epsilon) &= \frac{h-1}{h} \frac{\bar{x}-\bar{x}^2+\bar{x}^2 e_t}{[1+\bar{x}(\bar{r}+e_t)]^2}, \\
\frac{\partial x_{t,h}}{\partial r_{t-k}}(M_\epsilon) &= \frac{h-k-1}{h} \frac{\bar{x}-\bar{x}^{\,2}+\bar{x}^{\,2}\bar{e}}{(1+\bar{r}\,)^k\,[1+\bar{x}\,(\bar{r}+e_t)]^2}, \quad k = 1, \ldots, h-2, \\
\frac{\partial x_{t-1,h}}{\partial r_{t-k}}(M_\epsilon) &= \frac{h-k}{h} \frac{\bar{x}\,(1-\bar{x}\,)}{(1+\bar{r}\,)^k}, \quad k = 1, \ldots, h-1, \\
\frac{\partial x_{t,h}}{\partial e_{t-k}}(M_\epsilon) &= -\frac{h-k-1}{h} \frac{1+\bar{r}}{[1+\bar{x}(\bar{r}+e_t)]^2} \frac{\bar{x}^{\,2}}{(1+\bar{r}\,)^k}, \quad k = 1, \ldots, h-2, \\
\frac{\partial x_{t-1,h}}{\partial e_{t-k}}(M_\epsilon) &= -\frac{h-k}{h} \frac{\bar{x}^{\,2}}{(1+\bar{r}\,)^k}, \quad k = 1, \ldots, h-1.
\end{aligned} \quad (2.50)$$

We can now replace expressions (2.50) in the equations for partial derivatives. Let us denote

$$a_k = \sum_{h=k+2}^{H} \frac{h-k-1}{h}\xi_{0,h}\,,$$
$$b_k = \sum_{h=k+1}^{H} \frac{h-k}{h}\xi_{0,h}\,,$$
$$c = \sum_{h=1}^{H} \frac{h-1}{h}\xi_{0,h}\,.$$

It is straightforward that $0 < a_k < b_k < c < 1$. In these terms:

$$\frac{\partial F}{\partial r_{t-k}}(M_\epsilon) = \frac{a_k \frac{\bar{x}-\bar{x}^2+\bar{x}^2\bar{e}}{(1+\bar{r})^k[1+\bar{x}(\bar{r}+e_t)]^2}\left(\bar{x}+\bar{x}^2 e_t-\bar{x}^2\right)+b_k\frac{\bar{x}(1-\bar{x})}{(1+\bar{r})^k}\left(x_\epsilon{}^2-x_\epsilon\right)}{\bar{x}^2(1-x_\epsilon)^2} =$$
$$\frac{a_k\,\bar{x}^2(1-\bar{x}+\bar{x}\,\bar{e}\,)(1-\bar{x}+\bar{x}\,e_t)-b_k\,\bar{x}\,x_\epsilon\,(1-\bar{x}\,)(1-x_\epsilon\,)\left[1+\bar{x}\,(\bar{r}+e_t)\right]^2}{\bar{x}^2(1-x_\epsilon)^2(1+\bar{r}\,)^k\left[1+\bar{x}\,(\bar{r}+e_t)\right]^2},$$

$$\frac{\partial F}{\partial r_t}(M_\epsilon) = \sum_{h=1}^{H}\xi_{0,h}\,\frac{h-1}{h}\,\frac{\bar{x}-\bar{x}^2+\bar{x}^2e_t}{[1+\bar{x}(\bar{r}+e_t)]^2}\,\frac{\left(\bar{x}+\bar{x}^2e_t-\bar{x}^2\right)}{\bar{x}^2(1-x_\epsilon)^2}-1 =$$
$$\frac{c\,(1-\bar{x}+\bar{x}\,e_t)^2-(1-x_\epsilon)^2\left[1+\bar{x}(\bar{r}+e_t)\right]^2}{(1-x_\epsilon)^2\left[1+\bar{x}(\bar{r}+e_t)\right]^2},$$

$$\frac{\partial F}{\partial e_{t-k}}(M_\epsilon) =$$
$$\frac{-a_k\,(1+\bar{r}\,)\,\bar{x}^2(\bar{x}+\bar{x}^2e_t-\bar{x}^2)+b_k\,\bar{x}^2x_\epsilon\,(1-x_\epsilon\,)\left[1+\bar{x}(\bar{r}+e_t)\right]^2}{\bar{x}^2(1-x_\epsilon)^2(1+\bar{r}\,)^k\left[1+\bar{x}(\bar{r}+e_t)\right]^2}.$$

We can now give explicit expressions for all the terms of equation (2.41):

$$-\frac{F(M_\epsilon)}{\frac{\partial F}{\partial r_t}(M_\epsilon)} = -\left[\frac{x_\epsilon-\bar{x}+\bar{x}\,x_\epsilon\,e_t}{\bar{x}\,(1-x_\epsilon\,)}-\bar{r}\right]\times$$
$$\frac{(1-x_\epsilon)^2\left[1+\bar{x}(\bar{r}+e_t)\right]^2}{c\,(1-\bar{x}+\bar{x}\,e_t)^2-(1-x_\epsilon)^2\left[1+\bar{x}(\bar{r}+e_t)\right]^2},$$

$$-\frac{\frac{\partial F}{\partial r_{t-k}}(M_\epsilon)}{\frac{\partial F}{\partial r_t}(M_\epsilon)} =$$
$$-\frac{a_k\,\bar{x}^{\,2}(1-\bar{x}+\bar{x}\,\bar{e}\,)(1-\bar{x}+\bar{x}\,e_t)-b_k\,\bar{x}\,x_\epsilon\,(1-\bar{x}\,)(1-x_\epsilon\,)\,[1+\bar{x}\,(\bar{r}+e_t)]^2}{\bar{x}^{\,2}(1-x_\epsilon\,)^2(1+\bar{r}\,)^k\,[1+\bar{x}\,(\bar{r}+e_t)]^2}\times$$
$$\frac{(1-x_\epsilon\,)^2\,[1+\bar{x}(\bar{r}+e_t)]^2}{c\,(1-\bar{x}+\bar{x}\,e_t)^2-(1-x_\epsilon\,)^2\,[1+\bar{x}(\bar{r}+e_t)]^2} =$$
$$-\frac{a_k\,\bar{x}\,(1-\bar{x}+\bar{x}\,\bar{e}\,)(1-\bar{x}+\bar{x}\,e_t)-b_k\,x_\epsilon\,(1-\bar{x}\,)(1-x_\epsilon\,)\,[1+\bar{x}\,(\bar{r}+e_t)]^2}{\bar{x}\,(1+\bar{r}\,)^k\,\left[c\,(1-\bar{x}+\bar{x}\,e_t)^2-(1-x_\epsilon\,)^2\,(1+\bar{x}(\bar{r}+e_t))^2\right]},$$

$$-\frac{\frac{\partial F}{\partial e_{t-k}}(M_\epsilon)}{\frac{\partial F}{\partial r_t}(M_\epsilon)} =$$
$$-\frac{-a_k\,(1+\bar{r}\,)\,\bar{x}^{\,2}(\bar{x}+\bar{x}^{\,2}e_t-\bar{x}^{\,2})+b_k\,\bar{x}^{\,2}x_\epsilon\,(1-x_\epsilon\,)\,[1+\bar{x}(\bar{r}+e_t)]^2}{\bar{x}^{\,2}(1-x_\epsilon\,)^2(1+\bar{r}\,)^k\,[1+\bar{x}(\bar{r}+e_t)]^2}\times$$
$$\frac{(1-x_\epsilon\,)^2\,[1+\bar{x}(\bar{r}+e_t)]^2}{c\,(1-\bar{x}+\bar{x}\,e_t)^2-(1-x_\epsilon\,)^2\,[1+\bar{x}(\bar{r}+e_t)]^2} =$$
$$\frac{a_k\,(1+\bar{r}\,)\,(\bar{x}+\bar{x}^{\,2}e_t-\bar{x}^{\,2})-b_k\,x_\epsilon\,(1-x_\epsilon\,)\,[1+\bar{x}(\bar{r}+e_t)]^2}{(1+\bar{r}\,)^k\,\left[c\,(1-\bar{x}+\bar{x}\,e_t)^2-(1-x_\epsilon\,)^2\,(1+\bar{x}(\bar{r}+e_t))^2\right]}.$$

Now let us linearize the terms of (2.41) with respect to the disturbance term e_t. We will need to use the relation:

$$\frac{\partial x_\epsilon}{\partial e_t} = -\sum_{h=1}^{H}\xi_{0,h}\,\frac{h-1}{h}\,\frac{\bar{x}^{\,2}(1+\bar{r}\,)}{(1+\bar{x}\,(\bar{r}+\bar{e}\,))^2} = -\frac{c\,\bar{x}^{\,2}}{1+\bar{r}}. \tag{2.51}$$

The first term of (2.41) corresponds to the part of volatility that is independent of past realizations of the returns and dividend yields:

$$-\frac{F(M_\epsilon)}{\frac{\partial F}{\partial r_t}(M_\epsilon)} \approx \bar{V}\,\widetilde{e}_t$$

with

$$\bar{V} = \frac{c\,\bar{x}^{\,2}(1-\bar{e}\,)-c\,\bar{x}+\bar{x}\,(1-\bar{x}\,)(1+\bar{r}\,)}{(1-c\,)(1+\bar{r}\,)(1-\bar{x}\,)^2} =$$
$$\frac{\bar{x}\,(1-\bar{x}\,)(1+br)(1-c)}{(1-c\,)(1+\bar{r}\,)(1-\bar{x}\,)^2} = \frac{\bar{x}}{1-\bar{x}}$$

The second term gives both the autoregression coefficients in the equation for returns and the dependency of the volatility on past returns:

$$-\frac{\frac{\partial F}{\partial r_{t-k}}(M_\epsilon)}{\frac{\partial F}{\partial r_t}(M_\epsilon)} \approx A_k + B_k\,\widetilde{e}_t$$

with

$$A_k = \frac{a_k - b_k}{(1-c)(1+\bar{r}\,)^k},$$
$$B_k = \frac{\bar{x}\,(1-2\bar{x}\,)(c\,b_k - a_k)}{(1-\bar{x}\,)(1+\bar{r}\,)^{k+1}(1-c)}.$$

In the same way we find:

$$-\frac{\frac{\partial F}{\partial e_{t-k}}(M_\epsilon\,)}{\frac{\partial F}{\partial r_t}(M_\epsilon\,)} \approx C_k + D_k\,\widetilde{e}_t$$

with

$$C_k = -\frac{\bar{x}}{1-\bar{x}}A_k,$$
$$D_k = -\frac{\bar{x}}{1-\bar{x}}B_k.$$

The equation for returns thus reads:

$$\widetilde{r}_t = \sum_{k=1}^{H-1}(A_k + B_k\,\widetilde{e}_t)\,\widetilde{r}_{t-k} - \bar{V}\sum_{k=1}^{H-1}(A_k + B_k\,\widetilde{e}_t)\widetilde{e}_{t-k} + \bar{V}\,\widetilde{e}_t,$$

which can also be written as:

$$\widetilde{r}_t - \bar{V}\,\widetilde{e}_t = \sum_{k=1}^{H-1} A_k\left(\widetilde{r}_{t-k} - \bar{V}\,\widetilde{e}_{t-k}\right) + \sum_{k=1}^{H-1} B_k\left(\widetilde{r}_{t-k} - \bar{V}\,\widetilde{e}_{t-k}\right)\widetilde{e}_t.$$

Introducing the notation:

$$\hat{r}_{\,t} = \widetilde{r}_t - \bar{V}\,\widetilde{e}_t,$$

we re-write the dynamics in terms of the error correction with respect to the one-scale rational equilibrium path, for which $\hat{r}_{\,t} = 0$, as follows from theorem 2.5. In these terms we obtain:

$$\hat{r}_{\,t} = \sum_{k=1}^{H-1} A_k\,\hat{r}_{\,t-k} + \left(\sum_{k=1}^{H-1} B_k\hat{r}_{\,t-k}\right)\widetilde{e}_t, \tag{2.52}$$

which closes the proof of the theorem. □

Chapter 3

Volatility Models: from GARCH to Multi-Horizon Cascades

3.1 Introduction

Modeling stock prices is essential in many areas of financial economics, such as derivatives pricing, portfolio management and financial risk follow-up. One of the most criticized drawbacks of the so-called "modern portfolio theory" (MPT), including the diversification principle of Markowitz (1952) and the capital asset pricing model by Sharpe (1964) and Lintner (1965), is the non-realistic assumption about stock price variability. Clearly, stock returns are not *iid* distributed Gaussian random variables, but alternatives to this assumption are numerous, sometimes complicated and application-dependent. In this paper we review empirical properties of stock price dynamics and various models, proposed to represent it, focusing on the most recent developments, concerning mainly multi-horizon and multifractal stochastic volatility processes.

The subject of this study is the variability of stock prices, referred to as volatility. Usually introduction of scientific terminology aims at making a general concept more precise, but here we rather have an example of the contrary. Depending on the context and the point of view of the author, the term "volatility" in finance can stand for the variability of prices (in this sense we used it above), an estimate of standard deviation, financial risk in general, a parameter of a derivative pricing model or a stochastic process of particular form. We will continue using it in the most general sense, that is as a synonym of variability. Before reviewing volatility models, we examine in more detail the evolution of the notion itself. This will help for a better understanding of the logic of the evolution of the corresponding models.

One of the first interpretations of the term "volatility" is due to the fact that the name of variability phenomenon itself has been identified with the most elementary method of its quantitative measurement - standard deviation of stock returns. This interpretation is logically embedded in the concept of MPT, also called mean-variance theory, because under its assumptions these two parameters contain all relevant information about stock returns, distributed normally[1]. Note that in Markowitz (1952) risk is modeled statically: returns on each stock are characterized by constant volatility (variance or standard deviation) and covariances with the returns on other assets. So volatility can be seen as synonym for standard deviation, or as an estimate of a constant parameter in the simplest model of stock returns. This definition of volatility has deep roots and is still widely used among asset management professionals.

The appearance in 1973 of the option pricing models by Black and Scholes (1973) and Merton (1973) led to significant changes in the understanding of volatility. A continuous-time diffusion (geometric Brownian motion) is used to model stock prices:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{3.1}$$

with S_t the stock price, μ the drift parameter and W_t a Brownian motion. The parameter σ is called volatility because it characterizes the degree of variability. Since the log-returns, computed from stock prices that follow equation (3.1), are normally distributed, this model is also called a log-normal diffusion.

Very soon it became obvious that equation (3.1) poorly describes reality. Its parameters unambiguously define option prices for given exercise dates and strikes, so that volatility parameter can be inferred from observations of option prices by using the inverse of the Black-Scholes formula. An estimate obtained in this way is called implied volatility as opposed to historical volatility, measured as the standard deviation of returns. Contrary to the predictions of the Black and Scholes model empirical results show that implied volatility varies for option contracts with different parameters. This phenomenon is known as volatility smile. Its name is due to a characteristic convex form of the plot of the estimate of σ as function of the option exercise price.

The above remark does not mean that implied volatility is useless. It has been shown that it contains information about future variability of returns and thus it is often used in forecasting. In derivatives pricing implied volatility is important because it allows to extrapolate the observed market data, e.g. option prices, for the evaluation of other financial instruments, e.g. over-the-counter options (see Avellaneda et al., 1997;

[1] In MPT a simplifying assumption, alternative to the normality of returns, is the quadratic form of the utility function of investors. However, the latter can hardly be justified.

Dupire, 1993, 1994). Despite these partial successes, a more adequate model than log-normal diffusion could still be useful in both derivatives pricing and asset management applications. In Merton (1973) volatility parameter is already allowed to vary in time. Even earlier Mandelbrot (1963) points to the empirical properties of stock returns that to not correspond to the log-normal diffusion model and proposes a wider class of Levy-stable probability distributions. Further developments in the led to the understanding of volatility as a stochastic process and not merely as a parameter, even time-varying.

The meaning of the term "volatility" in finance has come full circle: from a general term for variability phenomenon to a statistical estimate, then a model parameter and finally a stochastic process, which again is supposed to characterize the whole structure of the stock price variability. More technically, the modern understanding of volatility can be characterized as a time structure of conditional second-order moments in the distribution of returns. In the simplest case of log-normal diffusion this structure is described by one parameter and in more complicated cases, by a separate stochastic process.

This paper starts with an overview of empirical properties of volatility, the so-called "stylized facts". Then we briefly discuss traditional approaches to its modeling - conditional heteroscedasticity and stochastic volatility, that reproduce empirical properties to some extent. Though many models are good enough to describe separate stylized facts, we show that none of them is quite sufficient to represent the whole structure of stock price variability. In particular, most traditional models do not allow for representing returns dynamics on multiple time horizons (e.g. from minutes to days and months) simultaneously, which is important both practically and theoretically. Stylized facts themselves have features specific to the frequency, at which price dynamics is observed. We analyze and compare recently proposed models of conditional heteroscedasticity and stochastic volatility, based on the multi-horizon approach, and discuss the main unsolved problems, related to them.

3.2 Empirical Properties of Volatility

Many empirical studies show that financial time series satisfy a number of general properties, referred to as stylized facts. A realistic model for prices is expected to reproduce these properties. We characterize them briefly, for a detailed survey on the subject see Cont (2001).

- **Excessive volatility.** The observed degree of variability in stock prices can hardly be explained by variations in fundamental economic factors. In particular, returns

of large magnitude (positive and negative) are often hard to explain by arrival of new information about future cash flows (Cutler et al., 1989).

- **Absence of linear correlations in returns**. Stock returns, computed over sufficiently long time periods (several hours and more) display insignificant linear correlation. This results is in accordance with the stock market efficiency hypothesis by Fama (1970) and the main results of MPT, using martingale measures.
- **Clustering of volatility and long memory in absolute values of returns**. Time series of absolute values of returns is characterized by important autocorrelation, and the autocorrelation function (ACF) decays slowly with time lags (slower than geometric decay). Long periods of high and low volatility are observed (Bollerslev et al., 1992; Ding and Granger, 1996; Ding et al., 1993).
- **The link between the trading volume and volatility.** Volatility of returns is positively correlated with the trading volume, and the latter time series displays the same long memory properties as in the absolute returns (Lobato and Velasco, 2000).
- **Asymmetry and leverage in the dynamic structure of volatility**. Positive and negative returns of the same magnitude, observed over the past period, have different effects on current volatility (asymmetry). Current returns and future volatility are negatively correlated (leverage). Presence of the leverage effect implies the asymmetry but the inverse does not hold (Black, 1976).
- **Heavy tails in the distribution of returns.** Unconditional probability distribution of daily returns is characterized by heavy tails, i.e. high probability of observing extreme values, compared to the normal distribution (Fama, 1965; Mandelbrot, 1963).
- **The form of the probability distribution of returns varies across time intervals, over which returns are computed** (Arneodo et al., 1998; Ghashghaie et al., 1996). Distributions of log-returns over long time intervals are relatively close to the normal law, while returns over short time intervals (5 - 30 minutes) have very heavy tails.

Among these stylized facts we shall be particularly interested in the properties related to the ACF of returns and the form of the probability distribution of returns and their magnitudes. We start with a definition of the above-mentioned long memory phenomenon in terms of ACF.

A stationary stochastic process X_t with finite variance has long memory (or long-range dependence) if its autocorrelation function $C(\tau) = \text{corr}(X_t, X_{t-\tau})$ at $\tau \to \infty$

decays with the time lag according to the power low (i.e. at hyperbolic speed):

$$C(\tau) \sim \frac{L(\tau)}{\tau^{1-2d}}, \tag{3.2}$$

where $0 < d < \frac{1}{2}$ and where $L(\cdot)$ is some continuous function that for $\forall x > 0$ and $\tau \to \infty$ satisfies $\frac{L(xt)}{L(t)} \to 1$. The process has short memory if its ACF decays exponentially (with geometric speed), so that:

$$\exists A > 0, c \in (0,1) : |C(\tau)| \leq Ac^{\tau} \tag{3.3}$$

In this definition the technical condition, imposed on the function $L(\cdot)$, implies that for infinite lag τ this function changes infinitely slowly. Notice that the definition refers to the theoretical ACF of the time series model and not to its sample estimate. As we will see, in many cases the sample ACF has properties, similar to those implied by definition (3.2), but the theoretical ACF does not satisfy this definition.

Alternatively, long memory can be characterized by the power law divergence of the spectral density of the time series X_t at the origin:

$$\Psi_x(u) \sim c_\Psi |u|^{-\alpha} \tag{3.4}$$

with $\Psi_x(\cdot)$ - spectral density function, α - scaling parameter and c_Ψ - a constant.

To illustrate the empirical properties of returns we use two types of stock index data: high frequency (intraday) observations over a relatively short time period (French CAC40 index) and daily observations for very long time period (Dow Jones Industrial Average Index, DJIA). We will see that the main empirical patterns are similar for these very different examples.

The return at time $t \in 1 \dots T$ over the interval τ is defined as the change in the logarithm of price S:

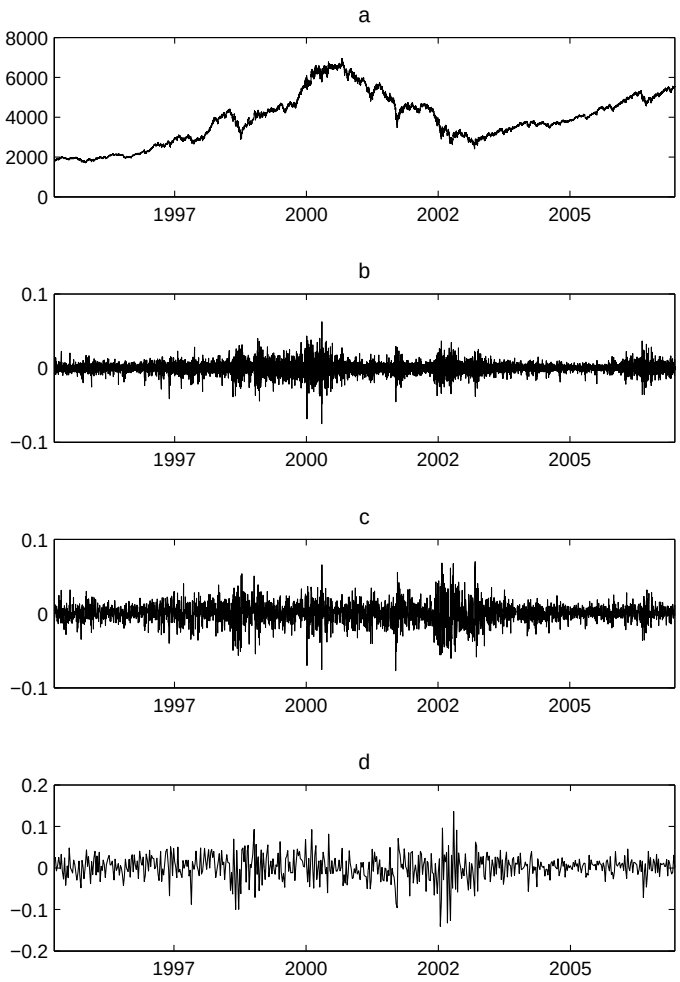
$$r_t = \ln(S_t) - \ln(S_{t-\tau}). \tag{3.5}$$

As a measure of volatility we take the magnitude of return $|r_t|$. Note that similar results could be obtained for squared returns and, most generally, for $|r_t|^\alpha$ (Bollerslev et al., 1992; Ding and Granger, 1996; Ding et al., 1993), but for $\alpha = 1$ the long memory properties are more pronounced (Ding et al., 1993; Forsberg and Ghysels, 2007; Ghysels et al., 2006).

Figures 3.1 and 3.2 represent the time series of index values and returns, computed over different time intervals. For the CAC40 index we compute 15-minutes, daily and weekly returns, and for the DJIA index - daily, monthly and quarterly returns. On both

data sets the phenomenon of volatility clustering can be easily identified: long-lasting and persistent periods of returns with high magnitude (positive and negative) alternate with low volatility periods. High volatility is rarely observed on up-going market trend. Large fluctuations are characteristic of trend reversals and slumps.

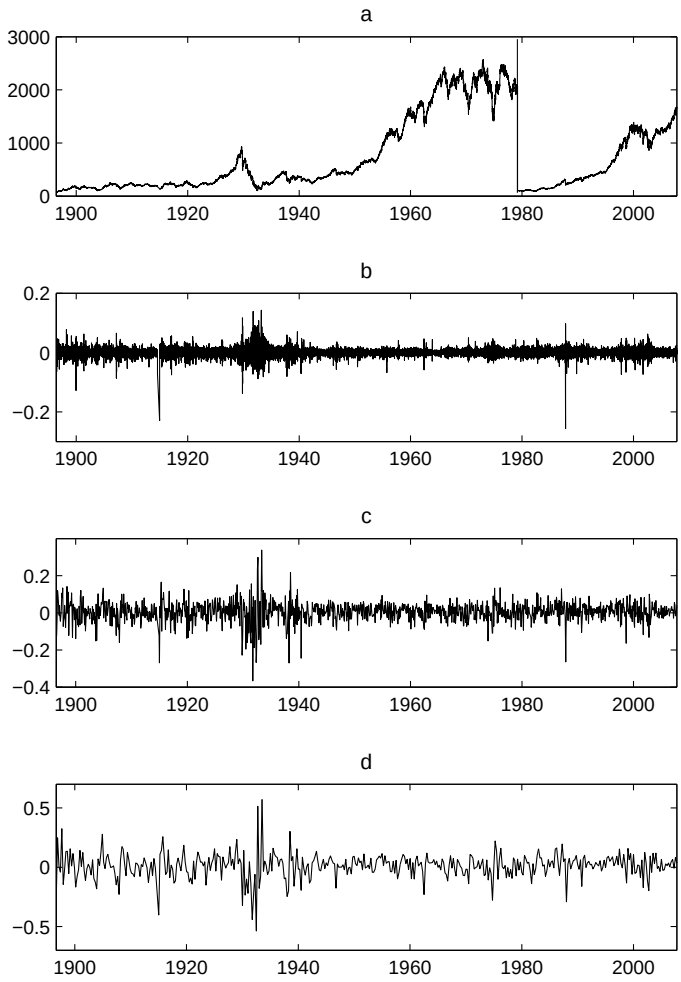
FIGURE 3.1: Returns on the CAC40 Index



Source: Euronext, CAC40 index from 20/03/1995 to 29/12/2006 at 15-minutes intervals, 100881 observations. the figure shows a: index values; b: 15-minute returns, 100880 observations; c: daily returns, 2953 observations; d: weekly returns, 590 observations.

Now consider the form of the probability distribution of returns, computed over different time intervals (Figures 3.3 and 3.4). For 15-minutes returns on the CAC40 index the distribution is clearly leptokurtic: the deviation from the normal curve in the tails is significant. As the frequency of observations is reduced this deviation decreases. This can be interpreted as an effect of the central limit theorem, though the the adequacy

FIGURE 3.2: Returns on the DJIA Index

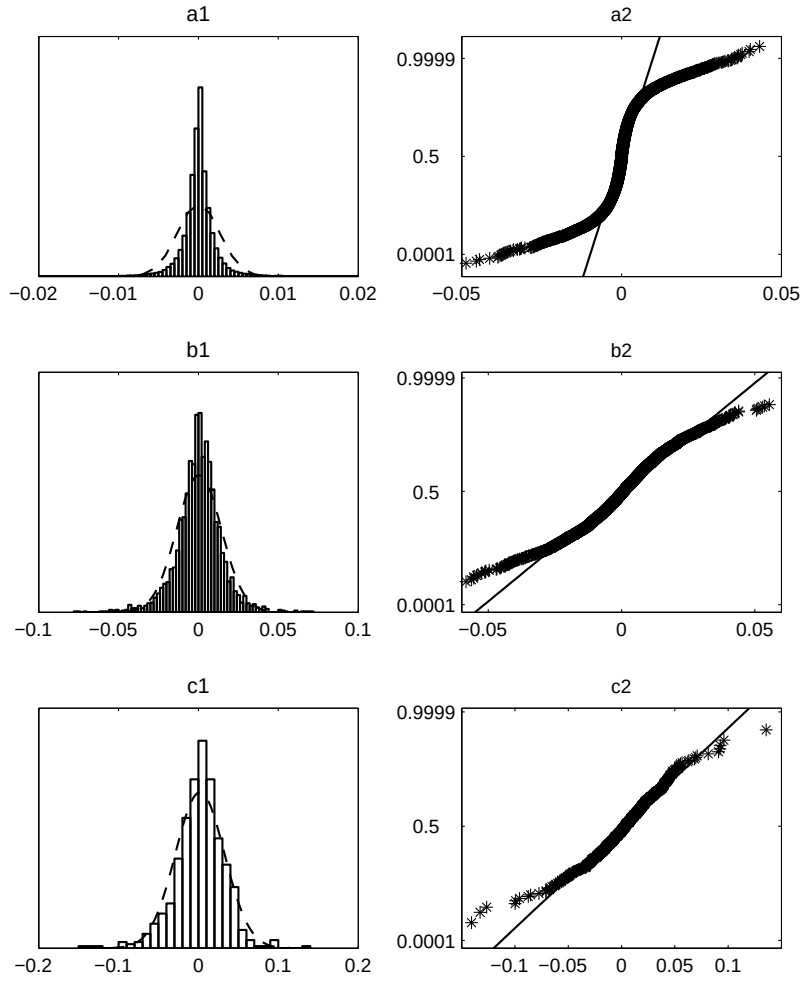


Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007, 28864 observations. The figure shows a: index values (for visualization purposes the values of index are reset to 100 at the beginning of the period and then again at 01/01/1979); b: daily returns, 28863 observations; c: monthly returns, 2953 observations; d: quarterly returns, 444 observations.

of hypotheses underlying its various forms is subject to debate among researchers[2]. For weekly returns fat tails are still observed, especially in the left side of the distribution, corresponding to negative returns. However, a relatively small number of observations at this frequency (590) does not allow a precise judgment about the distribution of extreme values in returns. For the DJIA case we have a larger sample (2953 observations). As in the previous case, extreme negative returns are observed much more frequently than the normal probability model predicts. For monthly returns the deviation in tails is smaller, but the size of the sample is not sufficient for final conclusions.

[2]The distribution of logarithmic returns at finite horizons can hardly be expected to follow the normal law exactly due to the fact that the support of normal distribution is the whole real line, while realizations of infinite prices of assets are impossible

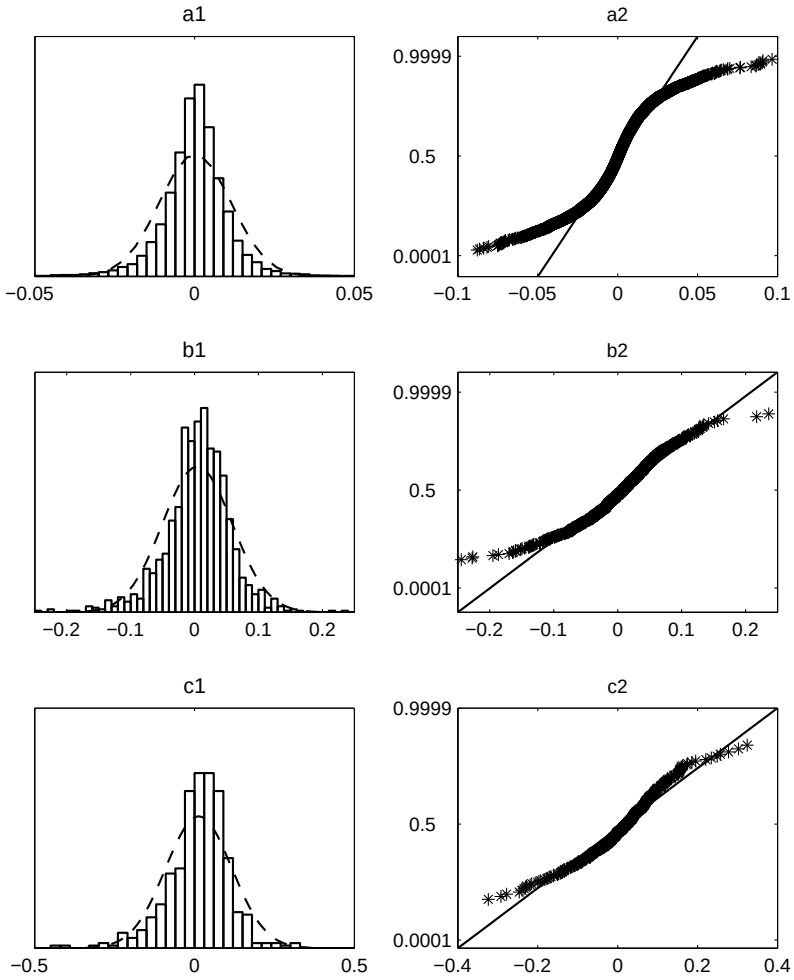
FIGURE 3.3: Probability Distribution of Returns on the CAC40 Index



Source: Euronext, values of index CAC40 from 20/03/1995 to 29/12/2006 at 15-minutes intervals, 100881 observations. The figure shows a1: histogram of the distribution density and its log-normal approximation for 15-minutes returns, 100880 observations; a2: probability plot for the same data, i.e. empirical cumulative distribution function (cdf), compared with the theoretical normal cdf (if the normal distribution perfectly approximates the empirical distribution, all points are on the diagonal straight line); b1,2: the same for daily returns, 2953 observations; c1,2: the same for weekly returns, 590 observations.

We find that, as the time horizon of returns increases, the distribution approaches to the normal law, but this convergence is very slow. Indeed, monthly logarithmic returns are obtained by summing up more than six hundred 15-minutes returns, so if assumptions of the classical central limit theorem were satisfied, the distribution would have been very close to Gaussian. But fat tails do not disappear even at that horizon. As we will show later, the question of whether a sufficiently long horizon, at which returns are normal, exists is important for building models of volatility at multiple horizons. Clearly, a strict empirical answer to this question cannot be obtained: if such horizon exists, it should be very long (longer than 3 month), but we do not dispose of sufficiently

FIGURE 3.4: Probability Distribution of Returns on the DJIA Index

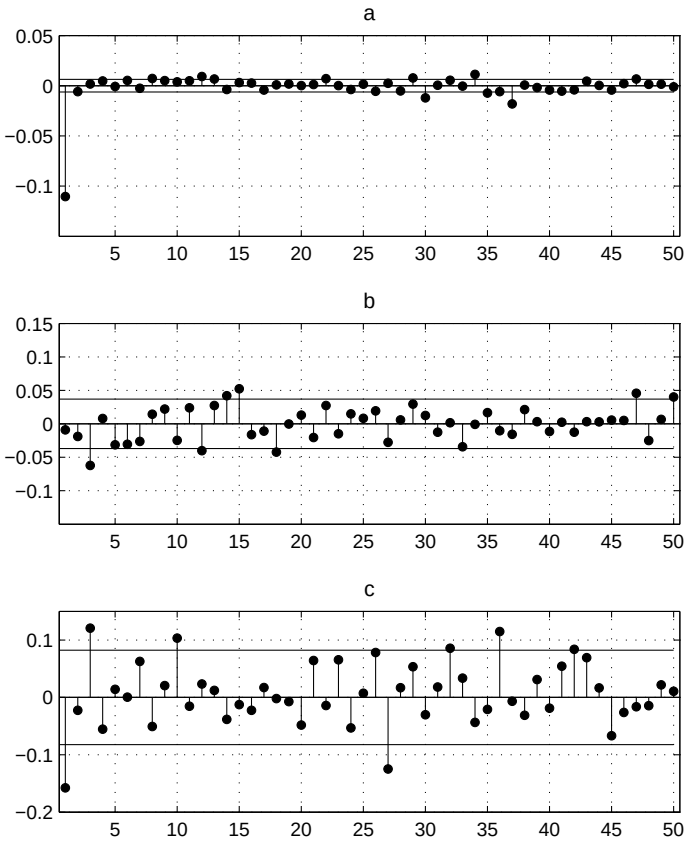


Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007, 28864 observations. The figure shows a1: histogram of the distribution density and its log-normal approximation for daily returns, 28863 observations; a2: probability plot for the same data, i.e. empirical cumulative distribution function (cdf), compared with the theoretical normal cdf (if the normal distribution perfectly approximates the empirical distribution, all points are on the diagonal straight line); b1,2: the same for monthly returns, 2953 observations; c1,2: the same for quarterly returns, 590 observations.

long samples to accurately carry out normality tests at such horizons. In fact, the DJIA time series is the longest time series currently available in financial economics.

The analysis of the dependence structure in returns confirms the intuitions from the visual observation of time series profiles. First, autocorrelations in returns are weak at all frequencies (Figures 3.5 and 3.6). We only notice significant positive autocorrelation between consecutive 15-minutes returns, which are induced by the microstructure effects, falling out of the scope of this study (see Zhou, 1996, for details). For the CAC40 index we also record small negative autocorrelation in consecutive weekly returns, which can

FIGURE 3.5: Sample ACF for the Returns on the CAC40 Index



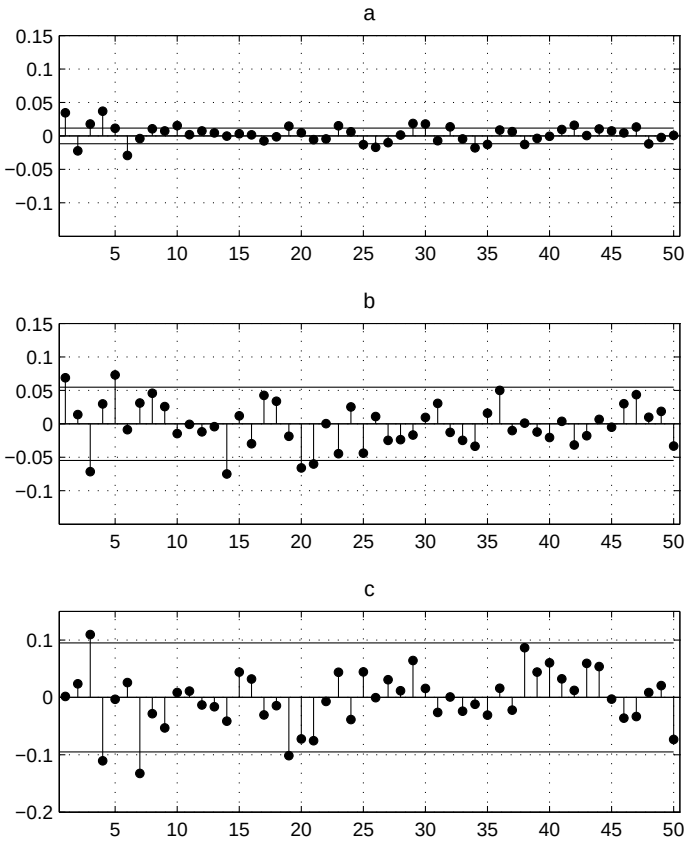
Source: Euronext, values of index CAC40 from 20/03/1995 to 29/12/2006 at 15-minutes intervals, 100881 observations. The figure shows a: ACF for 15-minutes returns, 100880 observations; b: the same for daily returns ,2953 observations; c: the same for weekly returns, 590 observations. Horizontal solid lines show confidence intervals for autocorrelations, computed under assumption that returns are normal white noise.

probably be explained by the "contrarian" effect[3], and positive correlation for lag 3 in weekly returns, which is probably a statistical artifact. For the returns on DJIA index no significant autocorrelations in returns are found.

The ACF computed for the absolute values of returns presents a big contrast (Figures 3.7 and 3.8). For magnitudes of returns on CAC40 positive autocorrelations are persistently significant up to very large lags at all frequencies of observation (15-minutes, daily and even weekly). Thus, at a 100-days lag correlations in daily volatilities are still significant, and for weekly returns they vanish no sooner than at lag 30 weeks (more than half of a year). The form of ACF can hardly be described by exponential decay, which characterizes the ARMA (autoregressive moving average) models. This illustrates long-range dependence in volatility. Daily volatilities of DJIA index display even stronger

[3]Contrarian strategy consists in selling stock that outperformed in past and buying those that underperformed, expecting trend reversal (see Conrad et al., 1997, and other behavioral finance literature)

FIGURE 3.6: Sample ACF for the Returns on the DJIA Index

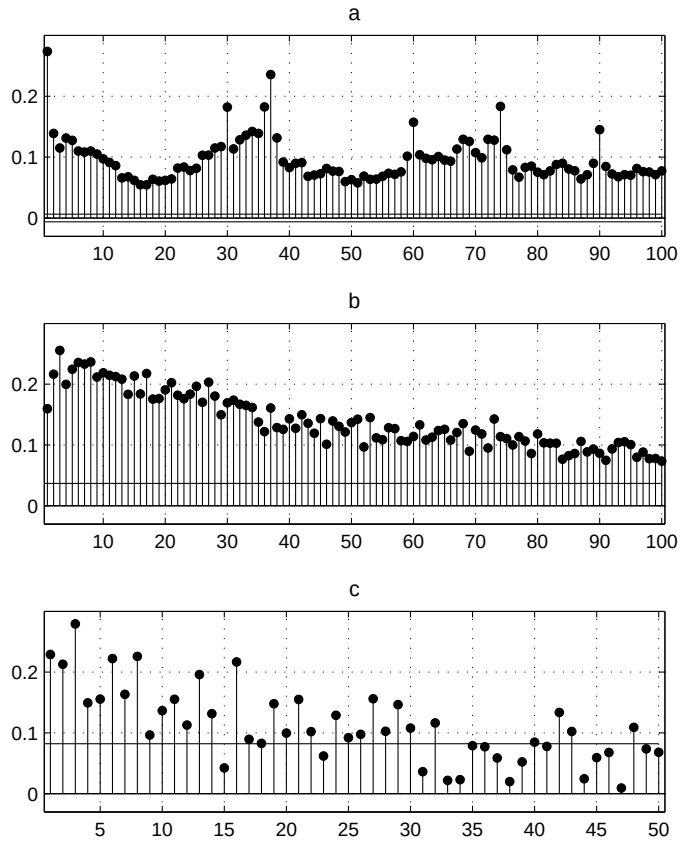


Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007, 28864 observations. The figure shows a: ACF for daily returns, 28863 observations; b: the same for daily returns ,2953 observations; c: the same for quarterly returns, 590 observations. Horizontal solid lines show confidence intervals for autocorrelations, computed under assumption that returns are normal white noise.

autocorrelations - they are still significant at 100-days lag and exceed 10% level. Autocorrelations in weekly absolute returns disappear at lags over 35 weeks, and in quarterly absolute returns - at 4 quarters. So long-range dependence can be observed both in high-frequency and in daily observations of volatility.

Figures 3.9 and 3.10 show the estimated spectrum of fluctuations of returns and their absolute values (data are taken at the highest available frequency). The spectral density is estimated by the eigenvectors of the correlation matrix method with maximum lag 10 (Marple, 1987, p. 373-378). Normalized frequencies (in radians per sample length) are shown on the X-axis and pseudospectrum values in decibels are on the Y-axis. The spectrum of fluctuations in returns' magnitudes (volatility) has a peak at frequency close to zero, so that a significant part of the variation in volatility corresponds to the fluctuations, whose duration is comparable with the sample length. This observation also

FIGURE 3.7: Sample ACF for the Magnitudes of Returns on the CAC 40 Index



Source: Euronext, values of index CAC40 from 20/03/1995 to 29/12/2006 at 15-minutes intervals, 100881 observations. The same as on Figure 3.5, but instead of returns their absolute values are used.

characterizes long memory: if the ACF decays at linear speed, the longest fluctuations' "cycle"[4] that can be observed equals the length of the sample.

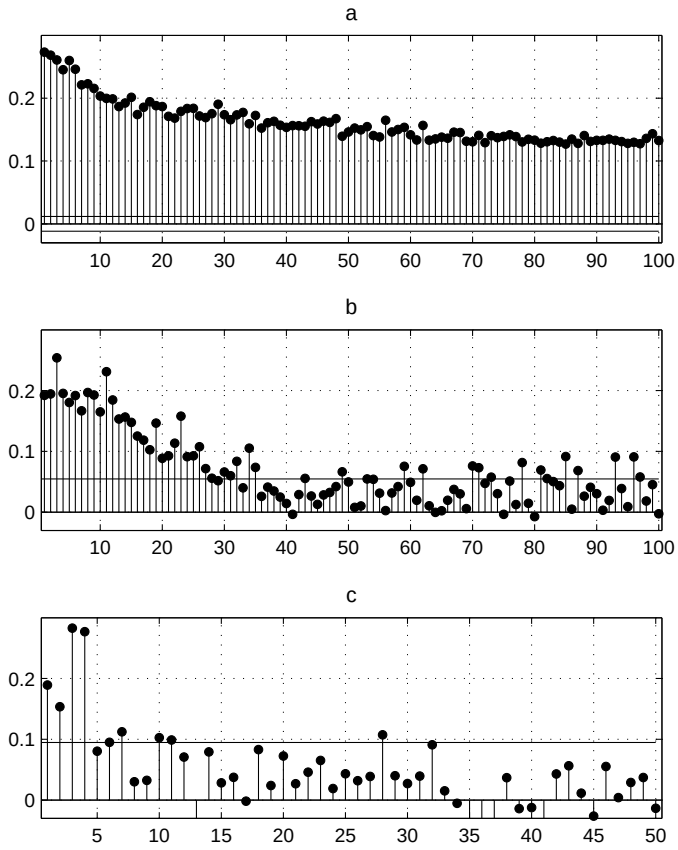
Our empirical results illustrate the presence of long memory in volatility time series and the non-Gaussian character of the distribution of returns, especially at high observation frequencies. In the next section we explain how these properties can be reproduced by the models proposed in financial literature.

