Univerzita Karlova v Praze
Matematicko-Fyzikální Fakulta

DIPLOMOVÁ PRÁCE

Róbert Chovanec

Zobecněný stabilní model ve financích

Katedra pravděpodobnosti a matematické statistiky
Vedoucí diplomové práce: Prof. Lev Borisovich Klebanov, DrSc.
Studijní program: Matematika
Studijní obor: Pravděpodobnost, matematická statistika a ekonometrie

Thanks: As the author of this text, on the first place I would like to thank to my supervisor, Prof. Lev Borisovich Klebanov, DrSc., for his great help and patience. I am very grateful to my parents for their lifetime support. For the help with the English grammar, I am thankful to Mgr. Michaela Roubalová.

I certify that I have written my diploma thesis just by myself and only by using the cited literature. I agree with lending of this work.

Prehlasujem, že som svoju diplomovú prácu napísal samostatne a výhradne s použitím citovaných prameňov. Súhlasím so zapožičiavaním práce.

In Prague on August 9, 2007 Róbert Chovanec

Contents

List of Figures **6**

List of Tables **7**

List of Abbreviations **8**

1 Introduction **9**

2 Univariate Stable Distribution **12**
2.1 Summation-Stable Distribution 12
2.1.1 Properties of Stable Random Variables 15
2.2 Geometric Stable Distribution 20

3 ARMA Sequences with Stable Innovations **23**

4 The Gaussian vs. Stable Paretian Distribution in ARCH/-GARCH Models **28**
4.1 ARCH and GARCH Models in Gaussian Distribution 28
4.2 ARCH and GARCH Models in non-Gaussian Distribution 30
4.2.1 Symmetric GARCH-stable Processes 31
4.2.2 Asymmetric GARCH-stable Processes 33

5 An Empirical Application **36**
5.1 Statistical Software 36
5.2 Data 36

6 Unconditional Fitting **40**

7 Conditional Gaussian Fitting **42**
7.1 Homoskedastic Model 42
7.2 Heteroskedastic Model 46

8 Conditional Stable Fitting **51**
8.1 Simulation of Stable Distributed Innovations 51
8.1.1 $AL^*(\kappa, \sigma)$ 52
8.1.2 A Strictly $GS_\alpha(\lambda, \tau)$ Generator 52
8.1.3 General $GS_\alpha(c, \beta, \delta)$ Generator 53
8.1.4 A Standard Stable $S_\alpha(1, \beta, 0)$ Generator 53
8.1.5 A Standard Stable $S_\alpha(c, \beta, \delta)$ Generator 54
8.2 Homoskedastic Model 54
8.3 Heteroskedastic Model 57

9 Conclusions **59**

Bibliography **61**

Abstrakt

Názov práce: Zobecnený stabilný model vo financiách
Autor: Róbert Chovanec
Katedra (ústav): Katedra pravdepodobnosti a matematickej štatistiky
Vedúci diplomovej práce: Prof. Lev Borisovich Klebanov, DrSc.
E-mail vedúceho: Lev.Klebanov@mff.cuni.cz
Abstrakt: V tejto práci je popísaný základný teoretický prístup k stabilnému rozdeleniu. Sú tu uvedené definície stabilných rozdelení, vlastnosti a správanie stabilne rozdelených náhodných veličín. Ďalej je tu rozobraté podmienené modelovanie za platnosti stabilných zákonov. Dajú sa tu nájsť homoskedastické (ARMA) ako aj heteroskedastické (GARCH) štruktúry. GARCH modely sú vysvetlené čiastočne aj pre prípad Gaussovho rozdelenia. Empirická časť tejto práce je založená na porovnávaní medzi modelmi, zavedenými v teoretickej časti, pod normálnym resp. stabilným rozdelením, budovaných na reálnych dátach z prostredia energetiky. Vychádza sa z nepodmieneného modelovania, potom sa prechádza na podmienené ARMA modely a nakoniec kombinované ARMA-GARCH modely. Výsledky prevedenej štatistickej analýzy ukazujú, že modely založené na stabilnom rozdelení vystihujú empirické rozdelenie lepšie ako modely založené na Gaussovom rozdelení.
Kľúčové slová: Stabilné rozdelenie, ARMA-GARCH v stabilnom rozdelení

Abstract

Title: General stable model in finance
Author: Róbert Chovanec
Department: Department of probability and mathematical statistics
Supervisor: Prof. Lev Borisovich Klebanov, DrSc.
Supervisor's e-mail address: Lev.Klebanov@mff.cuni.cz
Abstract: In this contribution, a basic theoretical approach to stable laws is described. There are mentioned some definitions of the stable distributions, properties and behavior of stable distributed random variables. Next, conditional modeling under the stable laws are analyzed. One can find homoskedastic (ARMA) and heteroskedastic (GARCH) structures. The GARCH models are explained partly for the Gaussian case too. An empirical application of this paper is based on comparison between the models, established in theoretical part, under the normal, and stable distribution respectively, built on real data from energetics. One issues from unconditional, then continues with conditional ARMA and finally, there are mixed ARMA-GARCH models. The results of interpreted statistical analysis demonstrate that the models based on the stable distribution matched the empirical distribution better than the the models based on the Gaussian distribution.
Keywords: Stable distribution, ARMA-GARCH in stable laws

List of Figures

2.1 Probability Density Functions for Standard Symmetric α-Stable Random Variables with Different Values of Parameter α . . . 14

5.1 January Levels of Time Series 37
5.2 January Returns of Time Series 38
5.3 Empirical Density Function of Returns 38

6.1 Fitted Unconditional Normal and α-Stable Densities 41

7.1 Sample Correlation Function of Returns 43
7.2 Sample Partial Correlation Function of Returns 44
7.3 Normal ARMA(2,1) Residuals 46
7.4 Sample Correlation Function of Squared Returns 47
7.5 Sample Partial Correlation Function of Squared Returns . . . 47
7.6 SACF of Conditional Heteroskedastic Residuals 49
7.7 SACF of Conditional Heteroskedastic Squared Residuals . . . 49
7.8 Fitted Normal ARMA(2,1)-GARCH(1,1) Densities 50

8.1 ARMA(2,1) Driven by α-Stable (steel blue line) and Normal (red line) Innovations 55
8.2 ARMA(2,1) Driven by GS (steel blue line) and Normal (red line) Innovations 56
8.3 ARMA(2,1)-GARCH(1,1) Driven by α-Stable (steel blue line) and Normal (red line) Innovations 58

List of Tables

5.1 Statistical Properties of Returns 39

6.1 ML Estimates of Unconditional Distributions 40

7.1 Parameter Estimates of ARMA(2,1) 45
7.2 ML Estimates of Conditional Homoskedastic Normal Distribution 45
7.3 Parameter Estimates of ARMA(2,1)-GARCH(1,1) 48
7.4 ML Estimates of Conditional Heteroskedastic Normal Distribution 50

8.1 ML Estimates of Conditional Homoskedastic Stable Distributions 56
8.2 ML Estimates of Conditional Heteroskedastic Stable Distributions 58

List of Abbreviations

AIC Akaike Information Criterion

AR Auto Regressive

ARMA Auto Regressive Moving Average

ARCH Auto Regressive Conditionally Heteroskedastic

a.s. Almost Surely

d.f. Distribution Function

FFT Fast Fourier Transformation

GARCH Generalized Auto Regressive Conditionally Heteroskedastic

GS Geometric Stable

IGARCH Integrated GARCH

ch.f. Characteristic Function

i.i.d. Independently, Identically Distributed

MA Moving Average

ML Maximum-Likelihood

Q.E.D. Quod Erat Demonstrandum

r.v.'s Random Variables

s.e. Standard Error

$\boldsymbol{S_\alpha S}$ Symmetric α-Stable

$\boldsymbol{S_\alpha S}$ **GARCH** Symmetric α-Stable GARCH

SACF Sample Autocorrelation Function

SPACF Sample Partial Autocorrelation Function

Chapter 1

Introduction

Modern finance still relies heavily on the assumption that the random variables under investigation follow a normal distribution. Distributional assumption for financial processes have important theoretical implications. Hence, solutions to such problems like portfolio selection, option pricing and risk management depend critically on distributional specifications.

However, time series observed in finance often deviate from the Gaussian model in that their marginal distributions are heavy-tailed and possibly asymmetric. In such situations, the appropriateness of the commonly adopted normal assumption is highly questionable. This is especially true in financial modeling.

Benoit Mandelbrot's fundamental work [26] in the 1960s strongly rejected normality as a distributional model for asset returns (see also Mandelbrot [25] and [27]). Examining various time series on commodity returns and interest rates, he conjectured that financial return processes behave like non-Gaussian stable processes. To distinguish between Gaussian and non-Gaussian stable distributions, the latter are often referred to as "stable Paretian" or "Lévy stable".

Only stable distributed returns encompass the property that linear combinations of different return series (e.g., portfolios) follow again a stable distribution (see Section 2.1.1). Indeed, the Gaussian law shares this feature, but it is only one particular member of a large and flexible class of distributions, which also allows for skewness and heavy-tailedness. Mandelbrot's contributions give rise to a new probabilistic foundation for financial theory and empirics.

An attractive distinctive feature of stable models – not shared by other distributional models – is that they allow to generalize Gaussian-based financial theories and thus, to build a coherent and more general framework for financial modeling. The generalizations are only possible because of specific probabilistic properties that are unique to stable laws, namely, the stability property, the Central Limit Theorem and the Invariance Principle for Lévy-stable processes. The central limit theorem provides a theoretically sound explanation for its emergence: whenever a financial variable can be regarded as the result of many microscopic effects, it can be described by a stable law, as this describes the fundamental "building blocks" (e.g., innovations) that drive asset return processes. In addition to describing these "building blocks", a complete financial model should be rich enough to encompass relevant stylized facts such as:

- Non-Gaussian, heavy-tailed and skewed distributions;
- Volatility clustering (ARCH-effects);
- Temporal dependence of the tail behavior;
- Short and long-range dependence.

Among many desirable properties, stable models are also highly versatile. They have many possible applications, ranging from equilibrium asset pricing to risk management. Moreover, with the current availability of computational power, stable models do not present serious numerical difficulties and should quickly repay the time and effort spent in their implementation.

An important desirable property is the fact that stable Paretian distributions have *domains of attraction* (see Section 2.1). In general, any decision (inference) based on observed data is a functional on the space of distributions that govern the data. Loosely speaking, any distribution in the domain of attraction of a specified stable distribution will have properties which are close to those of the stable distribution. Consequently, decisions will, in principle, not be affected by adopting an "idealizing" stable distribution as the distributional model instead of the true one.

A next attractive aspect of the stable Paretian assumption is the *stability property*. This is desirable because it implies that each stable distribution has an index of stability (shape parameter - see Definition 2.2), which remains the same regardless of the scale adopted. The index of stability can be regarded as an overall parameter, which can be employed for inference and decision making.

The main goal of this text is to approach stable theory and then to examine it on real data. The described theory is implemented on the concrete models, which are built and fitted over the data[1]. The consequences of relaxing the normality assumption are investigated and at last the models driven by normal and stable innovations are compared.