3.3 ARCH/GARCH Family of Volatility Models and Extensions

The key feature of the models proposed for stock price dynamics, has always been their capacity to reproduce the empirical properties of volatility in financial time series, and above all, the phenomenon of volatility clustering. It is appropriate to start the survey

[4]In this context the term "cycle" is used in stochastic sense rather than in strict deterministic sense.

FIGURE 3.8: Sample ACF for the Magnitudes of Returns on the DJIA Index



Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007, 28864 observations. The same as on Figure 3.6, but instead of returns their absolute values are used.

with autoregressive conditional heteroscedasticity (ARCH) models, used for the first time by Engle (1982) to represent inflation and later by Engle and Bollerslev (1986) for stock and FX market data. Returns in the ARCH model are represented as the sum of their conditional expectation and a Gaussian[5] disturbance of varying magnitude:

$$r_t = E(r_t|I_{t-1}) + \sigma_t \epsilon_t \tag{3.6}$$

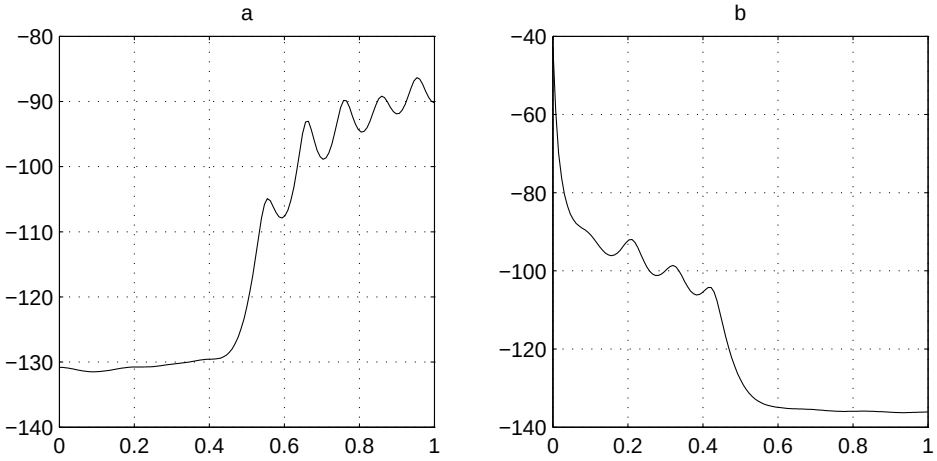
with $\varepsilon_t \sim iid\ N(0,1)$, I_t the information set at date t, defined as the natural filtration of the price process, and σ_t the magnitude of the disturbance term, satisfying:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \ldots + \alpha_q r_{t-q}^2 \tag{3.7}$$

with $\alpha_0 > 0, \alpha_i \geq 0$ for $\forall i > 0$ and $\sum_{i=1}^{q} \alpha_i < 1$. The parameter q specifies the depth of memory in the variance of the process.

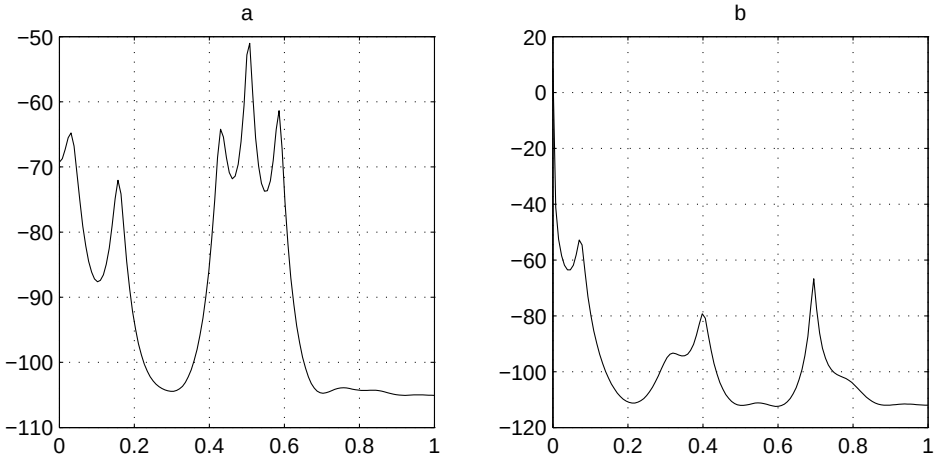
[5]In general, normality condition for the noise is not necessary

FIGURE 3.9: Sample Spectrum Density Function for the Returns on the CAC40 Index and their Magnitudes



Source: Euronext, values of index CAC40 from 20/03/1995 to 29/12/2006 at 15-minutes intervals, 100881 observations. the figure shows a: pseudospectrum for 15-minutes returns, 100880 observations; b: pseudospectrum for absolute values of returns. The spectral density is estimated by the eigenvectors of the correlation matrix method with maximum lag 10 (Marple, 1987, p.373-378). On the X-axis: normalized frequencies (in radians per sample length), on the Y-axis: pseudospectrum values in decibels.

FIGURE 3.10: Sample Spectrum Density Function for the Returns on the CAC40 Index and their Magnitudes



Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007, 28864 observations. The same as on Figure 3.9, but using daily returns.

A natural extension of ARCH is the generalized ARCH model (GARCH), first proposed in Bollerslev (1986) and widely used until know in the context of volatility forecasting (for example, see Bollerslev, 1987; Bollerslev et al., 1992; Hansen and Lunde, 2005). The model reads:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i r_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 = \alpha_0 + \alpha(L,q) r_t^2 + \beta(L,p)\sigma_t^2 \tag{3.8}$$

with L^n the lag operator of order n and $a(L,n)$ the operator of the form $\sum_{i=1}^{n} a_i L^i$, applied to a time series. So $a(L,q)X_t$ stands for $\sum_{i=1}^{q} a_i X_{t-i}$ and equation (3.8) can be rewritten:

$$[1 - \alpha(L,q) - \beta(L,p)]\, r_t^2 = \alpha_0 + [1 - \beta(L,p)]\left(r_t^2 - \sigma_t^2\right), \tag{3.9}$$

which corresponds to an ARMA model for the squared returns with parameters $\max\{p,q\}$ and p because $E(r_t^2 - \sigma_t^2 | I_{t-1})$ is an *iid* centered variable. To provide for the stability of the process, i.e. finite variation of the disturbances $\sigma_t \varepsilon_t$, all roots of the equations $\alpha(L_q) = 0$ and $1 - \alpha(L_q) - \beta(L_p) = 0$ must lie outside the unit circle. For GARCH(1,1) this constraint takes a simple form $\alpha + \beta < 1$. Sufficient and necessary conditions of strict stationarity, ergodicity and existence of moments of the GARCH-models are studied in Ling and McAleer (2002a,b).

GARCH models reproduce volatility clustering, observed empirically in financial time series (this is why volatility clustering is sometimes called GARCH-effect). The theoretical ACF of the process GARCH(1,1) decays at geometric speed, given by the sum $\alpha + \beta$. The closer this sum gets to unity, the more persistent autocorrelations are. In practice the estimates of $\alpha + \beta$ are often close to unity (Bollerslev et al., 1992). So the sample ACF for GARCH(1,1) is hard to distinguish from the long memory case, for which property (3.2) is verified.

The parameters of ARCH/GARCH models are usually estimated by the maximum likelihood method. The log-likelihood function for the Gaussian error case reads:

$$\ln L = -\frac{1}{2} \sum_{t=1}^{T} \left(2 \ln \sigma_t + \varepsilon_t^2\right) \tag{3.10}$$

If the normality assumption is violated, a quasi-maximum likelihood (QML) estimation procedure is possible (the prefix "quasi" means that statistical inference is made under possible model misspecification). QML estimates of parameters are consistent under finite variance of disturbances (i.e. if $\alpha + \beta < 1$) and asymptotically normal if the fourth moment of disturbances is finite (Ling and McAleer, 2003).

The main drawback inherent to GARCH(1,1) is that its memory is not long enough, because the ACF decreases too fast, though possibly from high values of autocorrelation. When $\alpha+\beta$ is not very different from one, GARCH(1,1) degenerates to a process, called integrated GARCH by Engle and Bollerslev (1986). This model is non-stationary and implies permanent (non-vanishing) effect of initial conditions on the price dynamics and thus can hardly pretend to correctly represent reality.

An alternative approach consists in using processes, whose theoretical properties imply the presence of long memory. An early example of such process is the fractal Brownian motion of Mandelbrot and Van Ness (1968). It is a continuous-time Gaussian process with zero drift, whose ACF has the form:

$$C(\tau) = E(W_t^H W_{t-\tau}) = \frac{1}{2}\left(|t|^{2H} + |t-\tau|2H - |\tau|^{2H}\right), \tag{3.11}$$

where W_t^H denotes a fractional Brownian motion with parameter $H \in (0,1)$ at time $t \in [0,T], t \in \Re$, such as $0 \leq \tau \leq t \leq T$. The spectral density of the process reads:

$$\Psi(x) = 4\sigma^2 c_H \sin^2(\pi x) \sum_{i=-\infty}^{\infty} (|x+i|)^{-2H-1} \tag{3.12}$$

with $-\frac{1}{2} \leq x \leq \frac{1}{2}$, σ^2 the variance of the process and c_H a positive constant. It is easy to notice that with $H = \frac{1}{2}$ the process degenerates to an ordinary Brownian motion and with $H > \frac{1}{2}$ it has stationary dynamics with long memory.

In the ARCH/GARCH framework the fractionally integrated process proposed in Granger and Joyeux (1980); Hosking (1981) is a discrete analogue of the fractional Brownian motion and it is defined by:

$$(l-L)^d X_t = \varepsilon_t \tag{3.13}$$

with $\varepsilon_t \sim iid\ N(0, \sigma_\varepsilon^2)$ and operator $(l-L)^d, 0 < d < 1$ is an infinite series of the below form:

$$(l-L)^d = \sum_{i=0}^{\infty} \frac{\Gamma(i-d)}{\Gamma(-d)\Gamma(k+1)} L^i, \tag{3.14}$$

where $\Gamma(\cdot)$ stands for the Gamma-function. The spectral density of the process reads:

$$\Psi(x) = \frac{\sigma_\varepsilon^2}{\left(4\sin^2(\pi x)\right)^d} c_H \sin^2(\pi x) \sum_{i=-\infty}^{\infty} (|x+i|)^{-2H-1} \tag{3.15}$$

with $-\frac{1}{2} \leq x \leq \frac{1}{2}$. For $|d| < \frac{1}{2}$ the process has stationary dynamics with hyperbolic decay of the ACF, thus displaying long memory.

Fractional Brownian motion was proposed as a model of price dynamics in Mandelbrot (1971) and later in many studies that aimed at estimating the parameter H in (3.11) empirically (see Mandelbrot and Taqqu, 1979)). But taking this approach means to accept the presence of long-range correlations in returns themselves and not only in their magnitudes. As shown in Heyde (2002), to generate long-range dependence in magnitudes of returns the memory parameter of the process must satisfy $\frac{3}{4} \leq H \leq 1$, which clearly contradicts empirical evidence.

As follows from the above discussion, models that straightforwardly exhibit long-range dependence in magnitudes of returns rather than in returns themselves could be more realistic. One of the most popular models of this kind is fractionally integrated GARCH (FIGARCH) proposed in Baillie et al. (1996) and Bollerslev and Mikkelsen (1996). The process for the variance of returns is given by:

$$[1 - \beta(L,p)]\,\sigma_t^2 = \alpha_0 + \left[1 - \beta(L,p) - \phi(L)(1-L)^d\right] r_t^2 \tag{3.16}$$

with $\phi(L) = [1 - \alpha(L,q) - \beta(L,p)]\,(1-L)^{-1}$. If d tends to one the model degenerates to IGARCH, discussed above.

A large number of other models, belonging to the GARCH family, were proposed to improve the forecasting power of GARCH(1,1). Among these models, the GARCH-in-mean first proposed by Engle et al. (1987) supposes that expected return increases with volatility and thus takes into account the effect of the varying risk premium. Other models include the effects of asymmetry and leverage, introduced in section 3.2. Among the most influential models we can mention the GJR model (from Glosten-Jagannathan-Runkle), proposed in Glosten et al. (1992), the exponential GARCH with leverage effect, in addition eliminating some undesirable constraints on the values of parameter estimates (Nelson, 1991) and generalized quadratic ARCH (GQARCH) by Sentana (1995). Non-linear extensions of GARCH (often called NGARCH) have also been proposed. They generalize the form of dependence of current variance on past observations of returns. This class includes models where volatility switches between "high" and "low" regimes (Higgins and Bera, 1992; Lanne and Saikkonen, 2005). The study by Hansen and Lunde (2005) of the predictive power of ARCH-models uses 330 various specifications. A more detailed description of some of them can be found in Morimune (2007). Derivatives pricing under the GARCH-like dynamics of the underlying asset is discussed in Barone-Adesi et al. (2008); Duan (1995); Ritchken and Trevor (1999).

Among all extensions including the jump component in the price dynamics is of particular importance (Bates, 1996; Eraker et al., 2003). This generates fat tails in the distribution of returns, a property that is characteristic of empirical data. As early as in 1960s, Mandelbrot proposed to use stable Levy processes (power law processes with

infinite variance) for this purpose (Mandelbrot, 1963). The properties of long memory processes, in which innovations are generated by Levy processes, are studied in Anh et al. (2002). Chan and Maheu (2002) proposed a rather general model, in which the intensity of price jumps is modeled by an ARMA process and volatility exhibits GARCH-effect.

All the above-mentioned extensions of GARCH are defined in discrete time. A continuous-time analogue of GARCH(1,1) was first studied by Drost and Werker (1996). They establish a link between GARCH and stochastic volatility models, which are are discussed in the next section. It is important that the estimates of the parameters of the discrete time GARCH(1,1) in, obtained for arbitrary chosen frequency of observation, can be converted to the parameters of a continuous process. This result is related to the time aggregation property of GARCH models that will be discussed in section 3.6. Continuous-time GARCH models with innovations driven by jump processes are described in Drost and Werker (1996) and more recently in (Klüppelberg et al., 2004).

Portfolio management and basket derivatives pricing applications motivate the study of multi-dimensional conditional heteroscedasticity models, accounting for correlations between assets. The first model of this kind, called constant conditional correlation model (CCC), was developed by Bollerslev (1990). The returns on each asset follow a one-dimensional GARCH process and conditional correlations are constant. So any conditional covariance is defined as the product of a constant correlation by the time-varying independent standard deviation of returns. The main advantage of CCC is the simplicity of estimation and interpretation. The main drawback is the absence of interdependence in conditional volatilities of assets. Besides, it does not account for leverage, asymmetry and, clearly, for possible changes in correlations. A more general model with constant correlations, introducing asymmetry, is studied in Ling and McAleer (2003). Engle (2002) further generalized CCC, allowing for GARCH-like dynamics in correlations. The model was named DCC, standing for dynamic conditional correlations. The dynamics of correlations in DCC is similar for all assets. This constraint is weakened in Billio et al. (2006).

3.4 Stochastic Volatility Models

Conditional heteroscedasticity models have only source of randomness. The variance of the returns process is some function of its past realizations (for example, a linear combination of lagged squared returns). An alternative approach is to set up a simple model for returns, for instance given by (3.1). Instead of considering σ as a parameter, one can model it as a separate stochastic process. Two sources of randomness thus emerge. This idea is the concept of stochastic volatility.

The first stochastic volatility model was proposed in Taylor (1982). It assumes that log-volatility is an AR(1) process:

$$\begin{aligned} r_t &= \mu\sigma_t\varepsilon_t \\ \ln\sigma_t^2 &= \phi\ln\sigma_{t-1}^2 + \nu_t, \end{aligned} \tag{3.17}$$

where μ is some positive constant, included in the model to get rid of the constant term in the volatility process, and ϕ is the autoregression parameter that determines memory in volatility. The properties of the autoregressive stochastic volatility (ARSV) models were studied by Andersen (1994); Capobianco (1996); Taylor (1994). In particular, under the constraint of the log-volatility process being stationary, the distribution of returns is fat-tailed and symmetric Bai et al. (2003). Returns are uncorrelated (but clearly not independent). The ACF for returns and squared returns decays at geometric speed, a characteristic of ARMA models.

Stochastic volatility have become popular in applications, related to pricing and hedging of financial derivatives. The returns are always given by a relation analogous to (3.1) where volatility is given by $\sigma_t = f(X_t)$. Usually X_t is an Ito process, so the whole model reads:

$$\begin{aligned} \frac{dS(t)}{S} &= \mu dt + \sigma dW(t) \\ \sigma_t &= f(X_t) \\ dX_t &= \theta(\psi - X_t)dt + g(X_t)dB_t \\ < W, B >_t &= \rho t \end{aligned} \tag{3.18}$$

with θ and ψ two constant parameters, $f(\cdot)$ and $g(\cdot)$ two continuous functions, verifying some regularity conditions (depending on the concrete specification), and ρ the correlation parameter, used to model the dependence between two Brownian motions that drive the price dynamics. Hull and White (1987) use the specification $f(X_t) = X_t$ with $\theta < 0, \mu = 0$ and $g(X_t) = \nu X_t$, which corresponds to the geometric Brownain motion for volatility. This model allows for easy derivation of closed-form formulas for option prices, but its properties are far from being realistic: the variance of returns is not bounded because the volatility process is not stationary.

An alternative specification proposed in Scott (1987) uses an Ornstein-Uhlenbeck (OU) process for volatility, taking $f(X_t) = X_t, g(X_t) = \nu$, so that, after a shock, volatility converges to its long-term average ψ at speed θ with "volatility of volatility" ν. Another possibility is the exponential OU model (Stein and Stein, 1991) with $f(X_t) = \exp X_t \;\; g(X_t) = \nu$, which is a continuous time analogue of ARSV(1) . Perhaps, the most popular is the Heston (1993) model, where $f(X_t) = \sqrt{X_t}, g(X_t) = \nu\sqrt{X_t}$. In

this case volatility is represented by a Cox-Ingersoll-Ross (CIR) model (see Cox et al., 1985).

The logic of the evolution of stochastic volatility models echoes the logic of GARCH extensions. Harvey and Shephard (1996) and later Jacquier et al. (2004) include the leverage effect in ARSV, letting two innovations in (3.17) be negatively correlated (in a continuous model of the form (3.18) this corresponds to the choice of $\rho < 0$). A stochastic volatility model with the effect of volatility on expected return, analogous to GARCH-M, is proposed in Koopman and Uspensky (2002). Jump component can be added to the stochastic volatility model by means of non-Gaussian processes. Instead of Brownian motion disturbances are generated by Levy processes (see Barndorff-Nielsen and Shephard, 2001; Chernov et al., 2003; Duffie et al., 2003; Eraker et al., 2003).

Various methods were proposed to incorporate long memory. Breidt et al. (1998); Harvey (1998) build discrete-time models with fractional integration, Comte and Renault (1998) propose a continuous time model with fractional Brownian motion. Chernov et al. (2003) considers models, in which stochastic volatility is driven by various factors (components). Such models generate price dynamics with slow decay in sample ACF, a characteristic of long memory models, though the data generating processes themselves do not possess this property (LeBaron, 2001a). In Barndorff-Nielsen and Shephard (2001) long memory effect is produced by superposition if an infinite number of non-negative non-Gaussian OU processes, which incorporates long-range dependence simultaneously with jumps. Besides, long-range dependence in stochastic volatility can be achieved using regime-switching models (Hwang et al., 2007; Liu, 2000; So et al., 1998).

Multi-dimensional extensions of stochastic volatility models are also available. Their comparative surveys can be found in Asai et al. (2006); Chib et al. (2006); Liesenfeld and Richard (2003). For some particular cases, notably for the Heston (1993) model, the problem of the optimal dynamic portfolio allocation is solved (Liu, 2007). Finally, similar to the GARCH literature, methods of derivatives pricing are developed for the case, when the underlying asset has stochastic volatility (Henderson, 2005; Heston, 1993; Hull and White, 1987; Maghsoodi, 2005).

Notice that realizations of volatility process, defined by models of type (3.17) and (3.18), are not observable (with reservations, discussed below), so that for their estimation we have to use returns and their transformations. Estimation methods can either be based on the statistical properties of returns (efficient method of moments, quasi-maximum likelihood method, etc.) or on building linear model for squared returns. A detailed survey of these methods can be found in Broto and Ruiz (2004).

The interest in SV models especially increased in recent years because an unobservable variable volatility turned to be an "almost observable" one. This occurred thanks to the availability of the intraday stock quotations, making possible precise non-parametric estimation of volatility. The concept of realized volatility (RV), defined as the square root of the sum of squared intraday returns (Andersen et al., 2001, 2003; Barndorff-Nielsen and Shephard, 2002b):

$$\hat{\sigma}_t^{RV} = \left(\frac{\sum_{i=1}^{M-1} r_{t,\delta}^2}{M-1} \right)^{\frac{1}{2}} \tag{3.19}$$

with $\hat{\sigma}_t^{RV}$ realized volatility of returns, $r_{i,\delta}$ logarithmic returns on the time interval $[i, i+\delta]$ c $\delta = \tau(M-1)^{-1}$, τ the length of period, over which volatility is computed (for example, one day) and M the number of price observations, available for that period. If in formula (3.19) we omit squared root and normalization on the number of observations, we obtain a realized variance estimation over the period τ, which is also often used in practice (Barndorff-Nielsen and Shephard, 2002a; Hansen, 2005).

Using realized volatility and variance is complicated by the correlation of returns at high frequencies, induced by market microstructure effects (also called microstructure noise, see Biais et al., 2005)). Methods of correction of realized variance for this noise and of the optimal choice of sampling frequency were proposed in (Bandi and Russel, 2008) and partially in some earlier studies. But the simplest method, most frequently used in practice, is to compute returns over sufficiently long time intervals, where correlations are negligible, but short enough to benefit from the information, contained in high-frequency data. A survey of the properties of realized volatility and its use in the context of stochastic volatility models is given in McAleer and Medeiros (2008).

A alternative non-parametric estimation of volatility can be obtained by aggregation of artificially computed returns, corresponding to the difference between the maximal $H_{t,i}$ and the minimal $L_{t,i}$ values of stock price over K intervals of time length $[i, i+\Delta]$, onto which a time period of interest τ is divided (Alizadeh et al., 2002; Christensen and Podolskij, 2007; Martens and van Dijk, 2007):

$$\hat{\sigma}_t^{RR} = \frac{1}{4\ln 2} \sum_{i=1}^{M-1} (\ln H_{t,i} - \ln L_{t,i}) , \tag{3.20}$$

where $\hat{\sigma}_t^{RR}$ is called realized range estimate. Clearly, the length of interval Δ must be chosen so as to contain several observations of prices. Statistical properties of the estimates, obtained in this way, can sometimes be better than those of realized variance. Another complement to realized variance is provided by the estimates with the process of bipower variation, which in particular allows estimation of the input of the jump

component to the integrated variance (Barndorff-Nielsen and Shephard, 2002c; Woerner, 2005).

One of the main challenges in building volatility models has always been its forecasting (Andersen and Bollerslev, 1998; Andersen et al., 1999; Christoffersen and Diebold, 2000; Ghysels et al., 2006; Granger and Poon, 2003; Hansen and Lunde, 2005; Hawkes and Date, 2007; Martens and Zein, 2004). The development of non-parametric methods of estimation with intraday returns allowed, on the one hand, to increase the quality of forecasts, based on the time series of historical prices, compared to implicit volatility methods, based on options prices calibration (Martens and Zein, 2004) and, on the other hand, made it possible to compare various SV models, taking non-parametric estimate of volatility for its actually observed values (Brooks and Persand, 2003; Corradi and Distaso, 2006).

3.5 Aggregation of Returns in Time

In section 3.2 we compared returns on stock indices CAC40 and DJIA, computed from observations at different frequencies. We showed that the form of the probability distribution of returns changes across frequencies of observation. At the same time dynamic properties of volatility, such as long memory in absolute returns and absence of linear correlations in returns themselves, are common for time series, corresponding to different frequencies. A series of practically important questions arises in this context. In what way the long memory phenomenon is related to the properties of returns at different horizons? Can volatility models, calibrated on data of some frequency, reproduce the properties of returns at other frequencies? Does it make sense to make estimations at several time horizons for the same time series of stock prices, and if yes, how to reconcile the results?

The answer to the first question was largely given by Mandelbrot and Van Ness in 1968. They pointed out that for some class of stochastic processes, their properties eswtablished on short horizons allow to completely describe the properties at longer horizons. A process X_t is called self-affine if there exists a constant $H > 0$[6], such as for any scaling factor $c > 0$ random variables X_{ct} and $c^H X_t$ are identically distributed:

$$X_{ct} \stackrel{L}{=} c^H X_t \tag{3.21}$$

[6]This parameter is called Hurst exponent. The name was given by Mandelbrot in honor of hydrologist Harold Hurst, who studied long-range dependence on the river Nile data.

Fractal Brownian motion, defined through the form of its ACF in (3.11) is an example of self-affine process. When the condition $\frac{1}{2} < H < 1$ is verified, this process possesses long memory, and for $H = \frac{1}{2}$ it is a standard Brownian motion with independent increments.

Notice that in general self-affinity with $H > \frac{1}{2}$ does not imply presence of long-range dependence and vice versa. As a counter-example we can evoke L-stable processes, verifying self-similarity condition (3.21), whose increments are independent and generated by stationary random variables whose probability distribution satisfies $P(X > x) \sim cx^{-\alpha}$, with $0 < \alpha < 2$. These processes have discontinuous paths and thus are helpful to represent heavy tails in returns. Thus two very different phenomena - long-range dependence and extreme fluctuations - can be observed within the class of the self-affine processes.

Intuitively, saying that a probability distribution is L-stable means that the form of distribution does not change (i.e. is invariant upto a scaling parameter) when independent random variables, following this probability law, are summed up. In particular, the normal distribution is L-stable and Brownian motion is an example of an L-stable process. It is the only L-stable process with continuous trajectory and independent increments. As explained above, the independence property is lost for fractional Brownian motion. But random variables with heavy tails (infinite variance) can also be used to generate self-affine processes.

A generalization of the class of self-affine processes is the class of multifractal processes, for which the self-affinity factor is no longer constant, so that the aggregation property reads:

$$X_{ct} \stackrel{L}{=} M(c)X_t \tag{3.22}$$

with $M(\cdot)$ - independent of X positive random function of scaling factor c, such as $M(xy) \stackrel{L}{=} M(x)M(y)$ for $\forall x, y > 0$. For strictly stationary (i.e. stationary in distribution) processes the following local scaling rule is verified:

$$X_{t+c\Delta t} \stackrel{L}{=} M(c)\left(X_{t+\Delta t} - X_t\right) \tag{3.23}$$

In the multifractal case we can define a generalized Hurst exponent as $H(c) = \log_c M(c)$ and rewrite (3.22) in the form:

$$X_{ct} \stackrel{L}{=} c^{H(c)}X_t \tag{3.24}$$

From (3.22) we can obtain scaling rules for the moments of X_t:

$$E(|X_t|^q) = c(q)t^{\zeta(q)+1} \tag{3.25}$$

with $c(q)$ and $\zeta(q)$ deterministic functions. The function $\zeta(q)$ is particularly important and is called scaling function. Substituting $q = 0$ in 3.25, it is straightforward to notice

that that the constant term in this function must be equal to one. For a self-affine process, which can also be called monofractal, the scaling function is linear and can be written $\zeta(q) = Hq - 1$. Applying Hölder inequality to (3.25) we can show that $\zeta(q)$ is always concave and that it becomes linear when $t \to \infty$. This implies that a multifractal process can only be defined for a finite time horizon, because beyond some horizon monofractal properties must prevail.

Alternatively (see Castaing et al., 1990) a multifractal process can be defined through the relation between the probability density functions of the increments of the process, computed for time intervals of different lengths l and L, such as $L = \lambda l, \lambda > 1$. This relation reads:

$$P_l(x) = \int G(\lambda, u) e^{-u} P_L(e^{-u} x) du \tag{3.26}$$

with $P_l(\cdot)$ the probability density function of the increments $\delta_l X_t$ of the process X_t at time horizon l, so that $x = \delta_l X_t = X_{t+l} - X_t$ (remember that for stationary processes $\delta_l X_t \stackrel{L}{=} X_l$). So if X_t is the logarithm of stock price, then the increments of the process represent returns at different time horizons. The function $G(\lambda, u)$, whose form depends exclusively on the relation between the lengths of two horizons, is called a self-similarity kernel. In the simplest case of a self-affine process it takes the form:

$$G(\lambda, u) = \delta(u - H \ln \lambda) \tag{3.27}$$

with $\delta(\cdot)$ the Dirac function[7]. In this monofractal case one point is enough to describe the evolution of the distributions, since P_l and P_L are different only by the scaling factor. This explains the degenerated form of (3.27).

In the general multifractal case equation (3.26) has a simple interpretation. The distribution P_l is a weighted superposition of scaled density functions P_L, with the weights defined by the self-similarity kernel. In other words, P_l is a geometric convolution between the self-similarity kernel and the density function P_L. Self-similarity kernel is also called propagator of a multi-fractal process. We will further need definition (3.26) to establish the multifractal properties of the multiplicative volatility cascade.

The scaling properties in stock prices and FX rates volatility have recently been studied in several papers. In particular, Schmitt et al. (2000) and Pasquini and Serva (2000) show that the non-linearity of the scaling function $\zeta(q)$, observed empirically, is incompatible with additive monofractal models of stochastic volatility, based on Brownian motion. So far this class of models has been most popular both among practitioners and researchers in finance. Multifractal properties can be due to a multiplicative cascade

[7]The Dirac function $\delta(x)$ is equal to 0 in all points except $x = 0$, and to infinity at $x = 0$, so that the integral of the function is equal to 1.

of disturbances (information flows or reactions to news), similar to the cascade used to model the turbulence in liquids and gazes. We discuss this issue later in more detail.

Interestingly, the time aggregation properties of simple models of the type GARCH and ARSV do not provide an adequate representation of stock returns at multiple horizons simultaneously. As regards the most popular GARCH(1,1) and its continuous time stochastic volatility analogue Drost and Nijman (1993) and Drost and Werker (1996) show that they verify the scale consistency property, i.e. if returns at some short scale follow GARCH(1,1), they must do so at any long scale with the same parameters. To prove this result the authors had to relax the assumption of the independence of errors in the model (3.8), assuming only that α and β are the best linear predictors of variance and that residuals ε_t are stationary (the so-called weak form of GARCH). Scale consistency is at the same time a strength and a weakness of the GARCH model. On the one hand, the results of statistic inference are independent of the frequency of observation. On the other hand, strict scale invariance does not allow reproducing the evolution in the form of the volatility distribution with time horizons and thus contradicts the empirical evidence.

The above arguments demonstrate the need for a model of volatility, that would not only reproduce long-range dependence and/or the presence of heavy tails in stock return, observed at some fixed frequency, but would give adequate results for other horizons. Ideally, this would give the possibility to model the change in the form of the probability distribution of returns at different time horizons and to reproduce the multifractal properties of the corresponding time series.

3.6 The Hypothesis of Multiple Horizons in Volatility

Up to now we discussed the time aggregation of returns from a purely statistical point of view. We noticed that the time series of returns, observed at different frequencies, have different properties. Can these properties be related to the real economic horizons, at which economic agents act?

The economic hypothesis of multiple horizons in volatility supposes that the heterogeneity in horizons of decision-taking by investors is the key element of explaining the complex dynamic of stock prices. For the first time the idea that price dynamics is driven by actions of investors at different horizons was advanced in Müller et al. (1997). They suppose that one can distinguish volatility components, corresponding to particular ranges of fluctuation frequencies, that are of unequal importance to different market participants. The latter include speculators that use intraday trades, daily traders,

portfolio managers and institutional investors, each having its own characteristic time of reaction to news and frequency of operations on the market. From the economic point of view, frequencies of price fluctuations are associated with the periods between asset allocation decisions, or frequencies of portfolio readjustments by investors.

A parametric model of volatility at multiple horizons in the spirit of ARCH approach has been proposed in Müller et al. (1997) and further studied in Dacorogna et al. (1998). Current volatility is represented as a linear function of squared returns over different time periods in the past:

$$\sigma_t^2 = c_0 + \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{j} r_{t-i} \right)^2 \tag{3.28}$$

with $c_k \geq 0$ for all $k = 0, ..., n$, so that for $k = 0$ and $k = n$ the inequality is strict, and with r_t the logarithmic return. Thus the expression $\sum_{i=1}^{j} r_{t-i}$ represents log-return over the period of length j. By construction the resulting heterogeneous ARCH (HARCH) model accounts for the hierarchical structure of the correlations in volatilities. The main problems of this model are a big number of parameters and high correlations between independent variables, that make its identification very complicated. The authors propose to reduce the dimension of the problem, using the principal components method. Later Corsi (2004) proposed a model, having the same form as HARCH, but using realized volatilities at different horizons (daily, monthly, weekly) as independent variables. This reduces correlations between regressors and the number of parameters.

Zumbach (2004) proposed to define current (or efficient) volatility as a weighted sum of several components, corresponding to different time horizons. He considers $n + 1$ representative horizons, whose length $\tau_k, k = 0 \dots n$ increases dyadically: $\tau_k = 2^{k-1}\tau_0$. The component of volatility, corresponding to horizon k, is defined by the exponential moving average:

$$\begin{gathered} \sigma_{t,k} = \mu_k \sigma_{k,t-\delta t}^2 + (1 - \mu_k) r_t^2 \\ \mu_0 = \exp(-\frac{\delta t}{\tau_0}) \;\; \mu_k = \exp(-\frac{\delta t}{\tau_0 2^{k-1}}), k = 1 \dots n \end{gathered} \tag{3.29}$$

with r_t current return at the minimum time interval δt, at which prices are observed ($\delta t \leq \tau_0$). Supposing that time is measured in units of length δt, we choose for simplicity $\delta t = 1$. Then, using (3.29), we can obtain the expressions for returns and volatility at different horizons:

$$\begin{gathered} r_{t,k} = \frac{1}{\sqrt{\tau_k}} \left[\ln(S_t) - \ln(S_{t-\frac{\tau_k}{\tau_0}}) \right] \\ \sigma_{t,k} = \mu_k \sigma_{k,t-1}^2 + (1 - \mu_k) r_{t,k}^2 \end{gathered} \tag{3.30}$$

with the return $r_{t,k}$ at horizon $k = 2^{k-1}$ defined as the change in the logarithm of price, scaled to the minimal time period $\delta t = 1$. Finally, the resulting (efficient) volatility,

corresponding to the unit time period, reads:

$$\sigma_t = \sum_{k=1}^{n} c\, 2^{-(k-1)\lambda}\, \sigma_{t,k} = \sum_{k=1}^{n} \omega_k \sigma_{t,k} \tag{3.31}$$

with $1/c = \sum_{k=1}^{n} 2^{-(k-1)\lambda}$, which provides $\sum_{k=1}^{n} \omega_k = 1$. The decay of weights in (3.31) according to the power law provides for long memory in the magnitudes of returns. This model is close to FIGARCH that uses the fractional differencing operator to create long-range dependence (see section 3.3), but Zumbach's model has a clear interpretation in terms of multiple horizons hypothesis. Compared to HARCH, it uses less parameters (only four). Note, however, that empirical tests of (3.31) showed only a very slight increase in the forecasting power of the model, compared to GARCH(1,1).

Another model of volatility at multiple horizons, this time based on a modification of the ARSV model, was proposed in Andersen (1996) and Andersen and Bollerslev (1997). Here the heterogeneity of time horizons is interpreted in terms of different persistence of information flows that influence price variability. These information flows can be seen as factors of volatility, important to different types of investors. Current return is defined through the latent volatility, which is assumed proportional to the intensity of the aggregated information flow V_t:

$$r_t = V_t^{\frac{1}{2}} \xi_t \tag{3.32}$$

where ξ_t is an *iid* random process with zero expectation and unit variance. The information flow V_t is the result of simultaneous action of n different information flows $V_{t,j}$, each following a log-normal ARSV model of the type (3.17):

$$v_{t,j} = \alpha_j + v_{t-1,j} + \varepsilon_{t,j} \tag{3.33}$$

with $v_{t,j} = \ln V_{t,j} - \mu_j$, $\mu_j = E(\ln V_{j,t})$ and $\varepsilon_{t,j} \sim iid\ N(0, \sigma_j^2)$. The parameter α_j represents the persistence of the information flow j, supposed to be stationary ($0 \leq \alpha_j < 1$). Aggregation of information flows is accomplished with the geometric mean rule:

$$\ln V_t = \sum_{i=1}^{N} v_{t,j} \sum_{i=1}^{N} \mu_j \tag{3.34}$$

According to this definition the spectrum of $\ln V_t$ is the mean spectrum of all autoregressive processes, defined by equations of the form (3.33).

Representing the heterogeneity of the parameter α_j by a standard β-distribution, the authors study the dynamics of returns' magnitudes and of the odd moments of returns, finding evidence in favor of long-range dependence. Besides, the process, obtained

through the mixture of distributions, is self-affine. In particular, this implies that the ACF of volatility process decays at the same hyperbolic speed, whatever the frequency of returns observation.

Andersen and Bollerslev (1997) model has mostly explicative character (the authors try to explain long-range dependence by the heterogeneity of information flows), unlike the models described earlier that suppose identification of parameters and practical use in forecasting. It still does not explain the multifractality property, which is empirically observed in stock price volatility. Besides, the model does not have a direct microeconomic justification, based on decision-taking behavior of investors.

Explanation of the properties of volatility in the market microstructure models with heterogeneous investors is proposed in several studies. In particular, Brock and Hommes (1997) introduce the notion of adaptive rational equilibrium which is reached by investors, rationally choosing the predicting functions for future prices. The set of predictive functions is specified *a priori* and the criterion of choice is the quality of the forecasts, obtained by using these functions on historical data. Artificial markets of this type are also studied in Lux and Marchesi (2000), Chiarella and He (2001) and Anufriev et al. (2006), where investors choose between chartist (extrapolating the past) and fundamentalist strategies. Reproducing some of the empirical properties of stock prices, these models explain the paradox of excessive price volatility and volatility clustering to some extent. However, none of them accounts for the heterogeneity of time horizons. In a similar context LeBaron (2001b) studies the choice between strategies, based on historical data collected over different horizons. However, he does not explicitly model the rational choice of agents that rebalance portfolios at different frequencies.

Subbotin et al. (2010a) study the effect of multiple investment horizons on the price dynamics in a context of a pure exchange economy with one risky asset, populated with agents maximizing expected utility of wealth over discrete investment periods. Investors' demand for the risky asset may depend on the historical returns, so a wide range of behaviorist patterns is exploited. They establish necessary conditions under which the risky return can be an *iid* stationary process and study the compatibility of these conditions with different types of demand functions in the heterogeneous agents' framework. It is explicitly shown that conditional volatility of returns on the risky asset cannot be constant in many generic situations, especially if agents with different investment horizons exist on the market. So volatility clustering can be seen as an inalienable feature of a speculative market, which can be present even if all investors are so-called "fundametialists". Thus it is demonstrated that heterogeneity of investment horizons is sufficient to generate many stylized facts in returns' volatility.

A general weak point of artificial market models is the *a priori* character of assumptions about economic agents' behavior (which apparently has impact on the form of resulting market dynamics), and absence or insufficiency of analytic relation with the specification of volatility processes, used in practice. Thus, almost simultaneously with the model of artificial market, mentioned above, LeBaron (2001a) proposes a simple model of stochastic volatility with three factors, each given by an OU process (see section 3.4) with different speed of mean reversion, which has no direct link to the former theoretical model. A similar stochastic volatility model with multiple horizons was proposed in Perello et al. (2004). Molina et al. (2004) study its estimation by the Monte Carlo Markov Chains method. Models with multiple factors, given by OU processes, can successfully reproduce long-range dependence and leverage effect, but are scale-inconsistent due to the finite (and small) number of factors and do not have any analytic relation to the economic microstructure models, which could justify multiple horizons. The model by Barndorff-Nielsen and Shephard (2001) that uses a superposition of an infinite number of OU processes avoids the first of these two problems.

3.7 Modeling Multiple Horizons in Volatility and Econophysics Approach

The models of volatility at multiple horizons, described above, represent current volatility as a result of impact of factors (or components), varying at different frequencies. Such description of volatility has straightforward analogy in physics of liquids and gases. Hydrodynamics studies the phenomenon of turbulence, characterized by the formation of eddies of different sizes in the flows of fluids and gases, leading to the random fluctuations in thermodynamic characteristics (temperature, pressure and density). Most of the kinetic energy of a turbulent flow is contained in the eddies at large scales. Energy cascades from large scales to eddies structures at smaller scales. This process continues, generating smaller and smaller eddies, having hierarchical structure. The condition, under which laminar (i.e. normal) flow becomes turbulent, is determined by the so-called Reynolds number that depends on the viscosity of the fluid and on the properties of the flow. A statistical theory of turbulence was developed by Kolmogorov (1941), and a contemporaneous survey can be found, for example, in Pope (2000).

For the first time analogy between turbulence and volatility on the financial market was proposed in Ghashghaie et al. (1996). The authors noticed that the relation between the density of distribution of returns at various horizons is analogous to the distribution of velocity differentials for two points of a turbulent flow, depending on the distance between these points (so instead of physical distance, in finance we use distance in time).

The cascade of volatility can be interpreted in terms of the multi-horizon hypothesis of Müller et al. (1997).

An analytical multiplicative cascade model (MCM) was proposed in Breymann et al. (2000). Volatility is represented as a product of disturbances at different frequencies. Denote S_t a discrete stochastic process for the stock price and $r_t = \ln S_t - \ln S_{t-1}$ the log-return. In MCM the returns are driven by equation:

$$r_t = \sigma_t \, \varepsilon_t, \tag{3.35}$$

with $\varepsilon(t)$ some *iid* noise, independent from the scale structure of volatility, and σ_t stochastic volatility process that can be decomposed for a series of horizons $\tau_1, \ldots, \tau_n$ (here we suppose that τ_1 is the longest horizon), so that volatility at horizon $k \in \{2, \ldots, n\}$ depends on volatility at the longer horizon $k-1$ and some renewal process $X_{t,k}$:

$$\sigma_{t,k} = \sigma_{k-1}(t) X_{t,k} \tag{3.36}$$

So the multiplicative cascade for volatility reads:

$$\sigma_t = \sigma_{t,n} = \sigma_0 \prod_{k=1}^{n} X_{t,k} \tag{3.37}$$

At the initial time period t_0 all renewal processes $X_{t,k}$ are initialized as *iid* lognormal random variables with expectation $\mathrm{E}(\ln X_{t,k}) = x_k$ and variance $\mathrm{Var}(\ln X_{t,k}) = \lambda_k^2$. For transition from time t_n to time $t_{n+1} = t_n + \tau_n$ (recall that τ_n is the shortest time scale) we define:

$$X_{t_{n+1},1} = \left(1 - \mathbf{I}\left\{A_{t_{n+1},1}\right\}\right) X_{t_n,1} + \mathbf{I}\left\{A_{t_{n+1},1}\right\} \xi_{t_{n+1},1} \tag{3.38}$$

with $A_{t_{n+1},1}$ an event, corresponding to the renewal of process $X_{t,1}$ at time t_{n+1}, $\mathbf{I}\{\cdot\}$ the indicator function and $\xi_{t,1}$ lognormal *iid* random variables with expectation μ and variance λ^2. At any moment t_n the event $\{A_{t_{n+1},1}\}$ happens with probability p_1. By analogy $\{A_{t_{n+1},k}\}$ is defined as the renewal of process $X_{t,k}$ at moment t_{n+1}. The dynamics at horizons $k = 2, ..., m$ is defined iteratively by means of equation:

$$\begin{aligned} X_{t_{n+1},k} = \left(1 - \mathbf{I}\left\{A_{t_{n+1},k-1}\right\}\right) \big[& (1 - \mathbf{I}\left\{A_{t_{n+1},k}\right\}) X_{t_n,k} + \\ & \mathbf{I}\left\{A_{t_{n+1},k}\right\} \xi_{t_{n+1},k} \big] + \mathbf{I}\left\{A_{t_{n+1},k-1} \xi_{t_{n+1},k}\right\}, \end{aligned} \tag{3.39}$$

where for any k the random variables $\xi_{t,k}$ are *iid* log-normal with parameters μ and λ^2.

It follows from equation (3.39) that renewal at horizon k at moment t_{n+1} occurs if it has already occurred at the preceding, longer horizon $k-1$, or in case of the event $\{A_{t_{n+1},k}\}$ that happens withy probability p_k. Probabilities of renewal p_k must

be calibrated so that the average interval between to renewal events would be equal to the length of the corresponding horizon τ_k. For simplicity we can consider only dyadic horizons, i.e. those satisfying $\tau_{k-1}/\tau_k = 2$ for $k \in \{2, \ldots, n\}$. Using the properties of Bernoulii process, one can easily show that:

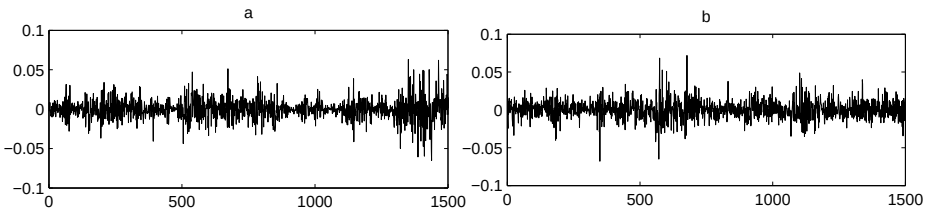
$$p_1 = 2^{1-n}, \quad p_k = \frac{2^{k-n} - 2^{k-n-1}}{1 - 2^{k-n-1}}, \quad k = 2, \ldots, n \tag{3.40}$$

Empirical adequacy of the model is confirmed by the properties of ACF of returns and their absolute values at different horizons, defined in a standard way: $r_{t,k} = \ln S_t - \ln S_{t-\tau_k}$. Arneodo et al. (1998) shows that under MCM assumptions the ACF of logarithms of absolute values of returns at all horizons decays at logarithmic speed:

$$\mathrm{Cov}(\ln |r_{(t+\Delta t),k}|,\ \ln |r_{t,k}|) \cong -\lambda^2 \ln \frac{\Delta t}{\tau_1},\ \ \Delta t > \tau_k \tag{3.41}$$

The last relationship can be used for identification of the "longest scale" in volatility (Muzy et al., 2001). From a practical point of view it is convenient to analyze MCM in an orthonormal wavelet basis, which simplifies simulations and allows to obtain analytical results of the the type of equation (3.41) (Arneodo et al., 1998).

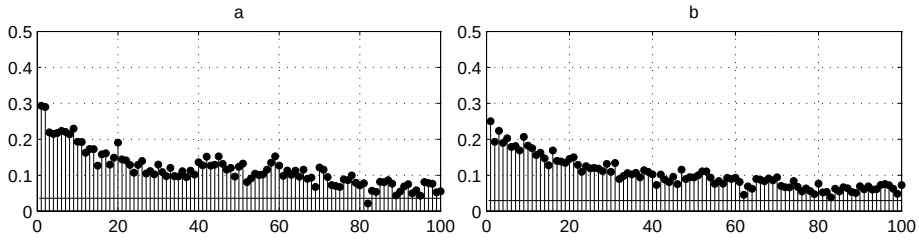
FIGURE 3.11: Simulation with Multiplicative Cascade Model and Real Data: Daily Returns



Left (a): daily returns on index CAC40 (source: Euronext, values of index CAC40 from 20/03/95 to 24/02/05). Right (b): daily returns, simulated with MCM at 14 horizons (from 15 minutes to 256 days). Returns are simulated for every 15 minutes and then are aggregated to daily time intervals.

Figures 3.11 and 3.12 show the results of simulation of MCM, compared with real data of index CAC40. The number of horizons in simulation is equal to 14, which allows to fit the speed of decay in the ACF, and other parameters are calibrated so as to match unconditional long-term estimates of the first two sample moments in the returns' distribution. Note that the figure shows the ACF for returns, aggregated into daily intervals, whereas the simulation itself was carried out at 15-minutes frequencies. This illustrates the most important property of the volatility cascade: clustering of volatility and long-range dependence robust to time aggregation, i.e. coexisting at multiple horizons.

FIGURE 3.12: Simulation with Multiplicative Cascade Model and Real Data: Sample ACF



Left (a): sample ACF for the magnitudes of daily returns on index CAC40 (source Euronext, daily values of index CAC40 from 20/03/95 to 24/02/05). Right (b): sample ACF for data, simulated with MCM at 14 horizons (from 15 minutes to 256 days). Returns are simulated for every 15 minutes and then are aggregated to daily time intervals. ACF is computed for daily data.

The MCM, described above, is called log-normal, because disturbances to volatility are log-normal. This does not mean that that the resulting distribution of returns is log-normal. Nothing prevents from specifying the model in a way that provides for fat tails at short horizons (see more about it below). In a form described above MCM allows to simulate data, corresponding to the observed financial time series in many properties. But its practical use is complicated because of the absence of strict parametrization and estimation methods.

The link between MCM (here we talk about multiplicative cascade in more general sense, not focused on Breymann et al. (2000) specification, described above) and multifractal processes is studied in Muzy et al. (2000). Consider dyadic horizons of length $\tau_n = 2^{-n}\tau_0$. The increment of some process X_t on interval τ_k, denoted $\delta_k X_t$, is linked to the increment on the longest scale through equation:

$$\delta_k X_t = \left(\prod_{i=1}^{k} W_i \right) \delta_0 X_t \tag{3.42}$$

with W_i some *iid* stochastic factor. In MCM the stochastic volatility process was defined in a similar way. The expression (3.42) can be rewritten in terms of a simple random walk in logarithms of local volatility:

$$\omega_{t,k+1} = \omega_{t,k} + \ln W_{k+1} \tag{3.43}$$

with $\omega_{t,k} = \frac{1}{2}\ln(|\delta_k X_t|^2)$. Notice that equation (3.41) with new notations corresponds to $\mathrm{Cov}\,(\omega_{t+\delta t,k}, \omega_{t,k})$. If disturbances $\ln W_i$ are normally distributed $N(\mu, \sigma^2)$, the distribution density $\omega_{t,k}$ denoted $P_k(\omega)$, satisfies:

$$P_k(\omega) = \left(N(\mu, \sigma^2)^{*k} * P_0 \right)(\omega) \tag{3.44}$$

with $*$ denoting the convolution operator, defined for two function $f(t)$ and $g(t)$ by the expression $(f * g)(t) = \int f(u)g(t-u)du$. Now it is straightforward to show that that equation (3.44) corresponds to the definition of multifractality in (3.26) with log-normal propagator of the form:

$$G_{\tau_k,\tau_0} = N(\mu, \sigma^2)^{*k} = N(k\mu, k\lambda^2) \tag{3.45}$$

In a similar way, a multifractal process, corresponding to (3.26), can be represented as a multiplicative cascade.

It follows from the above analysis that the MCM can be specified using a multi-fractal random walk. The class of these processes was proposed for volatility modeling in Bacry et al. (2001) and then generalized in Muzy and Bacry (2002); Pochart and Bouchaud (2002). A discrete version of MRW with step Δt can be obtained by summing up $t/\Delta t$ random variables:

$$X_{\Delta t}(t) = \sum_{k=1}^{t/\Delta t} \delta X_{\Delta t,k} \tag{3.46}$$

with $\delta X_{\Delta t,k}$ a noise, whose variance is given by stochastic process:

$$\delta X_{\Delta t,k} = \exp\left(\omega_{\Delta t,k}\right) \varepsilon_{\Delta t} \tag{3.47}$$

where $\omega_{\Delta t,k}$ is the logarithm of stochastic volatility, like in (3.43), and $\varepsilon_{\Delta t}$ is Gaussian noise, independent of ω. The definition of $\omega_{\Delta t,k}$ is based on the form of the autocovariance function, corresponding to the described above for MCM:

$$\begin{gathered} \mathrm{Cov}(\omega_{\Delta t,k}, \omega_{\Delta t,l}) = \lambda^2 \ln \rho_{\Delta t,|k-l|} \\ \rho_{\Delta t,m} = \frac{T}{(|m|+1)\Delta t}, \ |m| \leq \frac{T}{\Delta t} - 1 \\ \rho_{\Delta t,m} = 1, \ |m| > \frac{T}{\Delta t} - 1 \end{gathered} \tag{3.48}$$

Here T is the integral time, i.e. the longest horizon, after which multifractal properties are no more observed. To provide for finite variance of the increments of the process $X_{\Delta t}(t)$ at transition to the continuous time by taking $\Delta t \to 0$, we need to define the mean log-volatility in the following way:

$$\mathrm{E}(\omega_{\Delta t,k}) = -\mathrm{Var}(\omega_{\Delta t,k}) = -\lambda^2 \ln\left(\frac{T}{\Delta t}\right) \tag{3.49}$$

The MRW model is identified by three parameters: the variance of the process $X_{\Delta t}(t)$, the variance of logarithmic volatility process (λ^2) and the integral time T. These parameters can be easily clibrated using the form of the spectrum and of the ACF. MRW

can also be extended to multidimensional space (Muzy and Bacry, 2002). But the price to pay for the parsimony in parametrization of the model is the impossibility of direct and exact modeling of the interdependence between volatilities at different horizons, compared to the flexibility in this aspect, allowed by the HARCH or Zumbach model (see section 3.6). Lynch and Zumbach (2003) study the volatility cascade empirically through the correlations of historical and realized volatility and find that the structure of this cascade is different from the one observed in turbulence. This can be explained by the existence of "characteristic" horizons, corresponding to the frequencies of market operations, specific to investors of different types (daily traders, portfolio managers, pensions funds, etc.). Compared to traditional models of stochastic volatility of the form (3.18), MRW processes do not allow for leverage effect. Besides, the intuitively attractive property of volatility reversion to its mean level, present in OU processes, is lost. Anteneodo and Riera (2005) proposes an additive-multiplicative model of cascade that enriches the one described in this section by the mean-reversion effect. But its complexity is considerably higher.

An alternative approach to studying the properties of volatility, related to econophysics, consists in direct estimation of the evolution of probability distribution at different horizons. A necessary assumption for such analysis is Markov property of the cascade. Consider a series of horizons $\tau_0 < \tau_1 < \ldots < \tau_n$ (in this case, unlike the description of MCM, it is more convenient to numerate horizons in increasing order) and a process $\delta_k X$ of increments at horizons $\tau_k, k \in \{0, \ldots n\}$ of process X_t at some fixed time t (this can be a process of stochastic volatility or a trend-corrected price process, or its logarithm). By definition, Markovity for $\delta_k X$ means:

$$P_{k|k+1,\ldots,n}(x) = P_{k|k+1}(x), k = 0, \ldots, n-1, \tag{3.50}$$

where $P_{k|k+1}(\cdot)$ stands for the conditional density of the distribution of $\delta_k X$, given $\delta_{k+1} X$. Since

$$P_{k|k+1,\ldots,n}(x) = \frac{P_{k,\ldots,n}(x_k, \ldots, x_n)}{P_{k+1,\ldots,n}(x_{k+1}, \ldots, x_n)}, \tag{3.51}$$

it suffices to know the conditional densities at consecutive horizons and the distribution at the longest scale to define the joint distribution $P_{0,\ldots,n}$ of all increments. The last property is of special importance in finance. Using it, we can design an algorithm of simulation of a process with the same probability distribution of increments as in the empirical data (Nawroth and Peinke, 2006). Such algorithm can be useful for the implementation of a Monte Carlo algorithm in derivative pricing and portfolio management applications.

To verify if the Markov property is satisfied one can use the necessary condition, given by Chapman-Kolmogorov equation:

$$P_{m|k}(x) = \int P_{m|\delta_l X=u}(x)\ P_{l|k}(u)du, k < l < m \tag{3.52}$$

that can be checked for three different series of increments by direct comparison of the left and the right side of the equation. Empirical data for increments of exchange rates and stocks' volatilities are in good agreement with (3.52) and do not reject the Markovity hypothesis (Ausloos and Ivanova, 2003; Buchbinder and Chistilin, 2007; Cortines et al., 2007; Friedrich et al., 2000; Renner et al., 2001a). Notice that in MCM we did the Markov assumption implicitly, saying that volatility at each horizon is the result of adding multiplicative disturbance to the volatility at longer horizon.

For Markov processes conditional densities satisfy the Kramers-Moyal evolution equation (Risken, 1989, p.48-50):

$$-\tau\frac{\partial}{\partial\tau}P_{\tau|\tau_0}(x) = \sum_{k=1}^{\infty}\left(-\frac{\partial}{\partial x}\right)^k D_k(x,\tau)P_{\tau|\tau_0}(x) \tag{3.53}$$

Here we assume the length of horizons τ to be continuous. Coefficients $D_k(x,\tau)$ in Kramers-Moyal decomposition are defined as the limit at $\Delta_\tau \to 0$ of conditional moments $M_k(\delta_\tau X, \tau, \Delta\tau)$:

$$\begin{aligned} D_k(x,\tau) &= \lim_{\Delta\tau\to 0} M_k(x,\tau,\Delta\tau) \\ M_k(x,\tau,\Delta\tau) &= \frac{\tau}{k!\Delta\tau}\int (u-x)^k P_{\tau-\Delta\tau|\tau}(u)du \end{aligned} \tag{3.54}$$

In a general case all coefficients are different from zero, but according to Pawula's theorem, if $D_4(x,\tau) = 0$, then all coefficients in the decomposition starting from the third one are also equal to zero. This condition can also be verified empirically. If it is satisfied, (3.53) becomes a simple Fokker-Plank equation (also know as the second Kolmogorov equation):

$$-\tau\frac{\partial}{\partial\tau}P_{\tau|\tau_0}(x) = \left[-\frac{\partial}{\partial x}D_1(x,\tau) + \frac{\partial^2}{\partial x^2}D_2(x,\tau)\right]P_{\tau|\tau_0} \tag{3.55}$$

Unconditional density of the distribution of $\delta_\tau X$ at horizon τ satisfies the same differential equation.

Fokker-Plank equation describes the density of the stochastic process, given by Langevin equation:

$$-\tau\frac{\partial}{\partial\tau}x(\tau) = D_1(x,\tau) + \sqrt{D_2(x,\tau)}f(\tau) \tag{3.56}$$

with $f(\tau)$ the so-called Langevin force, which is usually modeled by Gaussian white noise. Thus, under a series of constraints, equation for stock prices and their volatilities can be obtained by estimation of Kramers-Moyal coefficients from (3.54) (Buchbinder and Chistilin, 2007; Cortines et al., 2007; Renner et al., 2001a). This unambiguously defines the evolution of the distribution from normal to fat tails. For example, Renner et al. (2001a) obtains the following form of coefficients, studying the increments in FX rates:

$$\begin{aligned} D_1(x,\tau) &= -\gamma x \\ D_2(x,\tau) &= \alpha\tau + \beta x^2 \end{aligned} \tag{3.57}$$

For a standard multifractal model of turbulent cascade (Castaing et al., 1990), described above, Kramers-Moyal coefficients take the form:

$$\begin{aligned} D_1(x,\tau) &= -\gamma(\tau)x \\ D_2(x,\tau) &= \beta(\tau)x^2 \end{aligned} \tag{3.58}$$

The resemblance of (3.58) and (3.57) evidences in favor of the analogy between turbulence and volatility. In Ausloos and Ivanova (2003) similar type of analysis is made for logarithmic returns on $S\&P500$ index. The results for D_2 are the same, but D_1 turned out to be very close to zero, which corresponds to the absence of the restoring force in terms of Langevin equation (i.e. no friction in the liquid). The last result is not confirmed in Cortines et al. (2007) on the logarithmic returns on the Brazilian index Ibovespa. Besides, the authors find significant linear trend in the equation for D_2 at horizons longer than one day. This is an important deviation from the classical multifractal model of turbulent cascade. Notice that the same deviation has been independently found on the empirical data for turbulence in liquids (Renner et al., 2001b).

In Buchbinder and Chistilin (2007) coefficients of Fokker-Plank equation are estimated for daily realized volatility of DJIA index, computed from 5-minutes returns. They find that the resulting estimates of Kramers-Moyal coefficients are well described by the equations:

$$\begin{aligned} D_1(\sigma,\tau) &= -\sigma(a_1 + a_2 \ln\sigma) \\ D_2(\sigma,\tau) &= b_1\sigma^2(\exp b_2\sigma) \end{aligned} \tag{3.59}$$

The first equation in (3.59) accounts for non-linearity in the coefficient of the restoring force at low volatility levels, the second models higher than quadratic speed of the increase in diffusion coefficient, observed at high volatility levels. With small σ replacing (3.59) in (3.55) results in the stochastic differential equation, corresponding to the exponential OU model of stochastic volatility. This model is also advocated in Masoliver and Perello (2006), based on entirely different considerations, related to the properties of ACF.

A huge number of methods and models that were proposed for describing volatility at multiple horizons evidences for rapid development of this research area in finance. It is too early to talk about a consistent theory, because for the moment there is no clear leadership among competing approaches. Besides, developing such a theory requires practical extensions, related to forecasting, optimal asset allocation and derivatives pricing. Some progress is made in each of these directions. Calvet and Fisher (2001) and Richrads (2004) propose methods of forecasting of multifractal time series. Some studies treat option prices under multi-horizon stochastic volatility, driven by a factor model (Fouque and Han, 2004; Fouque et al., 2003). Finally, solution of an asset allocation problem for the case when prices are driven by multifractal processes is given in Muzy et al. (2001).

Another direction of research, related to the multi-horizon models of volatility, deserves a special mention. Its aim is constructing indicators of volatility, that would represent the current state of the market, taking into account not only the magnitude of fluctuations, but also there frequency. As follows from the above theoretical arguments, considering volatility simultaneously at various horizons brings in important information, compared to measuring it at some particular horizon. This information can be used primarily for decision taking in dynamic portfolio management, based on volatility timing. Different multiple-horizon indicators, applicable to volatility measurement independently of the specification of the stochastic volatility process, were proposed in Maillet and Michel (2003); Maillet et al. (2007); Subbotin (2008); Zumbach et al. (2000). All of them are defined as probability transforms of volatility at different scales, based on an analogy with the Richter-Gutenberg scale in geophysics (Richter, 1958). Probability transform measures rareness of fluctuations of a given magnitude at the financial market. Thus by constructions they are universal in the sense that their values are comparable in time and over different assets. This is an important advantage from the practical point of view. The differences in indicators lie in how volatilities at multiple horizons are estimated, how the importance of each horizon is measured and how the results over different horizons are aggregated.

3.8 Conclusion

Modeling and measurement of stock price and exchange rate variability is one of the key elements of the theory and practice of investment portfolio management and other areas of finance. We discussed the notion of volatility and two approaches to its modeling in discrete and continuous time (conditional heteroscedasticity and stochastic volatility),

pointing to the differences in how they capture the changes in the parameters of conditional returns' distribution. Evolution of these models has always been directed to reproduce more exactly the empirical properties in time series of prices, such as long-range correlations in magnitudes of returns, their absence in returns themselves and fat tails in returns distributions at short horizons.

Among all models we draw special attention to those that represent volatility at multiple horizons, because they seem to be the most promising. Multi-horizon representation allows to take into account the properties of returns, that manifested themselves when the latter are aggregated in time. The challenge is to capture the evolution in the form of the probability distribution of returns, computed over time intervals of different length. We described several classes of multi-scale models, from heterogeneous ARCH to multiplicative cascades. An important role in multi-horizon analysis belongs to methods and techniques, borrowed from hydrodynamics and other areas of statistical physics. Such borrowing became possible thanks to the discovery of the analogy (though possible incomplete) between volatility and turbulence in liquids and gases.

The concept of volatility at multiple horizons suggests the development of methods of its measurement, that account not only for the magnitude of fluctuations, but also for their frequency. Information, obtained from measurement at different levels of time aggregation (i.e. at various horizons) can be used jointly. This can be helpful, in particular, in asset management applications and in forecasting. An interesting further development may include forecasting volatility at multiple horizons simultaneously. Another important issue is the study of derivatives hedging strategies with regards to the frequency of operators' interventions on the market.

Chapter 4

A Multi-Horizon Scale for Volatility

4.1 Introduction

Volatility of a stock market index is a natural candidate to characterize the state of the stock market and compare the impact of important events, such as financial crises. The need to compare events on the stock market arises in many applications: from structural analysis, studying the dynamics of the stock market in a general economic context, to asset allocation, for which measuring the severity and time bounds of extreme fluctuations is of interest. In principle, such analysis can be based on the values of some direct volatility estimates. But these estimates are hard to interpret and their magnitude varies for different assets. One way to overcome this inconvenience is to design a universal indicator, based on a probability transform of volatility measures.

Zumbach et al. (2000) proposed an intuitively appealing analogy with the Richter-Gutenberg scale in geophysics and constructed an indicator of market volatility, called a Scale of Market Shocks (SMS). The Richter-Gutenberg scale is a measure of the logarithm of the seismic wave amplitude, which is related to the total energy, liberated during an earthquake (Richter, 1958). The SMS maps the exchange rate volatility to a logarithmic scale, suggesting the analogy between volatility and mechanical work, i.e. the rate of change of energy in time. The indicator was originally applied to the currency market, but by construction it can be used for any traded asset, for which historical quotes are available.

The SMS accounts not only for the size but also for the scale of fluctuations. According to the heterogeneous market hypothesis, first proposed by Müller et al. (1997),

the price dynamics is driven by the actions of market participants, operating at different frequencies. To make this idea clear, compare an institutional investor who operates over medium and long term targets and a small private investor exploiting short term market moves. Hardly do they share the same opinion on what is trend and fluctuation on the market. However, sometimes they would agree to characterize the situation on the market as a crisis or a crash. The underlying intuition is to describe volatility in the time domain and in the spectral domain simultaneously, thus attributing fluctuations to particular ranges of frequencies (scales). Economically, these scales can be associated to the decision taking and portfolio readjustment horizons of investors. The SMS takes the form of an average realized volatility across different scales.

Maillet and Michel (2003, 2005) adapted the multiscale approach to the stock market. The new indicator, called Index of Market Shocks (IMS), is designed for the detection and comparison of severity of different crises. The authors change the way of computing and aggregating volatilities at multiple frequencies. Besides, instead of modeling their probability distributions, the principal components analysis is applied to estimate the hidden factors of volatility, affecting various scales. These factors are then subject to the probability transform.

Following the approach suggested in Zumbach et al. (2000) and Maillet and Michel (2003, 2005), we study the multiscale indicators of volatility and propose a new Market Volatility Scale (MVS). Though the general idea of such indicators is very appealing, the existing implementations suffer from several important drawbacks, which motivates our work. In our view, the main problem with SMS is the scheme of multi-resolution analysis which is based on sequential smoothing rather than on the scale-by-scale decomposition, which does not allow to determine the relative importance of different horizons. This problem subsists for the IMS, which in addition makes the computation of the indicator more complicated by introducing the latent factors of volatility.

We argue that modeling latent factors instead of volatilities themselves leads to the loss of simplicity and interpretation errors, and propose to work with the scale components of volatility directly. The MVS, which we introduce in this paper, uses the realized wavelet variance to represent volatilities at multiple horizons. The log-normality assumption, which is used for the probability transform in the IMS, is clearly too restrictive. This is why we apply the peak-over-threshold approach to fit the generalized Pareto density for the tails of the realized variances' distribution at each scale. This allows more accurate estimation of the probability of extreme events, such as stock market crashes.

We test our approach on two types of data: high-frequency (15-minute intervals) observations of the French CAC40 index from 1995 to 2006 and daily observations of the

Dow Jones Industrial Average (DJIA) index from 1896 to 2007. In these two examples the ranges of frequencies, to which we refer, are not the same, but the computation algorithm is quite similar. We detect periods of extreme volatility on the stock market, often provoked by financial crises, and study their structural characteristics using the scale-by-scale decomposition. This enables us to determine which component of volatility (short-term, mid-term or long-term) is more important in each detected period.

Potentially, the applications of multiscale decomposition of volatility can go far beyond event comparison and structural analysis. Most promising are volatility forecasting and asset allocation. These areas fall out of the scope of this paper, but the tools and ideas introduced here can be used in further studies.

The rest of the paper is organized as follows. In section 4.2 we briefly overview the heterogeneous market hypothesis, the existing multiscale indicators and discuss their main drawbacks. Section 4.3 introduces some wavelet formalism, used for the scale-by-scle decomposition of volatility. Section 4.4 defines the Market Volatility Scale and the associated computation algorithm. In section 4.4 we compute the MVS for the DJIA and CAC40 indexes and present the main results. Section 4.4 concludes.

4.2 Multiscale Indicators of Volatility

Multiscale indicators of volatility are based on the model of financial market with multiple horizons, known as the heterogeneous market hypothesis. First suggested in Müller et al. (1997), it attempts to explain well-known stylized facts about volatility of stock returns, such as long-range correlations in the time series of squared returns and absolute returns and volatility clustering (Andersen and Bollerslev, 1997; Ding et al., 1993; Lux, 1996; see Cont, 2001 for an overview on the subject). The central point is the scale-by-scale analysis of volatility which assumes that market data contains patterns specific to peculiar frequencies of observations and are thus of interest for different types of market agents. The latter may include intra-day speculators, daily traders, portfolio managers and institutional investors, each having their characteristic period of reaction to news and frequency of intervention in the market. A multiscale dynamic model for volatility was proposed in Müller et al. (1997) and Dacorogna et al. (1998), who represent current volatility as a function of squared returns over different horizons:

$$\sigma_t^2 = c_0 + \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{j} r_{t-i} \right)^2 \tag{4.1}$$

where $c_k \geq 0$ for all $k = 0, ..., n$ (for $k = 0$ and $k = n$ the inequality is strict), and r_t is logarithmic return. Thus the sum $\sum_{i=1}^{j} r_{t-i}$ represents the log-return for the period of length j. By construction, the resulting HARCH (Heterogeneous ARCH) model captures the hierarchical structure of volatility correlations. More recently Corsi (2004) proposed a simple additive cascade model, inspired by the HARCH, which uses realized volatilities over some representative scales (daily, weekly, monthly) to model current daily realized volatility.

From a more theoretical perspective Ghashghaie et al. (1996) and Breymann et al. (2000), inspired by the model of turbulence in hydrodynamics, designed a multiplicative cascade model which decomposes the stochastic volatility process into a set of scales, so that the volatility at each scale depends on the volatility at lower frequency and some stochastic renewal process. The probability of renewal is chosen so that the mean interval between two renewals is equal to the time length of the associated scale. Arneodo et al. (1998) shows that under the assumptions of the model the covariance between logarithms of absolute returns at each scale should decrease with time as logarithmic function, so that returns simulated from the model exhibit dynamic patterns fairly close to empirical evidence. For the moment this model has mainly theoretical interest because the methods of its identification on real data are not elaborated.

Zumbach et al. (2000) retain the multiscale concept, but their aim is to characterize the vector of volatilities across various scales rather than impose a dynamic model for predicting instantaneous volatility. The multiscale indicator, called the Scale of Market Shocks (SMS) is a probability transform of volatilities at different scales. It is based on an analogy with the Richter-Gutenberg scale in geophysics (Richter, 1958).

The Richter scale is a measure of the logarithm of the total energy E liberated during an earthquake, compared to a reference level E_0:

$$R \approx \ln\left(\frac{E_0}{E}\right) \tag{4.2}$$

According to recent evidence, the probability of occurrence of large earthquakes grouped within temporal clusters of high seismic activity obeys the inverse power law (see Christensen et al., 2002; Mega et al., 2003). The probability to observe an earthquake of energy E reads:

$$p(E) \approx \left(\frac{E_0}{E}\right)^{\kappa}, \tag{4.3}$$

with κ the scaling parameter. Using (4.3) the Richter scale can be rewritten:

$$R \approx \ln\left(\frac{E_0}{E}\right) = -\frac{1}{\kappa}\ln p(E) \tag{4.4}$$

.

By analogy with (4.4), Zumbach et al. (2000) map the exchange rate volatility to a logarithmic scale. Volatility stands as a counterpart of mechanical work, i.e. the rate of change of energy in time. The corresponding formula is:

$$SMS_t = -\alpha \ln P(\sigma_t) \tag{4.5}$$

with $P(\sigma_t)$ the distribution function associated with the volatility σ_t and α a scaling parameter. Note that the analogy is not direct because the probability distribution in (4.5) is not a pure power law and must be fitted.

The indicator is multiscale because the volatility used in (4.5) is computed separately for various "scales" of observations. An estimate of the realized price volatility for a given scale, used for that purpose, reads:

$$\widetilde{\sigma}_t^{(k)} = \left(\frac{\sum_{i=1}^{M_k - 1} r_{t,\delta}^2}{M_k - 1} \right)^{1/2} \tag{4.6}$$

with $\sigma_t^{(k)}$ the volatility of returns at scale k, $r_{t,\delta}$ the log-returns computed for time intervals δ of length $\tau_k (M_k - 1)^{-1}$, τ_k the time length of scale k and M_k the number of observations available for that period. As the indicator was constructed for inhomogeneous tick-by-tick data, (4.6) could not be applied directly and the authors used smoothed volatilities, computed over moving windows, of the form:

$$\widetilde{\sigma}_t^{(k)} = \int_{t-2\tau_k}^{t} K(\phi) \omega_\phi^{(k)} \sigma_\phi^{(k)} d\phi \tag{4.7}$$

with $K(\cdot)$ an appropriately chosen kernel function and $\omega_\phi^{(k)}$ a scale dependent weight. The final aggregated indicator is computed as the weighted average for all scales in the following way:

$$SMS_t = -\sum_{k=1}^{m} w_k \ln P\left(\sigma_t^{(k)}\right), \tag{4.8}$$

with w_k a convolution kernel, measuring the contribution of scale τ_k to the overall volatility, and $P(\cdot)$ is the complement of the cumulative density function.

A potential problem is that the scaling method suggested in Zumbach et al. (2000) gives no idea about the range of scales to be considered and their relative importance. Multi-resolution analysis consists in sequential smoothing rather than scale-by scale decomposition. The authors assert that the choice of the convolution kernel is not a crucial issue and it suffices to take a function which satisfies some regularity conditions and tends to zero at the edges of the range of frequencies, which is 15 minutes - 64 days

in their case. To our knowledge, this choice is made *a priori* and is not based on any economic notion or on statistical properties of the time series. The assumption that the mass point, or the most important frequency, is located in the center of the range is also doubtful.

Besides that, the SMS does not account for the possible interdependence between scales. The aggregated indicator is a function of transformed univariate cumulative density functions and not of a multivariate *cdf*.

Maillet and Michel (2003, 2005) account for the dependence between scales and propose a new indicator called the Index of Market Shocks (IMS). Instead of scaled volatilities themselves, principal components are used in its definition:

$$IMS_t = -\sum_{i=1}^{q} \omega_i \log_2 P(c_i) \tag{4.9}$$

with $c_i, \ldots, c_q$ normalized principal components (or factors underlying multiscale volatility), and ω_i the weights of each component determined as the portion of variance of the data explained by the corresponding component. The IMS also uses a different method of multi-resolution analysis. Instead of computing volatilities for different scales by successively moving a window of fixed length, the IMS uses different frequencies of sampling within each interval to obtain the scale components of volatility.

Several shortcomings of the IMS can be mentioned. First, the method of estimating volatilities at different scales by varying sampling frequency has one important drawback which becomes very important if applied to low-frequency data. If the length of computation window is fixed and volatilities are calculated from the samples with different number of points, the results of estimation will be different in terms of statistical error and become incomparable. For example, if the length of computation window is fixed at 60 business days, the daily estimate is an average of 59 points and the monthly volatility is an average of only 3 points. Estimated monthly volatilities become counter-intuitively less smooth than the times series of daily volatility because of statistical errors. On the contrary, the daily average computed over a 3-months interval smooths out all interesting dynamics, corresponding in fact to the daily frequencies. Hence such multi-resolution analysis becomes inadequate.

Second and more important, replacing volatilities by their linear combinations (factors) changes the meaning of the indicator, which is not explicited in Maillet and Michel (2003). The authors claim that the IMS has a clear economic sense: since a logarithm in base 2 is used in (4.9), a one point increase of the IMS corresponds to the volatilities vector twice as unlikely. This is obviously true for the baseline Richter transformation in the form analogous to (4.5), but not for (4.9), because the multivariate function of a

vector is not the same as the distribution function of a linear combination of its components. Hence a one point increase of the IMS corresponds to the factors vector half as likely. But the factors, obtained from the Principal Components Analysis (PCA) do not necessarily have positive or negative impact on the volatilities at all scales. They can have positive loadings in some scales and negative in others. Thus it is possible that an important change in some factor corresponds to a "redistribution" of volatility from one scale to another rather than increase in volatility. As the PCA is done on the variance-covariance matrix, the volatilities at different scales are treated as if they were all of equal importance to investors. All this makes the interpretation of the resulting indicator very problematic.

Modeling factors of volatility can be of interest if these factors are stable and meaningful. In our view, this is not the case. First, the definition of factors by the PCA is too restrictive: it relies on linear decorrelation and may not have the power to identify factors if they are not symmetrically distributed (Loretan, 1997) or if the underlying model is not a simultaneous linear mix. It is unrealistic to suppose that the underlying model for multiscale volatility is a simultaneous mix, because it contradicts to the very idea of the heterogeneous market.

Economic intuition suggests that investors acting on lower scales adjust their market positions according to the change in information since the previous observation. Investors acting at high frequencies react to information more often. But as the information set in the first case encompasses all the information that has already been used by the short-term investors (possibly accompanied by some long-term factors), it is natural to suppose that the long-scale volatility is correlated with the lagged short-scale volatilities. Yet the PCA allows for the instantaneous decorrelation only and the resulting factors can hardly correspond to the true hidden factors of volatilities. Studying the properties of the volatilities vector empirically, we find that the factor loadings in the PCA are unstable in time and principal components are dependent if lags are considered.[1] Unable to propose a simple and viable method of factor analysis, we prefer to model the volatility vector directly, but in a way different from that proposed in Zumbach et al. (2000).

Finally, the probability transform used to compute the indicator is an important issue. Maillet and Michel (2003) postulate the multi-normality for log-volatilities (and principal components) at all scales, though they report that the log-normality is violated for the high values of volatilities. This can affect the values of the IMS during extreme events. Meanwhile, these values are of great importance for analysis of financial crises.

[1]The discussion on factor analysis of multiscale volatility falls out of the scope of this paper, these preliminary results can be obtained from the author on request.

The SMS, in its turn, uses a more sophisticated modification of the log-normal *pdf*, used to account for the fat tails. In our view, this modification does not allow enough flexibility, because an *a priori* judgment about the scaling properties of financial time series has to be made. It can be accurate for some assets and not for others. We argue that to define a universal indicator we do not necessarily need an overall fit for volatility. Instead we focus on the extreme events only, using the peak-over-threshold approach to fit the generalized Pareto distribution for the tails (see section 4.4 for details).

As follows from the above discussion, the most important improvements to be made in multiscale indicators of volatility concern the multi-resolution analysis by itself and the fit of probability distribution for the volatilities vector. These improvements are discussed in the following section.

4.3 The Scale-by-Scale Decomposition of Volatility with Wavelets

In this section we propose a method of multiscale analysis of volatility, based on the wavelet filters. We introduce the realized variance of wavelet coefficients, which is used to define the Market Volatility Scale in the following section.

Wavelet analysis is widely used in signal processing in order to decompose a given time series called "signal" into a hierarchical set of approximations and details (multi-resolution analysis), and to decompose the energy of the signal on scale-by-scale basis. For a detailed description of the use of wavelets for time series analysis refer to Percival and Walden (2000) and Gençay et al. (2001a). Within the financial markets context, wavelets were used for multiscale analysis in several studies. An early application of wavelets can be found in Ramsey and Zhang (1995) who attempt to recover the dynamic structures of the foreign exchange rates. Capobianco (2004) identifies periodical patterns in intra-day stock returns. Fernandez and Lucey (2007) use wavelets to estimate value at risk for multiple investment horizons in a portfolio management perspective. Lee (2004) studies the effects of volatility spillovers from the US to the South Korean stock markets at multiple horizons.

The idea to use the wavelet decomposition for the analysis of volatility cascade can be found in Arneodo et al. (1998). Gençay et al. (2001a) used wavelets to investigate the scaling properties of foreign exchange volatility. They decompose the variance of a time series of absolute returns and the covariance between two time series on a scale by scale basis through the application of wavelet transforms and show that no unique scaling parameter can be adapted for these time series. In particular, there are two different

scaling parameters corresponding to intraday and longer scales. In a recent paper Fan et al. (2007) propose a wavelet-based method for removing the jump-component from high-frequency returns and estimating realized variance.

We use the Maximum Overlap Discrete Wavelet Transform (MODWT) to decompose variance both in time and scale. A brief overview of the wavelet formalism which we use can be found in the Appendix. Let $\widetilde{W}_{j,t}$ be the MODWT wavelet coefficients at level j and $\widetilde{V}_{J,t}$ the scaling coefficients corresponding to the maximum level of decomposition, obtained from filtering some process $x_t, t = 1, \ldots, T$. The scale-by-scale decomposition of variance then holds:

$$\|x\|^2 = \sum_{j=1}^{J} \|\widetilde{W}_j\|^2 + \|\widetilde{V}_J\|^2. \tag{4.10}$$

Now let $\overline{W}_{j,t}$ represent the random process resulting from filtering an infinite sequence x_t with the MODWT wavelet filter. The difference between $\overline{W}_{j,t}$ and $\widetilde{W}_{j,t}$ is that the latter is obtained by circularly filtering a finite sequence. The wavelet variance $\nu_{j,t}$ of the process x_t for the scale of length τ_j at time t is defined as the variance of $\overline{W}_{j,t}$:

$$\nu_{j,t} = \mathrm{Var}\{\overline{W}_{j,t}\} = \mathrm{E}\{\big(\overline{W}_{j,t} - \mathrm{E}\{\overline{W}_{j,t}\}\big)^2\} = \mathrm{E}\{\overline{W}^2_{j,t}\} \tag{4.11}$$

If wavelet variance is constant in time, its natural estimator is:

$$\widetilde{\nu}_j = \frac{1}{N} \sum_{t=0}^{N-1} \widetilde{W}^2_{j,t} \tag{4.12}$$

This estimator is biased because it includes wavelet coefficients depending on the boundary conditions (i.e. the circularity assumption, used to filter a finite-length time series). Its unbiased version can be computed by ignoring such coefficients:

$$\widehat{\nu}_j = \frac{1}{N - L_j + 1} \sum_{t=L_j-1}^{N-1} \widetilde{W}^2_{j,t} \tag{4.13}$$

with L_j the number of coefficients, affected by the circularity assumption, which clearly depends on the length of the filter. The main disadvantage of this unbiased estimator is that it does not allow the exact decomposition of the sample variance, given by (4.10), so we prefer to keep all the wavelet coefficients and use (4.12). The importance of each scale can be interpreted as the portion of total variance corresponding to each scale. It can be estimated from the decomposition of energy, defined in equation (4.10), by replacing wavelet variances by their estimates.

In our context we need to characterize volatility dynamically, so the estimate (4.12), which suggests that the wavelet variance is constant, is not quite appropriate. Dynamic

estimates of $\widetilde{\nu}_{j,t}$ can be obtained by applying (4.12) locally for some moving window in time. This is the approach we take. The length of the window, used to estimate the wavelet variances, is chosen depending on the frequency of available data. For the data sampled at 15-minute intervals we use daily windows (32 observations) without intersections, which allows computing one value per day (analogous to the way realized volatility is computed). We call this estimate "realized wavelet variance" by analogy to the usual realized variance estimator. For daily data, we use monthly moving windows and obtain monthly realized wavelet variance. The final formula reads:

$$\widetilde{\nu}_{j,t} = \frac{1}{K} \sum_{k=t-K+1}^{t} \widetilde{W}_{j,k}^2 \tag{4.14}$$

with K the computation window length.

This definition is close to the one proposed in Fan et al. (2007). One difference is that we do not remove jumps from the data. Such removal is desirable when the final goal is a non-parametric estimation of the continuous diffusion parameter and its forecasting. In the perspective of a multiscale structural analysis of volatility we prefer to include the variance related to jumps in its short-term component, because it may contain important information about extreme events on the market. Another difference with Fan et al. (2007) is that we consider wavelet variances at each scale separately instead of summing them over all horizons.

The practical implementation of the MODWT algorithm requires treating the following issues: (i) the choice of wavelet filter, (ii) handling boundary conditions, and finally (iii) the choice of the number of horizons, *i.e.* levels of decomposition (see Percival and Walden, 2000, pp.135 - 145).

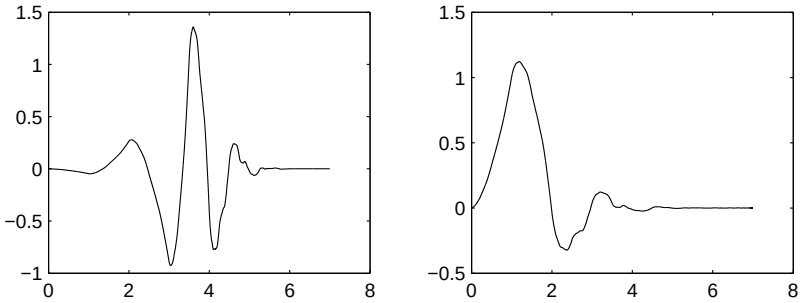
As for the first problem, the family of orthogonal and compactly supported wavelets is of particular interest. This family includes Daubechies extremal phase scaling filters (also called "daublet") and least asymmetric scaling filters (also called "symlets") that are used most often. For these wavelets, full reconstruction of the signal is possible and fast decomposition algorithms are available. The filters of each type can be of different width, usually taken from 2 to 20.

Though there are no universal criteria for the choice of the type of wavelets and their width, this choice must be dictated by the objective to balance two considerations. On the one hand, wavelet filters of too short width can introduce undesirable artifacts (unrealistic blocks) into the multi-resolution analysis. As the width of the wavelet filter increases, it can better match to the characteristics of the time series, but on the other hand, the influence of boundary conditions becomes more severe, the localization of

DWT coefficients decreases and the computation becomes more difficult. Percival and Walden (2000, p. 136) suggest that the best strategy is to choose the smallest width of wavelet that brings "reasonable" results. They also point out that this strategy often results in the choice of the least asymmetric filter with width equal to 8, denoted LA(8).

This wavelet was used in particular by Gençay et al. (2001b). They report however that the use of LA(4) did not significantly alter the results. In our case reducing the filter width had no significant influence on the results of the multiscale analysis. So we decided to apply the LA(4) wavelet for the multiscale decomposition and to use it as a matter of definition for the construction of the MVS. This choice was motivated by the parsimony arguments, but mostly by the need to minimize the impact of the boundary conditions (discussed below in details), so that the results for the latest observations of volatility could be adequate. The form of the wavelet and scaling functions for this wavelet is represented on Figure 4.1.

FIGURE 4.1: The form of the LA(4) Wavelet



The figure represents the form of the wavelet function ϕ (left) and scaling function ψ (right) for the wavelet filter LA(4).

The problem (ii) of handling boundary conditions arises because real time series are of finite length and we need to apply wavelet filters at their beginning and end. One possible solution is to use the so-called "circular filtering" operation, which treats the beginning and the end of the time series as if it were a portion of some periodic sequence. For financial time series the circularity assumption is hardly appropriate. Another solution which we retain in the absence of better alternatives is the "reflection" assumption: the latest observations are reflected around the boundary to obtain a time series of the necessary length. This assumption affects the results of scale by-scale decomposition for several latest observations of volatility. If a period of extremely high volatility is observed at the final date T and it started at some date $T-n$, we apply the $MODWT$ for the period $[T-n;T]$ supposing that this high volatility period will last till $T+n$ in future, thus possibly under- or overestimating the persistence of fluctuations and potentially attributing them to a wrong scale. Anyway, introducing a rule of this

kind is inevitable if we need to characterize currently observed fluctuations as short or long term, since we do not know *a priori* how long they will last.

The choice of the number of scales (iii) is a difficult issue. In some applications of multi-resolution analysis, it is dictated by some physical considerations of scale. For example, to decompose a cardiogram collected at a rate 180 samples per second, a natural level of decomposition can be 6, because is associated with the physical scale 0.36 seconds, which permits to isolate large scale fluctuations ("baseline drift") that are of little interest (Percival and Walden, 2000, p. 145). Unfortunately, no such easy rule can be established for financial time series. Apparently for the data sampled at 15-minute intervals (32 observations per day) the level of decomposition must be higher than 5 in order to isolate intraday fluctuations. Our choice of the number of horizons is a compromise between the high enough portion of energy, explained by the details, and the accuracy of the approximation, which declines as the number of scales increases for a given number of observations. Concrete explications on the choice of the maximum decomposition level for particular time series are presented in the following section.

4.4 A New Multi-Horizon Market Volatility Scale

The wavelet decomposition allows computing estimates of variance for a range of scales. They can be used to construct synthetic indicators by applying the transformation of the form given by equation (4.5). But unlike the SMS and the IMS, described in section 4.2, we first construct the indicator for particular scales rather than aggregate them directly.

To reduce the number of horizons we aggregate some dyadic scales in order to define three characteristic horizons, called Short (S), Medium (M) and Long (L) scales. For $s \in \{S, M, B\}$ the synthetic variance reads:

$$\upsilon_t^s = \sum_{j \in H_s} \nu_{j,t} \tag{4.15}$$

with $\nu(\tau_j))$ the realized wavelet variance for the scale of length τ_j, defined by equation (4.14) and H_s the set of horizons (scales) included in the range corresponding to s. The scale $s = L$ includes the highest level of decomposition and thus represents the realized variance of the scaling coefficients, corresponding to the approximation in terms of the MODWT. The scale-specific MVS is then defined by:

$$MVS_t^s = -\log_2 P(\upsilon_t^s) \tag{4.16}$$

with $P(v_t^s)$ the probability to observe the variance at the synthetic scale s above v_t^s.

We do not attempt to make an overall parameter fit for the volatilities' distribution (such as normal or log-normal). Instead we focus on the tails and propose to fit a Generalized Pareto distribution (GPD) for extremely high variances. For non-extreme variances we use a simple empirical probability transform. This makes the indicator more flexible and more accurate. The GPD density function reads:

$$f(x|\theta,\xi,\sigma) = \frac{1}{\sigma}\left(1+\xi\frac{x-\theta}{\sigma}\right)^{-1-\frac{1}{\xi}} \tag{4.17}$$

For $x > \theta$ with θ the threshold parameter, $\xi \neq 0$ the shape parameter and σ the scale parameter. For positive ξ, if σ and θ are equal, the GPD is equivalent to Pareto distribution. If $\xi < 0$ then x has zero probability to exceed $-\sigma/\xi$.

Subsequent realized variances are highly correlated in time, especially for the long scales. This leads to the clustering of extreme observations and the *iid* assumption, underlying the maximum likelihood estimation of parameters, is violated. A simple declustering procedure is envisaged (see Coles, 2001, Ch. 5). The details concerning the GPD fit can be found in the following section.