Generally speaking, this contribution is divided in three main parts. The first one describes basic definitions and properties of the stable theory, and answers the question of what are the stable laws for (see Chapter 2). The second part enlarges on theory of conditional modeling, such ARMA (see Chapter 3) and GARCH (see Chapter 4) models. Particularly GARCH models under stable assumptions are described in detail.

The first two parts build a basis for the third main part, simulation and the real data processing (see Chapter 5, 6, 7 and 8).

[1] In practice one cannot expect that observed data follow exactly the "ideal" distribution specified by the modeler. The distributional model represents only an approximation of the distribution underlying the observed data. This problem gives rise to the crucial question of what is the domain of applicability of the specified model.

Chapter 2

Univariate Stable Distribution

As it was discussed in the Introduction, in a search for satisfactory descriptive models for financial return data, the stable laws as a model for financial returns distribution were proposed. First the class of Paretian distributions is represented.

2.1 Summation-Stable Distribution

To define stable Paretian laws, suppose $X_1, X_2, \ldots$ are independently and identically distributed (i.i.d.), real-valued random variables (r.v.'s) with common distribution function (d.f.) $\mathcal{H}$. It is assumed that $\mathcal{H}$ is nondegenerate.

Definition 2.1 (Zolotarev [40] or Samorodnitsky and Taqqu [35]) *The d.f. $\mathcal{H}$ is said to be* ***stable*** *if there exist constants $a_n > 0$ and $b_n \in \boldsymbol{R}$, such that for any n*

$$a_n(X_1 + \cdots + X_n) + b_n \stackrel{\mathcal{D}}{=} X_1.^{1} \tag{2.1}$$

D.f. $\mathcal{H}$ is said to be ***strictly stable*** *if (2.1) holds with $b_n = 0$. A stable d.f. is called* **symmetric** *if $\mathcal{H}(x) = 1 - \mathcal{H}(-x)$.*

Note that a symmetric stable $\mathcal{H}$ is also strictly stable.

Definition 2.2 (Explicit representation of characteristic function) *The d.f. $\mathcal{H}$ is stable if there are parameters $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $0 \leq c$ and $\delta \in \boldsymbol{R}$ such that the characteristic function (ch.f.) of $\mathcal{H}$ has the following form:*

$$\int e^{itx} d\mathcal{H}(x) = \begin{cases} \exp\left\{-c^{\alpha}|t|^{\alpha}\left[1 - i\beta \text{sign}(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + i\delta t\right\}, & \alpha \neq 1, \\ \exp\left\{-c|t|\left[1 + i\beta\frac{2}{\pi}\text{sign}(t)\ln|t|\right] + i\delta t\right\}, & \alpha = 1, \end{cases} \tag{2.2}$$

[1]Notation "$\stackrel{\mathcal{D}}{=}$" stands for equality in distribution.

where $\operatorname{sign}(t)$ *is* 1 *if* $0 < t$, 0 *if* $t = 0$ *and* -1 *if* $t < 0$.

The parameters are unique (β is irrelevant when $\alpha = 2$). The characteristic exponent α is the *index of stability* and can also be interpreted as a *shape* parameter (or tail-index parameter), β is the *skewness* parameter, δ is a *location (shift)* parameter and c is the *scale* parameter. $\mathcal{H}$ is called **stable Paretian** or **α-stable** and is usually denoted by $S_\alpha = S_\alpha(c, \beta, \delta)$ and one writes

$$X \sim S_\alpha(c, \beta, \delta)$$

to indicate that X has the stable distribution $S_\alpha(c, \beta, \delta)$. One also writes

$$X \sim S_\alpha S$$

when X is symmetric α-stable ($\beta = \delta = 0$).

If $\alpha = 2$, the α-stable distribution coincides with the normal distribution – in this case c is proportional to the standard deviation σ, β can be taken to be zero and δ is the mean μ. If $\alpha \in (0, 2)$, it is fat tailed and has only moments of orders less then α. The tail thickness increases as α decreases. So, the distributions have infinite variance and when $\alpha \leq 1$, they have an infinite mean as well. If $\alpha < 2$ and $\beta \neq 0$, the distribution is asymmetric and the skewness increases as β moves away from 0 to ± 1. For example, when $\beta = 1$, the d.f. $\mathcal{H}$ or the stable r.v. with d.f. $\mathcal{H}$ is said to be totally right skewed.

See Figure 2.1 for a detailed exposition of the theory of stable laws. The solid steel blue line corresponds to $\alpha = 2$, the dotdashed magenta line corresponds to $\alpha = 1.5$, the dashed black line corresponds to $\alpha = 1$ and the dotted dark red line corresponds to $\alpha = 0.5$. The rest of the coefficients is fixed ($\beta = 0$, $c = 1$ and $\delta = 0$).

The probability densities of α-stable random variables exist and are continuous, but, with a few exceptions, they are not known in closed form (Zolotarev [40]). The exceptions are:

(a) the *Gaussian distribution* $S_2(\sigma, 0, \mu) = N(\mu, 2\sigma^2)$, whose density is

$$\frac{1}{2\sqrt{\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{(2\sigma)^2}},$$

(b) the *Cauchy distribution* $S_1(\sigma, 0, \mu)$, whose density is

$$\frac{\sigma}{\pi((x-\mu)^2 + \sigma^2)},$$

if $X \sim S_1(\sigma, 0, 0)$, then for $0 < x$ is $P(X \leq x) = \frac{1}{2} + \frac{1}{\pi}\arctan(\frac{x}{\sigma})$,

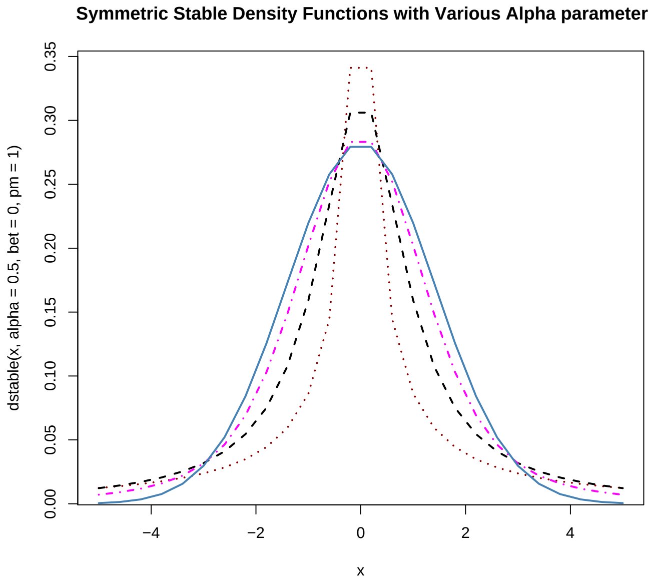

Figure 2.1: Probability Density Functions for Standard Symmetric α-Stable Random Variables with Different Values of Parameter α

(c) the *Lévy distribution* $S_{\frac{1}{2}}(\sigma, 1, \mu)$, whose density

$$(\frac{\sigma}{2\pi})^{1/2}\frac{1}{(x-\mu)^{3/2}}\exp\left\{-\frac{\sigma}{2(x-\mu)}\right\}$$

is concentrated on (μ, ∞),

if $X \sim S_{\frac{1}{2}}(\sigma, 1, 0)$, then for $0 < x$ is $P(X \leq x) = 2(1 - \Phi(\sqrt{\frac{\sigma}{x}}))$, where Φ denotes the cumulative distribution function of the $N(0, 1)$,

(d) a *constant* μ, which has the degenerate distribution $S_\alpha(0, 0, \mu)$ for any $0 < \alpha \leq 2$. A convention of this paper is to exclude degenerate distributions, because they have unusual properties. For example, all moments of a degenerate distribution are finite, whereas a non-degenerate α-stable distribution with $0 < \alpha < 2$ has infinite second moments.

Definition 2.3 (Domain of attraction of stable Paretian distributions)
The d.f. $\mathcal{H}$ is in **domain of attraction** *of the α-stable distribution S_α if for any sequence $X_1, X_2, \ldots$ of i.i.d. r.v.'s with common d.f. $\mathcal{H}$, there exist*

sequences of constants $a_n > 0$ *and* $b_n \in \boldsymbol{R}$ *such that*

$$Z_n \stackrel{df}{=} a_n(X_1 + \cdots + X_n) - b_n \xrightarrow{\mathcal{D}} Y_\alpha,^2 \tag{2.3}$$

where Y_α *in an* S_α*-distributed r.v.*

It follows from Definition 2.1 that S_α belongs to its own domain of attraction. Relation (2.1) states that if X_k is assumed to be α-stable distributed, then under certain normalization, the sum $X_1 + \cdots + X_n$ has the same distribution. On the other hand, (2.3) states that if the d.f. of X_k belongs to the domain of attraction of S_α the normalized sum Z_n is asymptotically S_α-distributed.

2.1.1 Properties of Stable Random Variables

A useful tool for studying α-stable distributions is the characteristic function (2.2). It will be used to derive some basic properties of stable random variables and to obtain an interpretation of the parameters α, c, β and δ.

Property 2.4 *Let* X_1 *and* X_2 *be independent random variables with* $X_i \sim S_\alpha(c_i, \beta_i, \delta_i)$, $i = 1, 2$. *Then* $X_1 + X_2 \sim S_\alpha(c, \beta, \delta)$, *with*

$$c = (c_1^\alpha + c_2^\alpha)^{1/\alpha}, \ \beta = \frac{\beta_1 c_1^\alpha + \beta_2 c_2^\alpha}{c_1^\alpha + c_2^\alpha}, \ \delta = \delta_1 + \delta_2.$$

PROOF:

- for $\alpha \neq 1$. By independence and from (2.2) follows:

$$\begin{aligned} &\ln(\mathrm{E}[\exp\{it(X_1 + X_2)\}]) = \ln(\mathrm{E}[\exp\{itX_1\}]) + ln(\mathrm{E}[\exp\{itX_2\}]) \\ &= \ -(c_1^\alpha + c_2^\alpha)|t|^\alpha + i|t|^\alpha \mathrm{sign}(t) \tan(\pi\alpha/2)(\beta_1 c_1^\alpha + \beta_2 c_2^\alpha) + it(\delta_1 + \delta_2) \\ &= \ -(c_1^\alpha + c_2^\alpha)|t|^\alpha \left[1 - i\frac{\beta_1 c_1^\alpha + \beta_2 c_2^\alpha}{c_1^\alpha + c_2^\alpha} \mathrm{sign}(t) \tan(\pi\alpha/2)\right] + it(\delta_1 + \delta_2) \end{aligned}$$

- the proof for $\alpha = 1$ is similar.

Q.E.D.

The parameter δ is a *shift* parameter because of:

[2]Notation "$\stackrel{df}{=}$" stands for "is defined", respectively "is set to be equal" and notation "$\xrightarrow{\mathcal{D}}$" stands for convergence in distribution.

Property 2.5 *Let* $X \sim S_\alpha(c, \beta, \delta)$ *and let* b *be a real constant. Then* $X + b \sim S_\alpha(c, \beta, \delta + b)$.