Once the MVS for the three scales has been computed, the rule for detecting periods of extreme volatility must be specified. Any definition of such periods is arbitrary, so we try to propose an intuitively plausible one. For that purpose we define two thresholds, noted Θ and Φ, which will be used to detect respectively the beginning and end of events. More precisely, the event starts when one of the three MVS components overcomes Θ and ends when all three components first fall below Φ. The thresholds can be chosen depending on the desired rareness of the events. Depending on applications, we use two values for Θ: $\log_2(100)$ and $\log_2(20)$, corresponding to the tail probabilities 0.01 and 0.05 respectively. For Φ we use the value $\log_2(10)$ for tail probability 0.1.

It is useful not only to characterize volatility at different scales, but to make an overall judgment about volatility on the market, which accounts for all scales. For that purpose we need to determine the importance of each scale. The additive decomposition of wavelet variance, described by (4.10), suggests that the magnitude of realized variances at different scales measures the impact of these scales onto the overall variance.

The problem is that a huge part of returns' variance can be attributed to the shortest scales, but the persistence of volatility at short scales is not necessarily the same as at long scales. The persistence, or the level of long-range dependence in some time series

x_t is characterized by the power-law divergence at the origin of its spectrum:

$$\Theta_x(u) \sim c_\Theta |u|^{-\alpha} \tag{4.18}$$

with α the scaling parameter and c_Θ a constant. The scaling parameter is related to the Hurst exponent, denoted H, by:

$$H = \frac{\alpha + 1}{2} \tag{4.19}$$

If the values of H for some stationary process lie in the range $(0.5, 1)$, the process is said to have long memory.

The Hurst exponent can be estimated using wavelet variances. The wavelet variance is closely related to the spectral density function of the time series. Namely, under some regularity conditions the wavelet variance for scale τ_j is approximately the mean value of the spectral density function within the interval of frequencies $[1/2^{j+1}, 1/2^j]$:

$$\nu_j \approx 2 \int_{1/2^{j+1}}^{1/2^j} \Theta_x(u) du \tag{4.20}$$

This gives us the possibility to study the scaling behavior separately for different ranges of frequencies. For a pure power law with spectral density $\Theta_X(u) \propto |u|^\alpha$, the wavelet variance is related to its scaling parameter α in the following way:

$$\nu_j \propto (\tau_j)^{-\alpha-1} \tag{4.21}$$

These properties underlie the estimation of the scaling parameter by regressing $\log \nu_j$ on $\log \tau_j$ (see Gençay et al., 2001a).

We use the so-called logscale diagram method, first proposed in Veitch and Abry (1999), to estimate the scaling parameter. The logscale diagram is essentially a log-log plot of variance estimates of the wavelet details at each scale, against scale, accompanied by the confidence intervals about these estimates at each scale. A weighted least estimate of the slope of the logscale diagram in some range of scales (alignment region) is transformed to the Hurst exponent using equation (4.19). The squared errors in the wavelet variance estimates are used as the weights. The estimator is semi-parametric, because prior to the estimation, an analysis phase is required to determine the lower and the upper cutoff scales, where linear alignment on the logscale diagram is adequate. It is unbiased even for samples of finite length and efficient. The method also provides an estimate of the constant c_Θ in equation (4.18). A detailed discussion of the statistical properties can be found in Veitch and Abry (1999).

Using a similar method, Gençay et al. (2001b) find two alignment regions on the logscale diagram for high frequency exchange rates volatilities. The scaling properties of the data change approximately at the daily horizon, so that intraday changes have significantly lower level of persistence. We reproduce their results with the 15-minute absolute returns on the CAC40 index and find a similar phenomenon in the daily returns on the DJIA index, analyzed in a very long sample. In the latter case the break is observed at approximately 3-weeks horizon (see the following section for details).

We propose to define the weights of scales in a way that makes them dependent on the portion of total variance, attributed to each scale, but also on the level of persistence associated with it. The latter can be measured by the Hurst exponent. Our definition is inspired by the practice of volatility "annualization", common in the financial industry: a volatility estimate, computed over a short period of time, is multiplied by the square root of the length of a reference period (usually one year). The square root corresponds to the Hurst exponent, equal to 0.5 for the Brownian motion process of log-returns. In our case the Hurst exponent is estimated for two ranges of scales. We use a 3-month reference scale instead of one year to reduce the impact of the errors in the Hurst exponent estimation on the weights of scales. The definition of the scale weights reads:

$$\begin{aligned}
\omega^S &= \frac{A_S R^{H_1}}{A_S R^{H_1} + A_M R^{H_2} + A_L R^{H_2}} \\
\omega^M &= \frac{A_M R^{H_2}}{A_S R^{H_1} + A_M R^{H_2} + A_L R^{H_2}} \\
\omega^L &= \frac{A_L R^{H_2}}{A_S R^{H_1} + A_M R^{H_2} + A_L R^{H_2}}
\end{aligned} \tag{4.22}$$

with $A_s, s \in \{S, M, L\}$ the mean variance for the corresponding scales, computed over a sufficiently long window (*i.e.* a direct measure of the role of the scales in the decomposition of variance), R - the time length of the reference scale (we use 3 months), and $H_{1,2}$ the scaling factors for the short and medium-long scales respectively.

The final indicator, which aggregates the three scales, is then computed in the following way:

$$MVS_t = \omega^S MVS_t^S + \omega^M MVS_t^M + \omega^L MVS_t^L = -\sum_{s \in S,M,L} \omega^s \log_2 P(v_t^s). \tag{4.23}$$

The resulting formula (4.23) resembles the definition of the IMS in equation (4.9), but this similarity is spurious: all terms in (4.23) are defined in a different way. The computation algorithm for the MVS, described above, can be summarized as follows:

- choose the number of decomposition levels and compute the wavelet and scaling coefficients from absolute logarithmic returns;

- compute realized variances of wavelet and scaling coefficients;
- estimate the Hurst exponent for different ranges of frequencies, determine if a break point is present (*i.e.* the limit between two alignment regions on the logscale diagram with different scaling parameters) and compute the weights of scales;
- aggregate the realized wavelet variances to three representative horizons (short, medium and long scale);
- fit the GPD probability distribution for the tails of realized wavelet variances at each of the three synthetic scales;
- compute the MVS for each of the three synthetic scales and the weighted MVS, representative of all scales.

This algorithm is applied in the following section to compute the MVS for the stock index data.

4.5 Application to the Stock Market Data

In this section we compute the MVS for real stock market data and use the indicator to detect and characterize periods of extreme volatility. Two different data sets are used: values of the French CAC40 Index from 20/03/1995 to 29/12/2006 sampled at 15-minute intervals (100, 881 observations) and daily values of the Dow Jones Industrial Average Index (DJIA) from 26/05/1896 to 10/10/2007 (28, 864 observations). The first sample is the same as in Maillet and Michel (2003, 2005), but extended till 2006. Absolute values of the logarithmic returns are used to measure volatilities.

Obviously, in these two applications different ranges of scales are considered. The choice of the number of decomposition levels is a compromise between the will to represent most of the overall variance of returns by the details (and thus to attribute this variance to specific ranges of frequency) and the constraints, imposed by the finite length of the sample. For the CAC40 data we chose the 9-level decomposition with the LA(4) wavelet, so that the approximation represents the scale 8 days and more (256 observations) and accounts for about 10% of the variance. For the DJIA sample we use an 8-level decomposition with the approximation for the scale 128 days and more which explains about 16% of the variance. The results of the variance decomposition by the MODWT according to equation (4.10) is shown in Tables 4.1 and 4.2. In both cases the biggest portion of variance is attributed to the shortest scale, and for longer scales it decreases rapidly.

TABLE 4.1: The MODWT variance decomposition for the CAC40 returns

Level	D1	D2	D3	D4	D5	D6	D7	D8	D9	A
τ_j	1	2	4	8	16	32	64	128	256	> 256
$\widetilde{\nu}(\tau_j) \times 10^5$	0.17	0.12	0.07	0.04	0.03	0.01	0.01	0.00	0.00	0.05
$\widetilde{\nu}(\tau_j / \sum_j \widetilde{\nu}(\tau_j)$	0.39	0.26	0.15	0.09	0.06	0.02	0.01	0.01	0.01	0.10

Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. The MODWT decomposition of variance with 9 levels is used (9 details and an approximation). τ_j is the time length of each scale (one unit of time corresponds to 15 minutes). The variance of the wavelet coefficients at the first level (detail D1) corresponds to the variance at physical scale 15 minutes. The variance at level $j > 1$ represents the range from $15 \times 2^{j-2}$ to $15 \times 2^{j-1}$ minutes. The variance of the scaling coefficients (A) at level 9 represents the range of scales superior to $15 * 2^8 = 3840$ minutes, which is roughly 8 trading days. $\widetilde{\nu}(\tau_j) \times 10^5$ is the biased estimate of wavelet variance according to formula (4.12). The last line of the table contains the portion of the total variance, attributed to each scale.

TABLE 4.2: The MODWT variance decomposition for the DJIA returns

Level	D1	D2	D3	D4	D5	D6	D7	D8	A
τ_j	1	2	4	8	16	32	64	128	> 128
$\widetilde{\nu}(\tau_j) \times 10^5$	2.52	1.27	0.66	0.44	0.32	0.24	0.21	0.14	1.12
$\widetilde{\nu}(\tau_j / \sum_j \widetilde{\nu}(\tau_j)$	0.36	0.18	0.10	0.06	0.05	0.03	0.03	0.02	0.16

Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007. The MODWT decomposition of variance with 8 levels is used (8 details and approximation). τ_j is the time length of each scale (one unit of time corresponds to one trading day). The variance of the wavelet coefficients at the first level (detail D1) corresponds to the variance at physical scale one day. The variance at level $j > 1$ represents the range from $\times 2^{j-2}$ to 2^{j-1} trading days. The variance of the scaling coefficients (A) at level 9 represents the range of scales superior to $2^7 = 128$ days. $\widetilde{\nu}(\tau_j) \times 10^5$ is the biased estimate of wavelet variance according to formula (4.12). The last line of the table contains the portion of the total variance, attributed to each scale.

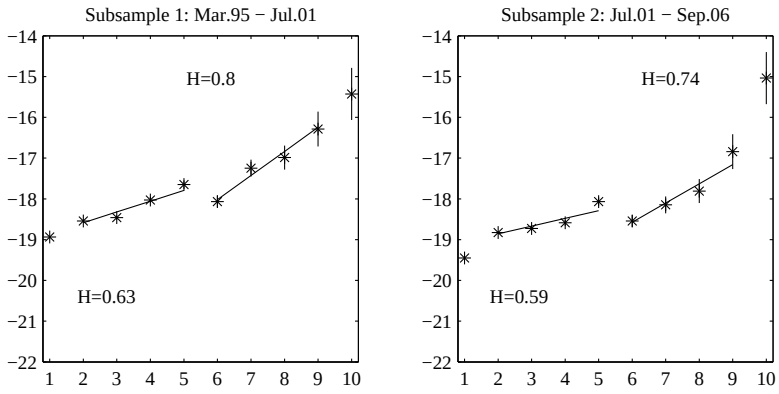
We then compute the realized variance of the wavelet and scaling coefficients at each scale. 32 observations are used to obtain daily realized variances of the CAC40 and 20 observations for the monthly realized variances of the DJIA[2]. For the CAC40 index a problem of treating overnight returns arises. We follow the procedure, suggested in Hansen (2005), to obtain an estimate of the realized variance of the whole day, which is a weighted average of the squared overnight return and the realized variance over the trading hours with the weights, minimizing the mean square error of the estimate. Optimal weights are determined for the conventional realized variance estimation, and then applied to the realized wavelet variances (i. e. the wavelet coefficient, corresponding to the overnight return, is taken with the same weight as would be attributed to the overnight return itself when computing the conventional realized variance). Overnight

[2]The exact number of observations may differ for weeks with less than 5 trading days.

returns are generally larger in absolute value than 15-minute returns, but this exceedance is filtered out by the first level wavelet detail, so the Hansen correction for the overnight returns is important only for the short-term component of volatility.

The next step in the algorithm is the study of the scaling properties of volatility over different ranges of frequencies. Figures 4.2 and 4.3 represent the logscale diagrams, used to estimate the Hurst exponent, as described in the previous section. Applying the logscale diagram method to different subsamples, we always find two alignment regions. For the CAC40 data: scales 2 to 5 and 6 to 9, and for the DJIA data: scales 2 to 4 and 5 to 8. For both datasets the scaling parameter is significantly higher for the longer scales, which means higher persistence. Similar results are reported by Gençay et al. (2001b) for the exchange rate volatility. On Figure 4.4 a logscale diagram for simulated log-normal returns is shown for comparison (the sample is of the same length as the actual DJIA data on Figure 4.3). Here the alignment is almost perfect for the scales 2-8 and no significant difference between the estimates of the scaling parameter for the short and long scale regions is present.

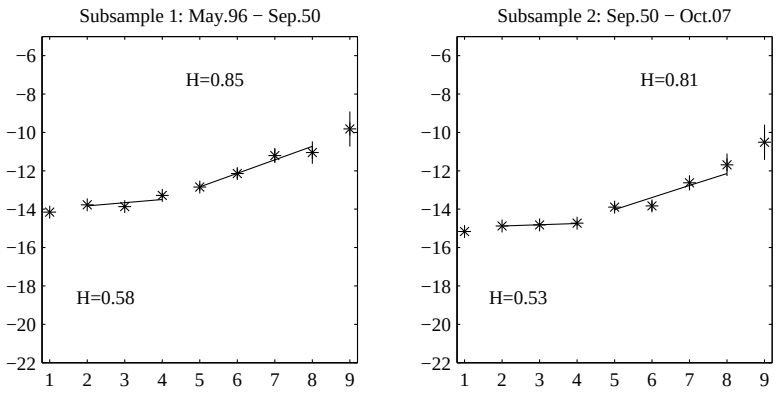
FIGURE 4.2: The Hurst exponent for the CAC40 index volatility for two subsamples



Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. The scaling factor is estimated by the Veitch-Abry logscale diagram method. The logscale diagram is generated for two subsamples of equal length (50,440 observations each) and examined to find a lower cutoff and upper cutoff scales, where alignment (a straight line) is observed. Two alignment regions are found: scales from 2 to 5 and scales from 6 to 9. The estimate of the scaling factor is obtained from the estimate of the slope of the of the WLS regression, as described in the text. Computation by the author.

In order to account for the possible changes in the scaling behavior of the time series we recompute the Hurst exponent on a rolling basis: for the CAC40 it is reestimated each 512 observations (approximately each 16 days) using a window of 32,512 observations (approximately 4 years), for the DJIA it is reestimated each 256 observations (approximately each year) on a window of 13,000 observations (approximately 50 years). The point estimates of the Hurst exponent and the associated confidence intervals are represented on Figures 4.5 and 4.6. For the CAC40 high frequency data, the value of the

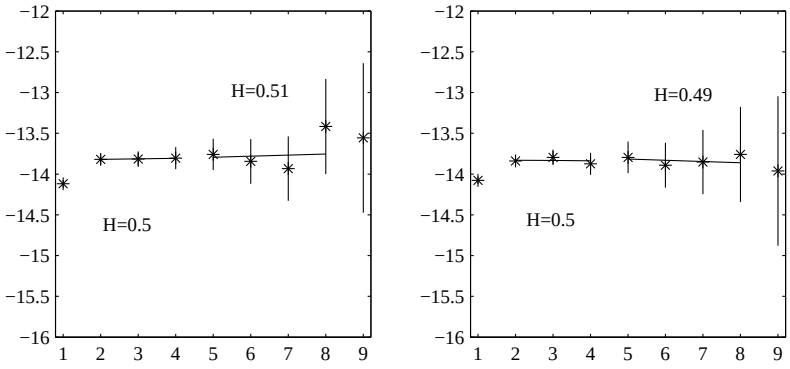
FIGURE 4.3: The Hurst exponent for the DJIA index volatility for two subsamples



Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007. The scaling factor is estimated by the Veitch-Abry logscale diagram method. The logscale diagram is generated for two subsamples of equal length (14,432 observations each) and examined to find a lower cutoff and upper cutoff scales, where alignment (a straight line) is observed. Two alignment regions are found: scales from 2 to 4 and scales from 5 to 8. The estimate of the scaling factor is obtained from the estimate of the slope of the of the WLS regression, as described in the text. Computation by the author.

parameter for the short scales is stable around 0.6, while the average for the long scales is about 0.8. The estimate of the Hurst exponent for the long scales falls after each period of extremely high volatility (2001, 2002-2003), and also during the recent period after the vanishing of the positive trend in returns, characteristic to years 2003-2005. For the construction of the aggregated MVS this implies that the relative importance of short scales increases for the periods of very high volatility and uncertainty.

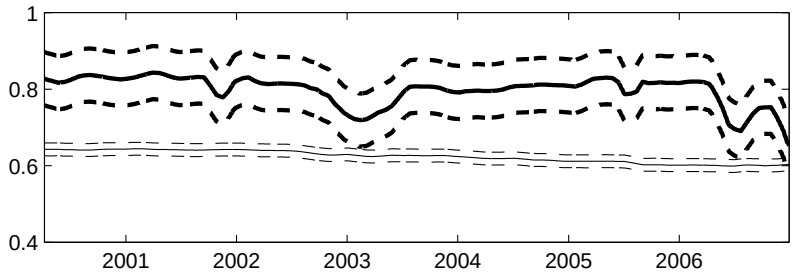
FIGURE 4.4: Estimation of the scaling parameter for the simulated Gaussian data



Logscale diagram for two simulated samples of Gaussian data of the sames size (14,432 observations) and of the same mean and variance as the data used for the DJIA index on figure (4.3). The scaling factor is then estimated by the Veitch-Abry logscale diagram method. The same two alignment regions as on figure (4.3) are explored. No significant change in slope of the WLS regression line is observed.

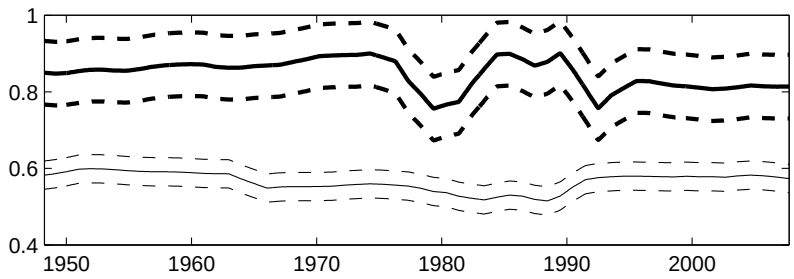
The realized wavelet variances are then aggregated. For the CAC40 data, the short-term component of volatility, which corresponds to the "intraday", is the sum of the

FIGURE 4.5: Dynamic estimation of the Hurst exponent for the CAC40 index volatility



Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. The scaling factor is reestimated each 512 observations (approximately each 16 days) on a window of 32,512 observations (approximately 4 years) length. The estimation is made by the Veitch-Abry logscale diagram method, described in the text. Computation by the author.

FIGURE 4.6: Dynamic estimation of the Hurst exponent for the DJIA index volatility



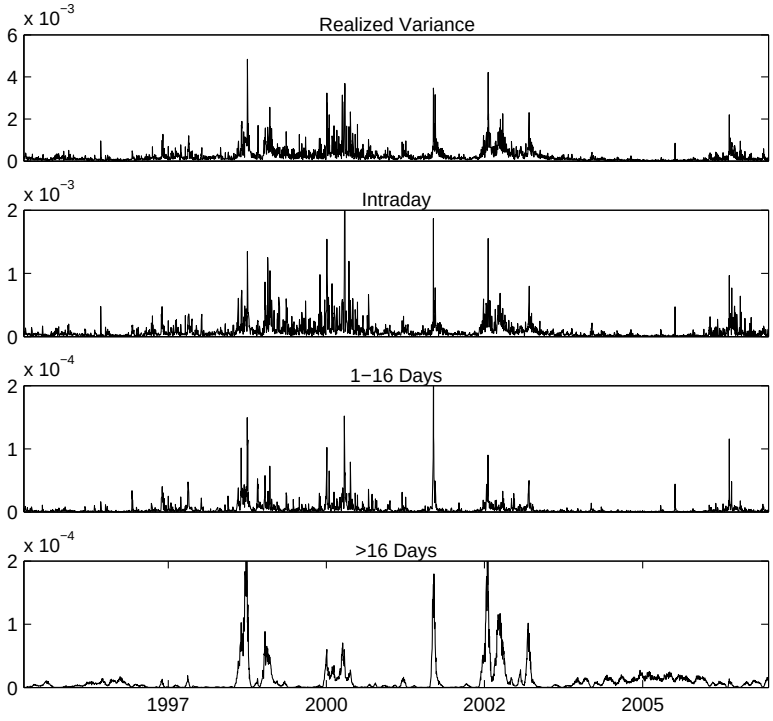
Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007. The scaling factor is reestimated each 256 observations (approximately each year) on a window of 13000 observations (approximately 50 years) length. The estimation is made by the Veitch-Abry logscale diagram method, described in the text. Computation by the author.

realized daily variances of wavelet coefficients at the scales of width inferior to 1 day (details 1-5), the medium term - at the scales of width between 1 and 16 days (details 6-9) and the long-term is the realized daily variance of the scaling coefficients (approximation at scale 9). For the DJIA data the short term aggregates the variances at the scales of less than 8 days width (details 1-4), the medium term is the range from 8 days to 128 days (details from 6 to 8) and the long term is the realized monthly variance of scaling coefficients (approximation at scale 8).

The aggregated realized variances for the three synthetic scales, defined above, are present on Figures 4.7 and 4.8 along with the traditional realized variance estimates. For both indexes, the profile of the short scale wavelet variance resembles almost perfectly the profile of the overall realized variance, though the magnitude of peaks is different. The peaks in the medium and long scale components generally coincide with the peaks

in the realized variance, but many short-term fluctuations are filtered out completely. For the CAC40 data we observe extreme long-term volatility in 1998, 2000, 2001 and 2002. Note that during the boom of 2003-2006, when the overall realized variance was relatively low, it was largely attributed to the long scale. The positive trend in returns translates into a gradual increase of the long-term variance, unlike the sharp and abrupt splashes of volatility observed during the stock market crises. For the DJIA data the most important period of extreme long-term variance is recorded during the Great Depression of 1929-1934. It largely surpasses in amplitude and duration the latest crashes in 1987 and 2001.

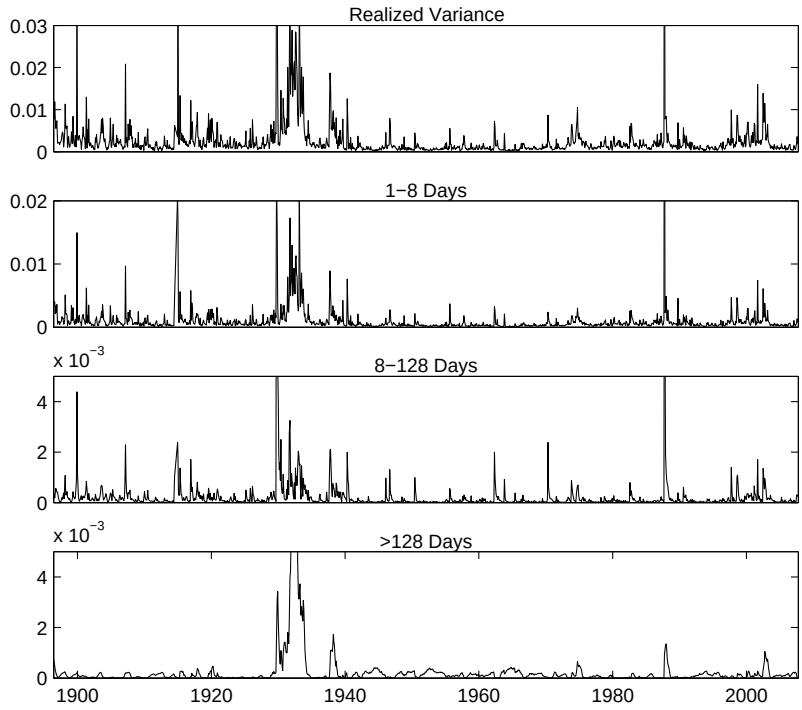
FIGURE 4.7: Realized wavelet variance of the CAC40 returns



Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. Realized daily variance includes overnight returns according to Hansen (2005) method. "Intraday" realized daily variance is the sum of the realized wavelet variances for the scales of width inferior to 1 day (details 1-5), "1-16 Days" realized daily variance is the sum of the realized wavelet variances for the scales of width between 1 and 16 days (details 6-9), "> 16 Days" is the realized daily variance of the scaling coefficients (approximation at scale 9). The LA(4) filter is used for the wavelet decomposition. Computation by the author.

We report the results of the GPD fit for the CAC40 index in Figure 4.9 and Table 4.3 and for the DJIA index in Figure 4.10 and Table 4.4. The parameters are estimated for the overall realized variance and for the three scales, based on the whole sample and on two subsamples of equal length. The threshold parameter θ in the probability

FIGURE 4.8: Realized wavelet variance of the DJIA returns

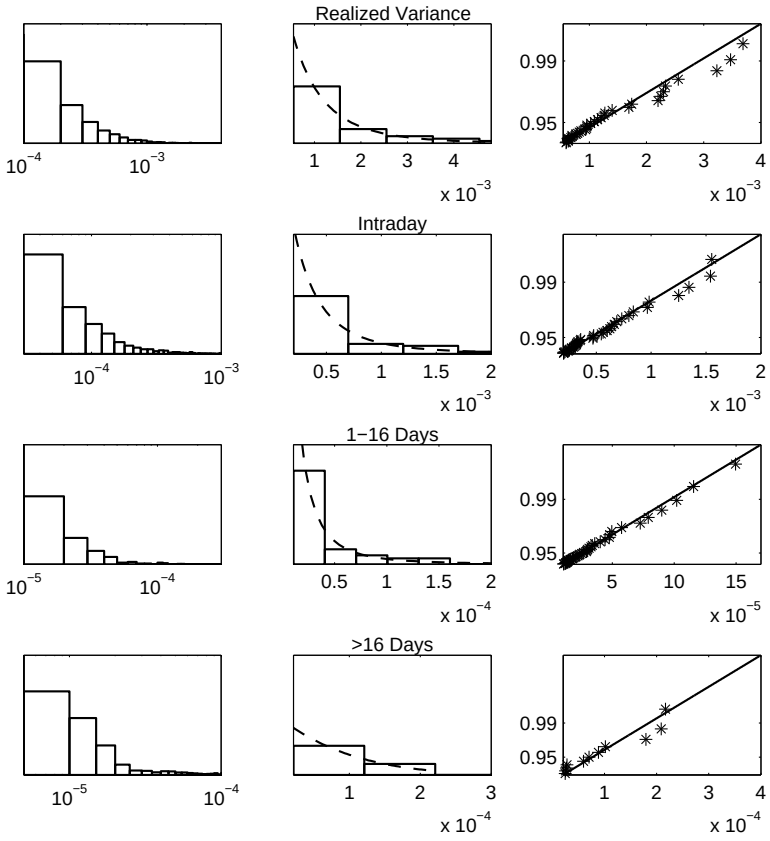


Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007. "1-8 Days" realized monthly variance is the sum of thye realized wavelet variances for scales of width inferior to 8 (details 1-4), "8 Days - 128 days" realized monthly variance is the sum of realized wavelet variances for details from 5 to 8, "> 128 Days" is the realized monthly variance of scaling coefficients (approximation at scale 8). The LA(4) filter is used for the wavelet decomposition. Computation by the author.

density function (4.17) equals to the 0.9 empirical quantile for the CAC40 and 0.8 empirical quantile for the DJIA. This choice allows both good quality of alignment and sufficient number of observations (after declustering procedure, described below) for the estimation. The study of the mean excess plot, commonly used to determine the threshold, results in roughly the same values for the threshold. The estimates of parameters σ and ξ are obtained by the maximum likelihood method. The associated confidence intervals are computed using the profile log-likelihood functions, as suggested in Gilli and E.Këllezi (2006). This method, based on reparameterizing the GPD distribution as a function of the unknown quantile, allows easy interval estimation of the quantile as well as of the parameters. Conventional bootstrap interval estimates for σ and ξ, not reported here, yield similar results.

Maximum likelihood estimation is based on the *iid* assumption for the data, which is obviously violated for the realized wavelet variances, especially at the medium and long scales. To overcome this problem we use a declustering procedure (see Coles, 2001). The cluster starts when the variance exceeds the threshold θ and ends when r consecutive

FIGURE 4.9: GPD fit for the realized wavelet variance of the CAC40 returns



Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. The figure represents the results of the GPD fit for the realized daily variance of the index and realized daily variances for different scale ranges, obtained by aggregation of realized wavelet variances as described in the text. Each line of figures shows (from left to right): the histogram of all data in logarithmic scale (2,954 observations), the fit of the GPD density for the upper tail of the distribution and the probability QQ-plot. Maximum likelihood estimation. The threshold parameter is chosen as the 0.9 empirical quantile of each distribution. Computation by the author.

observations fall below this threshold. We fix the value of r to be 20 daily observations of the realized wavelet variance for the high frequency CAC40 data and one monthly observation for the low frequency DJIA data. We then pick the maxima within each cluster and use them instead of raw data for the parameter estimation.

Note that the quality of the GPD fit is generally better for the variances at different scales than for the overall realized variances, especially for the quantiles above 0.97 (see the Q-Q probability plots on Figures 4.9 and 4.10). The probability law is relatively close to Pareto distribution for all scales and both subsamples (for the Pareto law we must have $\xi > 0$ and $\theta = \sigma$). For the CAC40 data, the Pareto conditions on parameters are accurately verified for the second subsample, while for the first subsample the point estimate of the scale parameter σ is considerably above the value of θ, though its 95% confidence interval estimate still contains θ. The same conslucions are globally true

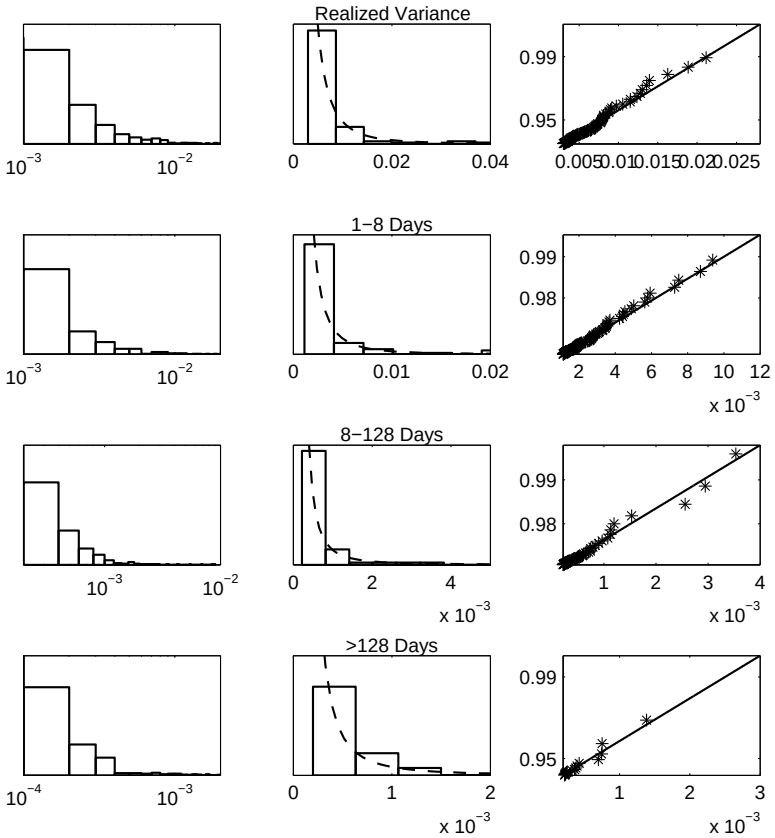
TABLE 4.3: Parameters of the GPD fit for the realized wavelet variance of the CAC40 returns

	$\Theta \times 10^5$	ξ	$\sigma \times 10^5$	$Q_{99} \times 10^5$
		All Sample: 03/1995 - 12/2006		
IV	33.8372	0.5236	47.7984	380.6894
		[0.2292; 1.0365]	[17.6859; 81.2570]	[380.6890; 381.0697]
Intraday	11.7858	0.501	20.8789	157.0575
		[0.2371; 0.9427]	[7.7256; 35.4937]	[157.0572; 157.2143]
$1-32$ days	0.5389	0.6782	1.2298	12.5549
		[0.3862; 1.1530]	[0.4554; 2.0903]	[12.5545; 12.5671]
>32 days	1.3778	1.2792	0.8117	30.0373
		[0.6486; 2.5726]	[0.3008; 2.0290]	[30.0368; 30.0669]
		Subample: 03/1995 - 06/2001		
IV	37.7046	0.2088	76.643	356.7373
		[-0.0447; 0.7812]	[28.3583; 130.2927]	[356.7370; 357.0937]
Intraday	13.4835	0.4493	24.2599	166.9273
		[0.1445; 1.0630]	[8.9766; 41.2415]	[166.9270; 167.0939]
$1-32$ days	0.7395	0.5163	1.4482	11.1065
		[0.1929; 1.1439]	[0.5363; 2.4614]	[11.1061; 11.1173]
>32 days	0.9093	1.5068	0.4341	26.9212
		[0.6417; 3.7712]	[0.1169; 1.7797]	[26.9207; 26.9477]
		Subample: 06/2001 - 12/2006		
IV	30.3373	0.9677	19.7181	379.8839
		[0.4517; 2.0316]	[7.2961; 49.2947]	[379.8835; 380.2634]
Intraday	10.1567	0.7716	11.9121	150.4925
		[0.3270; 1.6935]	[4.4079; 29.7800]	[150.4922; 150.6427]
$1-32$ days	0.4029	0.9508	0.8557	15.0357
		[0.4785; 1.8779]	[0.3169; 2.1388]	[15.0353; 15.0503]
>32 days	1.5412	1.3049	0.492	19.9626
		[0.6061; 2.9390]	[0.1823; 2.0168]	[19.9623; 19.9823]

Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. The fit is obtained by the maximum likelihood method. The threshold parameter is chosen as the 0.9 empirical quantile of each distribution. Computation by the author.

for the DJIA, except for the long term component, for which the estimate of σ is too low. But in the latter case the heterogeneity between two samples (prior and after year 1952) is much more important, which is obviously due to the evolution of the size and liquidity of the market since the beginning of the XXth century. Notably, the 0.8 empirical quantile for the realized monthly variance fell almost twice from 4.2×10^{-7} to 2.2×10^{-8}, which is mainly attributed to short and medium term. To account for evolution, we further estimate the parameters of the GPD distribution for the DJIA data using 50-year rolling windows, thus the recent values of the MVS are unaffected by the events before the mid 1950s.

FIGURE 4.10: GPD fit for the realized wavelet variance of the DJIA returns



Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007. The figure represents the results of the GPD fit for the realized monthly variance of the index and realized monthly variances for different scale ranges, obtained by aggregation of the realized wavelet variances, as described in the text. Each line of figures shows (from left to right): the histogram of all data in logarithmic scale (2,954 observations), the fit of the GPD density for the upper tail of the distribution and the probability QQ-plot. Maximum likelihood estimation. The threshold parameter is chosen as the 0.8 empirical quantile of each distribution. Computation by the author.

Our study of the probability distribution evidences in favor of the analogy between volatility and earthquakes: probability of extreme volatility clusters and high seismic activity clusters can both be described by the Pareto law. Thus a tool similar to the Richter-Gutenberg scale in geophysics, can be adequate to measure volatility in finance.

Finally we proceed with the computation of the MVS for each scale and its aggregation. Figure 4.11 shows the values of the CAC40 index, the realized variance and the aggregated MVS, computed as a weighted average over the three scale components, described above. The areas colored in gray correspond to the periods of extremely high volatility (probability less than 0.01), as detected by the MVS. The quantitative information about these periods is reported in Table 4.5. The beginning of each period is the date when the MVS for one of the scales overcomes the threshold $\log_2(100) = 6.6439$, which corresponds to 0.01 tail probability. The end of such period is the date when the

TABLE 4.4: Parameters of the GPD fit for the realized wavelet variance of the DJIA returns

	$\Theta \times 10^5$	ξ	$\sigma \times 10^5$	$Q_{99} \times 10^5$
		All Sample: 06/1896 - 10/2007		
IV	298.5244	0.5681	232.7840	789.4352
		[0.3649; 0.8599]	[160.6212; 395.7325]	[789.4348; 790.2242]
1-8 D	113.7658	0.6246	102.9759	340.7957
		[0.4107; 0.9305]	[71.0539; 175.0588]	[340.7954; 341.1362]
8-128 D	23.9888	0.8194	25.2972	89.2564
		[0.5448; 1.2205]	[17.4555; 43.0049]	[89.2560; 89.3453]
>128 D	21.3159	1.0740	5.9134	40.2136
		[0.6234; 1.8614]	[2.1884; 14.7833]	[40.2132; 40.2535]
		Subample: 06/1896 - 04/1952		
IV	410.5140	0.4377	375.9861	1127.3479
		[0.1977; 0.8500]	[259.4308; 639.1760]	[1127.3476; 1128.4750]
1-8 D	155.3517	0.6163	111.7318	400.0798
		[0.3554; 1.0292]	[77.0952; 189.9437]	[400.0794; 400.4795]
8-128 D	34.7769	0.8072	27.1548	104.1403
		[0.4696; 1.3586]	[10.0475; 46.1629]	[104.1399; 104.2440]
>128 D	22.9118	1.3701	3.9641	39.3502
		[0.7086; 2.7493]	[1.4670; 9.9099]	[39.3499; 39.3892]
		Subample: 05/1952 - 10/2007		
IV	216.6324	0.7106	121.6385	503.8929
		[0.4073; 1.2040]	[45.0065; 206.7851]	[503.8924; 504.3963]
1-8 D	81.5683	0.6619	60.7362	219.5061
		[0.3779; 1.1242]	[22.4727; 103.2512]	[219.5056; 219.7251]
8-128 D	14.0994	1.0602	10.3458	46.7705
		[0.6698; 1.6811]	[3.8284; 17.5875]	[46.7702; 46.8169]
>128 D	19.5670	0.9022	5.0805	33.6049
		[0.4055; 1.9608]	[1.8802; 12.7008]	[33.6046; 33.6382]

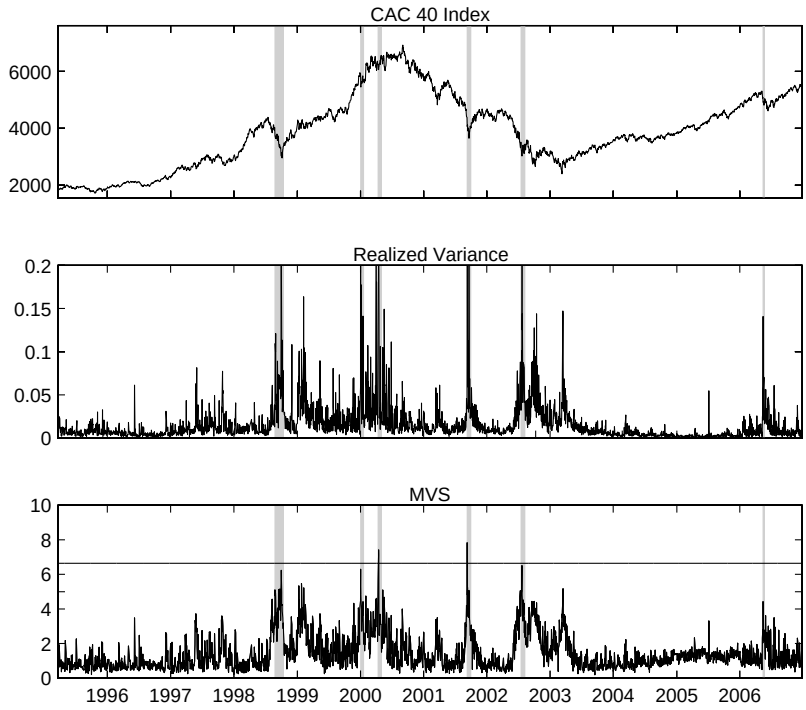
Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007. Maximum likelihood estimation. The threshold parameter is chosen as the 0.8 empirical quantile of each distribution. Computation by the author.

MVS for all scales falls below the threshold $\log_2(10) = 3.3219$, which corresponds to 0.1 tail probability.

We report the duration of the periods, determined as the number of trading days between the first and the last date, the maximum value of the MVS over the period (for the three scale and for the weighted average[3]) and the sum of the MVS values for the days, when one of the components was above the critical threshold of $\log_2(100)$, taken over the period of extreme volatility (for the three scales and for the weighted average). Each of the measures can be chosen to asses the gravity of the situation on the stock

[3]Note that the column reporting the maximum of the weighted MVS is not the mean of maxima at each scale, because the latter occur at different times.

FIGURE 4.11: Market Volatility Scale for the CAC40 index



Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. The figure shows (from top to bottom) the values of the index, the realized variance and the aggregated Market Volatility Scale (MVS), which is a weighted average of the three scale components, as described in the text. Regions colored in gray correspond to the periods of extremely high volatility (tail probability less than 0.01) as detected by the MVS (see Table 4.5). Computation by the author.

market. Another important characteristics is the scale of the MVS component, which first broke the critical threshold.

We detect six periods, corresponding to our criteria of extreme volatility. The period associated to the Asian crisis (from 27/8/98 to 14/10/98) was the longest in duration (34 open days) and the most severe in terms of the sum of the weighted MVS values above the threshold. The crisis is found to be the most "long-term" one, as evidenced by the values of the third component of the MVS. The beginning of the crisis was signaled by the medium-term component of the indicator. The period 19/7/02-9/8/02 is also characterized by persistently high values of the long-term volatility component, which was the first to break the critical threshold. The weighted average MVS is close to, but never breaks the threshold during these crises, which means that the peaks of volatility at different scales did not match in time.

The crisis following the September 11, 2001 events (11/9/01 - 1/10/01) ranks third in terms of the sum of the weighted MVS values, but it is structurally different from the two previously mentioned crises. The volatility is higher at the short and medium

TABLE 4.5: Periods of extreme volatility, detected by the MVS for the CAC40 index

Start	End	Length	Scale	Maximum MVS				Aggregated MVS			
				W	S	M	L	W	S	M	L
27/8/98	14/10/98	34	M	6.2	6.8	7.5	7.5	143.4	4.8	6.7	125.9
4/1/00	21/1/00	13	SM	6.3	7.2	6.7	4.0	53.6	7.2	6.7	55.3
14/4/00	3/5/00	10	M	7.4	9.5	7.6	4.0	42.2	5.9	7.6	44.4
11/9/01	1/10/01	14	SM	7.8	7.8	10.5	6.5	58.1	7.8	10.5	50.7
19/7/02	9/8/02	15	L	6.5	7.2	6.5	7.2	63.4	4.0	3.8	64.3
17/5/06	24/5/06	5	M	4.4	5.9	7.0	2.3	15.6	5.9	7.0	17.4

Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. The beginning of the period of extreme volatility is the date when the MVS for one of the scales overcomes the threshold $\log_2(100) = 6.6439$, which corresponds to 0.01 probability. The end of such period is the date when the MVS for all scales falls below the threshold $\log_2(10) = 3.3219$, which corresponds to 0.1 tail probability. Duration of the period is the number of trading days between the first and the last date. The scale refers to the component of the MVS, which first broke the critical threshold: "S" stands for the short scale (intraday), "M" stands for the medium scale ($1 - 16$ days) and "L" stands for the long scale (> 16 days). The following four columns report the maximum value of the MVS during the period of extreme volatility: "W" stands for the weighted average (aggregated) MVS over three scales, "S", "M" and "L" for the short, medium and long scales respectively. The next four columns contain the sum of the MVS values for the dates, when one of the scale components was above $\log_2(100)$, over the period of extreme volatility. The same notations for the scales are used. Computation by the author.

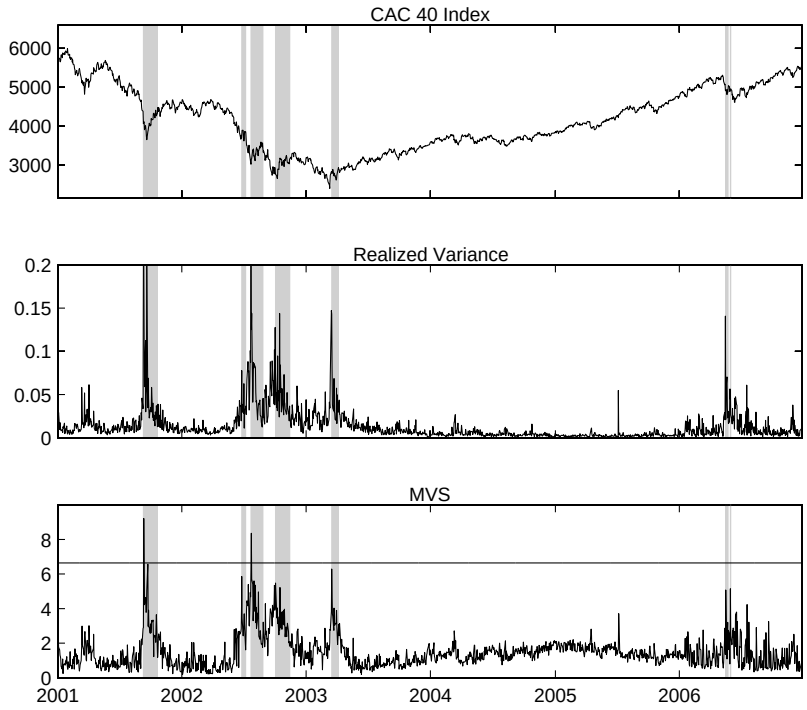
scales, clearly because the beginning of the crisis was sudden and driven by the factors, exogenous to the financial market. The shock was strong enough to affect the long-term component, though it does not overcome the critical value. The crisis of 2001 has the highest value of the weighted MVS over the whole sample (7.8, or tail probability about 0.005). Two detected periods in year 2000 are characterized by high volatility on the short and medium scale. The maximum aggregated MVS for the period 14/4/00 - 3/5/00, which marks the beginning of a long downturn in the market lasting till 2003 and is associated to the burst of the so-called "new economy" bubble, is 7.4 and ranks second after the 2001 crisis. Finally, fluctuations in May 2006 are labeled as extreme at the medium scale only. The weighted MVS stays below the critical threshold.

The MVS presented so far is suitable for the *a posteriori* structural analysis of market events, but it cannot be used for the real time applications, because the MVS values, as defined above, are not adapted. At each date we use all available data to perform the scale-by-scale wavelet transform and estimate the probabilities. Alternatively, we construct an adapted indicator for the CAC40 index. The part of the sample before 2001 (6 years) is used for initial estimates, which are updated dynamically as the new data arrives. So only the information prior to each computation date is used. At each iteration the last computed value of the MVS is retained. As we noted before, this value depends on the assumptions used for the boundary conditions. Here we use the reflection assumption, which implies that if at the date of analysis we observe extreme

volatility over some period of time in the past, we expect that it will last exactly the same period of time in future. This assumption affects the scale-by-scale decomposition and is the cost to pay for the non-anticipating property of the MVS.

The results, obtained for the CAC40 index with the adapted MVS, are reported in Figure 4.12 and Table 4.6. Clearly, using a smaller sample to estimate probabilities produces more periods of extreme volatility. We detect three periods in 2002 instead of one and a new 3-week period in 2003. But structurally the results are the same. The 2001 crash is still the biggest in terms of the average MVS and concentrates mainly on the medium and short scales (but it also surpasses the 2002 crisis by the sum of the MVS values). As before, the crisis in 2002 is characterized by the high long-term volatility: in July and October 2002 the long-term components of the MVS mark the beginning of the extreme volatility periods. Generally, the conclusions made with the adapted indicator are consistent with those presented before.

FIGURE 4.12: Adapted version of the MVS for the CAC40 index



Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15 minutes intervals. The same as in Figure 4.11, but the wavelet transform and the GPD fit for the MVS are recomputed dynamically at each date starting from 02/01/2001 till 29/12/2006 and using only previous observations. Only the last value of the MVS is kept each time. Computation by the author.

The results of similar analysis for the DJIA index are reported in Figure 4.13 and Table 4.7. In this case the definitions of the short, medium and long term are not the same - all the range of scales is shifted toward lower frequency and the realized variance is

TABLE 4.6: Periods of extreme volatility, detected by the adapted version of the MVS for the CAC40 index

Start	End	Length	Scale	Maximum MVS				Aggregated MVS			
				W	S	M	L	W	S	M	L
11/9/01	22/10/01	29	M	9.2	9.3	12.4	7.0	104.6	1.8	7.9	88.6
27/6/02	8/7/02	7	S	5.9	6.8	3.8	4.1	24.3	6.8	3.8	20.9
24/7/02	28/8/02	25	ML	8.4	9.0	7.8	8.0	96.2	5.5	7.8	83.8
4/10/02	15/11/02	30	L	5.5	5.7	4.6	7.6	95.6	5.7	2.4	85.9
18/3/03	7/4/03	14	SM	6.3	7.1	6.7	4.9	49.0	7.1	6.7	41.1
18/5/06	25/5/06	5	SM	5.1	7.7	8.9	0.5	15.8	7.7	8.9	16.0
1/6/06	2/6/06	1	S	5.2	7.8	3.6	1.0	5.2	7.8	3.6	1.0

Source: Euronext, values of the CAC40 index from 20/03/1995 to 29/12/2006 at 15-minute intervals. The same as in Table 4.5, but the adapted version of the MVS is used.

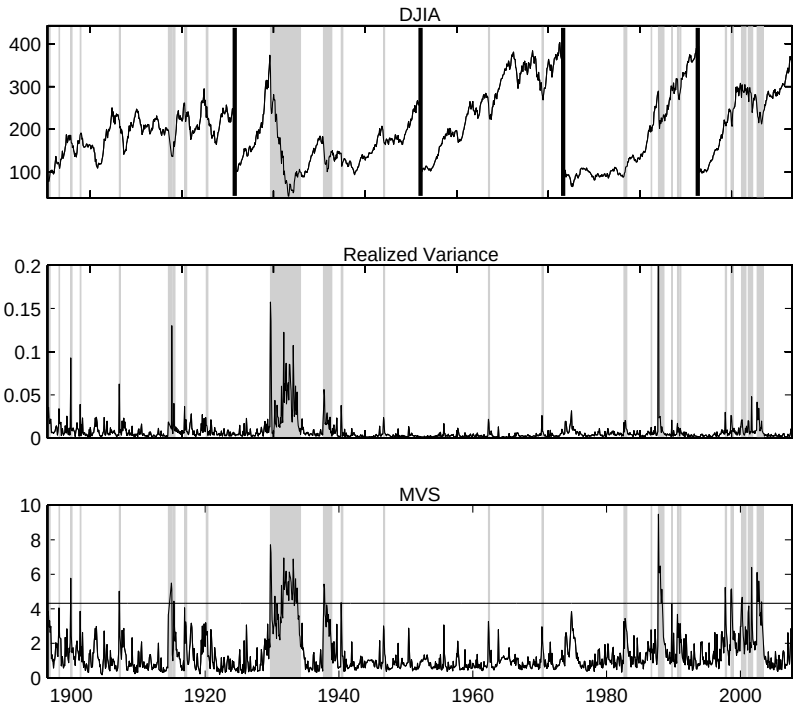
computed over monthly intervals. It is less variable than the daily realized variance, and critical levels of the MVS are chosen lower than for the high frequency data: $\log_2(20)$ for the beginning and $\log_2(10)$ for the end of the extreme volatility periods, which corresponds to 0.05 and 0.10 tail probabilities respectively.

As explained above, we estimate the distribution of variances over 50-year rolling windows (so the estimates till year 1944 are not adapted and further estimates are adapted). This convention allows to account for the changes in the structure of the financial market (and consequently, in the distribution of volatilities), but as a side-effect it produces an artificial shift in the 1990s - the period when the Great Depression falls out of the sample. Thus, the market volatility in late 1973 - early 1974, related to the oil crisis, Vietnam war and the Watergate scandal in the USA, followed by the resignation of president Nixon, is clearly higher than normal, as can be seen on Figure 4.13. But it does not reach the critical threshold, defined relatively to the extreme events of the 1930-s. Compared to the more recent events (say, from 1950 to 2000), this period would probably be detected as a period of extreme volatility.

This example demonstrates that the MVS is a relative indicator, since the underlying probability distribution is estimated from the historical data. Thus particular attention should be paid to the choice of the sample length and threshold values, especially when the sample is heterogeneous as in the DJIA case. However, this does not undermine the value of the indicator: it allows for the scale-by-scale structural analysis and comparison of events with the past, even though its absolute values and thresholds depend on arbitrary choices.

The results in Table 4.7 confirm that the Great Depression had an unprecedentedly strong affect on the stock market. The period of extreme volatility lasts for 54 months

FIGURE 4.13: Market Volatility Scale for the DJIA index



Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007. The figure shows (from top to bottom) the values of the index (bold vertical lines correspond to the dates when the index is rebased to 100 for the visualisation purpose), the realized variance and the aggregated MVS, which is a weighted average of the three scale components, as described in the text. Regions colored in gray correspond to the periods of extremely high volatility (tail probability less than 0.05) as detected by the MVS (see table 4.7). Computation by the author.

from the crash in 1929 till early 1934, when the recovery starts. The volatility shock associated to the Black Thursday (October 24, 1929) was so strong that all the three components of the MVS simultaneously break the critical threshold. The market downturn during the quick recovery (1937-1938), following the Great Depression, was first detected in the medium-term component of the MVS. The subsequent events, including those related to the beginning of the World War II, had a smaller impact on the volatility of the American stock market. Surprisingly, the panic provoked by the bankruptcy of the Penn Central Road in May 1970 had a stronger impact on the mid-term component of the MVS, than the following recession of 1973-1974, though, as it can be seen on Figure 4.13, the latter period is characterized by a persistently higher-than-average level of volatility, measured by the aggregated MVS.

The end of year 1982 was marked by the change in the trend of the index returns (from the bear market in 1967 - 1982 to the bull market in 1982-2000), accompanied by a period of extremely high short-term volatility. The following sharp rise in stock prices is interrupted by the crash of October 1987 (known by the Black Monday of October 19,

1987, though the first significant fall of 3.81% occured 5 days before). The weighted MVS in October 1987 has even higher value than the maximum recorded during the Great Depression, though the overall impact during the subsequent yearly period, measured by the sum of MVS extremes, is considerably lower.

The beginning of the Asian crisis in 1997 and its bottom in 1998, coupled with the Russian crisis, did not drive the long-term MVS component above the critical threshold. The medium and short-term volatility, however, was very high, resulting in extreme values of the weighted MVS. A rather long period from 2000 to 2003 is characterized by the burst of the "new economy" bubble. The exogenous shock of September 11, 2001 aggravates the situation, already featured by very high volatility. At the bottom of this crisis (late 2002 - early 2003) we detect the second period of extreme volatility since 1987, that affects the long-term component of the MVS in a significant way.

TABLE 4.7: Periods of extreme volatility, detected by the MVS for the DJIA index

Start	End	Length	Scale	Maximum MVS				Aggregated MVS			
				W	S	M	L	W	S	M	L
06/1896	11/1896	5	L	3.9	4.2	3.2	5.0	17.3	3.9	1.4	19.9
03/1898	04/1898	1	SM	4.1	4.7	4.4	2.3	4.1	4.7	4.4	2.3
12/1899	02/1900	2	SM	5.8	7.0	6.8	2.1	8.2	7.0	6.8	4.1
05/1901	06/1901	1	S	3.9	5.1	3.9	1.0	3.9	5.1	3.9	1.0
03/1907	05/1907	2	SM	5.0	6.1	5.7	2.1	7.5	6.1	5.7	4.1
07/1914	02/1915	3	S	5.5	7.7	5.8	1.2	14.3	4.5	4.3	16.1
05/1915	07/1915	2	SM	4.4	4.9	4.8	3.3	7.5	4.9	4.8	5.7
12/1916	04/1917	4	SM	4.1	5.0	5.2	1.9	12.3	5.0	5.2	10.2
03/1920	06/1920	3	L	2.7	2.9	1.3	4.4	7.7	2.5	0.5	9.3
10/1929	04/1934	54	SML	7.7	8.0	9.0	7.2	240.9	8.0	9.0	214.3
09/1937	12/1938	15	M	5.4	5.9	5.5	6.0	52.3	4.2	4.4	41.3
05/1940	08/1940	3	SM	4.4	5.5	5.4	0.6	8.2	5.5	5.4	4.4
09/1946	11/1946	2	M	3.0	3.4	4.7	0.9	5.8	3.4	4.7	4.1
05/1962	07/1962	2	M	3.3	3.9	5.4	0.6	5.9	3.9	5.4	4.0
05/1970	07/1970	2	M	3.0	3.7	5.6	0.3	5.1	3.7	5.6	3.2
08/1982	09/1982	1	SM	3.3	4.5	4.6	0.1	3.3	4.5	4.6	0.1
10/1982	03/1983	3	S	3.4	4.8	4.1	1.7	9.7	4.8	4.1	8.2
09/1986	10/1986	1	S	2.8	4.7	2.1	0.1	2.8	4.7	2.1	0.1
10/1987	08/1988	10	SML	9.5	10.8	9.2	8.0	52.8	10.8	9.2	43.6
10/1989	11/1989	1	S	4.3	6.4	4.2	0.3	4.3	6.4	4.2	0.3
08/1990	11/1990	3	SM	3.7	4.7	4.9	0.5	9.2	4.7	4.9	7.2
02/1991	02/1991	1	S	3.1	4.4	2.8	0.4	3.1	4.4	2.8	0.4
10/1997	12/1997	2	SM	5.2	6.4	6.3	1.0	8.4	6.4	6.3	4.4
08/1998	12/1998	4	SM	5.2	6.4	5.9	2.2	17.3	6.4	5.8	14.0
03/2000	11/2000	8	S	4.7	5.3	4.2	3.9	23.0	5.3	4.2	18.8
03/2001	07/2001	4	M	4.2	4.8	5.0	2.3	12.3	4.1	5.0	9.8
09/2001	11/2001	2	SM	6.4	7.2	6.4	3.8	9.9	7.2	6.4	7.1
07/2002	06/2003	11	SML	6.1	6.6	5.9	6.8	46.1	6.6	5.9	39.7

Source: Dow Jones Indexes, daily values of the DJIA index from 26/05/1896 to 10/10/2007. The beginning of the period of extreme volatility is the month when the MVS for one of the scales overcomes the threshold $\log_2(20) = 4.3219$, which corresponds to 0.05 probability. The end of such period is the month when the MVS for all scales falls below the threshold $\log_2(10) = 3.3219$, which corresponds to 0.1 probability. Duration of the period is the number of months between the first and the last date. The scale refers to the component of the MVS, which first broke the critical threshold: "S" stands for short scale (1-8 days), "M" stands for the medium scale (from 8 to 128 days) and "L" stands for the long scale (more than 128 days). The following four columns report the maximum value of the MVS over the period of extreme volatility: "W" stands for the weighted average (aggregated) MVS over the three scales, "S", "M" and "L" for the short, medium and long scales respectively. The next four columns contain the sum of the MVS values, when one of the scale components was above $\log_2(20)$, over the period of extreme volatility. The same notations for the scales are used. Computation by the author.

4.6 Conclusion

We design a new indicator of the stock market volatility, called the Market Volatility Scale, which characterizes the fluctuations both in time and frequency domains. Unlike the other multiscale indicators, proposed in the existing literature, it is based on the wavelet variances of returns, which are closely related to the spectrum of the time series of returns and allow an accurate scale-by-scale decomposition of the variance. We introduce a time-varying wavelet variance estimate, analogous to the conventional realized variance, but measuring variation in a particular range of frequencies.

The realized variances of the wavelet and scaling coefficients are aggregated to three representative scales (short term, medium term and long term). The MVS is computed at each date as the logarithm of the probability to observe the realized wavelet variance higher than its value at this date. This approach is inspired by the Richter-Gutenberg scale, used to measure the magnitude of earthquakes. Our scale is thus universal and easy to interpret, which is helpful for the comparison of events. When a logarithm to the base 2 is used for the definition, a one-point increase in the MVS corresponds to a volatility observation twice as unlikely.

We model the probability distribution of volatilities in a flexible way, using a generalized Pareto distribution fit for the clusters of extreme volatility. So we focus on the right tail, applying an empirical probability integral transform for the rest of the probability distribution. This choice is motivated by the need to be as accurate as possible in characterizing the rareness of extreme events.

The MVS is first computed separately for each scale, and then the results are aggregated to provide a general judgment on the market volatility at all scales. The aggregated MVS takes the form of the weighted average over the three scales. The weights are dependent on the scaling properties of the time series of returns, which are studied using the logscale diagrams. We show that the level of the long range dependence in volatility is not the same for different ranges of frequencies, supporting earlier findings reported for the high-frequency exchange rate volatility. Using the weights, depending on the Hurst exponent, we measure the importance of scales not only by the portion of the variance, attributed to them (clearly, most variance is due to the short scale), but also by the level of persistence, which is higher for the medium and long scales.

The MVS is computed for the daily Dow Jones Industrial Averages index data from 1896 to 2007 and for the intraday CAC40 data from 1995 to 2006. It is used for the detection and scale-by-scale analysis of the periods of extreme volatility. In particular, we report which scales of fluctuations (short, medium or long-term) were most important during each period, and at which scale the critical threshold, used to determine the

beginning of the period, was first overcome. The events can be ranked by the maximum value of the MVS at each scale and by the sum of the MVS values over the whole period of extreme volatility.

Clearly, the potential scope of application of the indicator and of the underlying variance decomposition goes far beyond the comparison of events. We mention several areas of further research which are of particular interest. From the theoretical viewpoint, it is important to understand the mechanics of the multiscale price dynamics. Wavelet variances can be an appropriate tool in testing market microstructure models of this type, both on simulation and real-data basis.

From a more practical perspective, the multiscale view of volatility can be suggestive in asset allocation and portfolio management. A multiscale indicator can be applied for conditioning the asset allocation process in a regime-switching framework. Another important area of application is volatility forecasting. In this paper we focused on structural rather than dynamic properties of volatility. But the interdependence of scales and the additivity property of wavelet variances make appealing the idea of the scale-by-scale autoregression, which can be used to forecast the components of volatility and then aggregate them.

4.7 Appendix. Wavelet Transforms

Any wavelet decomposition is based on the so-called wavelet filter and the associated scaling filter. A real-valued Discrete Wavelet Transform (DWT) filter $h_l : l = 0, \ldots, L-1$ of width L is an infinite sequence such as $h_l = 0$ for $l < 0$ and $l > L$, which sums to zero, has unit energy and is orthogonal to its even shifts:

$$\sum_{l=0}^{L-1} h_l = 0, \sum_{l=0}^{L-1} h_l^2 = 1, \sum_{l=0}^{L-1} h_l h_{l+2n} = 0 \tag{4.24}$$

The "quadrature mirror" filter corresponding to h_l is called the scaling filter and is given by:

$$g_l = (-1)^{l+1} h_{L-1-l}. \tag{4.25}$$

A Maximum Overlap Discrete Wavelet Transform (MODWT) wavelet filter and scaling filter are rescaled versions of the corresponding DWT filters defined by $\widetilde{h}_l = h_l/\sqrt{2}$ and $\widetilde{g}_l = g_l/\sqrt{2}$ respectively. Note that equation 4.25 also holds for the MODWT filters. The usage of the MODWT filters instead of the DWT aims mainly to define a transform which preserves the ability to carry out multi-resolution analysis and variance decomposition, but does not depend ont the choice of a starting point for a time series, which is the case when the outputs from the recursive application of the DWT filters are downsampled in the so-called "pyramid" algorithm. In the MODWT algorithm wavelet coefficients $\widetilde{W}_{1,t}$ and scaling coefficients $\widetilde{V}_{1,t}$ are obtained by filtering the circular shift of the vector X_t rather than X_t itself:

$$\widetilde{W}_{1,t} = \sum_{l=0}^{L-1} \widetilde{h}_l x_{t-l \bmod N}, \quad \widetilde{V}_{1,t} = \sum_{l=0}^{L-1} \widetilde{g}_l x_{t-l \bmod N} \tag{4.26}$$

The first stage of the MODWT pyramid algorithm consists in finding solution to the equation:

$$\begin{bmatrix} \widetilde{W}_1 \\ \widetilde{V}_1 \end{bmatrix} = \begin{bmatrix} \widetilde{B}_1 \\ \widetilde{A}_1 \end{bmatrix} x, \tag{4.27}$$

with $\widetilde{W}_1$ and $\widetilde{B}_1$ defined according to (4.26). The first level maximum overlap detail is given by $\widetilde{D}_1 = \widetilde{B}_1^T \widetilde{W}_1$ and the first level approximation is $\widetilde{S}_1 = \widetilde{A}_1^T \widetilde{V}_1$, so the signal admits the following additive (first-level) decomposition:

$$x = \widetilde{B}_1^T \widetilde{W}_1 + \widetilde{A}_1^T \widetilde{V}_1 = \widetilde{D}_1 + \widetilde{S}_1 \tag{4.28}$$

The energy of the signal, in its turn, can also be decomposed:

$$\|x\|^2 = \|\widetilde{W}_1\|^2 + \|\widetilde{V}_1\|^2. \tag{4.29}$$

The wavelet and scaling MODWT coefficients for higher levels of decomposition $j = 2, \ldots, J$ are obtained by applying recursive formulas of the pyramid algorithm without downsampling of the outputs:

$$\widetilde{W}_{j,t} = \sum_{l=0}^{L-1} \widetilde{h}_l \widetilde{V}_{j-1,t-2^{j-1}l \bmod N}, t = 0, \ldots, N-1; \tag{4.30}$$

$$\widetilde{V}_{j,t} = \sum_{l=0}^{L-1} \widetilde{g}_l \widetilde{V}_{j-1,t-2^{j-1}l \bmod N}, t = 0, \ldots, N-1. \tag{4.31}$$

The j-th level of the pyramid algorithm is defined by:

$$\begin{bmatrix} \widetilde{W}_j \\ \widetilde{V}_J \end{bmatrix} = \begin{bmatrix} \widetilde{B}_j \\ \widetilde{A}_j \end{bmatrix} \widetilde{V}_{j-1}, \tag{4.32}$$

Finally, the j-th level detail $\widetilde{D}_j$ and the J-th level smooth $\widetilde{S}_j$ are computed:

$$\widetilde{D}_j = \widetilde{A}_1^T \cdots \widetilde{A}_{j-1}^T \widetilde{B}_j^T \widetilde{W}_j, \quad \widetilde{S}_j = \widetilde{A}_1^T \cdots \widetilde{A}_{J-1}^T \widetilde{A}_J^T \widetilde{V}_J. \tag{4.33}$$

The multi-resolution analysis of signal X is thus its scale-by-scale additive decomposition:

$$x = \sum_{j=1}^{J} \widetilde{D}_j + \widetilde{S}_J \tag{4.34}$$

The scale-by-scale decomposition of energy reads:

$$\|x\|^2 = \sum_{j=1}^{J} \|\widetilde{W}_j\|^2 + \|\widetilde{V}_J\|^2. \tag{4.35}$$

Bibliography

S. Alizadeh, M. Brandt, and F. Diebold. Range-based estimation of stochastic volatility models. *The Journal of Finance*, 57(3):1047–1091, 2002.

T. Andersen. Stochastic autoregressive volatility: a framework for volatility modeling. *Mathematical Finance*, 4:75–102, 1994.

T. Andersen. Return volatility and trading volume: An information flow interpretation of stochastic volatility. *Journal of Finance*, 51(1):169–204, 1996.

T. Andersen and T. Bollerslev. Heterogeneous information arrivals and return volatility dynamics: Uncovering the long run in high frequency data. *Journal of Finance*, 52: 975–1005, 1997.

T. Andersen and T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905, 1998.

T. Andersen, T. Bollerslev, and S. Lange. Forecasting financial market volatility: Sample frequency vis-a-vis forecast horizon. *Journal of Empirical Finance*, 6(5):457–477, 1999.

T. Andersen, T. Bollerslev, F. Diebold, and P. Labys. The distribution of exchange rate volatility. *Journal of the American Statistical Association*, 96:42–55, 2001.

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

V. Anh, C. Heyde, and N. Leonenko. Dynamic models of long-memory processes driven by levy noise. *Journal of Applied Probability*, 39:730–747, 2002.

C. Anteneodo and R. Riera. Additive-multiplicative stochastic models of financial mean-reverting processes. *Physical Review E*, 72:026106, 2005.

M. Anufriev. Wealth-Driven Competition in a Speculative Financial Market: Examples with Maximizing Agents. *Quantitative Finance*, 8(4):363–380, 2008.

M. Anufriev and G. Bottazzi. Asset pricing model with heterogeneous investment horizons. *Laboratory of Economics and Management Working Paper Series, Sant'Anna School for Advanced Studies*, 2004.

M. Anufriev, G. Bottazzi, and F. Pancotto. Equilibria, stability and ssymptotic dominance in a speculative market with heterogeneous traders. *Journal of Economic Dynamics and Control*, 30(9-10):1787–1835, 2006.

A. Arneodo, J. Muzy, and D. Sornette. Casual cascade in stock market from the 'infrared' to the 'ultraviolet'. *European Physical Journal B*, 2:277–282, 1998.

M. Asai, M. McAleer, and J. Yu. Multivariate stochastic volatility: A review. *Econometric Reviews*, 25(2-3):145–175, 2006.

M. Ausloos and K. Ivanova. Dynamical model and nonextensive statistical mechanics of a market index on large time windows. *Physical Review E*, 68:046122, 2003.

M. Avellaneda, C. Friedmen, R. Holmes, and D. Samperi. Calibrating volatility surfaces via relative-entropy minimization. *Apllied Mathematical Finance*, 4(1):37–64, 1997.

E. Bacry, J. Delour, and J. Muzy. Multifractal random walk. *Physical Review E*, 64: 026103, 2001.

X. Bai, J. R. Russel, and G. Tiao. Kurtosis of garch and stochastic volatility models with non-normal innovations. *Journal of Econometrics*, 114:349–360, 2003.

R. Baillie, T. Bollerslev, and O. Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 74(1):3–30, 1996.

F. Bandi and J. Russel. Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, 75(2):339–369, 2008.

N. Barberis and A. Shleifer. Style investing. *Journal of Financial Economics*, 68(2): 161–199, 2003.

O. Barndorff-Nielsen and N. Shephard. Non-gaussian ornsteinuhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society B*, 63(2):167–241, 2001.

O. Barndorff-Nielsen and N. Shephard. Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17(5):457–477, 2002a.

O. Barndorff-Nielsen and N. Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B*, 64(2):253–280, 2002b.

O. Barndorff-Nielsen and N. Shephard. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37, 2002c.

G. Barone-Adesi, R. Engle, and L. Mancini. Garch option pricing model with filtered historical simulation. *Review of Financial Studies, forthcoming*, 2008.

D. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, 9:69–107, 1996.

B. Biais, L. Glosten, and C. Spatt. Market microstructure: A survey of microfoundations, empirical results and policy implications. *Journal of Financial Markets*, 8: 217264, 2005.

M. Billio, M. Caporin, and M. Gobbo. Applied financial economics letters. *Applied Financial Economics Letters*, 2(2):123–130, 2006.

F. Black. Noise. *Journal of Finance*, 41, 1976.

F. Black and M. Scholes. Pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 1973.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

T. Bollerslev. A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 69(3):542–547, 1987.

T. Bollerslev. Modeling the coherence in short-run nominal exchange rate: A multivariate generalized arch approach. *Review of Economics and Statistics*, 72:498–505, 1990.

T. Bollerslev and O. Mikkelsen. Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73(1):151–184, 1996.

T. Bollerslev, R. Chou, and K. Kroner. Arch modeling in finance : A review of the theory and empirical evidence. *Journal of Econometrics*, 52(1-2):5–59, 1992.

J. Boudoukh, M.P. Richardson, and R.F. Whitelaw. A tale of three schools: Insights on autocorrelations of short-horizon stock returns. *Review of Financial Studies*, 7: 539–573, 1994.

F. Breidt, N. Crato, and P. de Lima. The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83(1-2):325–34, 1998.

W. Breymann, S. Ghashghaie, and P. Talkner. A stochastic cascade model for fx dynamics. *International Journal of Theoretical and Applied Finance*, 3:357–360, 2000.

W. Brock and C. Hommes. A rational route to randomness. *Econometrica*, 65(5): 1059–1095, 1997.

W.A. Brock and C.H. Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22(8-9):1235–1274, 1998.

C. Brooks and G. Persand. Volatility forecasting for risk management. *Journal of Forecasting*, 22(1):1 – 22, 2003.

C. Broto and E. Ruiz. Estimation methods for stochastic volatility models: a survey. *Journal of Economic Surveys*, 18(5):613–649, 2004.

G. Buchbinder and K. Chistilin. Multiple time scales and the empirical models for stochastic volatility. *Physica A: Statistical Mechanics and its Applications*, 379(1): 168–178, 2007.

L. Calvet and A. Fisher. Forecasting multifractal volatility. *Journal of Econometrics*, 105(1):27–58, 2001.

E. Capobianco. State-space stochastic volatility models: a review of estimation algorithms. *Applied Stochastic Models and Data Analysis*, 12:265279, 1996.

E. Capobianco. Multiscale analysis of stock index return volatility. *Computational Economics*, 23(3):219–237, 2004.

B. Castaing, Y. Gagne, and E. J. Hopfinger. Velocity probability density functions of high reynolds number turbulence. *Physica D*, 46(2):177–200, 1990.

K. C. Chan. On the contrarian investment strategy. *Journal of Business*, 61(2):147–163, 1988.

W. Chan and J. Maheu. Conditional jump dynamics in stock market returns. *Journal of Business & Economic Statistics*, 20(3):377–389, 2002.

T. Chauveau and R. Topol. A unifying microstructure framework for modeling intraday and interday asset pricing dynamics: the case of exchange rates. *European Financial Management*, 5(3):341–368, 2002.

M. Chernov, R. Gallant, E. Ghysels, and G. Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2):225–257, 2003.

C. Chiarella and X.-Z. He. Asset price and wealth dynamics under heterogeneous expectations. *Quantitative Finance*, 1(5):509 – 526, 2001.

S. Chib, F. Nardari, and N. Shephard. Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 134(2):341–371, 2006.

P.-H. Chou, K.C. John Wei, and H. Chung. Sources of contrarian profits in the japanese stock market. *Journal of Empirical Finance*, 14(3):261–286, 2007.

K. Christensen and M. Podolskij. Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2):323–349, 2007.

K. Christensen, L. Danon, T. Scanlon, and P. Bak. Unified scaling law for earthquakes. *Physical Review Letters*, 88(17):178501–504, 2002.

P. Christoffersen and F. Diebold. How relevant is volatility forecasting for financial risk management? *The Review of Economics and Statistics*, 82(1):12–22, 2000.

S. Coles. *An Introduction to Statistical Modeling of Extreme Values.* Springer-Verlag, London, 2001.

F. Comte and E. Renault. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8(4):291–323, 1998.

J. Conrad, M. Gultekin, and G. Kaul. Profitability of short-term contrarian strategies: Implications for market efficiency. *Journal of Business & Economic Statistics*, 15(3): 379–386, 1997.

R. Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.

V. Corradi and W. Distaso. Semi-parametric comparison of stochastic volatility models using realized measures. *Review of Economic Studies*, 73(3):635–667, 2006.

F. Corsi. A simple long memory model of realized volatility. *Working Paper, University of Southern Switzerland*, 2004.

A. Cortines, R. Riera, and C. Anteneodo. From short to fat tails in financial markets: a unified description. *The European Physical Journal B*, 60(3):385–389, 2007.

J. Cox, J. Ingersoll, and S. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica*, 53:363–384, 1985.

D. Cutler, J. Poterba, and L. Summers. What moves stock prices? *Journal of Portfolio Management*, 15:4–12, 1989.

M. Dacorogna, U. Muller, R. Dave, R. Olsen, and O. Pictet. *Modelling Short-term Volatility with GARCH and HARCH models*, pages 161–176. John Wiley, N.-Y., 1998.

Z. Ding and C. Granger. Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics*, 73:185–215, 1996.

Z. Ding, C. Granger, and R. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106, 1993.

F. Drost and T. Nijman. Temporal aggregation of garch processes. *Econometrica*, 61 (4):909–927, 1993.

F. Drost and B. Werker. Closing the garch gap: Continuous time garch modeling. *Journal of Econometrics*, 74(1):31–57, 1996.

J.-C. Duan. The garch option pricing model. *Mathematical Finance*, 5(1):1332, 1995.

D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053, 2003.

B. Dupire. Pricing and hedging with smiles. *Proceedings of AFFI Conference, La Baule, June 1993*, 1993.

B. Dupire. Pricing with a smile. *Risk Magazine*, 7(1):18–20, 1994.

R. Engle. Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation. *Econometrica*, 50:987–1008, 1982.

R. Engle. Dynamic conditional correlation - a simple class of multivariate garch models. *Journal of Business and Economic Statistics*, 20(3):339–350, 2002.

R. Engle and T. Bollerslev. Modelling the persistence of conditional variances. *Econometric Reviews*, 5(1):1 – 50, 1986.

R. Engle, D. Lilien, and R. Robins. Estimating time varying risk premia in the term structure: The arch-m model. *Econometrica*, 55(2):391–407, 1987.

B. Eraker, M. Johannes, and N. Polson. The impact of jumps in returns and volatility. *Journal of Finance*, 53:1269–1300, 2003.

E. Fama. The behaviour of stock market prices. *Journal of Business*, 38:34–105, 1965.

E. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383–417, 1970.

J. Fan, Y. Fan, and J. Jiang. Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of American Statistical Association*, 102:618–631, 2007.

V. Fernandez and M. Lucey. Portfolio management under sudden changes in volatility and heterogeneous investment horizons. *Physica A: Statistical Mechanics and its Applications*, 375(2):612–624, 2007.

L. Forsberg and E. Ghysels. Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, 5(1):31–67, 2007.

J.-P. Fouque and C.-H. Han. Asian options under multiscale stochastic volatility. *Contemporary Mathematics*, 351:125–138, 2004.

J.-P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna. Multiscale stochastic volatility asymptotics. *Multiscale Modeling & Simulation*, 2(1):22–42, 2003.

R. Friedrich, J. Peinke, and C. Renner. How to quantify deterministic and random influences on the statistics of the foreign exchange market. *Physical Review Letters*, 84:5224 – 522, 2000.

R. Gençay, F. Selçuk, and B. Whitcher. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Harcourt Brace, San Diego, 2001a.

R. Gençay, F. Selçuk, and B. Whitcher. Scaling properties of foreign exchange volatility. *Physica A*, pages 249–266, 2001b.

S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge. Turbulent cascades in foreign exchange markets. *Nature*, 381:767–770, 1996.

E. Ghysels, P. Santa-Clara, and R. Valkanov. Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131 (1-2):59–95, 2006.

M. Gilli and E.Këllezi. An application of extreme value theory for measuring financial risk. *Computational Economics*, 27(2-3):207–228, 2006.

L. Glosten, R. Jagannathan, and D. Runkle. On the relation between the expected value and volatility of the nominal excess return on stocks. *Journal of Finance*, 46: 17791801, 1992.

C. Granger and R. Joyeux. An introduction to long memory time series models and fractional differencing,. *Journal of Time Series Analysis*, 1:15–29, 1980.

C. Granger and S.-H. Poon. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2):478–539, 2003.

P. Hansen. A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics*, 3(4):525–554, 2005.

P. Hansen and A. Lunde. A forecast comparison of volatility models: Does anything beat a garch(1,1). *Journal of Applied Econometrics*, 20(7):873 – 889, 2005.

A. Harvey. *Long Memory in Stochastic Volatility*, pages 307–320. Butterworth-Heineman, Oxford, 1998.

A. Harvey and N. Shephard. Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, 14(4):429–434, 1996.

R. Hawkes and P. Date. Medium-term horizon volatility forecasting: A comparative study. *Applied Stochastic Models in Business and Industry*, 23(6):465 – 481, 2007.

V. Henderson. Analytical comparisons of option prices in stochastic volatility models. *Mathematical Finance*, 15(1):49–59, 2005.

S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

C. Heyde. On modes of long-range dependence. *Journal of Applied Probability*, 39(4): 882–888, 2002.

M. Higgins and A. K. Bera. A class of nonlinear arch models. *International Economic Review*, 33(1):137–158, 1992.

C. Hommes. Heterogeneous agent models in economics and finance. In *Handbook of Computational Economics*, volume 2, pages 1109–1186, Amsterdam, 2006. Elsevier Science.

J. Hosking. Fractional differencing. *Biometrika*, 68:165–176, 1981.

J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300, 1987.

S. Hwang, S. Satchell, and P. Pereira. How persistent is stock return volatility? an answer with markov regime switching stochastic volatility models. *Journal of Business, Finance and Accounting*, 34(5-6):1002–1024, 2007.

E. Jacquier, N. Polson, and P. Rossi. Bayesian analysis of stochastic volatility models with fat tails and correlated errors. *Journal of Econometrics*, 122(1):185–212, 2004.

N. Jegadeesh and S. Titman. Overreaction, delayed reaction, and contrarian profits. *Review of Financial Studies*, 8(4):973993, 1995.

C. Klüppelberg, A. Lindner, and R. Maller. A continuous-time garch process driven by a levy process: Stationarity and second-order behaviour. *Journal of Applied Probability*, 41(3):601–622, 2004.

AN Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Proceedings of the USSR Academy of Sciences (in Russian)*, 30:299–303, 1941.

S. Koopman and E. Uspensky. The stochastic volatility in mean model: Empirical evidence from international stock markets. *Journal of Applied Econometrics*, 17(6): 667 – 689, 2002.

M. Lanne and P. Saikkonen. Non-linear garch models for highly persistent volatility. *The Econometrics Journal*, 8(2):251276, 2005.

B. LeBaron. Stochastic volatility as a simple generator of apparent financial power laws and long memory. *Quantitative Finance*, 1(6):621–631, 2001a.

B. LeBaron. Evolution and time horizons in an agent-based stock market. *Macroeconomic Dynamics*, 5:225–254, 2001b.

B. LeBaron. Agent-based computational finance. In *Handbook of Computational Economics*, volume 2, pages 1187–1233, Amsterdam, 2006. Elsevier Science.

D. Lee, H. Chan, R. Faff, and P. Kalev. Short-term contrarian investing is it profitable?... yes and no. *Journal of Multinational Financial Management*, 13(4-5):385–404, 2003.

H. Lee. International transmission of stock market movements: A wavelet analysis. *Applied Economics Letters*, 11:197–201, 2004.

R. Liesenfeld and J. Richard. Univariate and multivariate stochastic volatility models: Estimation and diagnostics. *Journal of Empirical Finance*, 10(4):505–531, 2003.

S. Ling and M. McAleer. Necessary and sufficient moment conditions for the garch(r, s) and asymmetric power garch(r, s) models. *Econometric Theory*, 18:722–729, 2002a.

S. Ling and M. McAleer. Stationarity and the existence of moments of a family of garch processes. *Journal of Econometrics*, 106:109–117, 2002b.

S. Ling and M. McAleer. Asymptotic theory for a vector arma-garch model. *Econometric Theory*, 19:278–308, 2003.

J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1):13–39, 1965.

J. Liu. Portfolio selection in stochastic environments. *Review of Finanical Studies*, 20 (1):1–39, 2007.

M. Liu. Modeling long memory in stock market volatility. *Journal of Econometrics*, 99 (1):139–171, 2000.

A. Lo and C. MacKinlay. When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3:175–250, 1990.

I. Lobato and C. Velasco. Long memory in stock market trading volume. *Journal of Business & Economic Statistics*, 18(4):410–427, 2000.

M. Loretan. Generating market risk scenarios using principal component analysis. *The Measurement of Aggregate Market Risk, Publications of the Committee on the Global Financial System No 7*, pages 23–60, 1997.

T. Lux. Long-term stochastic dependence in financial prices: Evidence from the german stock market. *Applied Economics Letters*, 3:701–706, 1996.

T. Lux and M. Marchesi. Volatility clustering in financial markets: A microsimulation of interacting agents. *International Journal of Theoretical and Applied Finance*, 3(4): 675 – 702, 2000.

P. Lynch and G. Zumbach. Market heterogeneities and the causal structure of volatility. *Quantitative Finance*, 3(4):320–331, 2003.

Y. Maghsoodi. Exact solution of a martingale stochastic volatility option problem and its empirical evaluation. *Mathematical Finance*, 17(2):249265, 2005.

B. Maillet and T. Michel. An index of market shocks based on multiscale analysis. *Quantitative Finance*, 3(2):88–97, 2003.

B. Maillet and T. Michel. The impact of the 9/11 events on the american and french stock markets. *Review of International Economics*, 13(3):597–611, 2005.

B. Maillet, T. Michel, and A. Subbotin. *A Revised Index of Market Shocks: A New Multi-horizon Richter Scale for Stock Markets*. 2007. JMA conference paper, Fribourg, Switzerland, 31 May - 1 June 2007.

B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36: 394–419, 1963.

B. Mandelbrot. When can price be arbitraged efficiently? a limit to the validity of the random walk and martingale models. *Review of Economics and Statistics*, 53(3): 225–236, 1971.

B. Mandelbrot and M. Taqqu. Robust r/s analysis of long run serial correlation. In *Bulletin of the International Statistical Institute, Vol. 48, Book 2. Proceedings of the 42nd session of the International Statistical Institute, Manila 1979*, pages 69–104, 1979.

B. Mandelbrot and J. Van Ness. Fractional brownian motion, fractional noises and applications. *SIAM Review*, 10:422–437, 1968.

H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

S. Marple. *Digital Spectral Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

M. Martens and D. van Dijk. Measuring volatility with the realized range. *Journal of Econometrics*, 138(1):181–207, 2007.

M. Martens and J. Zein. Predicting financial volatility: High-frequency time-series forecasts vis-á-vis implied volatility. *Journal of Futures Markets*, 24(11):1005–1028, 2004.

J. Masoliver and J. Perello. Multiple time scales and the exponential ornstein-uhlenbeck stochastic volatility model. *Quantitative Finance*, 6(5):423–433, 2006.

M. McAleer and M. Medeiros. Realized volatility: A review. *Econometric Reviews*, 27 (1-3):10–45, 2008.

M. Mega, P. Allegrini, P. Grigolini, V. Latora, L. Palatella, A. Rapisarda, and S. Vinciguerra. Power law time distributions of large earthquakes. *Physical Review Letters*, 90:188501, 2003.

R. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183, 1973.

G. Molina, C.H. Han, and J.P. Fouque. *MCMC Estimation of Multiscale Stochastic Volatility Models*. 2004. Unpublished manuscript. University of Calfornia.

K. Morimune. Volatility models. *The Japanese Economic Review*, 58(1):1–23, 2007.

U. Müller, M. Dacorogna, R. Dave, R. Olsen, O. Pictet, and J. Von Weizsacker. Volatilities of different time resolutions - analyzing the dynamics of market components. *Journal of Empirical Finance*, 4:213–239, 1997.

J. Muzy, J. Delour, and E. Bacry. Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *The European Physical Journal B*, 17 (3):537–548, 2000.

J.-F. Muzy and E. Bacry. Multifractal stationary random measures and multifractal random walks with log infinitely divisible scaling laws. *Physical Review E*, 66(5): 056121, 2002.

J.-F. Muzy, D. Sornette, J. Delour, and A. Arneodo. Multifractal returns and hierarchical portfolio theory. *Quantitative Finance*, 1(1):131–148, 2001.

A. Nawroth and J. Peinke. Multiscale reconstruction of time series. *Physics Letters A*, 360(2):234–237, 2006.

D. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59:347–370, 1991.

B. Ninness. Strong laws of large numbers under weak assumptions with application. *IEEE Transactions on Automatic Control*, 45(11):2117–2122, 2000.

M. Pasquini and M. Serva. Clustering of volatility as a multiscale phenomenon. *The European Physical Journal B*, 16(1):195–201, 2000.

D. Percival and A. Walden. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge, 2000.

J. Perello, J. Masoliver, and J.-P.Bouchaud. Multiple time scales in volatility and leverage correlations: A stochastic volatility model. *Applied Mathematical Finance*, 11: 27–50, 2004.

B. Pochart and J.-P. Bouchaud. The skewed multifractal random walk with applications to option smiles. *Quantitative Finance*, 2(4):303–314, 2002.

S.B. Pope. *Turbulent flows*. Cambridge university press, 2000.

J. Ramsey and Z. Zhang. The analysis of foreign exchange data using waveform dictionaries. *C.V. Starr Center for Applied Economics Working paper, New York University*, 1995.

C. Renner, J. Peinke, and R. Friedrich. Evidence of markov properties of high frequency exchange rate data. *Physica A*, 298(3):499–520, 2001a.

C. Renner, J. Peinke, and R. Friedrich. Experimental indications for markov properties of small-scale turbulence. *Journal of Fluid Mechanics*, 433:383–409, 2001b.

G. Richrads. A fractal forecasting model for financial time series. *Journal of Forecasting*, 23(8):586 – 601, 2004.

C. Richter. *Elementary Seismology*. Freeman, San Francisco, 1958.

H. Risken. *The Fokker-Planck equation: Methods of Solution and Applications*. Springer-Verlag, Berlin, 1989.

P. Ritchken and R. Trevor. Pricing options under generalized garch and stochastic volatility processes. *Journal of Finance*, 54(1):377–402, 1999.

F. Schmitt, D. Schertzer, and S. Lovejoy. Multifractal fluctuations in finance. *International Journal of Theoretical and Applied Finance*, 3(3):361–364, 2000.

L. Scott. Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative Analysis*, 22:419–438, 1987.

E. Sentana. Quadratic arch models. *Review of Economic Studies*, 62(4):639–661, 1995.

W. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

H. Simon. From substantive to procedural rationality. In S. J. Latsis, editor, *Method and Appraisal in Economics*, pages 129–148. Cambridge University Press, 1976.

M. So, K. Lam, and W. Li. A stochastic volatility model with markov switching. *Journal of Business & Economic Statistics*, 16(2):244–253, 1998.

E. Stein and J. Stein. Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4:727–752, 1991.

A. Subbotin. A multi-horizon scale for volatility. *CES Working Paper 2008.20, University of Paris-1*, 2008. 44 p.

A. Subbotin, T. Chauveau, and K. Shapovalova. Price volatility in a market with boundedly rational traders and heterogeneous investment horizons. *Forthcoming*, 2010a.

A. Subbotin, T. Chauveau, and K. Shapovalova. Volatility models: from garch to multi-horizon cascades. In B. Naas and J. Lysne, editors, *Financial Markets and the Global Recession*. Nova Science Publishers Inc, NY *(forthcoming)*, 2010b.