PROOF:
This follows trivially from the form of the characteristic function (2.2).
Q.E.D.

The parameter c is called the *scale* parameter. Observe that when $\alpha = 1$, multiplication by a constant affects the shift parameter in a non-linear way. Hence, the name "scale parameter" for c is a misnomer when $\alpha = 1$ and $\beta \neq 0$.

Property 2.6 *Let* $X \sim S_\alpha(c, \beta, \delta)$ *and let* b *be a non-zero real constant. Then*

$$\begin{array}{ll} bX \sim S_\alpha(|b|c, \mathrm{sign}(b)\beta, b\delta) & if\ \alpha \neq 1 \\ bX \sim S_1(|b|c, \mathrm{sign}(b)\beta, b\delta - \frac{2bc\beta}{\pi}\ln|b|) & if\ \alpha = 1 \end{array}$$

PROOF:

- By the ch.f. (2.2), it is for $\alpha \neq 1$

$$\begin{aligned} &\ln(\mathrm{E}[\exp\{it(bX)\}\,]) \\ &= -|bt|^\alpha c^\alpha(1 - i\beta \mathrm{sign}(bt)\tan(\pi\alpha/2)) + i\delta(bt) \\ &= -(c|b|)^\alpha |t|^\alpha (1 - i\beta \mathrm{sign}(b)\mathrm{sign}(t)\tan(\pi\alpha/2)) + i(\delta b)t. \end{aligned}$$

- the proof for $\alpha = 1$ is similar.

Q.E.D.

The following property identifies β as a skewness parameter:

Property 2.7 $X \sim S_\alpha(c, \beta, \delta)$ *is symmetric if and only if* $\beta = 0$ *and* $\delta = 0$. *It is symmetric about* δ *if and only if* $\beta = 0$.

PROOF: For a random variable to be symmetric, it is necessary and sufficient that its ch.f. is real. By (2.2), this is the case if and only if $\beta = 0$ and $\delta = 0$. The second statement follows from Property (2.5).
Q.E.D.

A symmetric stable random variable is strictly stable, but a strictly stable random variable is not necessarily symmetric. In fact:

Property 2.8 *Let* $X \sim S_\alpha(c, \beta, \delta)$ *with* $\alpha \neq 1$. *Then* X *is strictly stable if and only if* $\delta = 0$.

PROOF: Let X_1, X_2 be independent copies of X and let A and B be arbitrary positive constants. By Properties (2.4) and (2.6),

$$AX_1 + BX_2 \sim S_\alpha(c(A^\alpha + B^\alpha)^{1/\alpha}, \beta, \delta(A + B)).$$

However, the equation

$$AX_1 + BX_2 \stackrel{\mathcal{D}}{=} CX + D \tag{2.4}$$

is equivalent to equation from Definition (2.1). Set $C \stackrel{df}{=} (A^\alpha + B^\alpha)^{1/\alpha}$ in (2.4). By Properties (2.5) and (2.6),

$$CX + D \sim S_\alpha(c(A^\alpha + B^\alpha)^{1/\alpha}, \beta, \delta(A^\alpha + B^\alpha)^{1/\alpha} + D)$$

and therefore $D = 0$ holds for (2.4) if and only if $\delta = 0$.
Q.E.D.

Property 2.9 *Let $X \sim S_\alpha(c, \beta, \delta)$ with $\alpha \neq 1$. Then $X - \delta$ is strictly stable.*

PROOF: It follows directly from Properties (2.5) and (2.8). Q.E.D.

Thus, any α-stable random variable with $\alpha \neq 1$ can be made strictly stable by shifting. This is not true when $\alpha = 1$, as the next property indicates.

Property 2.10 *$X \sim S_1(c, \beta, \delta)$ is strictly stable if and only if $\beta = 0$.*

PROOF: Let X_1, X_2 be independent copies of X and let A and B be arbitrary positive constants. Then by Properties (2.4) and (2.6),

$$AX_1 + BX_2 \sim S_1((A + B)c, \beta, (A + B)\delta - \frac{2}{\pi}c\beta(A\ln(A) + B\ln(B))),$$

whereas

$$(A + B)X \sim S_1((A + B)c, \beta, (A + B)\delta - \frac{2}{\pi}c\beta(A + B)\ln(A + B)).$$

Therefore $D = 0$ in (2.4) if and only if $AX_1 + BX_2 \stackrel{\mathcal{D}}{=} (A + B)X$, i.e., if and only if

$$\beta(A\ln(A) + B\ln(B)) = \beta(A + B)\ln(A + B)$$

for any $0 < A$ and $0 < B$. It is thus necessary and sufficient that $\beta = 0$.
Q.E.D.

The parameter δ is the least important of the four parameters α, c, β and δ, because it affects only location. It is often assumed for simplicity that $\delta = 0$.

Let now focus on the skewness parameter β. For any $0 < \alpha < 2$ holds $X \sim S_\alpha(c, -\beta, 0) \Longleftrightarrow -X \sim S_\alpha(c, \beta, 0)$. The distribution $S_\alpha(c, \beta, 0)$ is said to be skewed to the right if $0 < \beta$ and to the left if $\beta < 0$. It is said to be *totally skewed* to the right if $\beta = 1$ and *totally skewed* to the left if $\beta = -1$. Random variables that are totally skewed to the right can be regarded as basic building blocks because of the following:

Theorem 2.11 *Let X has distribution $S_\alpha(c, \beta, 0)$ with $\alpha < 2$. Then there exist two i.i.d. random variables Y_1 and Y_2 with common distribution $S_\alpha(c, 1, 0)$ such that*

$$X \stackrel{\mathcal{D}}{=} \left(\frac{1+\beta}{2}\right)^{1/\alpha} Y_1 - \left(\frac{1-\beta}{2}\right)^{1/\alpha} Y_2$$

if $\alpha \neq 1$, and

$$X \stackrel{\mathcal{D}}{=} \left(\frac{1+\beta}{2}\right) Y_1 - \left(\frac{1-\beta}{2}\right) Y_2 + c\left(\frac{1+\beta}{\pi} \ln \frac{1+\beta}{2} - \frac{1-\beta}{\pi} \ln \frac{1-\beta}{2}\right)$$

if $\alpha = 1$.

PROOF:
This is a direct consequence of Properties (2.4), (2.5) and (2.6).
Q.E.D.

Note also that the support of $S_\alpha(c, \beta, 0)$ is the whole real line even for $\beta = \pm 1$. However, the tails of the distribution are affected by the skewness parameter β as the next Property (2.12) below indicates. The Property (2.12) concerns the asymptotic behavior of the tail probabilities $P(X > \lambda)$ and $P(X < -\lambda)$ as $\lambda \to \infty$.

In the Gaussian case ($\alpha = 2$),

$$P(X < -\lambda) = P(X > \lambda) \sim \frac{1}{2\sqrt{\pi}\sigma\lambda} e^{-\lambda^2/(4\sigma^2)}$$

as $\lambda \to \infty$, (Feller [14]). However, when $\alpha < 2$, the tail probabilities behave like $\lambda^{-\alpha}$.

Property 2.12 *Let $X \sim S_\alpha(c, \beta, \delta)$ with $0 < \alpha < 2$. Then*

$$\begin{aligned} lim_{\lambda \to \infty} \lambda^\alpha P(X > \lambda) &= C_\alpha \tfrac{1+\beta}{2} c^\alpha, \\ lim_{\lambda \to \infty} \lambda^\alpha P(X < -\lambda) &= C_\alpha \tfrac{1-\beta}{2} c^\alpha, \end{aligned}$$

where

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx\right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} & \text{if } \alpha \neq 1, \\ 2/\pi & \text{if } \alpha = 1. \end{cases}$$

PROOF:
One can apply a central limit theorem type argument, as in Feller [15, Theorem XVII.5.1.]
Q.E.D.

The tail behavior in Property (2.12) is widely used property of α-stable distributions.

Since $\mathrm{E}|X|^r = \int_0^\infty P(|X|^r > \lambda)d\lambda$, one has:

Property 2.13 *Let* $X \sim S_\alpha(c, \beta, \delta)$ *with* $0 < \alpha < 2$. *Then*

$$\begin{array}{ll} \mathrm{E}|X|^p < \infty & \text{for any } 0 < p < \alpha, \\ \mathrm{E}|X|^p = \infty & \text{for any } \alpha \leq p. \end{array}$$

The fact that α-stable random variables with $\alpha < 2$ have an infinite second moment means that many of the techniques valid for the Gaussian case do not apply. An added complication stems from the fact that even $\mathrm{E}|X|^\alpha$ is infinite. When $\alpha \leq 1$, one also has $\mathrm{E}|X| = \infty$, precluding the use of expectations.

Note that for fixed c and p and $\alpha \neq 1$, $\mathrm{E}|X|^p$ is an even function of β and increases in $|\beta|$. Note also that if $X \sim S_\alpha S$, $0 < \alpha \leq 2$, then the moment $\mathrm{E}|X|^p$ determines the scale parameter c of X, and therefore the whole distribution.

Property 2.14 *When* $1 < \alpha \leq 2$, *the shift parameter* δ *equals the mean.*

PROOF: Let $X \sim S_\alpha(c, \beta, \delta)$, $1 < \alpha \leq 2$. The r.v. X has finite mean (by Property 2.13 in the case $1 < \alpha < 2$ and because X is Gaussian when $\alpha = 2$). Moreover, $X - \delta$ is strictly stable by Property 2.9. Let X_1 and X_2 be two independent copies of X. By (2.4) the relation

$$A(X_1 - \delta) + B(X_2 - \delta) \stackrel{\mathcal{D}}{=} (A^\alpha + B^\alpha)^{1/\alpha}(X - \delta)$$

holds for any positive A and B. Taking expectations of both sides gives

$$A(\mathrm{E}X - \delta) + B(\mathrm{E}X - \delta) = (A^\alpha + B^\alpha)^{1/\alpha}(\mathrm{E}X - \delta),$$

and thus $\mathrm{E}X = \delta$.
Q.E.D.

2.2 Geometric Stable Distribution

The motivation underlying these geometric analogues is to enable to model processes that may, with some small probability, change in each period. In the context of asset returns we can, for example, think of major, unexpected new events, which could occur in any period and drastically affect investor's behavior and hence, the overall market.

To state this formally, let X_i denote the r.v. at period $t = t_0 + i$. R.v.'s $\{X_i\}$ are assumed to be i.i.d. with d.f.

$$\mathcal{H}(u) = P(X_i \leq u), \quad u \in \boldsymbol{R}. \tag{2.5}$$

With probability $p \in (0, 1)$ an investor may expect in any period the occurrence of an event altering the characteristics of the underlying asset-return process. Let $T(p)$ denote the period in which such an event is expected to occur. $T(p)$ is assumed to be independent on $\{X_i\}$ and to have a *geometric distribution*, i.e.,

$$P(T(p) = k) = (1-p)^{k-1}p, k = 1, 2, \ldots. \tag{2.6}$$

The time until random period $T(p)$ may be viewed as the time, for which an investor assumes the market fundamentals to remain unchanged and correspond to the investment horizon.