S. Taylor. *Financial Returns Modelled by the Product of two Stochastic Processes a Study of the Daily Sugar Prices 1961-75*, volume 1, pages 203–226. North-Holland, Amsterdam, 1982.

S. Taylor. Modeling stochastic volatility: A review and comparative study. *Mathematical Finance*, 4(2):183–204, 1994.

J. Vanden. Equilibrium analysis of volatility clustering. *Journal of Empirical Finance*, 12:374–417, 2005.

D. Veitch and P. Abry. A wavelet based joint estimator of the parameters of long-range dependence. *IEEE Transactions on Information Theory*, 45(3):878–897, 1999.

D. Weinbaum. Investor heterogeneity, asset pricing and volatility dynamics. *Journal of Economic Dynamics and Control*, 33:1379–1397, 2009.

J. Woerner. Estimation of integrated volatility in stochastic volatility models. *Applied Stochastic Models in Business and Industry*, 21(1):27 – 44, 2005.

B. Zhou. High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics*, 14(1):45–52, 1996.

G. Zumbach. Volatility processes and volatility forecast with long memory. *Quantitative Finance*, 4(1):70 – 86, 2004.

G. Zumbach, M. Dacorogna, J. Olsen, and R. Olsen. Measuring shocks in financial markets. *International Journal of Theoretical and Applied Finance*, 3(3):347–355, 2000.

Résumé de thèse en français

Horizons d'investissement multiples et dynamique des prix des titres

Annotation

Les prix de marché des actifs risqués sont dictés par les actions des agents économiques qui ont des horizons d'investissement différents : ils ajustent leurs portefeuilles à des fréquences différentes et observent les rendements à différentes échelles. Dans cette thèse, nous examinons trois aspects distincts du problème des horizons multiples d'investissement. En premier lieu, nous étudions les implications théoriques de l'hétérogénéité des horizons de décision des investisseurs pour la dynamique des prix. Cette analyse est effectuée dans le cadre d'une rationalité complète et bornée. Deuxièmement, nous testons la capacité de différents modèles de séries chronologiques de volatilité des prix à représenter les propriétés des rendements boursiers simultanément à différentes échelles de temps. Enfin, nous proposons une méthode de mesure de la volatilité à échelles multiples basée sur des filtres d'ondelettes, avec application à la détection des crises financières.

Mots clés : évaluation des actifs, agents hétérogènes, multiples horizons d'investissement, clustering de volatilité

J.E.L. Classification : G.10, G.14.

Chapitre 1

Introduction

Les thèmes centraux de cette thèse sont la modélisation et l'analyse des fluctuations des prix d'actions sur un marché, où de multiples échelles d'investissement co-existent. L'intuition derrière notre approche est que les différents types d'investisseurs (institutionnels, gestionnaires de portefeuille dans des fonds d'investissement, spéculateurs intrajournaliers, etc) ont des fréquences caractéristiques d'opérations sur le marché et des horizons différents de prise de décision. Leurs actions peuvent engendrer des fluctuations de prix dans les différentes bandes du spectre. Aussi, tous ces investisseurs observent les mêmes séries de prix à des échelles différentes, avec une résolution particulière.

Le concept de multiples échelles d'investissement n'est pas suffisamment étudié dans la littérature théorique existante. Il est important d'explorer si, en présence d'échelles multiples, le prix et le rendement d'équilibre d'un actif risqué peuvent être défini et comment. L'approche d'horizons multiples peut être utile pour expliquer certains faits stylisés concernant la dynamique des rendements, tels que le clustering de volatilité. D'ailleurs, les modèles économétriques de la volatilité des prix peuvent être analysés quant à leur compatibilité avec la modélisation multi-échelle. Les méthodes de mesure de volatilité, tenant compte des propriétés spectrales des fluctuations, pourrait également être utile. Ces questions font objet des trois essais, qui composent cette thèse.

Les implications de l'hypothèse de multiples échelles d'investissement pour les propriétés des séries chronologiques de rendements risqués sont d'abord étudiés du point de vue théorique dans un cadre simplifié d'un modèle à agents hétérogènes. Puis nous nous adressons à des questions pratiques et empiriques : le choix de modèles de volatilité des séries chronologiques, la prise en compte de diverses échelles de temps et la mesure de volatilité à des échelles multiples.

Dans le premier essai (Chapitre 2) nous étudions l'effet de multiples horizons d'investissement et de la rationalité limitée des investisseurs sur la dynamique des prix. Dans le cadre d'un modèle à agents hétérogènes, établi dans Anufriev et al. [2006], nous introduisons de multiples échelles d'investissement. Nous considérons une économie d'échange

pur avec un actif risqué et un actif sans risque, où les agents avec aversion relative constante pour le risque maximisent l'espérance d'utilité de la richesse sur des périodes d'investissement hétérogènes. La demande des investisseurs pour l'actif risqué dépend des rendements historiques, de sorte que notre modèle peut englober une vaste gamme d'hypothèses comportementales. Les conditions nécessaires, sous lesquelles les rendements risqués peuvent être un processus stationnaire *iid*, sont établies. La compatibilité de ces conditions avec différents types de fonctions de demande des agents hétérogènes est explorée.

Nous trouvons que la volatilité conditionnelle des rendements ne peut pas être constante dans de nombreuses situations génériques, en particulier si les agents avec des horizons d'investissement différents opèrent sur le marché. Dans ce dernier cas, le processus de rendements peut être hétéroscédastique, même si tous les investisseurs sont "fondamentalistes", mais leur demande de l'actif risqué est soumise à des chocs *iid* exogènes. Nous montrons que l'hétérogénéité des horizons d'investissement peut être une explication possible de différents faits stylisés, observés pour les rendements d'actions et, en particulier, du retour à la moyenne et du clustering de volatilité.

Dans le deuxième essai (Chapitre 3) , nous donnons un aperçu de différentes méthodes de modélisation de la volatilité des prix boursiers, en nous focalisant sur la capacité de ces méthodes à reproduire les propriétés empiriques des séries temporelles correspondantes. Les propriétés des fluctuations de prix varient selon les échelles temporelles d'observation. L'adéquation des modèles pour décrire la dynamique des prix simultanément à plusieurs horizons est le thème central de cette étude. Nous proposons une analyse détaillée des plus récents modèles de volatilité, qui sont fondés sur des concepts théoriques divers et parfois concurrents. Ils appartiennent soit à la famille de modèles GARCH, soit à la famille de modèles à volatilité stochastique, souvent empruntant des outils méthodologiques de la physique statistique. Nous comparons leurs propriétés et concluons sur leur utilité pratique et perspectives d'application.

Dans le troisième essai (Chapitre 4), la volatilité d'un indice boursier est décomposée en temps et en échelle avec un filtre d'ondelettes. Nous concevons un indicateur probabiliste de volatilité, analogue à l'échelle de Richter en géophysique. Le méthode des excès ("peak-over-threshold method") est utilisée pour calibrer la distribution généralisée de Pareto pour les valeurs extrêmes dans les variances réalisées des coefficients d'ondelettes. L'indicateur est calculé pour les observations quotidiennes de l'indice Dow Jones de 1896 à 2007 et pour les données intrajournalières de CAC40 de 1995 à 2006. Les résultats sont utilisés pour la comparaison et pour l'analyse structurelle des événements extrêmes sur le marché d'actions, ainsi que pour la détection des crises financières.

Chapitre 2

Dynamique des prix dans un marché à horizons d'investissement multiples et agents à rationalité bornée

2.1 Introduction

Jusqu'à présent, la littérature sur les marchés hétérogènes se focalise presque exclusivement sur les anticipations des agents concernant l'évolution du marché. Selon le mode suivant lequel ces anticipations sont formées, les investisseurs sont classés dans "fondamentalistes", "chartistes" et "noise traders". Il est montré que les interactions, le comportement imitatif de type "troupeau" et les changements de stratégies par les agents hétérogènes transforment les processus de bruit et provoquent la volatilité excessive des prix et les queues épaisses dans la distribution des rendements (voir Hommes [2006] et LeBaron [2006] pour une revue de ces modèles). Andersen [1996] interprète la volatilité agrégée comme la manifestation d'arrivées de nombreuses informations hétérogènes. Les limites à l'arbitrage, la "psychologie du marché", l'analyse heuristique et les biais qui font l'objet de la finance comportementale, peuvent également être utiles pour expliquer les données empiriques [voir Barberis and Shleifer, 2003].

Un certain nombre de modèles traitables analytiquement ont été proposés pour étudier la dynamique des marchés financiers avec une hétérogénéité de croyances des investisseurs à rationalité bornée sur les rendements futurs. Brock and Hommes [1998] proposent un modèle, où les investisseurs basculent entre un certain nombre de stratégies en fonction des profits excédentaires attendus ou réalisés. Les stratégies d'investissement

simples décrivent les modes de comportement des investisseurs qui sont souvent observés empiriquement - chartisme et la chasse aux tendances. Chiarella and He [2001] et Anufriev et al. [2006] étudient un marché artificiel avec les investisseurs, qui suivent des stratégies hétérogènes et maximisent l'espérance d'utilité à aversion relative constante pour le risque (CRRA). Par rapport aux études précédentes, utilisant l'aversion absolue constante pour le risque (CARA), ils prennent des décisions d'investissement dépendants de la richesse, ce qui est sans doute plus réaliste, mais techniquement plus difficile à modéliser. Vanden [2005] introduit une forme de dépendance plus sophistiquée, en escalier, de l'aversion au risque par rapport à la richesse. Il est montré que cela peut avoir des conséquences importantes pour la dynamique de rendements. Plus récemment Weinbaum [2009] montre que l'aversion au risque hétérogène et le partage des risques peuvent être à la source du clustering de volatilité.

À notre connaissance, tous les modèles mentionnés ci-dessus ne tiennent pas compte d'une des sources importantes de l'hétérogénéité - des horizons d'investissement différents. Par horizons d'investissement, nous désignons les périodes typiques entre deux ajustements consécutifs du portefeuille d'investissement, caractéristiques à un certain type d'investisseurs. L'hétérogénéité du marché par rapport aux fréquences des opérations des agents est appelée l'hypothèse des échelles d'investissement multiples (EIM). Nous supposons que les investisseurs maximisent l'espérance d'utilité de la richesse à la fin d'une période d'investissement.

Avant nos travaux, l'effet de l'hétérogénéité des horizons d'investissement a été étudié dans Anufriev and Bottazzi [2004]. Ils trouvent un point fixe dans la dynamique du prix d'un actif risqué sous l'hypothèse que les agents maximisent l'utilité CARA de la richesse sur des périodes différentes dans le futur. Mais leur modèle ne tient pas compte de l'effet de différentes fréquences d'ajustement du portefeuille et, en raison des contraintes de l'approche CARA, n'est pas réaliste dans la prise en compte de la dynamique de la richesse. Les auteurs concluent que l'hétérogénéité des horizons d'investissement à elle seule ne suffit pas à provoquer l'instabilité du prix fondamental et l'émergence des dynamiques non-triviales des rendements, avec clustering de volatilité ou corrélations temporelles. Dans cette étude, nous tirons la conclusion opposée, qui est proche de celle obtenue dans Chauveau and Topol [2002]. Dans un cadre différent, ils expliquent le clustering de volatilité des taux de change par des effets de la microstructure des marchés, afin d'unifier la dynamique intra- et inter-journalière.

Sans examiner l'hypothèse d'EIM analytiquement, plusieurs études antérieures évoquent l'hétérogénéité des horizons d'investissement comme une explication possible des faits stylisés dans la volatilité des prix boursiers. L'hypothèse que la dynamique des prix s'explique par les actions des investisseurs à différents horizons sert d'une base microéconomique au modèle de volatilité dans Müller et al. [1997]. Ils supposent l'existence des composantes de volatilité, qui correspondent aux différentes bandes de fréquences de fluctuations des prix et qui sont d'importance inégale pour les différents participants

du marché. Parmi ces participants, les spéculateurs intra-journaliers, les traders journaliers, les gestionnaires de portefeuilles et les investisseurs institutionnels, chacun ayant un temps caractéristique de réaction aux nouvelles et une fréquence caractéristique des opérations sur le marché. Ainsi, les fréquences des fluctuations des prix dépendent des périodes entre les décisions d'allocation d'actifs, et/ou de la fréquence des réajustements des portefeuille d'investissement.

Dans cette thèse, nous étudions si la présence des *(i)* investisseurs chasseurs de tendances ou "contrariants" et *(ii)* arrivées d'informations hétérogènes sur le marché sont des propriétés nécessaires pour qu'un modèle d'agents en interaction puisse reproduire les faits stylisés de la dynamique de clustering de volatilité. Nous montrons que, sous certaines conditions, le clustering de volatilité peut survenir dans un marché où seuls les investisseurs fondamentalistes agissent, s'ils ajustent leurs portefeuilles à des fréquences différentes. Nous proposons aussi une étude de l'effet conjoint de l'hypothèse d'EIM et de la rationalité limitée des stratégies d'investissement.

Dans la section suivante de ce chapitre, nous introduisons le cadre général du modèle. Dans la section 2.3 nous décrivons les équilibres dans un modèle à une échelle avec les investisseurs à rationalité bornée, réexaminons les conclusions de Anufriev et al. [2006] et préparons le terrain pour l'étude du cas multi-échelle. Dans la section 2.4, nous dérivons l'équilibre dans le cadre d'EIM et établissons les caractéristiques de la dynamique des rendements. Dans la section 2.5, nous illustrons nos résultats avec des exemples de simulation. En conclusion, les principaux résultats sont résumés et quelques extensions possibles du modèle sont décrites.

2.2 Modèle pour la dynamique jointe du prix boursier et de la richesse avec multiples horizons d'investissement

Dans cette section, nous formulons le modèle et ensuite discutons de ses différentes spécifications possibles et de ses hypothèses. La cadre général correspond aux travaux de Chiarella and He [2001] et Anufriev et al. [2006], auquel nous rajoutons l'hypothèse d'EIM et certaines contraintes sur le comportement des investisseurs, décrites plus tard. Quand c'est possible, nous gardons les mêmes notations que dans Anufriev et al. [2006], afin de permettre la comparaison facile des résultats.

Considérons deux actifs sur un marché où N agents opèrent à des dates distinctes. L'actif sans risque paie le taux R_f constant sur chaque période et l'actif risqué verse un dividende d_t au début de chaque période. Le prix de l'actif sans risque est normalisé à 1 et son offre est absolument élastique. La quantité de l'actif risqué est constante et normalisée à 1, tandis que son prix est déterminé par l'équilibre du marché par un mécanisme Walrasien. L'hypothèse de Walras signifie que tous les agents déterminent leur demande d'actif risqué en prenant le prix de l'actif risqué p_t en paramètre.

La demande de l'actif risqué est formulée en termes de parts de la richesse des agents, de sorte que $x_{t,i}$ représente la part de la richesse que l'investisseur i avec la richesse $W_{t,i}$ souhait d'investir dans l'actif risqué, ou $\frac{W_{t,i}\,x_{t,i}}{P_t}$ d'unités de l'actif. La condition d'équilibre du marché impose :

$$\sum_{i=1}^{N} x_{t,i} W_{t,i} = 1.$$

La richesse de chaque investisseur évolue selon l'équation ci-dessous :

$$W_{t,i} = (1 - x_{t-1,i}) W_{t-1,i}(1 + R_f) + \frac{x_{t-1,i} W_{t-1,i}}{P_{t-1}}(P_t + D_t) =$$
$$(1 - x_{t-1,i}) W_{t-1,i}(1 + R_f) + x_{t-1,i} W_{t-1,i}(1 + R_t + \varepsilon_t),$$

où D_t est un paiement de dividendes, dont le rapport au prix est assumé être une variable aléatoire *iid* notée ε_t, et R_t est le rendement sur l'actif risqué. Nous définissons le rendement total par

$$Y_t = \frac{P_t + D_t}{P_{t-1}}.$$

Suivant Anufriev et al. [2006], on réécrit le modèle en termes redimensionnés, ce qui permet d'éliminer l'expansion exogène du système liée à l'actif sans risque du modèle :

$$w_{t,i} = \frac{W_{t,i}}{(1 + R_f)^t},\ p_t = \frac{P_t}{(1 + R_f)^t},\ e_t = \frac{\varepsilon_t}{1 + R_f},\ y_t = \frac{Y_t}{1 + R_f}.$$

Par conséquent, le rendement de l'actif risqué est redéfini par :

$$r_t = \frac{p_t}{p_{t-1}} - 1 = \frac{1 + R_t}{1 + R_f} - 1. \tag{2.1}$$

En ces termes, la dynamique du système se simplifie :

$$\begin{aligned} p_t =& \sum_i x_{t,i} w_{t,i}, \\ w_{t,i} =& w_{t-1,i} \left[1 + x_{t-1,i} \left(r_t + e_t\right)\right]. \end{aligned} \tag{2.2}$$

Proposition 2.2.1. *La dynamique de prix, résolvant le système dynamique (2.2), vérifie :*

$$p_t = p_{t-1} \frac{\sum_i w_{t-1,i} \left(x_{t,i} - x_{t-1,i} x_{t,i}\right) + e_t \sum_i x_{t,i} x_{t-1,i} w_{t-1,i}}{\sum_i w_{t-1,i} \left(x_{t-1,i} - x_{t,i} x_{t-1,i}\right)},$$

La proposition 2.2.1 décrit la dynamique des prix d'équilibre Walrasien sur le marché à deux actifs. Il est facile de voir que le rendement d'équilibre doit satisfaire :

$$r_t = \frac{\sum_i w_{t-1,i} \left(x_{t,i} - x_{t-1,i} + x_{t,i} x_{t-1,i} e_t\right)}{\sum_i w_{t-1,i} \left(x_{t-1,i} - x_{t,i} x_{t-1,i}\right)}, \tag{2.3}$$

L'équation (2.3) spécifie explicitement le rendement r_t conditionnellement à l'information de la période $t-1$ si et seulement si sont imposées des hypothèses supplémentaires : à la fois la demande $x_{t,i}$ et le taux de dividendes e_t doivent être indépendants du niveau actuel des prix p_t.

La plus simple hypothèse relative aux dividendes, qu'on peut suggérer pour rendre le modèle (2.2) analytiquement soluble, est de supposer que le taux de dividendes est un processus stochastique *iid* à valeurs non-négatives. Suivant Chiarella and He [2001] et Anufriev et al. [2006], on maintient cette hypothèse, tout en restant conscients des contraintes qu'elle impose.

Jusqu'à présent, rien n'a été dit à propos de la manière dont les agents déterminent les proportions souhaitées de l'investissement dans l'actif risqué. L'hypothèse de EIM, étudiée dans cette thèse, implique que certains investisseurs ne participent pas au marché à toutes les périodes de temps mais peuvent rester passifs. Au cours de la période où un investisseur est hors du marché, sa part de l'investissement dans l'actif risqué n'est plus un résultat de ses décisions, mais une conséquence des mouvements de prix et de sa richesse, indépendants de sa volonté. La proposition suivante décrit l'évolution de la part de richesse investie dans un actif risqué pour un investisseur passif.

Proposition 2.2.2. *Soit $x_{t,i}^{-k}$ la part d'investissement dans l'actif risqué de l'investisseur i, qui a participé aux échanges il y a k périodes, $k = 1, \ldots, h$ avec h son horizon d'investissement. La part d'investissement vérifie la relation récurrente suivante :*

$$x_{t,i}^{-k} = \frac{x_{t-1,i}^{-k+1}(1 + r_t)}{1 + x_{t-1,i}^{-k+1}(r_t + e_t)} \tag{2.4}$$

Aux périodes où l'investisseur i réajuste son portefeuille, sa demande pour l'actif risqué $x_{t,i}^0$ est déterminée par sa fonction d'investissement. Dans ce document, nous supposons que les fonctions d'investissement sont données sous forme de la dépendance de la part de la richesse à investir dans l'actif risqué, par rapport aux croyances concernant les bénéfices futurs. Nous supposons également que les fonctions d'investissement sont déterministes et ne changent pas au fil du temps pour le même investisseur [1]. Les croyances sont fondées sur les observations antérieures des prix et des dividendes, sans autres informations privées qui pourraient être utilisées pour prévoir les rendements futurs. En outre, chaque fonction d'investissement est indépendante de la richesse actuelle, ce qui découle de l'hypothèse CRRA. Donc, la fonction d'investissement de l'investisseur i est :

$$x_{t,i}^0 = f_i(r_{t-1}, \ldots, r_{t-L_i}, e_{t-1}, \ldots, e_{t-L_i}) \tag{2.5}$$

où L_i est le nombre maximal d'observations historiques, utilisés par l'agent i pour la prise de décisions.

1. Notons que cela n'exclut pas de fonctions, correspondant à des stratégies d'investissement qui évoluent en fonction de règles prédéfinies.

En particulier, nous allons étudier le cas qui correspond à la maximisation de la fonction d'utilité de type "moyenne-variance" [2]. Supposons que les investisseurs, opérant éventuellement sur différentes échelles de temps, optimisent l'espérance d'utilité de la forme :

$$\max_{x^0_{t,i}} \left\{ \mathrm{E}_{t-1,i}(W_{t+h,i}) - \frac{\gamma_i}{2W_{t,i}} \mathrm{Var}_{t-1,i}(W_{t+h}) \right\} \tag{2.6}$$

avec les opérateurs $\mathrm{E}_{t-1,i}(\cdot)$ et $\mathrm{Var}_{t-1,i}(\cdot)$ représentant les croyances de l'agent i concernant la moyenne et la variance de richesse, étant donné l'information au temps $t-1$. Cette information comprend des prix de l'actif risqué et les dividendes des périodes $t-1$ et antérieures. Le coefficient γ_i est une constante positive qui mesure l'aversion au risque de l'investisseur i. L'horizon temporel de la prise de décision, noté H, correspond à la période de temps où l'investisseur i ne réajuste pas son portefeuille. Le nombre d'unités d'actif risqué en possession de l'investisseur reste constant en $[t; t+h]$, tandis que la part d'investissement dans l'actif risqué évolue. Nous supposons que les dividendes, versés par l'actif risqué au cours de cette période, sont accumulés sur le compte bancaire, qui paie le taux sans risque.

Proposition 2.2.3. *La solution $x^{0*}_{t,i}$ du problème de maximization (2.6) est approximativement donné par :*

$$x^{0*}_t \approx \frac{\mathrm{E}_{\,t-1,i}\left[\sum_{k=1}^{h}(e_{t+k}+r_{t+k})\right]}{\gamma_i \mathrm{Var}_{\,t-1,i}\left[\sum_{k=1}^{h}(e_{t+k}+r_{t+k})\right]} \tag{2.7}$$

Chiarella and He [2001] montrent que l'expression similaire à (2.7) avec $h = 1$, que nous avons présentée comme la solution du problème d'optimisation "moyenne - variance" (2.6), apparaît aussi comme une solution approximative au problème de la maximisation avec l'utilité de richesse qui est une fonction puissance. Cette approximation, cependant, utilise la discrétisation d'un processus en temps continu avec des incréments Gaussiens et peut être loin de la vraie solution pour les unités de temps non-infime. Donc, nous préférons travailler avec la maximisation de type "moyenne - variance" directement. Alternativement, une fonction d'investissement de la forme (2.7) peut être définie sur par une règle *a priori* car elle décrit le comportement d'un investisseur "moyenne-variance" avec l'aversion relatif au risque constant.

Il est à noter que ci le processus de rendement est *iid*, nous avons $\mathrm{E}_{\,t-1}[y_{t,t+h}] = h\,\mathrm{E}_{\,t-1}[r_{t+1}+e_{t+1}]$ et $\mathrm{Var}_{\,t-1}[y_{t,t+h}] = h\,\mathrm{Var}_{\,t-1}[r_{t+1}+e_{t+1}]$. Cela garantit que si, en outre, l'aversion au risque est homogène pour les investisseurs à toutes les échelles ($\gamma_i = \gamma$), la demande de l'actif risqué ne dépend pas de l'horizon d'investissement. Nous maintenons l'hypothèse d'aversion au risque homogène tout au long de ce document.

2. La plupart de nos résultats sont également valables dans le cas de la fonction d'investissement générale donnée par (2.5), n'appartenant pas nécessairement au type "moyenne - variance".

Dans l'équation (2.7) la portion de la richesse à investir dans l'actif risqué dépend exclusivement des croyances des agents au sujet des rendements futurs. Dans les modèles à agents hétérogènes, ces croyances sont fondées sur les prix historiques de l'actif risqué. Le problème est avec le prix p_t de la période t. Dans un marché Walrasien, le prix courant est pris comme paramètre pour déterminer la demande pour l'actif risqué. Il n'y a donc aucune raison formelle d'exclure ce prix de l'ensemble de l'information des agents.

Chiarella and He [2001] affirment que, bien que le prix de la période courante peut être inclut de l'équation de croyances, il n'y est pas indispensable, mais il induit des pertes dans la "tractabilité" du modèle. Si p_t est simplement utilisé par les investisseurs à côté d'autres prix historiques pour calculer certaines quantités, telles que la moyenne et la variance, il peut facilement être exclut de l'ensemble d'information. Il s'agit d'une hypothèse appropriée pour un système dynamique, dans lequel le dividende est la seule source d'innovation incertaine. Si le système était par ailleurs innové par des signaux d'information, la situation deviendrait plus compliquée, car le prix de la période courante pourrait être révélateur de l'information, perçue par d'autres agents, étant un complément à l'information privée.

Dans notre modèle sans signaux nous excluons p_t de l'ensemble d'information utilisée pour former les croyances, afin d'éviter une complexité inutile. Néanmoins, la demande de la période t dépend de p_t pour le cas multi-échelle. Supposons que l'investisseur i a participé aux échanges à la période $t-k$. A cette date, la part de la richesse investie dans l'actif risqué $x^0_{t-k,i}$ était déterminée selon (2.5). Ensuite, il découle de (2.4) que sa part d'investissement courante $x^{-k}_{t,i}$ dépend des rendements et des dividendes historiques jusqu'au lag $L_i + k - 1$, mais aussi du rendement courant, qui est inconnu avant les échanges de la date t. Donc l'équation (2.3) ne spécifie pas explicitement la dynamique du rendement.

Dans la section suivante, nous étudions la dynamique des prix et de la richesse dans un modèle à un horizon d'investissement, qui est un cas particulier du modèle, introduit dans la section précédente. Par la suite, nous nous référons à ce cas comme à un modèle de référence. Nous prolongeons l'analyse de Anufriev et al. [2006] à plusieurs égards, qui sont aussi importants dans le cas de multiples horizons, étudié plus tard.

2.3 Equilibres dans un modèle à un horizon avec rationalité bornée

Comme nous l'avons mentionné auparavant, dans le cas d'un seul horizon, l'équation (2.3) décrit complètement et explicitement la dynamique du rendement de l'actif risqué, sous la condition d'équilibre du marché. En précisant la fonction de demande, on peut déterminer le prix et le rendement d'équilibre. La dynamique d'équilibre a été étudié dans

Anufriev et al. [2006], qui remplacent le taux de dividende par sa moyenne et travaillent avec la "squelette déterministe" du système. Dans le cas déterministe, le rendement est constant : $r_t = r$. Les auteurs montrent que deux types d'équilibres sont possibles : soit un seul agent survit [3], ou de nombreux agents survivent, mais dans les deux cas, en équilibre les parts d'investissement dans l'actif risqué et le rendement sont déterminés de même manière. Ils doivent satisfaire la relation, qui est facilement obtenue à partir de (2.2) pour les cas d'un seul agent, lorsque nous posons $x_t = x_{t-1}$ pour tout t. Cette relation, appelée la "ligne d'équilibre du marché" (Equilibrium Market Line, EML) est ainsi :

$$x = \frac{r}{r + \bar{e}} \tag{2.8}$$

où $\bar{e}$ est le taux moyen de dividendes.

Les fonctions de demande des investisseurs dépendent d'une seule variable et sont de la forme :

$$x = f(r_{t-1}, \ldots, r_{t-L}) = f(r, \ldots, r) = \tilde{f}(r) \tag{2.9}$$

Les points d'équilibre sont déterminés comme les intersections de la courbe de demande $\tilde{f}(r)$ et la EML. Il est montré que, si plusieurs agents survivent, leurs fonctions de demande doivent toutes croiser la LME au même point d'équilibre. Les conditions de stabilité, dépendant des propriétés des dérivés de $f_i(\cdot)$ par rapport aux rendements passés, sont établies. Nous renvoyons le lecteur à l'article de Anufriev et al. [2006] pour plus de détails. Dans notre approche, la principale différence est que nous sommes intéressés par les propriétés stochastiques du processus de rendements. En particulier, nous établissons analytiquement, sous quelles conditions la dynamique des rendements est "simple"(*iid*) et quand elle possède des propriétés "intéressantes" (hétéroscédasticité conditionnelle et / ou corrélations temporelles).

Dans le cas de plusieurs agents avec des fonctions d'investissement hétérogènes, Anufriev et al. [2006] déterminent, quelle forme de la fonction de demande "domine" les autres. Par exemple, si un chasseur de tendances (investisseur qui extrapole fortement les rendements passés) rencontre un investisseur fondamentaliste, dont la fonction de demande est indépendante de l'historique des prix, on peut prédire lequel d'entre eux survit, en fonction de la forme respective de leurs fonctions d'investissement. Sous certaines conditions, le chasseur de tendances surpasse le fondamentaliste et survit. Une caractéristique frappante de ce modèle est que les équilibres sont possibles pour presque toutes, même complètement irraisonables, fonctions de demande et peuvent même être stable.

Le problème ici est dans la rationalité limitée. Plus précisément, il est important de savoir, dans quelle mesure la rationalité est limitée. Dans Anufriev et al. [2006] et Chiarella and He [2001], les fonctions d'investissement sont données *a priori* et bien qu'elles dépendent formellement des croyances des agents à propos de la moyenne et de

3. *i.e.* sa part dans la richesse totale ne diminue pas à zéro en un temps infini

la variance des rendements futurs, il n'y a aucune restriction quant à la façon dont ces croyances devraient être liées aux vrai quantités.

La rationalité limitée signifie que les agents peuvent ne pas connaître le vrai modèle. Mais en équilibre, lorsque le rendement de l'actif risqué est censé être constant, il est difficile d'admettre que les croyances n'ont rien à voir avec la réalité. D'ailleurs, la stabilité de tels équilibres ne fait guère de sens du point de vue économique, puisque les agents pourraient être incités à changer leurs stratégies, si ils y étaient autorisés.

Dans Brock and Hommes [1998], les agents sont autorisés à choisir des stratégies en fonction des profits qu'elles auraient dégagé dans le passé. Les agents peuvent donc être appelés procéduralement rationnels, puisse qu'ils tentent de choisir rationnellement leur stratégies en fonction de certains critères. Dans notre cas, une définition plus exacte de la rationalité procédurale peut être utile pour étudier le modèle analytique. Nous allons restreindre la classe des fonctions d'investissement admissibles, en réduisant considérablement les possibilités pour la non-rationalité des agents économiques, sans nécessairement imposer les anticipations rationnelles.

Les fonctions d'investissement déterminent essentiellement la façon dont les croyances des agents au sujet des rendements futurs sont formées, c'est à dire qu'elles sont des descriptions concises des résultats de la procédure de prise de croyances. La rationalité d'une telle procédure peut être testé dans certains cas de référence, où l'issue de la procédure est censée correspondre au comportement rationnel. Dans le cas d'un investisseur "moyenne-variance", nous exigeons que pour les rendements *iid* les croyances à propos de la moyenne et de la variance du processus soient non-biaisées. Ceci est formalisé dans la définition suivante.

Definition 2.3.1. *Une fonction d'investissement de la forme*

$$x^0_{t,i} = \tilde{f}_i \left(\mathrm{E}_{t-1,i}[y_{t,t+h}], \mathrm{Var}_{t-1,i}[y_{t,t+h}] \right) \tag{2.10}$$

est appelée procéduralement rationnelle si les croyances $\mathrm{E}_{t-1,i}(y_{t-1,t+h})$ *et* $\mathrm{Var}_{t-1,i}(y_{t,t+h})$*, concernant la moyenne et la variance du rendement futur sont des estimations non-biaisées avec une erreur finie de ses mêmes quantités pour le vrai processus* $y_{t,t+1}$*, si ce dernier est iid.*

Cette définition est une adaptation de la rationalité procédurale de Simon [Simon, 1976, p.131] à notre contexte. Elle affirme que, si les observations précédentes ne contiennent pas de dynamiques non-triviales, les croyances des investisseurs au sujet de la moyenne et de la variance ne devraient pas avoir d'erreurs systématiques. Notez que dans aucun cas nous déclarons que les rendements doivent effectivement suivre un processus *iid*, nous décrivons seulement le comportement de la fonction d'investissement dans ce cas hypothétique afin d'imposer certaines contraintes sur la rationalité de la procédure de prise de décisions, utilisée par les investisseurs.

Notre définition ne contredit pas à la notion de la rationalité limitée, mais elle exige un certain degré modéré de cohérence dans les croyances des investisseurs. Les investisseurs peuvent être effectivement chasseurs de tendances ou contrariants. Considérons, par exemple, les spécifications suivantes pour les croyances à propos de la moyenne des rendements futurs :

$$\begin{aligned} \mathrm{E}_{t,i}(y_{t+1}) &= c_i + \frac{d_i}{l}\sum_{k=1}^{l} y_{t-k}, \qquad (A) \\ \mathrm{E}_{t,i}(y_{t+1}) &= \frac{1-d_i}{L}\sum_{k=1}^{L} y_{t-k} + \frac{d_i}{l}\sum_{k=1}^{l} y_{t-k}. \quad (B) \end{aligned} \tag{2.11}$$

Une fonction de type (2.11A) est utilisée dans Chiarella and He [2001] pour représenter le comportement des agents hétérogènes. Ici c_i est un paramètre comportemental, qui spécifie comment l'investisseur i extrapole la performance de l'actif risqué de l périodes dans le passé. Si $d_i = 0$, l'investisseur est fondamentaliste, si $d_i > 0$ il est chasseur de tendances, sinon il est contrariant. Il est facile de montrer que cette spécification ne correspond pas à notre définition de rationalité procédurale, sauf si $c_i = 0$ et $d_i = 1$ simultanément. La fonction (2.11B) elle aussi extrapole les rendements passés via le paramètre d. Si $l < L$, d_i positif correspond à la chasse de tendances. Mais cette fonction vérifie la condition de la rationalité procédurale : dans le cas *iid* l'espérance de la différence entre la moyenne "court terme" et "long terme" et nulle.

Supposons maintenant que les agents ont des fonctions d'investissement de type "moyenne-variance", décrites par (2.7), satisfaisant la définition 2.3.1. Ayant ainsi restreint l'ensemble des fonctions d'investissement admissibles, nous nous tournons vers l'étude de la dynamique des prix dans le modèle de référence. Dans le théorème suivant nous établissons les conditions qui doivent être vérifiées par la fonction d'investissement pour assurer que la dynamique des rendements est *iid*, ce qui peut être interprété comme une trajectoire régulière de croissance de prix. nous montrons que supposer les anticipations rationnelles des agents est équivalent à supposer la dynamique *iid* des rendements.

Theorem 2.3.2. *Dans le modèle de référence avec des agents homogènes procéduralement rationnels le processus de rendements peut être iid avec moyenne et variance finies si et seulement si les investisseurs ont des anticipations rationnelles. Dans ce cas, la moyenne et la variance du processus de rendement sont uniquement définis par le taux de dividendes moyen et l'aversion au risque des investisseurs.*

Une conséquence importante du théorème de 2.3.2 est que dans le modèle de référence avec des agents homogènes procéduralement rationnels, à moins que les investisseurs ont des anticipations rationnelles, le rendement de l'actif risqué n'a jamais une dynamique simple *iid*.

Il est à noter que l'équation

$$\mathrm{Var}_{t-1}(r_t) = \sigma_e^2 \frac{x_t^2}{(1-x_t)^2}, \tag{2.12}$$

obtenue dans la preuve du théorème, décrit la dynamique de la volatilité conditionnelle dans le modèle. Il résulte de (2.12) que pour $0 < x_t < 1$, la variance conditionnelle augmente toujours avec x_t. Si la fonction d'investissement dépend positivement de la moyenne historique des rendements et négativement de la variance historique, il est de même pour la volatilité conditionnelle des rendements. En même temps, la volatilité a la même mémoire que le carré de la part d'investissement dans l'actif risqué, qui est déterminé par les croyances des investisseurs. Si ces dernières sont ajustées lentement, l'ajustement de la volatilité est également lent.

Considérons maintenant la dynamique de rendements dans un cas plus général d'agents homogènes ayant des fonctions d'investissement arbitraires (pas nécessairement de type moyenne-variance) de la forme

$$x_t = f(r_{t-1}, \ldots, r_{t-L}, e_{t-1}, \ldots, e_{t-L}).$$

Le processus stochastique de rendements est non-linéaire. Nous tenons à étudier les propriétés de sa linéarisation de Taylor de premier ordre au voisinage de l'espérance de rendement, $r_{t-k} = \bar{r}$ pour tout k. Notons $\widetilde{e}_t = e_t - \bar{e}$ et $\widetilde{r}_t = r_t - \bar{r}$ les déviations du taux de dividendes et des rendements de leur valeurs moyennes. Notons aussi f'_k les dérivés de premier ordre de $f(\cdot)$ par rapport au r_{t-k} pour $k = 1, \ldots, L$. La forme du processus de rendement est donnée par le théorème suivant.

Theorem 2.3.3. *Dans le modèle de référence avec des agents homogènes, si le processus de rendement est stationnaire de deuxième ordre, il vérifie :*

$$\begin{aligned} r_t &= \frac{\bar{x}}{1-\bar{x}}\bar{e} + \widetilde{r}_t \\ \widetilde{r}_t &= \sum_{k=1}^{L+1} a_k \widetilde{r}_{t-k} + v_t \widetilde{e}_t \\ v_t &= \frac{\bar{x}}{1-\bar{x}} + \sum_{k=1}^{L} b_k \widetilde{r}_{t-k} \end{aligned} \tag{2.13}$$

avec :

$$\begin{aligned} a_1 =& \frac{f_1'[1-\bar{x}\,(1-\bar{e}\,)]}{\bar{x}\,(1-\bar{x}\,)^2} \\ a_k =& \frac{f_k'[1-\bar{x}\,(1-\bar{e}\,)]+f_{k-1}'(\bar{x}\,-1)}{\bar{x}\,(1-\bar{x}\,)^2},\ k\in\{2,\ldots,L\} \\ a_{L+1} =& \frac{f_L'}{\bar{x}\,(1-\bar{x}\,)} \\ b_k =& \frac{f_k'}{(1-\bar{x}\,)^2},\ k\in\{1,\ldots,L\} \\ \bar{x}\ =& \end{aligned} \tag{2.14}$$

Les équations (2.13) peuvent être écrites sous forme équivalente :

$$\begin{aligned} \widetilde{r}_t &= \sum_{k=1}^{L+1} a_k\,\widetilde{r}_{t-k} + \sigma_e\left(u_t + \frac{\bar{x}}{1-\bar{x}}\right)\varepsilon_t, \\ u_t^2 &= \sum_{k=1}^{L} b_k^2\,\widetilde{r}_{t-k}^2 + 2\sum_{\substack{i,j\in\{1,\ldots,L\}\\ i\neq j}} b_i b_j \widetilde{r}_{t-i}\widetilde{r}_{t-j} \end{aligned} \tag{2.15}$$

avec ε_t un bruit blanc. Cela souligne la nature ARCH du processus stochastique de rendements. Notons que la moyenne de ce processus est décrite par une expression, équivalente à la définition de l'équilibre sur la EML dans Anufriev et al. [2006].

Maintenant nous pouvons nous tourner au cas des agents hétérogènes, c'est à dire le cas où les $x_{t,i}$ sont déterminés d'une manière différente par chaque investisseur. Le théorème 2.3.4 montre que la dynamique *iid* n'apparait pas d'une manière générique si les investisseurs sont hétérogènes.

Theorem 2.3.4. *Dans le modèle de référence avec des agents hétérogènes le processus de rendements peut être iid avec moyenne et variance finies que si la part globale de la richesse investie dans l'actif risqué reste constante. Dans ce cas, la moyenne et la variance du processus de rendements sont proportionnelles à la moyenne et la variance du taux de dividendes.*

Fondamentalement ce théorème dit que si la part globale de l'investissement dans l'actif risqué est soumise à des chocs ou à des fluctuations stochastiques, la dynamique de rendements est presque sûrement pas triviale et contient des dépendances temporelles. La situation où la fonction de l'investissement global est constante et les rendements sont *iid*, ne peut survenir que lorsque la dépendance de chacune des fonctions d'investissement par rapport aux rendements passés n'est pas caractérisée par des schémas de décisions communes à plusieurs investisseurs. Plus précisément, les écarts individuels $\nu_{t,i} = x_{t,i} - \bar{x}$ de la part globale d'investissement dans l'actif risqué $\bar{x}$, pondérés par les portions de la

richesse des agents $\xi_{t,i}$, sont éliminés par agrégation avec probabilité 1 :

$$P(\sum_{i=1}^{N} \xi_{t,i}\nu_{t,i} = 0) = 1$$

pour tout t. Pour que cette condition soit remplie, une certaine forme de la loi des grands nombres doit être satisfaites et, en outre, l'espérance de $\nu_{t,i}$, conditionnellement aux rendements passés, doit être constante. Ceci est improbable dans une situation où tous les investisseurs fondent leurs attentes sur le même vecteur des rendements passés et ce vecteur n'est pas constant.

2.4 Equilibres avec multiples horizons d'investissement

Dans la section précédente, nous avons examiné le cas des investisseurs qui ont le même horizon d'investissement, mais les fonctions d'investissement qui peuvent être différentes. Maintenant, nous revenons à l'hypothèse EIM et étudions une autre source d'hétérogénéité, liée à des horizons d'investissement. Supposons maintenant qu'il existe H échelles d'investissement avec des périodes de réajustement de portefeuille $h = 1, \ldots, H$ unités de temps, de sorte que chaque agent a un horizon d'investissement caractéristique qui ne change pas. Supposons que les investisseurs au sein de chaque échelle d'investissement sont homogènes, *i.e.* ont la même spécification de la fonction d'investissement. Enfin, supposons que, à chaque date la richesse des investisseurs, ayant la même échelle d'investissement, est répartie de sorte qu'une partie constante de cette richesse, égale à $1/H$ appartient à des investisseurs, réajustant leurs portefeuilles à la date actuelle.

Cette dernière hypothèse ne signifie pas nécessairement que la richesse peut être redistribuée entre les différents groupes d'investisseurs dans une période donnée. Cela signifie plutôt qu'il existe un grand nombre d'investisseurs, entrant et sortant du marché, qui ont des dates aléatoires d'intervention sur le marché, mais les fréquences de ses interventions sont fixes. Ainsi, la composition de chaque cohorte d'investisseurs peut changer, mais la part moyenne de la richesse, investie dans l'actif risqué, reste constante.

Sous ces hypothèses simplificatrices, on peut agréger l'ensemble des investisseurs, ayant la même échelle H, et les remplacer par un agent représentatif, dont la part de la richesse, investie dans l'actif risqué, est :

$$x_{t,h} = \frac{1}{h}\sum_{k=0}^{h-1} x_t^{-k} \tag{2.16}$$

Les équations (2.2), décrivant la dynamique du système, sont toujours vrai, mais maintenant l'indice i correspond à l'échelle d'investissement et la richesse $w_{t,i}$ est la richesse totale d'une catégorie d'investisseurs, dont l'horizon est le même. Dans la section 2.2 nous avons démontré l'équation (2.4) qui décrit l'évolution de la part de la richesse,

investie dans l'actif risqué, pour des investisseurs passifs. Alors, le système d'équations, décrivant la dynamique du rendement risqué, est :

$$\begin{aligned} r_t =& \frac{\sum_{h=1}^{H} w_{t-1,h}\left(x_{t,h} - x_{t-1,h}\right) + e_t \sum_{h=1}^{H} x_{t,h} x_{t-1,h} w_{t-1,h}}{\sum_{h=1}^{H} w_{t-1,h} x_{t-1,h}\left(1 - x_{t,h}\right)} \\ x_{t,h} =& \frac{1}{h}\sum_{k=0}^{h-1} x_{t,h}^{-k} \\ x_{t,h}^{-k} =& \frac{x_{t-1,h}^{-k+1}(1+r_t)}{1 + x_{t-1,h}^{-k+1}(r_t + e_t)} \end{aligned} \tag{2.17}$$

Comme indiqué plus haut, une caractéristique importante de l'équation (2.17) est qu'elle décrit la dynamique de rendements d'une manière implicite, car la part de l'investissement pour tous, sauf les investisseurs à la plus courte des échelles, dépend du rendement courant. La relation entre le prix et le taux de dividendes devient non-linéaire. Pour l'équation générale de la dynamique des prix, nous démontrons le théorème d'existence d'équilibre :

Theorem 2.4.1. *Quel que soit le nombre d'horizons H, il existe toujours au moins un prix positif, pour lequel le rendement r_t satisfait (2.17).*

Il est important de préciser les conditions, sous lesquelles la dynamique multi-échelle ne dégénère pas, c'est à dire les parts de la richesse, détenues par les agents, investissant à chaque échelle, ne tendent pas vers zéro quand le temps tend vers l'infini. Plus précisément, notons $\xi_{t,h}$ la part de la richesse globale, appartenant aux investisseurs du type h.

Definition 2.4.2. *La dynamique multi-échelle, décrite par l'équation (2.17), est non-dégénérée, si pour tout horizon d'investissement h tel que $\xi_{0,h} > 0$ nous avons :*

$$P(\xi_{t,h} = 0) = 0,$$

quand t tend vers infini.