In the following one replaces the deterministic variable n in the probabilistic schemes considered in Section 2.1 by the geometric random variable $T(p)$ and investigates their properties given by the geometric randomization. The geometric sum

$$G_p \stackrel{df}{=} \sum_{i=1}^{T(p)} X_i \tag{2.7}$$

represents the accumulation of the X_i's up to the event at time $t_0 + T(p)$, i.e., the total return of an asset over that period. Operational time is non-random in the geometric-summation model, i.e., the number of X_i's per unit of calendar time is fixed, but the lengths of the time intervals are random (i.e., $T(p)$ is a random variable).

Definition 2.15 *D.f. $\mathcal{H}$ is said to be strictly geometric-stable with respect to the summation scheme – in short,* **strictly (geo,sum)-stable** *if for any $p \in (0, 1)$, there exist constants $a = a(p) > 0$, such that*

$$aG_p \stackrel{\mathcal{D}}{=} X_1. \tag{2.8}$$

According to Klebanov et al. [21], one has:

Definition 2.16 (General geometric stable distribution) *A random variable Y is said to be geometric stable with respect to the summation scheme (in short, geometric stable, or GS) if there exists a sequence of i.i.d. random variables $X_1, X_2, \ldots$, a geometric (2.6) random variable $T(p)$ independent of all X_i and constants $a = a(p) > 0$ and $b = b(p) \in \boldsymbol{R}$ such that*

$$a(p)\sum_{i=1}^{T(p)}(X_i + b(p)) \xrightarrow{\mathcal{D}} Y, \quad \text{as } p \to 0. \tag{2.9}$$

If $f(t)$ is the ch.f. of $\mathcal{H}$ (i.e., $f(t) = \mathrm{E}\left[\exp\{itX_1\}\right]$), then by (2.8) it follows that $f(t)$ satisfies

$$f(t) = \frac{pf(at)}{1-(1-p)f(at)}, \tag{2.10}$$

see Klebanov et al. [19] and [20]. In this case, setting $\varphi(t) \stackrel{df}{=} \exp\{1 - 1/f(t)\}$ in (2.10), one has

$$\varphi(t) = [\varphi(at)]^{1/p} \quad p \in (0,1), \tag{2.11}$$

i.e., φ is the ch.f. of a strictly stable distribution. This gives rise to the following result of Klebanov et al. [19].

Definition 2.17 (Explicit representation of ch.f. of strictly GS distribution) *A nondegenerate d.f. $\mathcal{G} = \mathcal{G}_p$ is strictly (geo,sum)-stable if and only if, its ch.f., f, has the form*

$$f(t) = \left[1 + \lambda|t|^\alpha \exp\left\{-i\frac{\pi}{2}\theta\alpha \mathrm{sign}(t)\right\}\right]^{-1}, \tag{2.12}$$

where $0 < \alpha \leq 2$, $|\theta| \leq \theta_\alpha = \min(1, 2/\alpha - 1)$ and $0 < \lambda$.

If $\alpha = 2$ and $\theta = 0$, $\mathcal{G}$ is a symmetric Laplace distribution, i.e.,

$$\mathcal{G}(t) = \frac{\lambda}{2}\int_{-\infty}^{t} e^{-\lambda|u|}du. \tag{2.13}$$

The Laplace distribution plays among the (geo,sum)-stable distributions a role that is analogous to that of the normal distribution in the class of the stable Paretian distributions. The exponential distribution $\mathcal{G}(t) = 1 - e^{-\lambda t}, 0 < t$ is a strictly (geo,sum)-stable distribution and plays the role of the discrete distribution in the class of stable Paretian distributions. Stable Paretian and (geo,sum)-stable distributions have one and the same tail behavior for $0 < \alpha < 2$ (see Mittnik and Rachev [29]).

Definition 2.18 (Domain of attraction of strictly GS distributions)
The d.f. $\mathcal{H}$ is in the domain of attraction of the strictly (geo,sum)-stable distribution $\mathcal{G}$ if for any sequence $X_1, X_2, \ldots$ of i.i.d. r.v.'s with common d.f. $\mathcal{H}$ and geometric r.v. $T(p)$, which is independent on $\{X_i\}$, there exist constants $a(p) > 0$ such that, as $p = 1/\mathrm{E}[T(p)] \to 0$,

$$Z_p = a(p) \sum_{i=1}^{T(p)} X_i \xrightarrow{\mathcal{D}} Y, \tag{2.14}$$

where Y is $\mathcal{G}_p$-distributed.

Chapter 3

ARMA Sequences with Stable Innovations

Autoregressive-moving average (ARMA) processes are often used for modeling empirical time series. Let p and q be non-negative integers. The sequence $\{X_n, n = \ldots, -1, 0, 1, \ldots\}$ is called ARMA(p,q) if it satisfies the equations

$$X_n - \phi_1 X_{n-1} - \cdots - \phi_p X_{n-p} = \varepsilon_n + \theta_1 \varepsilon_{n-1} + \cdots + \theta_q \varepsilon_{n-q}. \tag{3.1}$$

The *innovations* ε_n are i.i.d. random variables. In the classical time series literature, the ε_n are either Gaussian or non-Gaussian with finite variance and therefore the probability that they take large values is very small.

Here, it will be supposed that the ε_n are i.i.d. α-stable with $0 < \alpha \leq 2$. In other words $\varepsilon_n \sim S_\alpha(c, \beta, \mu)$ if $0 < \alpha < 2$ and $N(\mu, 2c^2)$ if $\alpha = 2$. The finite-dimensional distributions of the X_n depend on the coefficients $\theta_1, \ldots, \theta_q$ and $\phi_1, \ldots, \phi_p$.

Consider the system (3.1) with real coefficients $\phi_0 = 1, \phi_1, \ldots, \phi_p$ and $\theta_0 = 1, \theta_1, \ldots, \theta_q$ and define the polynomials

$$\begin{aligned} \Phi(z) &= 1 - \phi_1 z - \cdots - \phi_p z^p, \\ \Theta(z) &= 1 + \theta_1 z + \cdots + \theta_q z^q, \end{aligned}$$

where z is complex variable. One can write (3.1) symbolically as

$$\Phi(B) X_n = \Theta(B) \varepsilon_n, \; n = \ldots, -1, 0, 1, \ldots \;, \tag{3.2}$$

where B is the *backward operator*, formally, $B(X_n) = X_{n-1}, \; B^2(X_n) = X_{n-2}, \ldots$. As in the Gaussian case, one solves (3.1) by showing that $X_n = \Phi(B)^{-1}\Theta(B)\varepsilon_n$ is well defined. It is natural to suppose:

Property 3.1 *The polynomials $\Phi(z)$ and $\Theta(z)$ do not have common roots.*

The following theorem shows that, as in the Gaussian case, a non-anticipating solution exists if and only if $\Phi(z)$ has no roots in the closed unit disk $\{z : |z| \leq 1\}$.

Theorem 3.2 *The system (3.1) has a unique solution of the form*

$$X_n = \sum_{j=0}^{\infty} c_j \varepsilon_{n-j},\ n \in \boldsymbol{Z},\ \ a.s. \tag{3.3}$$

with real c_j satisfying $|c_j| < Q^{-j}$ eventually,[1] $1 < Q$, *if and only if $\Phi(z)$ has no roots in the closed unit disk $\{z : \ |z| \leq 1\}$. The sequence $\{X_n,\ n \in \boldsymbol{Z}\}$ is then stationary and α-stable.*

The c_j are the coefficients in the series expansion of $\Theta(z)/\Phi(z),\ |z| < 1$.

PROOF: Suppose $\Phi(z)$ has no roots in $\{z : |z| \leq 1\}$. The function

$$C(z) = \frac{\Theta(z)}{\Phi(z)}$$

is, therefore, analytic in the disk $\{z : |z| < R\}$, where $R > 1$ is the radius of convergence of the series $C(z) = \sum_{j=0}^{\infty} c_j z^j$. Since $1/R = \limsup_{j\to\infty} |c_j|^{1/j}$, for any $1 < Q < R$, $|c_j| < Q^{-j}$ eventually. Using the relation $\Phi(z)C(z) = \Theta(z)$, which holds for $|z| \leq 1$, and the fact that the series $\sum_{j=0}^{\infty} c_j z^j$ converges absolutely for $|z| \leq 1$, one obtains the following system of equations:

$$\begin{aligned}
&c_0 = 1,\\
&c_1 - \phi_1 c_0 = \theta_1,\\
&c_2 - \phi_1 c_1 - \phi_2 c_0 = \theta_2,\\
&\vdots\\
&c_q - \phi_1 c_{q-1} - \phi_2 c_{q-2} - \cdots - \phi_q c_0 = \theta_q,\\
&c_s - \phi_1 c_{s-1} - \phi_2 c_{s-2} - \cdots - \phi_s c_0 = 0,\ \ q < s,
\end{aligned} \tag{3.4}$$

with the understanding that $\phi_i = 0$ if $p < i$. It follows from (3.4) that the c_j are real and since $|c_j| < Q^{-j}$ eventually, the series (3.3) is well defined and, in fact, converges *absolutely* a.s.

To see that the process (3.3) with the c_j, uniquely defined by (3.4), satisfies (3.1), just use relation (3.4) and the fact that the series (3.3) converges absolutely a.s. Rearranging the terms in

$$\sum_{j=0}^{\infty} c_j \varepsilon_{n-j} - \phi_1 \sum_{j=0}^{\infty} c_j \varepsilon_{n-1-j} - \cdots - \phi_p \sum_{j=0}^{\infty} c_j \varepsilon_{n-p-j}$$

[1] "$a_j - b_j$ eventually" means that there is a j_0 such that $a_j < b_j$ for all $j > j_0$

yields (3.1).

To prove the converse suppose that the system of equations (3.1) has a solution of the form (3.3) with the c_j satisfying $|c_j| < Q^{-j}$ eventually for some $1 < Q$. One wants to show that $\Phi(z) \neq 0$ for $|z| \leq 1$.

Consider the series $C(z) = \sum_{j=0}^{\infty} c_j z^j$, which, under assumptions, converges absolutely and uniformly in the closed unit disk $\{z : |z| \leq 1\}$. Setting

$$\widetilde{\Theta}(z) \stackrel{df}{=} \Phi(z)C(z) \stackrel{df}{=} \sum_{j=0}^{\infty} \tilde{\theta}_j z^j, \; |z| \leq 1, \tag{3.5}$$

one obtains

$$\begin{aligned} \tilde{\theta}_0 &= c_0 \\ \tilde{\theta}_s &= c_s - \phi_1 c_{s-1} - \cdots - \phi_s c_0, \; 1 \leq s. \end{aligned} \tag{3.6}$$

Since, for any n, the series $\sum_{j=0}^{\infty} c_j \varepsilon_{n-j}$ converges absolutely a.s., (3.1) and (3.6) imply

$$\sum_{j=0}^{q} \theta_j \varepsilon_{n-j} = X_n - \phi_1 X_{n-1} - \cdots - \phi_p X_{n-p} = \sum_{j=0}^{\infty} \tilde{\theta}_j \varepsilon_{n-j} \quad a.s.,$$

which, in turn, yields $\tilde{\theta}_j = \theta_j$ for $j = 0, 1, \ldots, q$ and $\tilde{\theta}_j = 0$ for $q < j$. Thus, $\widetilde{\Theta}(z) = \Theta(z)$ and by (3.5)

$$\Phi(z) = \frac{\Theta(z)}{C(z)}, \; |z| \leq 1.$$

As $C(z)$ is bounded on $\{z : |z| \leq 1\}$, $\Phi(z) = 0$ implies $\Theta(z) = 0$. But $\Phi(z)$ and $\Theta(z)$ do not have common roots, so $\Phi(z) \neq 0$, for all $|z| \leq 1$, proving the converse.