Dans le théorème suivant nous établissons les conditions nécessaires et suffisantes pour que la dynamique multi-échelle soit non-dégénérée. Notons $g_{t,h}$ le taux de croissance de la richesse des investisseurs de type h à la période t :

$$g_{t,h} = \frac{w_{t,h}}{w_{t-1,h}} = 1 + x_{t-1,h}(r_t + e_t).$$

Supposons que le processus stochastique $\ln(g_{t,h})$ est stationnaire de deuxième ordre. Supposons aussi qu'il vérifie les conditions suivantes sur sa mémoire :

$$N^{-1}||\{\mathrm{Cov}\left(\ln(g_{t+i,h}), \ln(g_{t+j,h})\right)\}_{\{i=1,\ldots,N, j=1,\ldots,N\}}||_2 \leq C \tag{2.18}$$

pour tout N positif et un C fini. Cette condition technique implique que $\ln(g_{t,h})$ est un processus stochastique à densité spectrale bornée et que le taux de croissance moyen de la richesse converge presque sûrement vers son espérance quand le temps tend vers l'infini. Ce résultat est démontré dans Ninness [2000].

Theorem 2.4.3. *La dynamique multi-échelle, décrite par l'équation (2.17), est non-dégénérée si et seulement si pour tout h :*

$$\mathrm{E}\left[\ln(g_{t,i})\right] = \mathrm{E}\left[\ln(g_{t,j})\right],\ \forall\ i,j \in \{1,\ldots,H\}.$$

Pour interpréter le théorème, remarquez que le taux logarithmique de croissance de la richesse est approximativement égal au produit du rendement total de l'actif risqué avec la part de la richesse, investie dans l'actif risqué à la période précédente. Ainsi, pour un modèle non-dégénéré, les investisseurs doivent avoir soit la même part moyenne d'investissement dans l'actif risqué, soit les parts plus basses doivent être compensées par la corrélation positive entre la part d'investissement et les rendements futurs. Un cas particulier d'un système non-dégénéré est le cas des parts de richesse non-prédictives et égales en loi :

$$\begin{aligned} x_{t,i} &\overset{L}{=} x_{t,j}\ \forall\ i,j \in \{1,\ldots,H\},\ \forall\ t \\ \mathrm{Cov}\left(x_{t-i,h}, r_t\right) &= 0,\ \forall\, h \in \{1,\ldots,H\},\ \forall\, i,t. \end{aligned} \tag{2.19}$$

Il est à noter que dans un modèle à horizons multiples l'existence des autocorrélations dans les rendements implique la corrélation entre les parts d'investissement dans l'actif risqué et les rendements futurs. De plus, cete corrélation est plus élevée pour les investisseurs à longs horizons, car à chaque période il y a plus d'investisseurs passifs, dont les parts d'investissement dépendent des rendements passés, même si les fonctions élémentaires d'investissement sont constantes. Ainsi la condition (2.19) est liée à l'absence de corrélations temporelles dans les rendements.

Par analogie avec le cas d'un seul horizon, nous analysons la dynamique d'équilibre du système (2.17). Etudions d'abord la dynamique moyenne, en supposant $e_t = \bar{e}$. Le théorème suivant montre qu'il existe une trajectoire d'équilibre $r_t = \bar{r}$ qui résout l'analogue déterministe de (2.17).

Theorem 2.4.4. *Le système dynamique (2.17) avec $e_t = \bar{e}$ a une solution unique d'équilibre avec rendement constant :*

$$\begin{aligned} \bar{r} &= \frac{\bar{x}}{1-\bar{x}}\bar{e}\,, \\ f_h(\bar{r}\,,\ldots,\bar{r}\,) &= \bar{x} \end{aligned} \tag{2.20}$$

Nous pouvons maintenant étudier les propriétés du processus stochastique des rendements risqué et de comparer les résultats avec ceux obtenus pour le cas d'un horizon.

Comme précédemment, nous allons procéder par la linéarisation du système dynamique. Définissons la fonction suivante : $F : \Re^{t-1} \times \Re^{t-1} \times \Re \times \Re \to \Re$:

$$\begin{gathered} F(r_1, \ldots, r_{t-1}, e_1, \ldots, e_{t-1}, r_t, e_t) = \\ \frac{\sum_{h=1}^{H} w_{t-1,h} \left[x_{t,h} - x_{t-1,h} + e_t x_{t,h} x_{t-1,h} \right]}{\sum_{h=1}^{H} w_{t-1,h} x_{t-1,h} \left(1 - x_{t,h} \right)} - r_t \end{gathered} \tag{2.21}$$

avec $x_{t,h}$ donné par (2.17). Le théorème suivant décrit la dynamique d'équilibre au voisinage de $\bar{r}$.

Theorem 2.4.5. *Dans le modèle avec agents rationnels homogènes et échelles multiples d'investissement, le processus de rendements est approximativement décrit par :*

$$\begin{aligned} r_t &= \bar{r} + \hat{r}_t + \bar{V} \widetilde{e}_t, \\ \hat{r}_t &= \sum_{k=1}^{H-1} A_k \, \hat{r}_{t-k} + V_t \widetilde{e}_t, \\ V_t &= \sum_{k=1}^{H-1} B_k \hat{r}_{t-k} \end{aligned} \tag{2.22}$$

où :

$$\begin{aligned} A_k &= \frac{a_k - b_k}{(1-c)(1+\bar{r})^k}, \\ B_k &= \frac{\bar{x}\,(1-2\bar{x})(c\,b_k - a_k)}{(1-\bar{x})(1+\bar{r})^{k+1}(1-c)}, \\ \bar{V} &= \frac{\bar{x}}{1-\bar{x}}, \\ a_k &= \sum_{h=k+2}^{H} \frac{h-k-1}{h} \xi_{0,h}\,, \\ b_k &= \sum_{h=k+1}^{H} \frac{h-k}{h} \xi_{0,h}\,, \\ c &= \sum_{h=1}^{H} \frac{h-1}{h} \xi_{0,h}\,. \end{aligned}$$

Le résultat du théorème 2.4.5 montre que la dynamique de rendements dans le modèle à horizons multiples avec des investisseurs rationnels est très proche de celle dans le cas d'anticipations rationnelles, la seule différence étant le terme $\hat{r}_t$. Il représente la déviation par rapport à la trajectoire hypothétique de rendements, qui serait réalisé dans un marché à un horizon, et peut être interprété comme un terme de correction d'erreur. Il est à noter qu'il n'y a pas de constante avant la volatilité dans le terme de perturbation, ce qui signifie que le terme de correction disparaît ou explose, en fonction des valeurs des coefficients A_k, B_k et de la variance de $\widetilde{e}_t$. Nous allons étudier son comportement pour des valeurs plausibles des paramètres dans la section suivante. Notez que la dynamique dans

le cas multi-échelle est considérablement différente de celle à une échelle avec fonction d'investissement générale, décrite par le théorème 2.3.3. Dans cette dernière, les termes impliquant des corrélations temporelles et l'hétéroscédasticité ne disparaissent pas, alors que dans le cas multi-échelle avec investisseurs rationnels, leur présence est temporelle après un choc, en l'absence duquel ils disparaissent complètement de la dynamique de rendements.

Le théorème 2.4.5 correspond au cas où les fonctions de la demande des investisseurs aux moments des réajustements de portefeuille sont triviales : les parts d'investissement dans l'actif risqué sont constantes au niveau, correspondant à l'équilibre rationnel, qui coïncide avec l'équilibre à un horizon. En pratique, les décisions d'investissement peuvent dépendre des rendements historiques, donc le cadre de la rationalité procédurale serait plus adéquat pour la modélisation. On pourrait établir une représentation analytique générale de la dynamique de rendements dans ce cas. Cependant, à notre avis, une telle représentation aurait peu de valeur pratique. Plutôt, nous explorons la dynamique de rendements à l'aide de simulations, correspondant à des exemples concrètes stylisées des fonctions d'investissement. Ce sujet est abordé dans la section suivante.

2.5 Etude de simulation

Nous avons déterminé les équations de la dynamique de rendements risqués dans le cas d'un marché avec des acteurs rationnels, agissant sur un ou plusieurs horizons. Nous avons également étudié le cadre des fonctions d'investissement plus générales, qui puisse intégrer les modèles de comportement, tels que l'extrapolation de tendances et les stratégies contrariantes. Notre objectif dans cette section est d'explorer les propriétés empiriques de la série de rendements, générée par les différentes versions de notre modèle, et d'associer les propriétés du modèle avec les motifs stylisés, observés sur les données réelles du marché : les rendements contrariants, la formation de tendances et l'hétéroscédasticité conditionnelle.

De l'équation (2.22) il est clair que l'introduction d'horizons multiples modifie la façon dont le système dynamique réagit aux chocs. Ces chocs pourraient être totalement exogènes, ou de nature comportementale. Nous allons d'abord étudier le cas où, à côté du terme de perturbation, interprété comme taux de dividendes, le modèle est parfois perturbé par des chocs exogènes sur le rendement, sans rapport avec les fonctions d'investissement. Ces rendements anormaux peuvent refléter les écarts des prix par rapport à l'équilibre Walrasien à certaines périodes de temps.

Les trajectoires des rendements sont simulées pour un marché avec cinq horizons, où les rendements anormaux se produisent à des périodes aléatoires, en moyenne une fois par 50 périodes. Nous sommes intéressés par les valeurs des coefficients A_k and B_k, qui déterminent la façon dont le choc à la période t est répercuté à des dates futures. Notez

que dans le modèle à un horizon ce type de chocs n'a absolument aucune incidence sur les rendements futurs. Les coefficients mentionnés ci-dessus dépendent de a_k, b_k et c, qui caractérisent comment la richesse initiale est répartie entre les investisseurs.

Sur le Graphique 2.1 nous présentons les résultats pour le cas de la forme en $\cup$ de le distribution de richesse initiale. La première ligne des graphiques représente la répartition de richesse et les coefficients a_k, b_k, c, A_k et B_k. La deuxième contient des exemples des trajectoires du terme de la correction d'erreurs $\hat{r}_t$ pour 10 périodes après un choc, ainsi que la fonction d'autocorrélation empirique pour $\hat{r}_t$ et $|\hat{r}_t|$, estimés sur 10 000 périodes de simulation. La troisième ligne représente la composante volatilité du terme d'erreur, V_t.

Les coefficients A_k sont toujours négatifs et leur amplitude diminue avec le retard, la vitesse de diminution dépendant de la répartition de richesse initiale. Dans tous les cas, cela conduit à une anti-corrélation significative d'ordre 1. Ce résultat est spectaculaire car il montre que même sans hypothèses comportementales ou autres sur les fonctions d'investissement le modèle multi-échelles peut générer les rendements anti-corrélés.

La présence des anti-corrélations, petites mais significatives, dans les séries des rendements boursiers est l'un des faits stylisés sur la dynamique des prix des actions, connu depuis Fama [1965]. En pratique, la présence de l'arbitrage statistique peut réduire ces autocorrélations, mais ces possibilités sont limitées par divers coûts de transaction, de manière à ce que les anticorrélations à des fréquences quotidiennes restent souvent visibles [voir Jegadeesh and Titman, 1995]. Dans des études plus récentes, les anti-corrélations dans les rendements sont mises en évidence pour de nombreux marchés, autres que les Etats-Unis [see Lee et al., 2003]. Nos résultats fournissent une explication possible de ce phénomène.

Non-techniquement, l'effet de correction d'erreurs dans notre modèle peut être décrit comme suit. Supposons qu'à la période t un rendement anormalement élevé (ou faible) tire vers le haut (ou vers le bas) la part de richesse, investie dans l'actif risqué par les agents passifs. À la période $t+1$ ceux d'entre eux, qui participent au marché, réajustent leur portefeuilles afin de rétablir la répartition cible. Cela engendre le mouvement du prix de l'actif risqué dans les sens opposé et la richesse des investisseurs passifs évolue dans la "bonne" direction, effaçant l'effet du choc initial. A la prochaine période, les investisseurs, qui ont été passifs dans les deux périodes précédentes et participent au marché, réajustent leurs portefeuilles en fonction de l'effet composé des deux fluctuations antérieures, et ainsi de suite.

L'effet d'hétéroscédasticité conditionnelle est également présent et son importance dépend de la répartition de richesse à travers les échelles, avec plus de richesse affectée aux échelles plus longues signifiant plus de mémoire dans la volatilité. Cependant, cet effet est relativement faible, il n'est que légèrement reflété dans les autocorrélogrammes de $|\hat{r}_t|$. Il est pratiquement impossible à déceler dans la série r_t, à laquelle le bruit blanc

de variance $\frac{\bar{x}}{1-\bar{x}}\sigma_e^2$ est ajouté (dans notre exemple, l'écart type de bruit standard est 3×10^{-3}).

Considérons maintenant un autre type de chocs - la déviation de la fonction d'investissement du niveau d'équilibre. Nous simulons des modèles avec 1 à 5 horizons, perturbés par les fluctuations aléatoires de petite amplitude et non-persistantes, appliquées à $x^0_{t,h}$ à des temps aléatoires en moyenne une fois par $10\,h$ périodes. Dans le Tableau 2.1, nous rapportons les moyennes et les écarts types des rendements pour différents modèles. Le rendement moyen reste constant, on observe que le niveau de volatilité baisse lorsque le nombre d'horizons augmente. Dans le modèle multi-échelle les impacts des chocs sont dilués, car beaucoup d'investisseurs ne participent pas au marché au moments où les chocs se produisent. Ces investisseurs sont toutefois affectés par les rendements anormaux, générés à ces périodes.

TABLE 2.1: Volatilités des rendements dans un modèle multi-échelles avec des chocs à la fonction d'investissement

	1 scale	2 scales	3 scales	4 scales	5 scales
Rendement moyen	0.009	0.009	0.009	0.009	0.009
Ecart type	0.025	0.021	0.017	0.015	0.014

Modèles avec $H = 1, \ldots, 5$, où la richesse initiale est repartie selon la loi $\beta(2,2)$ discrétisée, $e \sim N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Les chocs Gaussiens de variance 0.01^2 sont appliqués aux fonctions d'investissement à des dates aléatoires avec les fréquences $\frac{1}{10\,H}$, i.e. en moyenne chaque $10 - 50$ points selon l'horizon d'investissement. Les paramètres sont estimés avec 10 000 simulations.

Sur le graphique 2.2 nous représentons les propriétés dynamiques des rendements dans les modèles avec différent nombre d'horizons. Pour le cas d'un horizon, nous observons les rendements anormaux importants aux périodes des chocs et des rendements d'amplitude importante aux périodes juste après les chocs, ce qui s'explique par le retour à la normale des parts de l'actif risqué dans les portefeuilles des investisseurs. Sur l'autocorrélogramme des rendements, nous trouvons une forte autocorrélation négative d'ordre 1, caractéristique du processus MA(1). Ce n'était pas le cas pour les chocs dans les rendements sans rapport avec la demande, qui ont été décrits précédemment. Dans le modèle à un horizon, ces chocs ne déclenchent pas de correction ultérieure.

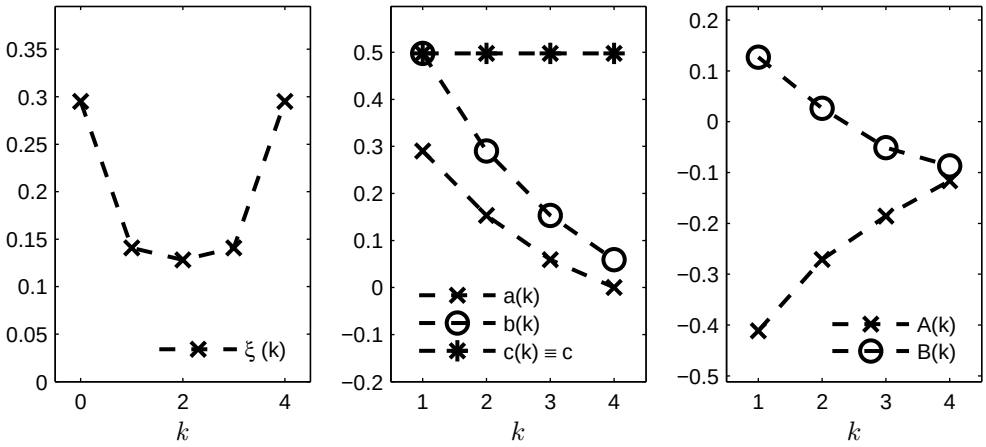
Dans un modèle à horizons multiples avec h échelles, les chocs à $x^0_{t,h}$ impactent la fonction d'investissement agrégée $x_{t,h}$ au cours de $h-1$ périodes, le temps nécessaire de rééquilibrer les portefeuilles après un choc. Cela provoque un effet durable sur le terme de la volatilité $\frac{x_t}{1-x_t}$ et crée des écarts par rapport à la trajectoire d'équilibre, défini par le théorème 2.4.5 avec $\hat{r}_t = 0$. L'écart, à son tour, déclenche le mécanisme de correction d'erreur, décrit ci-dessus en détail. Nous constatons que, dans le cas des chocs à la fonction de la demande, l'effet de l'hétéroscédasticité conditionnelle n'est plus

négligeable. Il se manifeste par l'apparition des autocorrélations significatives dans les rendements absolus jusqu'à $h-1$ retards.

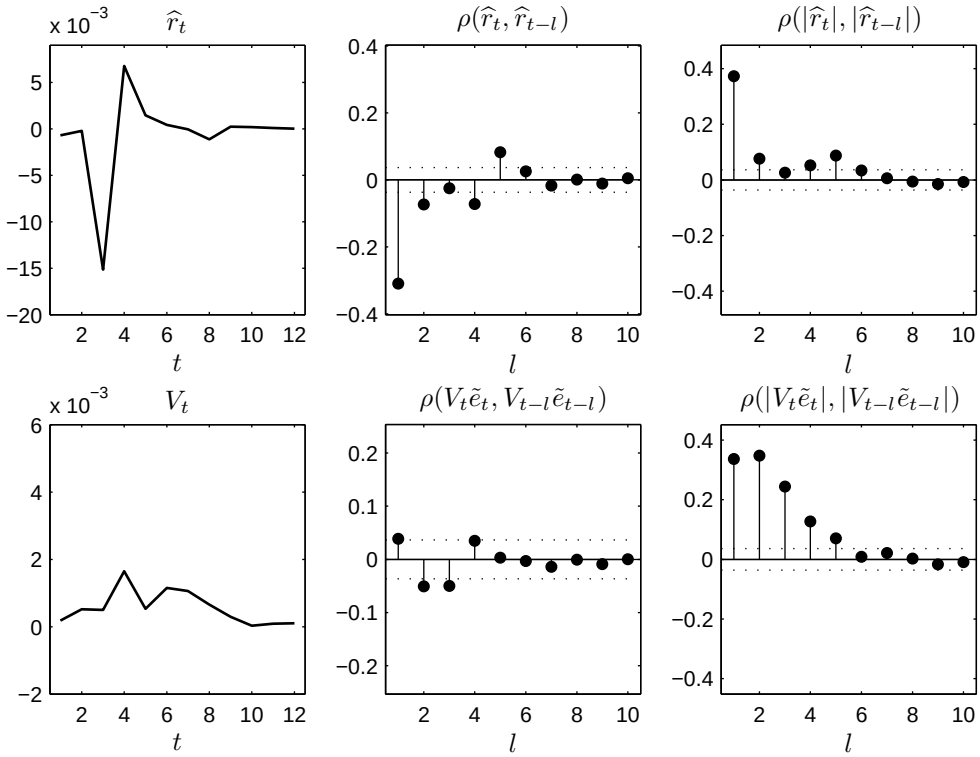
Pour le moment, nous avons supposé que les perturbations du système ont été purement exogènes. Nous n'avons pas analysé les types particuliers de comportement, tels que l'extrapolation des tendances du marché ou l'inverse, et nous n'avons pas induit de chocs dépendant de l'histoire passée des rendements. Il est important que même dans ce cas simple, nous constatons que notre modèle peut générer des dynamiques de rendements "intéressantes". Certainement, les déviations de l'équilibre rationnel dans les fonctions d'investissement peuvent être associées au comportement des investisseurs et peuvent être présentes à toutes les périodes, contrairement à notre exemple stylisée, conçue à des fins d'illustration.

FIGURE 2.1: Chocs aux rendements, répartition de richesse en ∪

(a) Paramètres du modèle

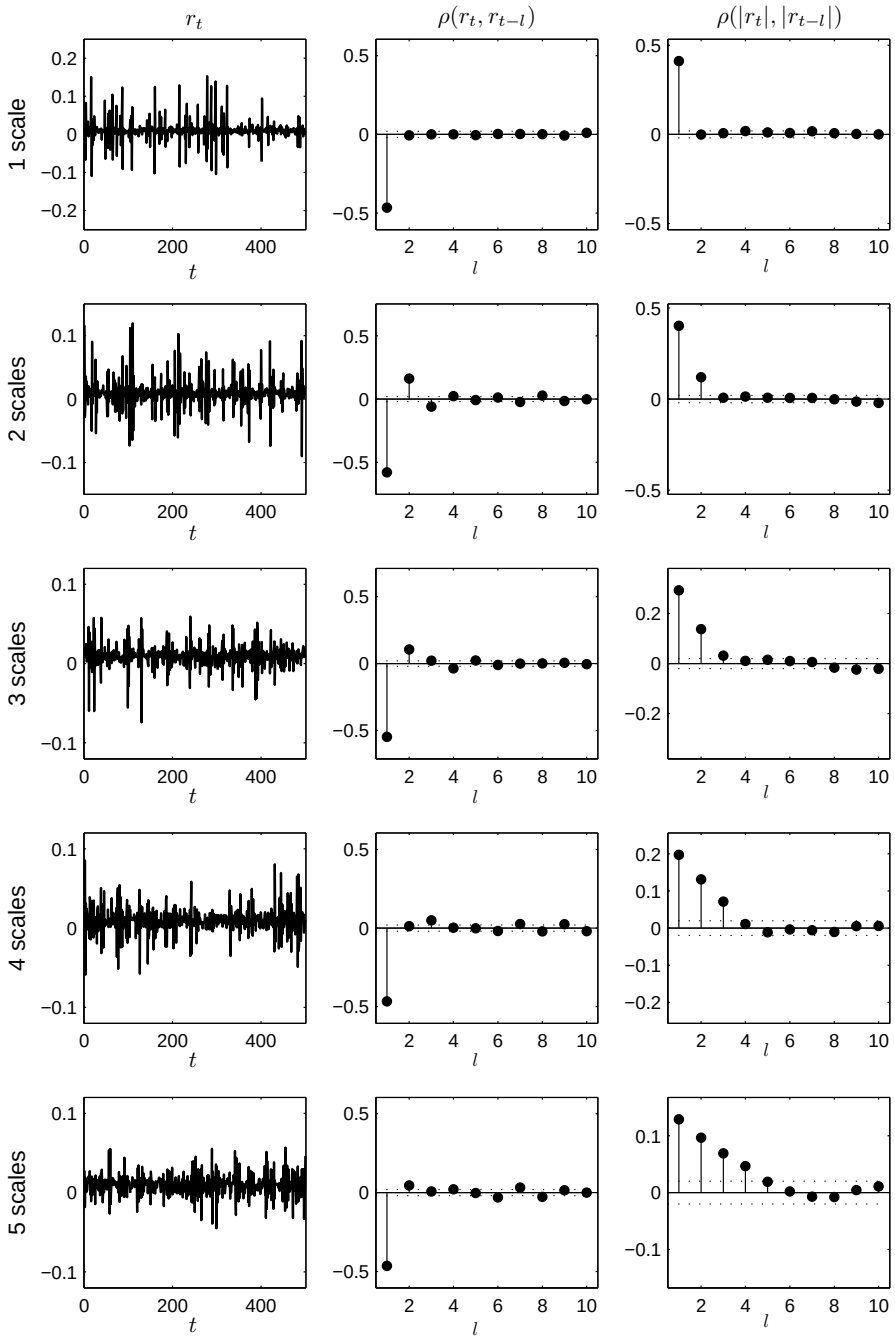


(b) Réaction au chocs exogènes



Modèle avec $H = 5$, où la richesse initiale est repartie selon la loi $\beta(2, 2)$ discrétisée, $e \sim N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Les chocs Gaussiens de variance $\frac{\bar{x}^2}{(1-\bar{x})^2}\sigma_\epsilon^2$ sont appliqués à la série $\widetilde{r}_t$ à des dates aléatoires avec les fréquences $\frac{1}{10\,H}$, i.e. en moyenne chaque 50 points. Les autocorrélations sont estimées avec 10 000 simulations.

FIGURE 2.2: Chocs aux fonctions d'investissement



Modèles avec $H = 1, \ldots, 5$, où la richesse initiale est repartie selon la loi $\beta(2,2)$ discrétisée, $e \sim N(0.03, 0.02^2)$, $\bar{x} = 0.75$. Les chocs Gaussiens de variance 0.01^2 sont appliqués aux fonctions d'investissement à des dates aléatoires avec les fréquences $\frac{1}{10\,H}$, i.e. en moyenne chaque $10 - 50$ points selon l'horizon d'investissement. 10 000 simulations.

2.6 Conclusion

Nous avons montré que le processus de rendements en équilibre rationnel pour les investisseurs avec l'aversion au risque relatif constant est le même pour les cas d'un et de plusieurs horizons d'investissement. Cependant, ce résultat ne tient pas si le système est soumis à des chocs exogènes ou liés à des déviations du comportement rationnel. En fait, la principale différence entre les modèles à un ou plusieurs horizons est la façon dont le système dynamique qu'on en déduit réagit aux chocs.

Nous avons d'abord démontré que le modèle multi-échelle à des chocs exogènes appliqués aux rendements génère les anti-corrélations temporelles, ce qui est en accord avec les faits stylisés sur les "profits contrariants". Notre modèle ne génère pas de réactions excessives des investisseurs, mais inclut la correction d'erreur qui consiste à la tendance des acteurs du marché temporairement passifs de réajuster leurs portefeuilles pour revenir à la répartition cible de leur richesse, après que cette dernière a dévié suite aux fluctuations des prix. Nos résultats théoriques contribuent à la littérature sur le retour à la moyenne, offrant une explication plausible pour l'anti-corrélation des rendements.

Nous démontrons aussi que le modèle multi-échelle avec des chocs à des fonctions d'investissement génère l'hétéroscédasticité conditionnelle. Jusqu'à présent, les explications de l'hétéroscédasticité conditionnelle dans la littérature théorique ont été presque exclusivement fondées sur la variation des stratégies d'investissement [Brock and Hommes, 1998, Chiarella and He, 2001, Anufriev et al., 2006, Weinbaum, 2009] ou sur l'aversion au risque relative dépendant de la richesse [Vanden, 2005]. Contrairement à la première classe de ces modèles, aucune propriété particulière dans le comportement des investisseurs n'a besoin d'être posée. Même les perturbations exogènes *iid* des fonctions de la demande génèrent les effets GARCH. La mécanique de cet effet est méthodologiquement proche du modèle de Vanden [2005]. Dans les deux modèles la volatilité dépend de la demande pour l'actif risqué, exprimée en tant que part de la richesse. Dans Vanden [2005] la volatilité varie à causse de la dépendance en escalier de l'aversion relative au risque par rapport à la richesse. Ainsi, la relation entre le prix de l'actif risqué et la valeur de l'investissement dans ce dernier ne peut rester constante. Dans notre modèle, la demande ne suit pas parfaitement les fluctuations de prix, car une partie des agents reste passive à chaque date d'échanges.

Notre cadre d'analyse est parfaitement compatible avec l'introduction des modèles comportementaux de l'investissement. Nous avons étudié en détail les stratégies d'investissement "moyenne-variance" procéduralement rationnelles, stipulant que dans le cas des rendements *iid*, les participants du marché ne font pas d'erreurs systématiques dans l'estimation de la moyenne et de la variance. Nous montrons que dans le cas d'une échelle, qui a été étudié avant dans Anufriev et al. [2006], les rendements d'équilibre ne sont jamais *iid* à moins que les agents ont des anticipations rationnelles et leurs fonctions d'investissement sont constantes. Ce résultat est également valable pour le cas

multi-échelle. Une des conséquences de l'introduction des échelles multiples est que les chocs comportementaux à des fonctions d'investissement sont lissés dans le temps et réduits en ampleur, ce qui ajoute à la stabilité du système.

Nous avisons le lecteur que, bien que nos résultats sont assez généraux, ils sont néanmoins fondés sur une série d'hypothèses fortes. Les plus importantes d'entre eux incluent *(i)* l'absence de signaux d'information, liés aux rendements futurs, autres que contenus dans l'historique des prix ; *(ii)* constance et exogénéité des fréquences d'interventions des investisseurs au marché ; *(iii)* l'hypothèse que le taux de dividendes est une variable aléatoire *iid*. Ces hypothèses peuvent être sujets d'extensions et de nouveaux modèles. L'importance des résultats, présentés ici, c'est qu'ils établissent un cadre général, qui peut être utilisé pour l'étude des problèmes plus spécifiques.

Chapitre 3

Modèles de volatilité : de GARCH aux cascades à horizons multiples

3.1 Introduction

Dans ce chapitre nous étudions la variabilité des prix des actions, appelée “la volatilité”. Habituellement l’introduction de la terminologie scientifique vise à rendre un concept général plus précis, mais nous avons ici plutôt une exemple du contraire. Selon le contexte et le point de vue de l’auteur, le terme “volatilité” en finance peut référer à la variabilité des prix (dans ce sens, on l’a utilisé ci-dessus), une estimation de l’écart-type, le risque financier en général, un paramètre d’un modèle de l’évaluation des produits dérivés ou d’un processus stochastique de forme particulière. Nous allons continuer à l’utiliser au sens large, comme synonyme de la variabilité. Avant de revoir les modèles de volatilité, nous examinons plus en détail la genèse de la notion elle-même. Cela contribuera à une meilleure compréhension de la logique de l’évolution des modèles correspondants.

Une des premières interprétations de la “volatilité” est due au fait que le nom du phénomène même de la variabilité a été identifié avec la méthode la plus élémentaire de sa mesure quantitative - l’écart type des rendements boursiers. Cette interprétation s’intègre logiquement dans le concept de la théorie moderne du portefeuille, également appelée théorie de moyenne-variance, parce que sous ses hypothèses ces deux paramètres contiennent toute l’information pertinente sur les rendements boursiers, suivant la loi normale. Ainsi la volatilité peut être considérée comme synonyme de l’écart-type ou comme une estimation d’un paramètre constant dans le modèle le plus simple des rendements boursiers. Cette définition de la volatilité a des racines profondes et est toujours largement utilisée par les professionnels de la gestion d’actifs.

L'apparition en 1973 des modèles d'évaluation d'options par Black and Scholes [1973] et Merton [1973] conduit à des changements significatifs dans la compréhension de la volatilité. Une diffusion en temps continu (mouvement brownien géométrique) est utilisée pour modéliser les cours des actions :

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{3.1}$$

avec S_t le prix, μ la dérive et W_t un mouvement brownien standard. Le paramètre σ est appelé volatilité car il caractérise le degré de variabilité. Très vite, il est devenu évident que l'équation (3.1) décrit mal la réalité.

Ses paramètres définissent sans ambiguité les prix des options pour les dates et les prix d'exercice donnés, de sorte que le paramètre de volatilité peut être déduit à partir des observations de prix d'options en utilisant l'inverse de la formule de Black-Scholes. Une estimation obtenue de cette façon est appelée volatilité implicite par opposition à la volatilité historique, mesurée par l'écart-type des rendements. Contrairement aux prédictions du modèle de Black et Scholes, les résultats empiriques montrent que la volatilité implicite varie selon les contrats optionnels avec des paramètres différents.

La remarque précédente ne signifie pas que la volatilité implicite est inutile. Il a été démontré qu'elle contient des informations sur la variabilité des rendements futurs, et donc elle est souvent utilisée pour les prévisions. Aussi, la volatilité implicite est importante car elle permet d'extrapoler les données de marché observées, par exemple, les prix d'options, pour évaluer les autres instruments financiers, par exemple, les options traitées de gré à gré [voir Dupire, 1993, 1994, Avellaneda et al., 1997]. En dépit de ces succès partiels, un modèle plus adapté que la diffusion log-normale serait utile à la fois dans l'évaluation des produits dérivés et les applications de la gestion d'actifs. Dans Merton [1973] le paramètre de volatilité est déjà variable dans le temps. Bien plus tôt Mandelbrot [1963] met en évidence les propriétés empiriques des rendements des actions qui correspondent pas au modèle de la diffusion log-normale et propose une catégorie plus large de distributions de probabilité Lévy-stable. Les évolutions suivantes ont conduit à la modélisation de la volatilité comme un processus stochastique et non comme un simple paramètre, même variable dans le temps.

Le sens du terme "volatlité" en finance n'a cessé d'évoluer : à partir d'un terme général pour la variabilité vers un estimateur statistique, puis paramètre d'un modèle et, enfin, un processus stochastique, qui caractérise la structure entière de la la variabilité des cours d'actions. Plus techniquement, la volatilité dans sa compréhension moderne peut être définie comme une structure temporelle des moments conditionnels du second ordre dans la distribution des rendements. Dans le cas le plus simple de la diffusion log-normale, cette structure est décrite par un paramètre et dans des cas plus complexes, par un processus stochastique séparé.

Ce chapitre commence par un aperçu des propriétés empiriques de la volatilité, des soi-disants "faits stylisés". Ensuite, nous discuterons brièvement des approches traditionnelles à sa modélisation - de la hétéroscédasticité conditionnelle et de la volatilité stochastique, qui reproduisent les propriétés empiriques dans une certaine mesure. Bien que de nombreux modèles sont assez bonnes pour décrire les différents faits stylisés, nous montrons qu'aucun d'entre eux n'est tout à fait suffisant pour représenter toute la structure de la variabilité des prix boursiers. En particulier, la plupart des modèles ne permettent pas de représenter la dynamique des rendements sur des horizons de temps multiples (par exemple, de quelques minutes à plusieurs jours et mois) simultanément. Les faits stylisés eux-mêmes ont des caractéristiques spécifiques à la fréquence, à laquelle la dynamique des prix est observée. Nous analysons et comparons des modèles récents d'hétéroscédasticité conditionnelle et de la volatilité stochastique, basés sur l'approche multi-horizon, et discutons des principaux problèmes non résolus, liés à ces modèles.

3.2 Propriétés empiriques de la volatilité

De nombreuses études empiriques montrent que les séries chronologiques financières satisfont un certain nombre de propriétés générales, appelées faits stylisés. Un modèle réaliste pour les prix est censé reproduire ces propriétés. Nous les caractérisont brièvement, pour une revue approfondie sur le sujet voir Cont [2001].

- **Volatilité excessive.** Le degré de la variabilité observée dans les cours d'actions ne peut guère s'expliquer par des variations dans les facteurs économiques fondamentaux [Cutler et al., 1989].
- **Absence de corrélations linéaires dans les rendements**. Les rendements, calculés sur les périodes de temps suffisamment longues (plusieurs heures et plus) ne contiennent pas de corrélation linéaire significative.
- **Clustering de la volatilité et mémoire longue dans les valeurs absolues des rendements**. Les séries chronologiques des valeurs absolues des rendements sont caractérisées par une autocorrélation importante, et la fonction d'autocorrélation (ACF) décroit plus lentement que la vitesse géométrique. De longues périodes de forte volatilité et de faible volatilité sont observées [Bollerslev et al., 1992, Ding et al., 1993, Ding and Granger, 1996].
- **Le lien entre le volume des transactions et la volatilité.** La volatilité des rendements est positivement corrélée avec le volume de transactions [Lobato and Velasco, 2000].
- **Asymétrie et effet de levier dans la structure dynamique de la volatilité**. Les rendements passés positifs et négatifs de la même ampleur ont des effets différents sur la volatilité des cours (asymétrie). Les rendements actuels et la volatilité future sont corrélés négativement (effet de levier) [Black, 1976].

– **Queues lourdes dans la distribution des rendements.** La distribution de probabilité des rendements quotidiens est caractérisée par des queues lourdes, c'est-à-dire une forte probabilité d'observer des valeurs extrêmes, par rapport à la distribution normale [Mandelbrot, 1963, Fama, 1965].
– **La forme de la distribution de probabilité des rendements varie selon les intervalles de temps, sur lesquelles les rendements sont calculés** [Ghashghaie et al., 1996, Arneodo et al., 1998].

Parmi ces faits stylisés, nous serons particulièrement intéressés par les propriétés, liées à l'ACF des rendements et la forme de la distribution de probabilité des rendements et de leur valeurs absolues. Nous commençons par la définition du phénomène de la mémoire longue en termes d'ACF.

Un processus stochastique stationnaire X_t avec variance finie est caractérisé par la mémoire longue si sa fonction d'autocorrélation $C(\tau) = \text{corr}(X_t, X_{t-\tau})$ à $\tau \to \infty$ décroit avec le retard selon la loi de puissance (i.e. à la vitesse hyperbolique) :

$$C(\tau) \sim \frac{L(\tau)}{\tau^{1-2d}}, \tag{3.2}$$

où $0 < d < \frac{1}{2}$ et $L(\cdot)$ est une fonction continue telle que $\forall x > 0$ et $\tau \to \infty$ satisfait $\frac{L(xt)}{L(t)} \to 1$. Le processus a une mémoire courte si son ACF décroit exponentiellement, de sorte que :

$$\exists A > 0, c \in (0,1) : |C(\tau)| \leq Ac^{\tau}. \tag{3.3}$$

Autrement, la mémoire longue peut être caractérisée par la divergence de type puissance de la densité spectrale de la série chronologique X_t à l'origine :

$$\Psi_x(u) \sim c_\Psi |u|^{-\alpha} \tag{3.4}$$

avec $\Psi_x(\cdot)$ la densité spectrale, α le paramètre d'échelle et c_Ψ une constane.

Pour illustrer les propriétés empiriques des rendements, nous utilisons deux types de données sur les indices boursiers : les observations en haute fréquence sur une période de temps relativement courte (indice CAC40 français) et les observations quotidiennes sur une très longue période (Dow Jones Industrial Average Index, DJIA) . Nous montrons que les principaux constats empiriques sont similaires pour ces exemples très différents.

Nos résultats empiriques montrent la présence de la mémoire longue dans les séries chronologiques de volatilité et le caractère non-gaussien de la distribution des rendements, en particulier à des fréquences d'observation élevées. Dans la prochaine section, nous expliquons comment ces propriétés peuvent être reproduites par les modèles proposés dans la littérature financière.

3.3 Famille de modèles ARCH/GARCH et les extensions

Le modèle de l'hétéroscédasticité conditionnelle, le plus souvent utilisé en partique, est le modèle GACRH, proposé dans Bollerslev [1986] :

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i r_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 = \alpha_0 + \alpha(L,q) r_t^2 + \beta(L,p)\sigma_t^2 \tag{3.5}$$

avec L^n l'operator de retard d'ordre n et $a(L,n)$ l'opérateur de la forme $\sum_{i=1}^{n} a_i L^i$, appliquée à la série temporelle. Donc $a(L,q)X_t$ correspond à $\sum_{i=1}^{q} a_i X_{t-i}$ et l'équation(3.5) devient :

$$[1 - \alpha(L,q) - \beta(L,p)]\, r_t^2 = \alpha_0 + [1 - \beta(L,p)] \left(r_t^2 - \sigma_t^2\right), \tag{3.6}$$

ce qui correspond à un modèle ARMA pour les rendements au carré avec les paramètres $\max\{p,q\}$ et p car $E(r_t^2 - \sigma_t^2 | I_{t-1})$ est une variable *iid* centrée. Pour que le processus soit stable, i.e. la variance des innovations $\sigma_t \varepsilon_t$ soit finie, toutes les racines des équations $\alpha(L_q) = 0$ et $1 - \alpha(L_q) - \beta(L_p) = 0$ doivent être hors du cercle unitaire. Pour GARCH(1,1) cette contrainte prend une forme simple $\alpha + \beta < 1$.

La fonction ACF théorique d'un processus GARCH(1,1) décroit à la vitesse géométrique, donnée par la somme $\alpha + \beta$. Le défaut principal de GARCH est que sa mémoire n'est pas assez longue, car l'ACF décroit trop vite. Quand $\alpha + \beta$ s'approche de 1, GARCH(1,1) dégénère dans un processus, appelé GARCH intégré par Engle and Bollerslev [1986]. Ce modèle est non-stationnaire et implique l'effet permanent des conditions initiales sur la dynamique des prix ce qui le rend difficilement acceptable pour représenter la réalité.

Une autre approche consiste à utiliser les processus, dont les propriétés théoriques impliquent la présence de la mémoire longue. Dans le cadre GARCH, le processus FIGARCH proposé dans Baillie et al. [1996] et Bollerslev and Mikkelsen [1996] est défini par :

$$[1 - \beta(L,p)]\, \sigma_t^2 = \alpha_0 + \left[1 - \beta(L,p) - \phi(L)(1-L)^d\right] r_t^2 \tag{3.7}$$

avec $\phi(L) = [1 - \alpha(L,q) - \beta(L,p)]\,(1-L)^{-1}$.

3.4 Modèles de volatilité stochastique

Les modèles GARCH ont une seule source d'innovations aléatoires. La variance conditionnelle du processus de rendements est une fonction de ses réalisations passées. Une autre approche consiste à mettre en place un modèle simple pour les rendements, par exemple, donné par (3.1), et au lieu de considérer σ en tant que paramètre, essayer de le modéliser comme un processus distinct. Deux sources de hasard vont ainsi apparaître. Cette idée est le concept de la volatilité stochastique.

Le premier modèle de la volatilité stochastique a été proposé dans Taylor [1982]. Il s'agit de modéliser la log-volatilité en temps discret par un processus AR(1) :

$$\begin{aligned} r_t &= \mu\sigma_t\varepsilon_t \\ \ln\sigma_t^2 &= \phi\ln\sigma_{t-1}^2 + \nu_t, \end{aligned} \tag{3.8}$$

où μ est une constante positive et ϕ est un paramètre autorégressif, déterminant la mémoire dans la volatilité. Les propriétés de ce modèle autoregressif (ARSV) ont été étudié dans Andersen [1994], Taylor [1994], Capobianco [1996]. En particulier, si la log-volatilité est stationnaire, la distribution des rendements est symétrique et a des queues épaisses [Bai et al., 2003]. Les rendements sont non-corrélés mais dépendants. L'ACF des rendements au carré décroit à la vitesse géométrique, comme pour les processus GARCH.

Diverses méthodes ont été proposées pour intégrer la mémoire longue. Breidt et al. [1998], Harvey [1998] construisent des modèles en temps discret avec l'intégration fractionnaire, Comte and Renault [1998] proposent un modèle en temps continu avec un mouvement brownien fractionnaire. Chernov et al. [2003] et [LeBaron, 2001] considèrent des modèles où la volatilité stochastique est composée de plusieurs facteurs, de manière à obtenir une décroissance lente de la fonction ACF, même si le processus générateur de données ne possède pas la propriété de mémoire longue.

3.5 Agrégation des rendements dans le temps

Dans la section 3.2 nous avons montré que la forme de la distribution des rendements dépend des fréquences d'observation. En même temps, les propriétés dynamiques de la volatilité, telles que la mémoire longue dans les rendements absolus et l'absence de corrélations linéaires dans les rendements eux-mêmes, sont communes pour les séries chronologiques à des fréquences différentes. Une série de questions se pose. De quelle façon le phénomène de la mémoire longue est lié aux propriétés des rendements à différents horizons ? Est-ce que les modèles de volatilité, calibrés sur des données à une certaine fréquence, peuvent reproduire les propriétés des rendements à d'autres fréquences ? Est-il judicieux de faire des estimations à plusieurs horizons pour la même série et si oui, comment concilier les résultats ?

La réponse à la première question a été donnée par Mandelbrot et Van Ness en 1968. Ils ont montré que pour une certaine classe de processus stochastiques, leurs propriétés établies sur des horizons courts permettent de décrire complètement les propriétés à des horizons longs. Un processus X_t est appelé auto-affine s'il existe une constante $H > 0$, telle que pour tout facteur d'échelle $c > 0$ les variables aléatoires X_{ct} et $c^H X_t$ sont égales en loi :

$$X_{ct} \stackrel{L}{=} c^H X_t \tag{3.9}$$

Quand la condition $\frac{1}{2} < H < 1$ est vérifiée, le processus possède la mémoire longue, et pour $H = \frac{1}{2}$ c'est un mouvement brownien avec incréments indépendants. Une généralisation de la classe des processus auto-affines est la classe des processus multifractaux, pour lesquels le facteur d'auto-affinité n'est plus constant, de sorte que la propriété d'agrégation se lit :

$$X_{ct} \stackrel{L}{=} M(c)X_t \tag{3.10}$$

avec $M(\cdot)$ une fonction aléatoire positive de c, indépendante de X, telle que $M(xy) \stackrel{L}{=} M(x)M(y)$ pour $\forall x, y > 0$. Pour un processus strictement stationnaire, la loi d'échelle locale est vérifiée :

$$X_{t+c\Delta t} \stackrel{L}{=} M(c)\left(X_{t+\Delta t} - X_t\right) \tag{3.11}$$

Autrement [see Castaing et al., 1990] un processus multifractal peut être défini par la relation entre les fonctions de densité de probabilité des incréments du processus, calculées pour des intervalles de temps de longueurs différentes l et L, tels que $L = \lambda l, \lambda > 1$. Cette relation s'écrit :

$$P_l(x) = \int G(\lambda, u)e^{-u}P_L(e^{-u}x)du \tag{3.12}$$

avec $P_l(\cdot)$ la densité de probabilité des incréments $\delta_l X_t$ du processus X_t à l'horizon l, i.e. $x = \delta_l X_t = X_{t+l} - X_t$. La fonction $G(\lambda, u)$, est appelée le noyau d'auto-affinité. Dans le cas le plus simple d'un processus auto-affine elle prend la forme :

$$G(\lambda, u) = \delta(u - H\ln\lambda) \tag{3.13}$$

avec $\delta(\cdot)$ la fonction de Dirac. Dans ce cas monofractal un point est suffisant pour décrire l'évolution de la distribution, car P_l and P_L ne sont différentes que par le facteur d'échelle. Ceci explique la forme dégénérée de (3.13).

Il est intéressant que les propriétés d'agrégation du temps des modèles simples de type GARCH et ARSV ne fournissent pas une représentation adéquate des rendements boursiers à des horizons multiples simultanément. En ce qui concerne GARCH (1,1), Drost and Werker [1996] montrent que ce modèle vérifie la propriété d'invariance d'échelle, c'est à dire si les rendements à un certain horizon suivent GARCH (1,1), c'est vrai pour n'importe quelle échelle de temps avec les mêmes paramètres. C'est à la fois une force et une faiblesse du modèle GARCH. D'une part, les résultats de l'inférence statistique sont indépendants de la fréquence d'observation. D'autre part, l'invariance d'échelle stricte ne permet pas de reproduire l'évolution de la forme de la distribution de volatilité selon les horizons de temps d'observation.

Les arguments ci-dessus montrent la nécessité d'un modèle de volatilité, qui permettrait non seulement de reproduire la mémoire longue et/ou la présence de queues lourdes dans les distributions des rendements, observés à une certaine fréquence fixe, mais qui donnerait aussi des résultats satisfaisants pour d'autres horizons.

3.6 L'hypothèse d'horizons multiples dans la volatilité

L'hypothèse économique des horizons multiples consiste à supposer que les propriétés de la volatilité sont affectées par l'hétérogénéité dans les horizons temporels des décisions, prises par les investisseurs. Un modèle paramétrique de la volatilité à des horizons multiples dans l'esprit de l'approche ARCH a été proposé dans Müller et al. [1997] et Dacorogna et al. [1998]. La volatilité courante est représentée comme une fonction linéaire des rendements au carré sur des périodes différentes dans le passé :

$$\sigma_t^2 = c_0 + \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{j} r_{t-i} \right)^2 \tag{3.14}$$

avec $c_k \geq 0$ pour tout $k = 0, ..., n$, et pour $k = 0$ et $k = n$ l'inégalité est stricte. Le principal problème de ce modèle est le grand nombre de paramètres, induisant de très fortes corrélations entre les variables indépendantes, qui rendent l'identification très compliquée.

Zumbach [2004] définit la volatilité comme une somme pondérée de plusieurs composantes, correspondant à différents horizons temporels. Ils considèrent $n+1$ horizons, dont la longueur $\tau_k, k = 0, \ldots, n$ augmente avec k : $\tau_k = 2^{k-1}\tau_0$. La composante de la volatilité, ce qui correspond à l'horizon k, est définie par la moyenne mobile exponentielle :

$$\begin{gathered} \sigma_{t,k} = \mu_k \sigma_{k,t-\delta t}^2 + (1-\mu_k) r_t^2 \\ \mu_0 = \exp(-\frac{\delta t}{\tau_0}) \ \mu_k = \exp(-\frac{\delta t}{\tau_0 2^{k-1}}), k = 1 \ldots n \end{gathered} \tag{3.15}$$

où r_t est le rendement à l'échelle δt d'observation des prix ($\delta t \leq \tau_0$). Ce modèle est proche de FIGARCH, mais le modèle de Zumbach a une interprétation claire en termes de l'hypothèse d'horizons multiples. Il est à noter, cependant, que les tests empiriques du modèle de Zumbach [2004] n'ont montré qu'une très légère augmentation de la puissance de prévision du modèle, par rapport à GARCH(1,1).

3.7 Modélisation à horizons multiples et l'approche éconophysique

Les modèles de volatilité à horizons multiples, décrits ci-dessus, représentent la volatilité courante comme un résultat des impacts des facteurs (ou composantes), qui varient à des fréquences différentes. Cette description de la volatilité a une analogie en physique des liquides et des gaz. L'hydrodynamique étudie le phénomène de turbulence, caractérisé par la formation de tourbillons de différentes tailles dans les écoulements de fluides et des gaz, conduisant à des fluctuations aléatoires dans les caractéristiques thermodynamiques (température, pression et densité). La plupart de l'énergie cinétique d'un

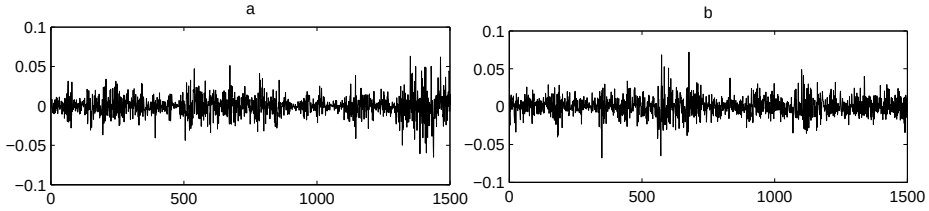
écoulement turbulent est contenue dans les tourbillons à grande échelle. L'énergie circule à partir de grandes échelles de tourbillons à des échelles plus petites. Une théorie statistique de la turbulence a été développée par Kolmogorov [1941], et une revue récente peut être trouvée, par exemple, dans Pope [2000].

Un modèle analytique du cascade multiplicatif (Multiplicative Cascade Model, MCM) a été proposé dans Breymann et al. [2000]. La volatilité est représentée comme un produit des perturbations à des fréquences différentes. La pertinence empirique du modèle est confirmée par les propriétés de l'ACF des rendements et de leurs valeurs absolues à différents horizons. Arneodo et al. [1998] montrent que, dans MCM, l'ACF des logarithmes des valeurs absolues des rendements à tous les horizons décroit à une vitesse logarithmique :

$$\mathrm{Cov}(\ln|r_{(t+\Delta t),k}|,\ \ln|r_{t,k}|) \cong -\lambda^2 \ln\frac{\Delta t}{\tau_1},\ \ \Delta t > \tau_k \tag{3.16}$$

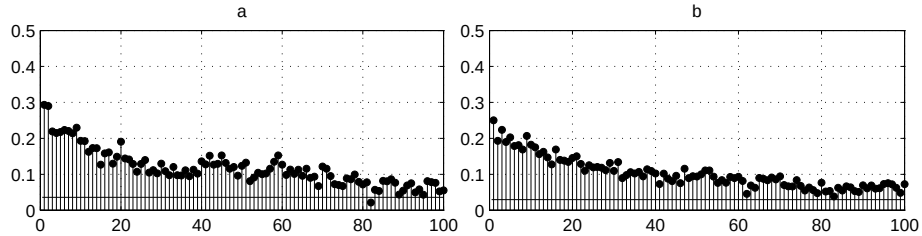
Cette relation peut être utilisée pour l'identification de l'échelle la plus longue dans la modélisations de la volatilité.

FIGURE 3.1: Simulation avec MCM et les données réelles : rendements journaliers



A gauche (a) : rendements journaliers de CAC40 (source : Euronext, valeurs de CAC40 de 20/03/95 à 24/02/05). A droite (b) : rendements journaliers, simulé avec MCM à 14 horizons (de 15 minutes à 256 jours). Les rendements sont simulés pour chaque 15 minutes et ensuite agrégés.

FIGURE 3.2: Simulation avec MCM et les données réelles : ACF empirique



A gauche (a) : ACF estimée des valeurs absolues des rendements journaliers de CAC40 (source : Euronext, valeurs de CAC40 de 20/03/95 à 24/02/05). A droite (b) : ACF estimée des valeurs absolues des rendements, simulés avec MCM à 14 horizons (de 15 minutes à 256 jours). Les rendements sont simulés pour chaque 15 minutes et ensuite agrégés. ACF est calculée pour les données journalières.

Les Graphiques 3.1 et 3.2 montrent les résultats d'une simulation de MCM, comparés aux données réelles de l'indice CAC40. Le nombre d'horizons dans la simulation est estimé à 14, ce qui permet de s'adapter à la vitesse de décroissance de l'ACF. Les

autres paramètres sont calibrés de manière à reproduire les deux premiers moments de la distribution des rendements observés. Notez que le graphique représente l'ACF des rendements journaliers, alors que l'estimation de MCM et la simulation elle-même ont été effectué pour les données échantillonnées à la fréquence de 15 minutes. Nos résultats illustrent la propriété la plus importante des cascades de la volatilité : clustering de volatilité et la présence de la mémoire longue, robustes à l'agrégation des rendements dans le temps.

3.8 Conclusion

La modélisation de la volatilité des prix boursiers est l'un des éléments clés de la théorie et de la pratique de la gestion de portefeuille et d'autres domaines de la finance. Les évolutions dans cette modélisation ont souvent pour but de mieux reproduire les propriétés empiriques des séries chronologiques des prix, telles que les autocorrélations dans les valeurs absolues des rendements, leur absence dans les rendements eux-mêmes et les queues épaisses dans les distributions des rendements à des horizons courts.

Parmi tous les modèles nous traitons avec attention particulière ceux qui représentent la volatilité à des horizons multiples. La représentation multi-horizon permet de prendre en compte les propriétés des rendements, qui se manifestent avec l'agrégation dans le temps. L'objectif pratique est de reproduire la forme de la distribution des rendements, calculés sur des intervalles de temps de longueur différente. Nous décrivons plusieurs classes de modèles multi-échelle, des ARCH hétérogènes aux cascades multiplicatifs. L'analyse multi-horizon s'appuie largement sur des méthodes et des techniques empruntées à la physique statistique.

Le concept de la volatilité à horizons multiples nécessite le développement des méthodes de sa mesure, qui tiendraient compte non seulement de l'ampleur des fluctuations, mais aussi de leurs fréquences. Les informations, obtenues à différents niveaux d'agrégation temporelle, peuvent être utilisés conjointement, en particulier, dans les applications de la gestion d'actifs et dans les prévisions. La construction des indicateurs multi-échelle de la volatilité est sujet du prochain chapitre.

Chapitre 4

Indicateur multi-échelle de la volatilité

4.1 Introduction

La volatilité d'un indice boursier est un candidat naturel pour caractériser l'état du marché et pour comparer l'impact sur le marché des événements importants, tels que les crises financières. La tache de comparer les événements se pose dans de nombreuses applications : de l'analyse structurelle et l'étude de la dynamique du marché financier dans un contexte économique général, à l'allocation d'actifs, où la mesure de la sévérité et de la durée des fluctuations extrêmes des prix est d'intérêt particulier. En principe, une telle analyse peut être fondée sur les estimations traditionnelles de la volatilité historique. Mais ces estimations sont difficiles à interpréter et leur importance varie selon les différents actifs. Une façon de surmonter cet inconvénient est de concevoir un indicateur universel, basé sur une transformation probabiliste de la volatilité.

Zumbach et al. [2000] proposent une analogie avec l'échelle de Richter-Gutenberg en géophysique et construisent un indicateur de la volatilité, appelé "échelle des chocs du marché" (Scale of Market Shocks, SMS). L'échelle de Richter-Gutenberg est une mesure du logarithme de l'amplitude des ondes sismiques, liée à l'énergie totale, libérée pendant un tremblement de terre. Le SMS reporte la volatilité du taux de change à une échelle logarithmique, ce qui suggère une analogie entre la volatilité et le travail mécanique, i.e. le taux de variation de l'énergie dans le temps. L'indicateur a été appliqué au marché de devises, mais par construction, il peut être utilisé pour tout instrument financier côté.

Le SMS représente non seulement l'amplitude, mais aussi l'échelle des fluctuations. Selon l'hypothèse du marché hétérogène, formulée par Müller et al. [1997], la dynamique des prix est expliquée par les actions des acteurs du marché à des fréquences différentes. L'intuition des indicateurs multi-horizon tels que SMS est de décrire la volatilité dans le

domaine temporel et dans le domaine spectral simultanément, reportant ainsi les fluctuations à des gammes de fréquences (échelles) particulières . Sur le plan économique, ces échelles peuvent être associées aux horizons de la prise de décisions avant chaque prochain réajustement du portefeuille. Le SMS prend la forme de la moyenne des volatilités, réalisées à différentes échelles.

Maillet and Michel [2003, 2005] adaptent l'approche multi-échelle pour le marché d'actions. Le nouvel indicateur, appelé "l'indice des chocs de marché" (Index of Market Shocks, IMS), est conçu pour la détection et la comparaison de la gravité des événements extrêmes sur le marché. Les auteurs modifient le mode du calcul et de l'agrégation des volatilités à des fréquences multiples. D'ailleurs, l'analyse en composantes principales est appliquée pour estimer les facteurs inobservables de la volatilité, qui affectent différentes échelles. Ces facteurs sont ensuite soumis à la transformation probabiliste.

Suivant Zumbach et al. [2000] et Maillet and Michel [2003, 2005], nous étudions les indicateurs multi-échelle de la volatilité et proposons un nouvel indicateur - échelle de la volatilité de marché (Market Volatility Scale, MVS). Bien que l'intuition générale des indicateurs multi-horizon est très attrayante, les implémentations existantes souffrent de plusieurs inconvénients importants, ce qui motive notre travail. A notre avis, le principal problème avec le SMS est le système d'analyse multi-échelle qui utilise le lissage séquentiel plutôt qu'une décomposition échelle par l'échelle, ce qui ne permet pas de déterminer l'importance relative de différents horizons. Ce problème subsiste pour l'IMS, qui en plus fait le calcul de l'indicateur plus complexe en introduisant les facteurs latents de la volatilité.

Nous pensons que la modélisation des facteurs latents plutôt que de la volatilité elle-même induit une perte de simplicité et mène aux erreurs d'interprétation. Nous proposons de travailler avec les volatilités multi-horizons directement. Le MVS, que nous présentons ici, utilise la variance d'ondelettes réalisée pour représenter la volatilité à des horizons multiples. L'hypothèse de log-normalité, qui est utilisée pour la transformation probabiliste dans IMS, est clairement trop restrictive. Nous utilisons la loi de Pareto généralisée pour les queues de la distribution de la volatilité réalisée à chaque échelle. Cela permet une estimation plus précise de la probabilité d'événements extrêmes, tels que les krachs boursiers.

Nous testons notre approche sur deux types de données : à haute fréquence (intervalles de 15 minutes) des observations de l'indice CAC40 de 1995 à 2006 et les observations quotidiennes de l'indice Dow Jones (DJIA) de 1896 à 2007. Dans ces deux exemples, les bandes de fréquences ne sont pas les mêmes, mais l'algorithme de calcul est assez semblable. Nous détectons les périodes de volatilité extrême sur le marché d'actions et étudions leurs caractéristiques structurelles en utilisant la décomposition multi-échelle. Cela nous permet de déterminer quelle composante de la volatilité (court terme, moyen terme ou long terme) est la plus importante à chaque période détectée.

4.2 Construction des indicateurs multi-horizons de la volatilité

Les indicateurs de volatilité multi-échelles se basent sur une analogie avec l'échelle de Richter-Gutenberg en géophysique. L'échelle de Richter est une mesure du logarithme de l'énergie totale E libérée lors d'un tremblement de terre, par rapport à un niveau de référence E_0 :

$$R \approx \ln\left(\frac{E_0}{E}\right) \tag{4.1}$$

La probabilité d'occurrence de tremblements de terre importants, regroupés dans des clusters temporels de forte activité sismique, obéit une loi puissance [voir Christensen et al., 2002, Mega et al., 2003]. La probabilité d'observer un tremblement de terre de l'énergie E est alors :

$$p(E) \approx \left(\frac{E_0}{E}\right)^{\kappa}, \tag{4.2}$$

avec κ le paramètre d'échelle. En utilisant (4.2) l'échelle de Richter peut être exprimé :

$$R \approx \ln\left(\frac{E_0}{E}\right) = -\frac{1}{\kappa}\ln p(E). \tag{4.3}$$

Par analogie avec (4.3), Zumbach et al. [2000] calculent la transformée probabiliste du logarithme de la volatilité :

$$SMS_t = -\alpha \ln P(\sigma_t) \tag{4.4}$$

avec $P(\sigma_t)$ la fonction de répartition associée à la volatilité σ_t et α le paramètre d'échelle. Notez que l'analogie n'est pas directe, car la loi de la distribution de probabilité dans (4.4) n'est pas forcément une loi de puissance et doit être estimée. L'indicateur est multi-échelle puisse que la volatilité utilisée dans (4.4) est calculée séparement pour différents horizons. L'estimateur de la volatilité réalisée pour une échelle donnée, est définie par :

$$\widetilde{\sigma}_t^{(k)} = \left(\frac{\sum_{i=1}^{M_k-1} r_{t,\delta}^2}{M_k - 1}\right)^{1/2} \tag{4.5}$$

avec $\sigma_t^{(k)}$ la volatilité des rendements à l'échelle k, $r_{t,\delta}$ les log-rendements calculés pour les intervalles de temps δ de longueur $\tau_k(M_k-1)^{-1}$, τ_k la longueur de l'échelle k et M_k le nombre d'observations disponibles pour cette période.

La méthode d'estimation des volatilités, utilisée dans Zumbach et al. [2000], ne donne aucune indication sur les bords de la gamme des échelles à prendre en considération et sur l'importance relative de différentes échelles. Par ailleurs, les auteurs affirment que le choix du noyau de convolution n'est pas une question cruciale et il suffit de prendre une fonction qui satisfait certaines conditions de régularité et tend vers zéro sur les bords de la

gamme de fréquences (15 minutes - 64 jours dans leur cas). A notre connaissance, ce choix est fait *a priori* et n'a pas de fondements économiques ni statistiques. L'hypothèse que le point de masse, ou la fréquence la plus importante, est située au centre de la gamme est également douteuse. En outre, le SMS ne tient pas compte de l'interdépendance possible entre les échelles.

Maillet and Michel [2003, 2005] tiennent compte de la dépendance entre les échelles. Au lieu des volatilités elles-mêmes, les composantes principales sont utilisées dans la définition de leur indicateur IMS :

$$IMS_t = -\sum_{i=1}^{q} \omega_i \log_2 P(c_i) \tag{4.6}$$

avec $c_i, \ldots, c_q$ facteurs de la volatilité (composantes principales), et ω_i le poids de la composante i, determiné comme la part de la variance des données, expliquée par cette composante.

Le principal inconvénient de l'IMS est que le remplacement des volatilités par leurs combinaisons linéaires (facteurs) change le sens de l'indicateur, ce qui n'est pas explicité dans Maillet and Michel [2003]. Les auteurs affirment que l'IMS a un sens économique évident : comme un logarithme en base 2 est utilisé dans (4.6), une augmentation d'un point de l'IMS correspond à un vecteur des volatilités deux fois moins probable. C'est le cas pour la transformation de base type Richter, définie dans (4.4), mais pas pour (4.6), car la fonction de répartition multivariée d'un vecteur n'est pas la même que la fonction de répartition d'une combinaison linéaire de ses composantes. Donc une augmentation d'un point de l'IMS correspond à un vecteur des facteurs deux fois moins probable. Mais les facteurs, obtenus à partir de l'analyse en composantes principales (ACP), n'ont pas nécessairement d'impact positif ou négatif sur les volatilités à toutes les échelles. Une hausse d'un facteur peut engendrer une hausse de la volatilité dans une bande de fréquences et une baisse dans une autre. Par ailleurs, comme l'analyse en composante principales se fait sur la matrice de variance-covariance, les volatilités à des échelles différentes sont traitées comme si elles étaient toutes d'égale importance pour les investisseurs. Tout cela rend l'interprétation de l'indicateur résultant très problématique.

La transformée probabiliste, utilisée pour calculer l'indicateur, est un aspect important. Maillet and Michel [2003] postulent la loi gaussienne des log-volatilités (et des composantes principales) à toutes les échelles, mais ils signalent que l'hypothèse de normalité est probablement violée pour les valeurs extrêmes des volatilités. Cependant, ces valeurs sont d'une grande importance pour l'analyse des crises financières.

Ainsi qu'il ressort de la discussion ci-dessus, les améliorations les plus importantes à faire dans les indicateurs de volatilité à plusieurs échelles concerneraient l'analyse multi-résolution par lui-même et la transformation probabiliste du vecteur des volatilités. Ces améliorations sont proposées dans les sections suivantes.

4.3 La décomposition de la volatilité par échelles avec les ondelettes

Dans cette section, nous proposons une méthode d'analyse multi-échelle de la volatilité, basée sur les filtres d'ondelettes. Nous introduisons la variance réalisée de coefficients d'ondelettes, qui est utilisée pour définir l'indicateur MVS.

Nous utilisons la transformée en ondelettes discrète avec chevauchement maximal (Maximum Overlap Discrete Wavelet Transform, MODWT) pour décomposer la variance dans le temps et dans l'échelle. Soit $\widetilde{W}_{j,t}$ les coefficients d'ondelette MODWT à l'échelle j et $\widetilde{V}_{J,t}$ les coefficients d'échelle qui correspondent au niveau maximal de la décomposition, issus de la filtration du processus $x_t, t = 1, \dots, T$. La décomposition de la variance par échelles s'écrit alors :

$$\|x\|^2 = \sum_{j=1}^{J} \|\widetilde{W}_j\|^2 + \|\widetilde{V}_J\|^2. \tag{4.7}$$

Soit $\overline{W}_{j,t}$ le processus aléatoire, résultant de la filtration d'une suite infinie x_t avec le filtre MODWT. La différence entre $\overline{W}_{j,t}$ et $\widetilde{W}_{j,t}$ est que ce dernier est obtenu par filtrage circulaire d'une séquence finie. La variance d'ondelettes $\nu_{j,t}$ du processus x_t pour une échelle de longueur τ_j au moment t est définie comme la variance de $\overline{W}_{j,t}$:

$$\nu_{j,t} = \mathrm{Var}\{\overline{W}_{j,t}\} = \mathrm{E}\{\Big(\overline{W}_{j,t} - \mathrm{E}\{\overline{W}_{j,t}\}\Big)^2\} = \mathrm{E}\{\overline{W}_{j,t}^2\} \tag{4.8}$$

Si la variance d'ondelettes est constante dans le temps, son estimateur naturel est :

$$\widetilde{\nu}_j = \frac{1}{N} \sum_{t=0}^{N-1} \widetilde{W}_{j,t}^2. \tag{4.9}$$

Nous adaptons cet estimateur pour la variance locale, en l'appliquant par fenêtres de taille prédéfinie (journaliers ou mensuels dans nos applications).

4.4 Nouvel indicateur de volatilité

La décomposition en ondelettes permet de calculer les estimations de la variance pour une gamme d'échelles. Ces estimations peuvent être utilisées pour construire des indicateurs synthétiques, en appliquant la transformation de la forme donnée par l'équation (4.4). Mais contrairement au SMS et IMS, décrits dans la section 4.2, nous allons d'abord construire des indicateurs pour différentes échelles séparément.

Afin de réduire le nombre d'horizons, nous regroupons certaines échelles dyadiques afin de définir trois horizons caractéristiques, appelés Court (C), Moyen (M) et Long

(L). Pour $s \in \{C, M, L\}$ la variance synthétique s'écrit :

$$\upsilon_t^s = \sum_{j \in H_s} \nu_{j,t} \tag{4.10}$$

avec $\nu(\tau_j)$ la variance d'ondelettes réalisée pour l'échelle de longueur τ_j, et H_s l'ensemble des échelles, inclues dans la gamme qui correspond à l'horizon synthétique s. L'horizon $s = L$ comprend le plus haut niveau de la décomposition et représente donc la variance réalisée des coefficients d'échelle, correspondant à l'approximation en termes de MODWT. Le MVS à un horizon $s = \{C, M, L\}$ est alors défini par :

$$MVS_t^s = -\log_2 P(\upsilon_t^s) \tag{4.11}$$

avec $P(\upsilon_t^s)$ la probabilité d'observer la variance au-dessus de υ_t^s.

Nous ne tentons pas de calibrer la distribution globale de la volatilité (telle que log-normale). Nous nous concentrons sur la queue de la distribution en estimant la loi de Pareto généralisée (GPD) pour les fluctuations extrêmes. Pour le reste de la distribution, nous utilisons une simple transformée probabiliste empirique. Cela rend l'indicateur plus souple et plus précis.

Une fois que le MVS pour les trois échelles est calculé, on doit spécifier la règle pour détecter les périodes de volatilité extrême. Toute définition de ces périodes est arbitraire, donc nous essayons d'en proposer une intuitivement plausible. Nous définissons deux seuils, notés Θ et Φ, qui sont utilisés pour détecter respectivement le début et la fin d'un événement. Plus précisément, un événement commence lorsque l'une des trois composantes de MVS dépasse Θ et se termine lorsque toutes les trois composantes se retrouvent en-dessous de Φ. Les seuils peuvent être choisis en fonction de la rareté des événements qu'on souhaite étudier. Dans nos applications, nous utilisons deux valeurs de Θ : $\log_2(100)$ et $\log_2(20)$, correspondant à la probabilité de 0.01 et 0.05 dans la queue. Pour Φ, nous utilisons la valeur $\log_2(10)$ pour la probabilité 0.1 dans la queue.

Il est utile non seulement de caractériser la volatilité à des échelles différentes, mais de pouvoir porter un jugement global sur la volatilité du marché, qui représente toutes les échelles. A cette fin, nous devons déterminer l'importance de chaque échelle. La décomposition de la variance dans (4.7) suggère que l'amplitude de la volatilité réalisée à chaque échelle mesure l'impact de ce cette échelle sur la variance globale. Une grande partie de la variance des rendements est attribuée à la plus courte échelle, mais la persistance de la volatilité à des échelles courtes n'est pas nécessairement la même qu'à des échelles plus longues. La persistance peut être caractérisée par la divergence en loi de puissance à l'origine du spectre de la série chronologique :

$$\Theta_x(u) \sim c_\Theta |u|^{-\alpha} \tag{4.12}$$

avec α le paramètre d'échelle et c_Θ une constante. Le paramètre d'échelle est lié à l'exposant de Hurst, noté H, par l'équation :

$$H = \frac{\alpha + 1}{2} \tag{4.13}$$

Si la valeur de H pour un processus stationnaire est entre 0.5 et 1, ce processus possède la mémoire longue. L'exposant de Hurst est estimé à partir des variances d'ondelettes [voir Gençay et al., 2001].

Nous proposons de définir les poids des échelles d'une manière à les rendre dépendants de la partie de la variance totale attribuée à chaque échelle, mais aussi du niveau de la persistance associée. Celle-ci peut être mesurée par l'exposant de Hurst. Notre définition est inspirée par la pratique de l'annualisation de la volatilité, courante dans l'industrie financière : une estimation de la volatilité, calculée pour une courte période de temps, est multipliée par la racine carrée de la longueur d'une période de référence (généralement un an). La racine carrée correspond à l'exposant de Hurst, égal à 0.5 pour le processus du mouvement brownien. Nous trouvons que l'exposant de Hurst est différent entre l'horizon court et les deux autres horizons. Nous utilisons une échelle de référence de 3 mois au lieu de un an, afin de réduire l'impact des erreurs dans l'estimation de l'exposant de Hurst sur le poids des échelles. La définition des poids dans l'indicateur MVS global devient :

$$\begin{aligned} \omega^S &= \frac{A_S R^{H_1}}{A_S R^{H_1} + A_M R^{H_2} + A_L R^{H_2}} \\ \omega^M &= \frac{A_M R^{H_2}}{A_S R^{H_1} + A_M R^{H_2} + A_L R^{H_2}} \\ \omega^L &= \frac{A_L R^{H_2}}{A_S R^{H_1} + A_M R^{H_2} + A_L R^{H_2}} \end{aligned} \tag{4.14}$$

avec $A_s, s \in \{S, M, L\}$ la variance moyenne à l'horizon correspondant, calculé sur une fenêtre suffisamment longue, R la longueur de l'horizon de référence, et $H_{1,2}$ les exposants de Hurst pour différentes échelles.

L'indicateur agrégé, qui regroupe les trois échelles, est ensuite calculé de la manière suivante :

$$MVS_t = \omega^S MVS_t^S + \omega^M MVS_t^M + \omega^L MVS_t^L = -\sum_{s \in S,M,L} \omega^s \log_2 P(v_t^s). \tag{4.15}$$

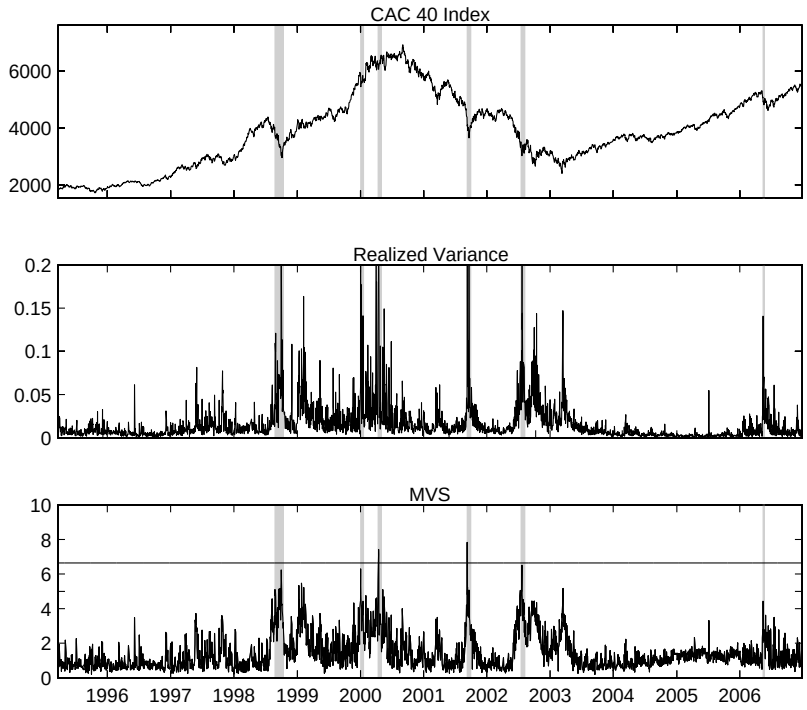
4.5 Application aux données réelles

Dans cette section, nous calculons le MVS pour les données réelles du marché boursier et utilisons l'indicateur pour détecter et caractériser les périodes de la volatilité extrême. Deux séries de données sont utilisées : les valeurs de l'indice CAC40 de 20/03/1995 au 29/12/2006 échantillonnées à des intervalles de 15 minutes (100 881 observations) et les

valeurs quotidiennes de l'indice Dow Jones (DJIA) de 26/05/1896 à 10/10/2007 (28 864 observations). Le premier échantillon est la même que dans Maillet and Michel [2003, 2005], mais étendu jusque'à 2006. Ici nous présentons uniquement les résultats pour l'échantillon CAC40. Les valeurs absolues des rendements logarithmiques sont utilisées pour mesurer la volatilité.

Le Graphique 4.1 présente les valeurs de l'indice CAC40, la variance réalisée et le MVS agrégé, décrit ci-dessus. Les zones colorées en gris correspondent à des périodes de la volatilité très élevée (probabilité inférieure à 0.01), détectées par le MVS. Les informations quantitatives sur ces périodes sont indiquées dans le Tableau 4.1. Le début de chaque période est la date, à laquelle le MVS pour l'une des échelles dépasse le seuil de $\log_2(100) = 6.6439$. La fin de telle période est la date, à laquelle le MVS pour toutes les échelles est en dessous du seuil $\log_2(10) = 3.3219$.

FIGURE 4.1: MVS pour l'indice CAC40



Source : Euronext, les valeurs de l'indice CAC40 de 20/03/1995 à 29/12/2006 aux intervalles de 15 minutes.

Nous détectons six périodes, correspondant à nos critères de la volatilité extrême. La période associée à la crise asiatique (de 27/8/98 au 14/10/98) a été la plus longue (34 jours ouvrés) et la plus sévère, mesurée en termes de la somme des valeurs de MVS agrégé au-dessus du seuil. La troisième composante du MVS qui correspond au long terme est ici la plus importante. Le début de la crise est signalé par la composante de moyen terme. La période 19/7/02-9/8/02 est également caractérisée par des valeurs très

TABLE 4.1: Les périodes de la volatilité extrême, détectées par le MVS pour l'indice CAC40

Start	End	Length	Scale	Maximum MVS				Aggregated MVS			
				A	C	M	L	A	C	M	L
27/8/98	14/10/98	34	M	6.2	6.8	7.5	7.5	143.4	4.8	6.7	125.9
4/1/00	21/1/00	13	CM	6.3	7.2	6.7	4.0	53.6	7.2	6.7	55.3
14/4/00	3/5/00	10	M	7.4	9.5	7.6	4.0	42.2	5.9	7.6	44.4
11/9/01	1/10/01	14	CM	7.8	7.8	10.5	6.5	58.1	7.8	10.5	50.7
19/7/02	9/8/02	15	L	6.5	7.2	6.5	7.2	63.4	4.0	3.8	64.3
17/5/06	24/5/06	5	M	4.4	5.9	7.0	2.3	15.6	5.9	7.0	17.4

La durée d'une période de la volatilité extrême est le nombre de jours ouvrés entre la première et la dernière date. L'échelle correspond à la composante du MVS, qui est la première à franchir le seuil critique (C, M ou L). Les quatre colonnes suivantes présentent les valeurs maximales de l'MVS au cours de la période (pour chaque échelle et en agrégé). Les quatre colonnes suivantes contiennent la somme des valeurs de MVS pour les dates, où une des composantes était au-dessus du seuil $\log_2(100)$, au cours de la période de volatilité extrême.

élevées de la composante de volatilité de long terme, qui a été la première à franchir le seuil critique. Le MVS agrégé est proche du seuil, mais ne le franchit jamais au cours de ces crises, ce qui signifie que les pics de la volatilité aux différentes échelles ne coïncident pas dans le temps.

La crise qui a suivi les événements de 11 Septembre 2001 (11/9/01 - 1/10/01) est la troisième en termes de la somme des valeurs de MVS agrégé, mais elle est structurellement différente des deux crises mentionnée précédemment. La volatilité est la plus élevée aux horizons courts et moyens, clairement parce que le début de la crise a été soudain et provoqué par les facteurs, exogènes au marché financier. Le choc a été assez fort pour influer sur la composante de long terme, mais elle n'atteint pas le seuil critique. Pendant la crise de 2001 on enregistre la valeur la plus élevée de MVS agrégé sur la totalité de l'échantillon (7.8, la probabilité dans la queue environ 0.005). Deux périodes détectées en 2000 sont aussi caractérisées par une forte volatilité sur les échelles de court et moyen terme. Pour la période 14/4/00 - 3/5/00, qui marque le début d'une récession de longue durée sur le marché jusqu'en 2003 et est associé à l'éclatement de la bulle de l'économie dite "nouvelle", la valeur maximale de MVS est 7.4. C'est le deuxième rang après la crise de 2001. Enfin, les fluctuations de mai 2006 sont caractérisées comme extrêmes à horizon moyen seulement. La valeur de MVS agrégé reste en-dessous du seuil critique.

4.6 Conclusion

Nous concevons un nouvel indicateur de la volatilité des marchés boursiers (MVS), qui caractérise les fluctuations des prix dans le temps et dans le domaine spectral. Contrairement aux autres indicateurs multi-échelle, proposés dans la littérature existante, il est basé sur les variances des coefficients de la transformée d'ondelettes, qui sont étroitement liées au spectre de la série temporelle des rendements et permettent une décomposition exacte échelle par l'échelle de la variance. Le MVS est d'abord calculé séparément pour chaque échelle, puis les résultats sont agrégés pour donner un jugement général sur la volatilité du marché à toutes les échelles. Le MVS agrégé prend la forme de la moyenne pondérée des composantes à trois échelles. La pondération dépend de la persistance de la volatilité et de la portion de la variance, attribuée à chaque échelle.

Le MVS est utilisé pour la détection et l'analyse structurelle des périodes de volatilité extrême. En particulier, nous déterminons quelles échelles des fluctuations (court, moyen ou long terme) ont été les plus importantes au cours de chaque période, et à quelle échelle a été initialement franchi le seuil critique, utilisé pour déterminer le début de la période. Les événements peuvent être classés selon la valeur maximale de la MVS à chaque échelle et selon la somme des valeurs MVS sur toute la période de la volatilité extrême.

De toute évidence, le champ d'application potentiel de l'indicateur même et de la décomposition de la variance par échelle va au-delà de la comparaison des événements. Nous mentionnons plusieurs domaines de recherches complémentaires qui sont d'intérêt particulier. Du point de vue théorique, il est important de comprendre les mécanismes de la dynamique des prix à plusieurs échelles. Les variances d'ondelettes peuvent être un outil approprié pour tester les modèles de la microstructure de marché de ce type.

Du point de vue plus pratique, l'analyse multi-échelle de la volatilité peut être suggestive en allocation d'actifs et en gestion de portefeuille. Un indicateur multi-échelle peut être appliqué pour conditionner le processus d'allocation d'actifs du type "changement de régime". Un autre domaine important d'application est la prévision de la volatilité. Dans cette thèse, nous nous sommes concentrés sur la structure plutôt que sur les propriétés dynamiques de la volatilité. Mais l'interdépendance des échelles et la propriété d'additivité des variances d'ondelettes font appel à l'idée de l'autorégression vectorielle échelle par échelle, qui peut être utilisée pour prévoir les composantes de la volatilité et leurs agrégats.

Bibliographie

T. Andersen. Stochastic autoregressive volatility : a framework for volatility modeling. *Mathematical Finance*, 4 :75–102, 1994.

T. Andersen. Return volatility and trading volume : An information flow interpretation of stochastic volatility. *Journal of Finance*, 51(1) :169–204, 1996.

M. Anufriev and G. Bottazzi. Asset pricing model with heterogeneous investment horizons. *Laboratory of Economics and Management Working Paper Series, Sant'Anna School for Advanced Studies*, 2004.

M. Anufriev, G. Bottazzi, and F. Pancotto. Equilibria, stability and ssymptotic dominance in a speculative market with heterogeneous traders. *Journal of Economic Dynamics and Control*, 30(9-10) :1787–1835, 2006.

A. Arneodo, J. Muzy, and D. Sornette. Casual cascade in stock market from the 'infrared' to the 'ultraviolet'. *European Physical Journal B*, 2 :277–282, 1998.

M. Avellaneda, C. Friedmen, R. Holmes, and D. Samperi. Calibrating volatility surfaces via relative-entropy minimization. *Apllied Mathematical Finance*, 4(1) :37–64, 1997.

X. Bai, J. R. Russel, and G. Tiao. Kurtosis of garch and stochastic volatility models with non-normal innovations. *Journal of Econometrics*, 114 :349–360, 2003.

R. Baillie, T. Bollerslev, and O. Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 74(1) :3–30, 1996.

N. Barberis and A. Shleifer. Style investing. *Journal of Financial Economics*, 68(2) : 161–199, 2003.

F. Black. Noise. *Journal of Finance*, 41, 1976.

F. Black and M. Scholes. Pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 1973.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3) :307–327, 1986.

T. Bollerslev and O. Mikkelsen. Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73(1) :151–184, 1996.

T. Bollerslev, R. Chou, and K. Kroner. Arch modeling in finance : A review of the theory and empirical evidence. *Journal of Econometrics*, 52(1-2) :5–59, 1992.

F. Breidt, N. Crato, and P. de Lima. The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83(1-2) :325–34, 1998.

W. Breymann, S. Ghashghaie, and P. Talkner. A stochastic cascade model for fx dynamics. *International Journal of Theoretical and Applied Finance*, 3 :357–360, 2000.

W.A. Brock and C.H. Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22(8-9) :1235–1274, 1998.

E. Capobianco. State-space stochastic volatility models : a review of estimation algorithms. *Applied Stochastic Models and Data Analysis*, 12 :265279, 1996.

B. Castaing, Y. Gagne, and E. J. Hopfinger. Velocity probability density functions of high reynolds number turbulence. *Physica D*, 46(2) :177–200, 1990.

T. Chauveau and R. Topol. A unifying microstructure framework for modeling intraday and interday asset pricing dynamics : the case of exchange rates. *European Financial Management*, 5(3) :341–368, 2002.

M. Chernov, R. Gallant, E. Ghysels, and G. Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2) :225–257, 2003.

C. Chiarella and X.-Z. He. Asset price and wealth dynamics under heterogeneous expectations. *Quantitative Finance*, 1(5) :509 – 526, 2001.

K. Christensen, L. Danon, T. Scanlon, and P. Bak. Unified scaling law for earthquakes. *Physical Review Letters*, 88(17) :178501–504, 2002.

F. Comte and E. Renault. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8(4) :291–323, 1998.

R. Cont. Empirical properties of asset returns : Stylized facts and statistical issues. *Quantitative Finance*, 1(2) :223–236, 2001.

D. Cutler, J. Poterba, and L. Summers. What moves stock prices ? *Journal of Portfolio Management*, 15 :4–12, 1989.

M. Dacorogna, U. Muller, R. Dave, R. Olsen, and O. Pictet. *Modelling Short-term Volatility with GARCH and HARCH models*, pages 161–176. John Wiley, N.-Y., 1998.

Z. Ding and C. Granger. Modeling volatility persistence of speculative returns : a new approach. *Journal of Econometrics*, 73 :185–215, 1996.

Z. Ding, C. Granger, and R. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1) :83–106, 1993.

F. Drost and B. Werker. Closing the garch gap : Continuous time garch modeling. *Journal of Econometrics*, 74(1) :31–57, 1996.

B. Dupire. Pricing and hedging with smiles. *Proceedings of AFFI Conference, La Baule, June 1993*, 1993.

B. Dupire. Pricing with a smile. *Risk Magazine*, 7(1) :18–20, 1994.

R. Engle and T. Bollerslev. Modelling the persistence of conditional variances. *Econometric Reviews*, 5(1) :1 – 50, 1986.

E. Fama. The behaviour of stock market prices. *Journal of Business*, 38 :34–105, 1965.

R. Gençay, F. Selçuk, and B. Whitcher. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Harcourt Brace, San Diego, 2001.

S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge. Turbulent cascades in foreign exchange markets. *Nature*, 381 :767–770, 1996.

A. Harvey. *Long Memory in Stochastic Volatility*, pages 307–320. Butterworth-Heineman, Oxford, 1998.

C. Hommes. Heterogeneous agent models in economics and finance. In *Handbook of Computational Economics*, volume 2, pages 1109–1186, Amsterdam, 2006. Elsevier Science.

N. Jegadeesh and S. Titman. Overreaction, delayed reaction, and contrarian profits. *Review of Financial Studies*, 8(4) :973993, 1995.

AN Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Proceedings of the USSR Academy of Sciences (in Russian)*, 30 :299–303, 1941.

B. LeBaron. Stochastic volatility as a simple generator of apparent financial power laws and long memory. *Quantitative Finance*, 1(6) :621–631, 2001.

B. LeBaron. Agent-based computational finance. In *Handbook of Computational Economics*, volume 2, pages 1187–1233, Amsterdam, 2006. Elsevier Science.

D. Lee, H. Chan, R. Faff, and P. Kalev. Short-term contrarian investing is it profitable ?... yes and no. *Journal of Multinational Financial Management*, 13(4-5) : 385–404, 2003.

I. Lobato and C. Velasco. Long memory in stock market trading volume. *Journal of Business & Economic Statistics*, 18(4) :410–427, 2000.

B. Maillet and T. Michel. An index of market shocks based on multiscale analysis. *Quantitative Finance*, 3(2) :88–97, 2003.

B. Maillet and T. Michel. The impact of the 9/11 events on the american and french stock markets. *Review of International Economics*, 13(3) :597–611, 2005.

B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36 : 394–419, 1963.

M. Mega, P. Allegrini, P. Grigolini, V. Latora, L. Palatella, A. Rapisarda, and S. Vinciguerra. Power law time distributions of large earthquakes. *Physical Review Letters*, 90 :188501, 2003.

R. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4 :141–183, 1973.

U. Müller, M. Dacorogna, R. Dave, R. Olsen, O. Pictet, and J. Von Weizsacker. Volatilities of different time resolutions - analyzing the dynamics of market components. *Journal of Empirical Finance*, 4 :213–239, 1997.

B. Ninness. Strong laws of large numbers under weak assumptions with application. *IEEE Transactions on Automatic Control*, 45(11) :2117–2122, 2000.

S.B. Pope. *Turbulent flows*. Cambridge university press, 2000.

H. Simon. From substantive to procedural rationality. In S. J. Latsis, editor, *Method and Appraisal in Economics*, pages 129–148. Cambridge University Press, 1976.

S. Taylor. *Financial Returns Modelled by the Product of two Stochastic Processes a Study of the Daily Sugar Prices 1961-75*, volume 1, pages 203–226. North-Holland, Amsterdam, 1982.

S. Taylor. Modeling stochastic volatility : A review and comparative study. *Mathematical Finance*, 4(2) :183–204, 1994.

J. Vanden. Equilibrium analysis of volatility clustering. *Journal of Empirical Finance*, 12 :374–417, 2005.

D. Weinbaum. Investor heterogeneity, asset pricing and volatility dynamics. *Journal of Economic Dynamics and Control*, 33 :1379–1397, 2009.

G. Zumbach. Volatility processes and volatility forecast with long memory. *Quantitative Finance*, 4(1) :70 – 86, 2004.

G. Zumbach, M. Dacorogna, J. Olsen, and R. Olsen. Measuring shocks in financial markets. *International Journal of Theoretical and Applied Finance*, 3(3) :347–355, 2000.