The solution (3.3) is α-stable, because it is a linear combination of α-stable random variables. It is clearly stationary.
Q.E.D.

The condition that $\Phi(z)$ has no roots in the closed unit disk $\{z : |z| \leq 1\}$ is a natural one for it ensures that the system of equations (3.1) has the *stationary non-anticipating solution* (3.3). It will be supposed from now on that this condition holds.

The c_j are obtained by identifying the coefficients of $C(z) = \sum_{j=0}^{\infty} c_j z^j$ with those in the power series expansion of $\Theta(z)/\Phi(z)$. Since this is the same procedure as in the Gaussian case, the explicit form of the c_j for specific ARMA(p,q) models can be readily found in the time series literature.

Example 3.3 Consider the autoregressive process $\{X_n\}$ of order 2 defined by

$$X_n - \phi_1 X_{n-1} - \phi_2 X_2 = \varepsilon_n.$$

If $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2$ has two different roots z_1 and z_2, satisfying $1 < |z_i|,\ i = 1, 2$, then

$$\Phi(z) = \frac{1}{z_1 z_2}(z - z_1)(z - z_2)$$

and

$$\frac{1}{\Phi(z)} = \frac{z_1 z_2}{z_2 - z_1}\left(\frac{1}{z_1 - z} - \frac{1}{z_2 - z}\right) = \frac{z_1 z_2}{z_2 - z_1}\left(\frac{z_1^{-1}}{1 - (z/z_1)} - \frac{z_2^{-1}}{1 - (z/z_2)}\right).$$

The coefficients c_j in the series expansion of $1/\Phi(z)$ are, therefore,

$$c_j = \frac{z_1 z_2}{z_2 - z_1}\left(z_1^{-j-1} - z_2^{-j-1}\right),\ 0 \leq j.$$

If $\Phi(z)$ has complex conjugate roots $\rho e^{\pm i\mu},\ \mu \neq k\pi$, then it is not difficult to see that

$$c_j = \frac{\sin \mu(j+1)}{\sin \mu}\rho^{-j}.$$

The ARMA(p,q) time series is *invertible* if there exists a sequence of constants $\{c_j\}$ such that $\sum_{j=0}^{\infty} |\tilde{c}_j| < \infty$ and $\sum_{j=0}^{\infty} \tilde{c}_j X_{n-j} = \varepsilon_n,\ n \in \boldsymbol{Z}$, where convergence holds in probability. Invertibility is particularly useful for prediction, because it allows X_n to be expressed in terms of the previous observations $X_j,\ j \leq n$. The following theorem provides a condition for invertibility.

Theorem 3.4 *Suppose that $\Theta(z)$ has no roots in the closed unit disk $\{z : |z| \leq 1\}$. Then ARMA(p,q) is* invertible, *i.e.,*

$$\sum_{j=0}^{\infty} \tilde{c}_j X_{n-j} = \varepsilon_n,\ n \in \boldsymbol{Z},\ a.s.\ .$$

The $\tilde{c}_j$ are the coefficients in the series expansion of $\Theta^{-1}(z)\Phi(z),\ |z| < 1$.

PROOF: See Samorodnitsky and Taqqu [35]. Q.E.D.

Since the coefficients c_j of the moving average (3.3) satisfy $|c_j| < Q^{-j}$ eventually with $1 < Q$, they lie within two exponentially decreasing functions.

As for the codifference $\tau(n) = \tau_{X_n, X_0}$, one has:

Theorem 3.5 *Suppose that $\Phi(z)$ has no roots in the closed unit disk $\{z : |z| \leq 1\}$ and let $1 < Q$ be as in Theorem 3.2. Then there are constants K_1 and K_2 depending on α, Q and the c_j such that*

$$\limsup_{n\to\infty} Q^n|\tau(n)| \leq K_1 \quad for\ 1 \leq \alpha \leq 2$$

$$\limsup_{n\to\infty} Q^{\alpha n}|\tau(n)| < K_2 \quad for\ 0 < \alpha < 1.$$

PROOF: See Kokoszka and Taqqu [22]. Q.E.D.

If $1 < \alpha$ a similar result holds for the covariation, namely

$$\limsup_{n\to\infty} Q^n|\,[X_n, X_0]_\alpha| < K_3,$$

because

$$\begin{aligned} Q^n|[X_n, X_0]_\alpha| &\leq \sum_{j=0}^{\infty} Q^n|c_{j+n}||c_j|^{\alpha-1} \\ &\leq \text{ const.} \sum_{j=0}^{\infty} Q^n Q^{-(j+n)} Q^{-(\alpha-1)j} = \text{const.} \sum_{j=0}^{\infty} Q^{-\alpha j} < \infty. \end{aligned}$$

Chapter 4

The Gaussian vs. Stable Paretian Distribution in ARCH/GARCH Models

In this chapter, alternative ways to model changing volatility of a time series y_{t+1} will be considered. Section 4.1 presents univariate Autoregressive Conditionally Heteroskedastic (ARCH) and Generalized Autoregressive Conditionally Heteroskedastic (GARCH) models under Gaussian distribution. The next Section 4.2 presents this models under stable distribution.

4.1 ARCH and GARCH Models in Gaussian Distribution

First, in order to concentrate on volatility, it is assumed that y_{t+1} is an innovation, which has mean zero conditional on time t information. In a finance application, y_{t+1} might be the innovation in an asset return. Let σ_t^2 be denoted as the time t conditional variance of y_{t+1} or equivalently the conditional expectation of y_{t+1}^2. One assumes that conditional on time t information, the innovation is normally distributed, i.e.,

$$y_{t+1} \sim N_t(0, \sigma_t^2). \tag{4.1}$$

The unconditional variance of the innovation y_{t+1} is denoted σ^2 and is just the unconditional expectation of σ_t^2,

$$\sigma^2 \stackrel{df}{=} \mathrm{E}[\sigma_t^2] = \mathrm{E}[\mathrm{E}_t[y_{t+1}^2]] = \mathrm{E}[y_{t+1}^2]. \tag{4.2}$$

Thus variability of σ_t^2 around its mean does not change the unconditional variance σ^2.

The variability of σ_t^2 does, however, affect higher moments of the unconditional distribution of y_{t+1}. In particular, with time-varying σ_t^2 the unconditional distribution of y_{t+1} has fatter tails than a normal distribution. A useful measure of tail thickness for the distribution of a random variable y is the normalized fourth moment, respectively kurtosis, defined by $K(y) \stackrel{df}{=} \mathrm{E}[y^4]/(\mathrm{E}[y^2])^2$. One first writes (see Campbell et al. [8]):

$$y_{t+1} = \sigma_t \varepsilon_{t+1},$$

where ε_{t+1} is an i.i.d. random variable with zero mean and unit variance that is normally distributed. It is well known that the kurtosis of a normal random variable is 3, hence $K(\varepsilon_{t+1}) = 3$. But for innovation y_{t+1}, one has

$$\begin{aligned} K(y_{t+1}) &= \frac{\mathrm{E}[\sigma_t^4]\mathrm{E}[\varepsilon_{t+1}^4]}{(\mathrm{E}[\sigma_t^2])^2} \\ &= \frac{3\mathrm{E}[\sigma_t^4]}{(\mathrm{E}[\sigma_t^2])^2} \\ &\geq \frac{3(\mathrm{E}[\sigma_t^2])^2}{(\mathrm{E}[\sigma_t^2])^2} = 3, \end{aligned}$$

where the first equality follows from the independence of σ_t and ε_{t+1}, and the inequality is implied by Jensen's inequality.

However intuitively, the unconditional distribution of y_{t+1} is a mixture of normal distributions, some with small variances that concentrate mass around the mean and some with large variances that put mass in the tails of the distribution. Thus the mixed distribution has fatter tails than the normal one.

A basic observation about asset return data is that large returns (of either sign) tend to be followed by more large returns (of either sign). In other words, the volatility of asset returns appears to be serially correlated. To capture the serial correlation of volatility, Engle [12] proposed the class of Autoregressive Conditionally Heteroskedastic, or ARCH, models. These write conditional variance as a distributed lag of past squared innovations:

$$\sigma_t^2 = \delta + \alpha(L) y_t^2, \tag{4.3}$$

where $\alpha(L)$ is a polynomial in the lag operator. To keep the conditional variance positive, δ and the coefficients in $\alpha(L)$ must be nonnegative.

As a way to model persistent movements in volatility without estimating a very large number of coefficients in a high-order polynomial $\alpha(L)$, Bollerslev [2] suggested the Generalized Autoregressive Conditionally Heteroskedastic, or GARCH, model

$$\sigma_t^2 = \delta + \beta(L)\sigma_{t-1}^2 + \alpha(L) y_t^2, \tag{4.4}$$

where $\beta(L)$ is also a polynomial in the lag operator. By analogy with ARMA models, this is called a GARCH(p,q) model when the order of the polynomial $\beta(L)$ is p and the order of the polynomial $\alpha(L)$ is q.

The most commonly used model in the GARCH class is the simple GARCH(1,1), which can be written as:

$$\begin{aligned} \sigma_t^2 &= \delta + \beta\sigma_{t-1}^2 + \alpha y_t^2 \\ &= \delta + (\alpha+\beta)\sigma_{t-1}^2 + \alpha(y_t^2 - \sigma_{t-1}^2) \\ &= \delta + (\alpha+\beta)\sigma_{t-1}^2 + \alpha\sigma_{t-1}^2(\varepsilon_t^2 - 1), \end{aligned}$$

where in the second equality, the term $(y_t^2 - \sigma_{t-1}^2)$ has mean zero, conditional on time $t-1$ information and can be thought of as the shock to volatility. The coefficient α measures the extent, to which a volatility shock today feeds through into next period's volatility, while $(\alpha+\beta)$ measures the rate at which this effect dies out over time. The third equality rewrites the volatility shock as $\sigma_{t-1}^2(\varepsilon_t^2 - 1)$, the square of a standard normal less its mean – that is, a demeaned $\chi^2(1)$ random variable – multiplied by past volatility σ_{t-1}^2.

The GARCH(1,1) model can also be written in terms of its implications for squared innovations y_{t+1}^2:

$$y_{t+1}^2 = \delta + (\alpha+\beta)y_t^2 + (y_{t+1}^2 - \sigma_t^2) - \beta(y_t^2 - \sigma_{t-1}^2).$$

Hence, the GARCH(1,1) model can be represented as an ARMA(1,1) for squared innovations. However, it is needed to pay attention, because a standard ARMA model has homoskedastic shocks, while here the shocks $(y_{t+1}^2 - \sigma_t^2)$ are themselves heteroskedastic.

4.2 ARCH and GARCH Models in non-Gaussian Distribution

The GARCH models, which are considered in the previous section, imply that the distribution of returns, conditional on the past history of returns, is normal. Equivalently, the standardized residuals of these models, $\varepsilon_{t+1}(\theta) = y_{t+1}/\sigma_t(\theta)$, should be normal. Unfortunately, in practice, there is an excess kurtosis in the standardized residuals of GARCH models, albeit less than in the raw returns. To handle this problem, one can explicitly model the fat-tailed distribution of the innovations driving a GARCH process.

In the next sections, the necessary and sufficient conditions for existence and uniqueness of a stationary solution of the stable GARCH equation will be provided.

4.2.1 Symmetric GARCH-stable Processes

One restricts itself to symmetric, strictly stable random variables, that is one assumes $\beta = \mu = 0$, see (2.2) in Section 2.1. A symmetric α-stable random variable X with scale parameter σ is denoted by $S_\alpha S(\sigma)$, i.e., its ch.f. equals $\exp\{-|\sigma t|^\alpha\}$.

Recall first the definition of the stable GARCH process.

Definition 4.1 *A sequence of random variables $\{Y_n, n \in \boldsymbol{Z}\}$ is said to be a stable GARCH(α,p,q) process if:*

(i) $Y_n = \sigma_n S_n$ *, where S_n are i.i.d. r.v.'s distributed as $S_\alpha S(1)$, $1 < \alpha \leq 2$,*

(ii) *there exist nonnegative constants α_i, $i = 1, \ldots, q$ and β_j, $j = 1, \ldots, p$ and $0 < \delta$, such that*

$$\sigma_n = \delta + \sum_{i=1}^{q} \alpha_i |Y_{n-i}| + \sum_{j=1}^{p} \beta_j \sigma_{n-j},\ n \in \boldsymbol{Z}. \tag{4.5}$$

The assumption $1 < \alpha \leq 2$ is not very restrictive in practice, because most of the financial time series have finite mean. For $\alpha = 2$, one obtains a L_1-version of the classical Gaussian GARCH(p,q) model.

To show the existence and uniqueness of strictly stationary solutions of equation (4.5), one uses the results of Bougerol and Picard [5]. Following their notation one defines:

- $B = [\delta, 0, \ldots, 0]' \in \boldsymbol{R}^{p+q-1}$,
- $X_n = [\sigma_{n+1}, \ldots, \sigma_{n-p+2}, |Y_n|, \ldots, |Y_{n-q+2}|\,] \in \boldsymbol{R}^{p+q-1}$,
- $t_n = [\beta_1 + \alpha_1 |S_n|, \beta_2, \ldots, \beta_{p-1}] \in \boldsymbol{R}^{p-1}$,
- $z_n = [\,|S_n|, 0, \ldots, 0] \in \boldsymbol{R}^{p-1}$,
- $\alpha = [\alpha_2, \ldots, \alpha_{q-1}] \in \boldsymbol{R}^{q-2}$,
- $$A_n = \begin{bmatrix} t_n & \beta_p & \alpha & \alpha_q \\ I_{p-1} & 0 & 0 & 0 \\ z_n & 0 & 0 & 0 \\ 0 & 0 & I_{q-2} & 0 \end{bmatrix} \quad (p+q-1 \times p+q-1), \tag{4.6}$$

where I_{p-1} and I_{q-2} are identity matrices of size $(p-1)$ and $(q-2)$, respectively.

Then, one has

$$X_{n+1} = A_{n+1}X_n + B, \quad n \in \boldsymbol{Z}. \tag{4.7}$$

Clearly, existence of solutions of (4.7) is equivalent to existence of solutions of (4.5). The major role in what follows is played by a Lyapunov exponent associated with matrices $\{A_n, n \in \boldsymbol{Z}\}$. The definition of the top Lyapunov exponent will be recalled next.

Let $||\cdot||$ be any norm on $\boldsymbol{R}^d$ and define an operator norm on the set $M(d)$ of $(d \times d)$ matrices by $||M|| = \sup\{||Mx||/||x||, x \in \boldsymbol{R}^d, x \neq 0\}$, for any $M \in M(d)$. Then, the top Lyapunov exponent associated with the sequence $\{A_n, n \in \boldsymbol{Z}\}$ of i.i.d. random matrices, is defined by

$$\gamma = \inf \mathrm{E}\left[\frac{1}{n+1}\log||A_0A_{-1}\cdots A_{-n}||\right], \quad n \in \boldsymbol{N},$$

when $\mathrm{E}[\log^+||A_0||] < \infty$ (where $\log^+ x = \max(\log x, 0)$). It follows, see Bougerol and Picard [4], that, almost surely,

$$\gamma = \lim_{n\to\infty}\frac{1}{n}\log||A_0A_{-1}\cdots A_{-n}||.$$

Given that r.v.'s S_n are i.i.d., random matrices A_n defined in (4.6) are also i.i.d. Because $\gamma \leq \mathrm{E}[\log||A_0||]$ and $\mathrm{E}[\log||A_0||] < \infty$ (because $\mathrm{E}|S_0| < \infty$), γ is well defined for the sequence $\{A_n, n \in \boldsymbol{Z}\}$. Clearly, $\mathrm{E}[\log^+||B||] < \infty$. Thus, by Bougerol and Picard [4, Theorem 3.2], one has the following theorem:

Theorem 4.2 *The stable GARCH(α,p,q) process with* $0 < \delta$*,* $2 \leq p$ *and* $2 \leq q$ *has a stationary solution if the Lyapunov exponent of* $\{A_n, n \in \boldsymbol{Z}\}$ *is strictly negative. The series* $X_n = B + \sum_{i=1}^{\infty} A_nA_{n-1}\cdots A_{n-i+1}B$ *converges almost surely for all* n*, and the process* $\{X_n, n \in \boldsymbol{Z}\}$ *is the unique strictly stationary and ergodic solution of (4.7).*

To obtain the conditions on the coefficients α_i, $i = 1, \ldots, q$ and β_j, $j = 1, \ldots, p$ for an existence of the strictly stationary solution, the characteristic polynomial of $\mathrm{E}[A_0]$ is considered:

$$f(z) = \det(Iz - \mathrm{E}[A_0]) = z^{p+q-1}\left(1 - a\sum_{i=1}^{q}\alpha_i z^{-i} - \sum_{j=1}^{p}\beta_j z^{-j}\right),$$

where $a = \mathrm{E}|S_1| = \Gamma(1-1/\alpha)/\int_0^\infty u^{-2}\sin^2(u)du$.

Theorem 4.3 (Bougerol and Picard [4, Collary 2.2])
If $0 < \delta$ *and* $a \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j \leq 1$, *then the stable GARCH(α,p,q) with* $2 \leq p$ *and* $2 \leq q$, *has a unique strictly stationary solution.*

The case when $p = q = 1$ has to be treated separately, because then the matrices $\{A_n, n \in \boldsymbol{Z}\}$ from (4.6) have zero columns and the results of Kesten and Spitzer [18] do not apply. The largest Lyapunov exponent can be computed for stable GARCH(α,1,1) directly as follows:

$$||A_0 A_{-1} \cdots A_{-n}|| = \prod_{i=1}^{n} (\beta_1 + \alpha_1 |S_i|) \max(\beta_1 + \alpha_1 |S_0|, 1, |S_0|),$$

and by the Law of Large Numbers one obtains

$$\begin{aligned} \gamma &= \frac{1}{n+1} [\log(\max(\beta_1 + \alpha_1 |S_0|, 1, |S_0|)) \\ &+ \sum_{i=0}^{n} \log(\beta_1 + \alpha_1 |S_i|)] \xrightarrow{p} \mathrm{E}[\log(\beta_1 + \alpha_1 |S_0|)]. \end{aligned}$$

If $\mathrm{E}[\log(\beta_1 + \alpha_1 |S_0|)] < 1$, then stable GARCH($\alpha$,1,1) has a strictly stationary solution. A stationary condition $\beta_1 + a\alpha_1 < 1$ compatible with the results of Theorem 4.3 is obtained by using Jensen's inequality.

4.2.2 Asymmetric GARCH-stable Processes

Generalizing the stable GARCH process to the asymmetric case.

Definition 4.4 *Sequence* $y_t, t \in \boldsymbol{Z}$ *is said to be a stable Paretian Asymmetric GARCH process, in short, an* $S_{\alpha,\beta}$*GARCH process, if*

$$y_t = \mu_t + c_t \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} S_{\alpha\beta}, \tag{4.8}$$

where

$$c_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i |y_{t-i} - \mu_{t-i}| + \sum_{j=1}^{s} \beta_j c_{t-j}, \tag{4.9}$$

and where $S_{\alpha\beta}$ *denotes the standard asymmetric stable Paretian distribution with stable index* α, *skewness parameter* $\beta \in [-1, 1]$, *zero location parameter and unit scale parameter.*

By letting $\mathrm{E}(y_t) = \mu_t$ in (4.8) to be time-varying, one allows for a broad range of mean equations, including e.g., regression or/and ARMA structures.

According to Liu and Brorsen [24] and Panorska et al. [31], it is assumed that $\alpha \in (1, 2]$, to avoid a number of technical problems arising from the fact that $\alpha \leq 1$ implies such fat tails that ε_t and thus y_t and c_t do not even possess the first moments. However, the restriction does not seem to have practical relevance in empirical work, because the existence of the first moments is hardly rejected in financial or economic time series.

Liu and Brorsen [24] state the conditional volatility equation 4.9 more generally as

$$|c_t|^\delta = \alpha_0 + \sum_{i=1}^{r} \alpha_i |y_{t-i} - \mu_{t-i}|^\delta + \sum_{j=1}^{s} \beta_j |c_{t-j}|^\delta.$$

In their empirical applications, they experiment by different settings of δ.

Observe that, as α approaches 2, model (4.8) becomes the so-called absolute value GARCH model with normal innovations, originally proposed by Taylor [37] and Schwert [36]. Nelson and Foster [30] have shown that, compared to GARCH models in c_t^2, the absolute value GARCH model is a more efficient filter of the conditional variance in the presence of leptokurtic error distributions (i.e., compared to the normal distribution they are typically fat-tailed and more peaked around the center, a phenomenon, which is commonly observed with asset-return data).

Bougerol and Picard [4] have shown that a mean-corrected GARCH(r,s) process driven by *normally* distributed innovations has a unique strictly stationary solution if

$$\sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j \leq 1. \tag{4.10}$$

The following result implies that condition (4.10) does not apply to $S_{\alpha,\beta}$-GARCH models with $\alpha < 2$.

Theorem 4.5 *An $S_{\alpha,\beta}$GARCH process defined by (4.8) and (4.9) with $1 < \alpha < 2$ has a unique strictly stationary solution if $0 < \alpha_i$, $i = 0, \ldots, r$, $0 < \beta_j$, $j = 1, \ldots, s$ and*

$$\lambda_{\alpha,\beta} \sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j \leq 1 \tag{4.11}$$

where $\lambda_{\alpha,\beta} \stackrel{df}{=} \mathrm{E}|\varepsilon_t|$ is given by

$$\lambda_{\alpha,\beta} = \begin{cases} \frac{2}{\pi}\Gamma(1 - \frac{1}{\alpha})(1 + \tau_{\alpha,\beta}^2)^{\frac{1}{2\alpha}} \cos(\frac{1}{\alpha} \arctan \tau_{\alpha,\beta}), & 1 < \alpha < 2, \\ \sqrt{\frac{2}{\pi}}, & \alpha = 2, \end{cases} \tag{4.12}$$

with $\tau_{\alpha,\beta} \stackrel{df}{=} \beta \tan \frac{\alpha\pi}{2}$.
In the symmetric case, i.e. $\beta = 0$, *(4.12) reduces to*

$$\lambda_{\alpha,\beta} = \begin{cases} \frac{2}{\pi}\Gamma(1 - \frac{1}{\alpha}), & 1 < \alpha < 2, \\ \sqrt{\frac{2}{\pi}}, & \alpha = 2. \end{cases} \tag{4.13}$$

The value of parameter $\lambda_{\alpha,\beta}$ defined in (4.12) and (4.13) depends on the stable index α and the skewness parameter β. For $\alpha < 2$ one has $1 < \lambda_{\alpha,\beta}$. This implies that stationary condition (4.11) is more restrictive than in the normal case. Observe also that, in the normal case $\lambda_{2,0} = \sqrt{2/\pi} < 1$, implying that $\sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j$ could be greater than one and not violate the stationarity condition.

In case of normally distributed innovations for the variance-GARCH model, Engle and Bollerslev [13] referred to GARCH processes satisfying the borderline condition

$$\sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j = 1 \tag{4.14}$$

as an *integrated* GARCH or IGARCH process. Condition (4.14) implies that autoregressive polynomial $1-(\alpha_1+\beta_1)L-\cdots-(\alpha_n+\beta_n)L^n$ with $n = \max(r, s)$, when rewriting the GARCH process in an ARMA form in terms of ε_t^2, has a unit root.

In the $S_{\alpha,\beta}$GARCH case, the analogue to (4.14) is

$$\lambda_{\alpha,\beta} \sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j = 1, \tag{4.15}$$

giving rise to *integrated* $S_{\alpha,\beta}$GARCH or $S_{\alpha,\beta}$IGARCH processes. The implied autoregressive polynomial for the ARMA representation of $|c_t|$ is

$$1 - (\lambda_{\alpha,\beta}\alpha_1 + \beta_1)L - \cdots - (\lambda_{\alpha,\beta}\alpha_n + \beta_n)L^n,$$

in which case one has the persistence of the conditional volatility.

The GARCH(1,1) model is most commonly specified in empirical work. Stationarity of an $S_{\alpha,\beta}$GARCH(1,1) process requires

$$\lambda_{\alpha,\beta}\alpha_1 + \beta_1 \leq 1.$$

Thus, when estimating an $S_{\alpha,\beta}$IGARCH(1,1) model the restriction $\beta_1 = 1 - \lambda_{\alpha,\beta}\alpha_1$ needs to be imposed during estimation.

Note that all proofs of results from sections 4.2.1 and 4.2.2 can be found, e.g., in Rachev and Mittnik [34].

Chapter 5

An Empirical Application

5.1 Statistical Software

All statistical computations were performed with the statistical software R 2.5.0 for Mac OS X Aqua GUI, Copyright © 2004-2007 The R Foundation for Statistical Computing (http://www.R-project.org). This software package is distributed as a free software under the terms of the Free Software Foundation's GNU General Public License in source code form.

Statistical tests, which are not incorporated in "R", were written by the author of this text directly in "R" environment.

5.2 Data

As an illustration of modeling of returns, it is reported on a single time series, whose behavior will be described bellow. These real data become from energetic field of activity. The data set relates on the consumption of energy of an economical subject during one year period (2006). The consumption of energy was recorded every hour in sequential mapping. So there is a sample of 8760 observations in equidistant time spots. The January levels of this time series are shown in Figure 5.1 (only January, because the levels for longer period would not be comprehensive).

One can observe everyday peaks during rush hours, decays and saddles of current consumption. It is visible that this subject has his main activity during working days – there is obvious decrement at the weekends.

As a first step, a test based on reversion points was performed (see Cipra [10]). This test rejects the hypothesis, that the investigated time series

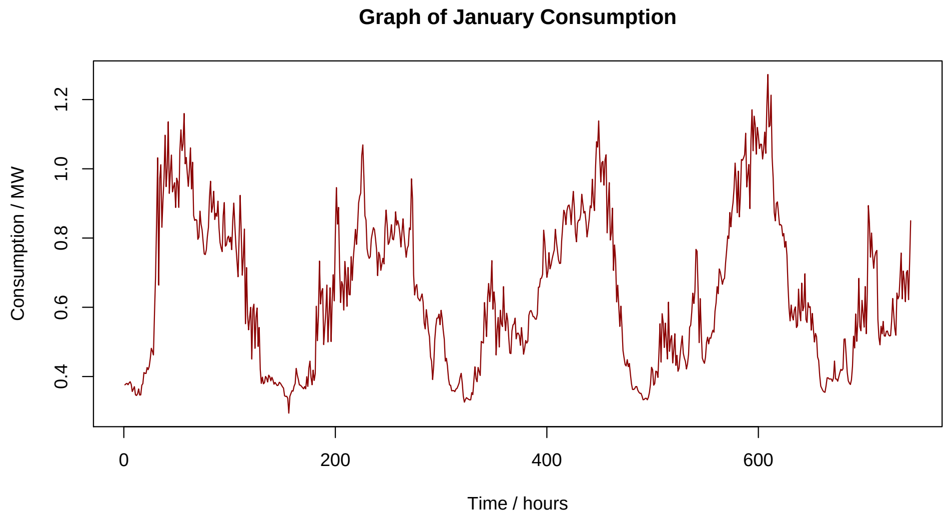

Figure 5.1: January Levels of Time Series

is a pure stochastic process, with probability 95%. The term, that a time series is stochastic, means, that the time series does not have any patterns, like cyclic or seasonal. It means that it behaves only as a pure white noise.

For historical reasons, the standard convention was accepted and return r_t in period t was counted by

$$r_t \stackrel{df}{=} \ln\left(\frac{P_t}{P_{t-1}}\right) \times 100,$$

where P_t is the consumption of the subject at time t. The January return series is shown in Figure 5.2.

To provide a visual impression, Figure 5.3 shows the empirical density of the returns obtained via kernel density estimation. The graph is consistent. From the Figures 5.1, 5.2 and 5.3 extreme kurtosis, as well as volatility clustering, are clearly visible.

Table 5.1 summarizes the basic statistical properties of the return series. The returns show the evidence of fat tails. A negative skewness statistic indicates that the distribution is skewed to the left, i.e., compared to the right tail, the left tail is elongated. The kurtosis statistic reflects the significant peakedness of the center compared to that of the normal distribution (a value near three would be indicative of normality). Although formal tests could, in principle, be conducted, it should be kept in mind that under the non-Gaussian stable hypothesis, second and higher moments do not exist, rendering such

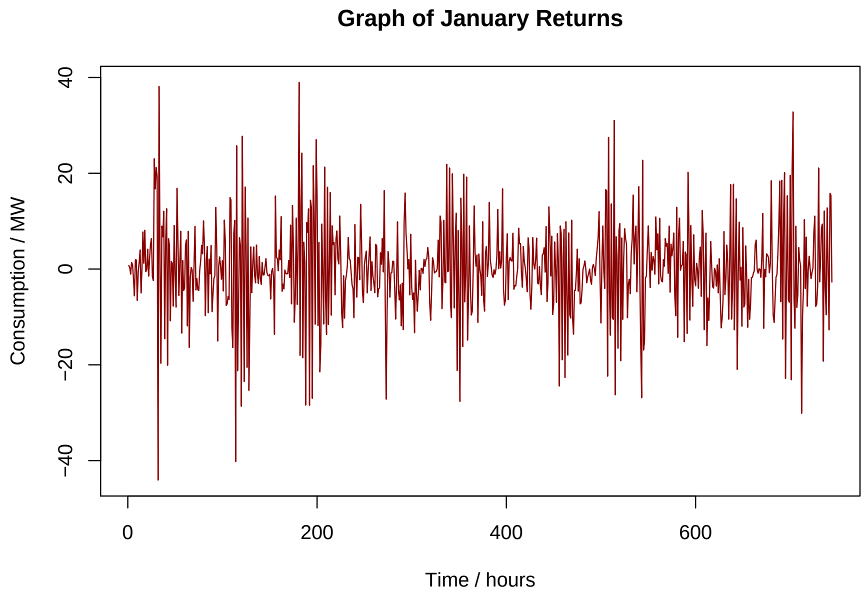

Figure 5.2: January Returns of Time Series

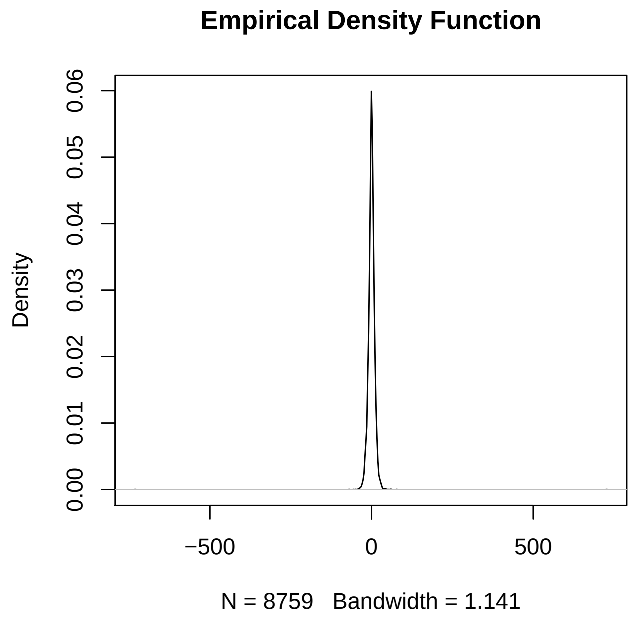

Figure 5.3: Empirical Density Function of Returns

Mean	Standard Deviation	Skewness	Kurtosis
-0.0004212325	14.47665	-0.2227454	1469.197

Table 5.1: Statistical Properties of Returns

tests useless. The numbers in the Table 5.1, as well as the Figure 5.3, indicate considerable deviation from normality. Formal tests, such as Shapiro-Wilk or Kolmogorov-Smirnov normality test, confirmed this too.

Chapter 6

Unconditional Fitting

Considering the unconditional case, the ML estimate of $\theta = (\alpha, \beta, c, \delta)$ is obtained by maximizing the logarithm of the likelihood function

$$L(\theta) = \prod_{t=1}^{T} S_{\alpha,\beta}\left(\frac{r_t - \delta}{c}\right) c^{-1}.$$

Evaluation of the probability density function and thus, the likelihood function of the $S_{\alpha,\beta}$ distribution is nontrivial, because it lacks an analytic expression. The estimation of α-stable models is *approximate* in the sense that the α-stable density function $S_{\alpha,\beta}((r_t - \delta)/c)$ needs to be approximated. One numerically approximates the α-stable density via FFT (fast Fourier transformation) of the ch.f. rather than some series expansion.

The ML estimation of the unconditional return distribution led to the estimates given in Table 6.1, where the parameter labels correspond to ch.f. (2.2). Output of the "R" software does not provide standard deviations of this estimations. Note that the scale for the normal case is given by the estimated standard deviation of the return series and not by parameter c in (2.2), so that the entries cannot be compared directly (between stable and normal case). The same situation is with the shift parameter δ, which is given by the estimation of mean in Gaussian case. The skewness parameter for normal

Distribution	$\hat{\alpha}$ (index)	$\hat{\beta}$ (skewness)	$\hat{c}$ (scale)	$\hat{\delta}$ (location)
NORMAL	2	0	14.4758284435	-0.0004212325
α-STABLE	1.8023219	-0.1254743	5.8689991	0.1587154

Table 6.1: ML Estimates of Unconditional Distributions

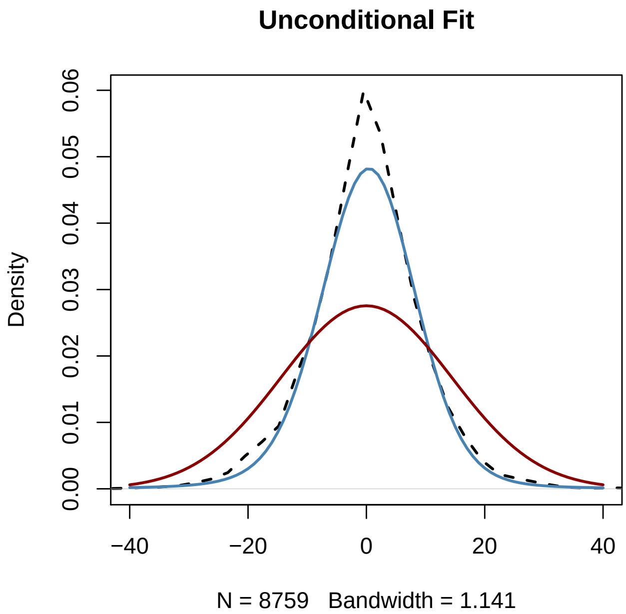

Figure 6.1: Fitted Unconditional Normal and α-Stable Densities

case is zero and the shape parameter two by the definition. The estimated shape parameter of the α-stable distribution is well below $\alpha = 2$, the Gaussian case. However, one can observe in Table 6.1 the substantial differences between the numbers for the normal and stable distribution.

Figure 6.1 clearly indicates that the fitted α-stable distribution dominates that of the normal. This picture shows the empirical density obtained via kernel density estimation (black dashed line), along with the fitted normal distributions (dark red solid line) and fitted α-stable distribution (steel blue solid line). The graph is consistent with the fit estimations reported in Table 6.1. As one can see, the α-stable approach provides a much closer approximation to the empirical density function.

Chapter 7

Conditional Gaussian Fitting

7.1 Homoskedastic Model

The unconditional results ignore possible temporal dependencies in the return series. Typical serial dependence in a time series is modeled by ARMA structures. Such models allow conditioning of the process mean on past realizations and have been proven successful for the short-term prediction of time series.

To obtain conditional models, as the first an ARMA model under normal assumptions will be demonstrated, employing (approximate) conditional maximum-likelihood (ML) estimation. The ML estimation is conditional, in the sense that, when estimating, for example, an ARMA(p,q) model, one conditions on the first p realizations of the sample, $r_p, r_{p-1}, \ldots, r_1$, and set innovations $\varepsilon_p, \varepsilon_{p-1}, \ldots, \varepsilon_{p-q+1}$ to their unconditional mean $\mathrm{E}[\varepsilon_t] = 0$.

A peculiar feature of ARMA models is that the conditional (prediction error) variances are, in fact, independent of past realizations, i.e., they are conditionally homoskedastic (i.e., constant-conditional-volatility). In view of the facts, the assumption of conditional homoskedasticity is commonly violated in financial data, where volatility clusters are typically observed, implying that a large return is often followed by more large returns, which more or less slowly decay. Such behavior can be captured by AutoRegressive Conditional Heteroskedastic – ARCH and GARCH models (see Engle [12] and Bollerslev [2]), possibly in combination with ARMA model, referred to as an ARMA-GARCH model. They express the conditional variance as an explicit function of past information and permit conditional heteroskedasticity (i.e., varying-conditional-volatility).

The homoskedastic case will be considered first, before moving to the more interesting conditional heteroskedastic case. An ARMA model of au-

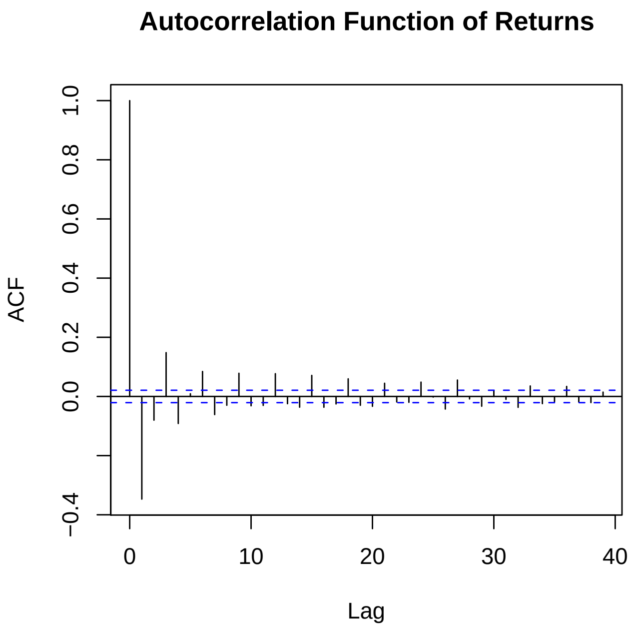

Figure 7.1: Sample Correlation Function of Returns

toregressive order p and moving average order q is of the form

$$r_t = \mu + \sum_{i=1}^{p} a_i r_{t-i} + \varepsilon_t + \sum_{j=1}^{q} b_j \varepsilon_{t-j}, \tag{7.1}$$

where $\{\varepsilon_t\}$ is a white noise process. To specify the orders p and q in (7.1), standard Box-Jenkins identification techniques are followed (see, for example, Box and Jenkins [6], Wei [38], Brockwell and Davis [7] or Johnston [17]) and inspected sample autocorrelation function (SACF) and sample partial autocorrelation function (SPACF) of the return series, as shown in Figures 7.1 and 7.2.

The exponentially decaying SAFC and the second large spike in the SPACF strongly suggest the appropriateness of an AR(2) structure. The formal statistical test based on Quenouille approximation confirmed that the point of truncation in SPACF is really at lag two. The test did not reject the hypothesis that the lag-values of the SPACF — from the third one — are not distinguished from zero. However, both the SACF and SPACF exhibit relatively large spikes at lag one, possibly suggesting either a subset AR(1) or MA(1) or the combinations ARMA(1,1), or ARMA(2,1) respectively. So, these options were analyzed:

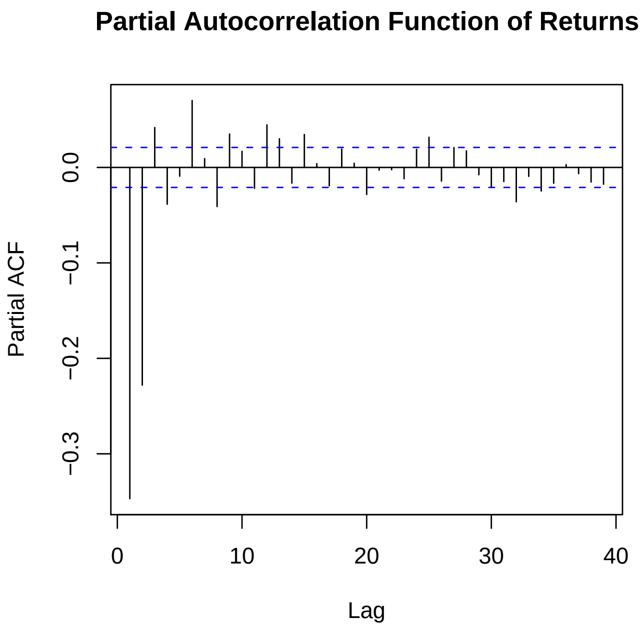

Figure 7.2: Sample Partial Correlation Function of Returns

- AR(2)
- AR(1)
- MA(1)
- ARMA(1,1)
- ARMA(2,1)

As the decision criteria, AIC (Akaike-Information-Criterion, see Akaike [1]), the estimated variance of the white noise (variance of innovations) and the maximum-likelihood value were used.

As the next step, the estimations of all parameters for all suggested models and comparison of all decision criteria were proceeded. The conditional ML estimations were used to establish the normal ARMA models. The lowest AIC, the lowest estimated variance of the white noise and the highest likelihood value for the return series had the model ARMA(2,1). So, it was opted for this variant. The intercept parameter μ of the chosen model seemed to be redundant (notable standard deviation of its estimation), therefore was

	$\hat{a}_1$	$\hat{a}_2$	$\hat{b}_1$
ESTIMATIONS	-0.6326	-0.2992	0.2186
S.E.	0.0107	0.0160	0.0464

Table 7.1: Parameter Estimates of ARMA(2,1)

$\hat{\alpha}$ (index)	$\hat{\beta}$ (skewness)	$\hat{c}$ (scale)	$\hat{\delta}$ (location)
2	0	5.9338265147	0.0001858718

Table 7.2: ML Estimates of Conditional Homoskedastic Normal Distribution

excluded from the next computation. Other parameters were statistically significant. Table 7.1 shows all estimations of the parameters and their standard errors (standard deviations).

Finally, the best model, which was obtained and statistically confirmed, has this form:

$$r_t = -0.6326 r_{t-1} - 0.2992 r_{t-2} + \varepsilon_t + 0.2186 \varepsilon_{t-1}. \quad (7.2)$$

The estimation of the variance of the white noise is $\hat{\sigma}_\varepsilon^2 = 174.3$.

The parameter estimates of the fitted conditional distribution are reported in Table 7.2 (output of the "R" does not provide standard deviations of this estimations). Comparing the results to those of the fitted unconditional *normal* distributions, one can see considerable difference in the scale parameter $\hat{c}$ and location parameter $\hat{\delta}$. This is not surprising in light of the relatively strong ARMA components.

To provide a visual impression, Figure 7.3 shows the empirical density obtained via kernel density estimation (dashed lines), along with fitted ARMA model under normal distributions. The graph is consistent with the values reported in Table 7.1. One can observe how much close is the approximation of the estimated model to the empirical one.

For comparison reasons, this ARMA model driven by the stable distributed innovations (not Gaussian like in this case) will be investigated in the next text.

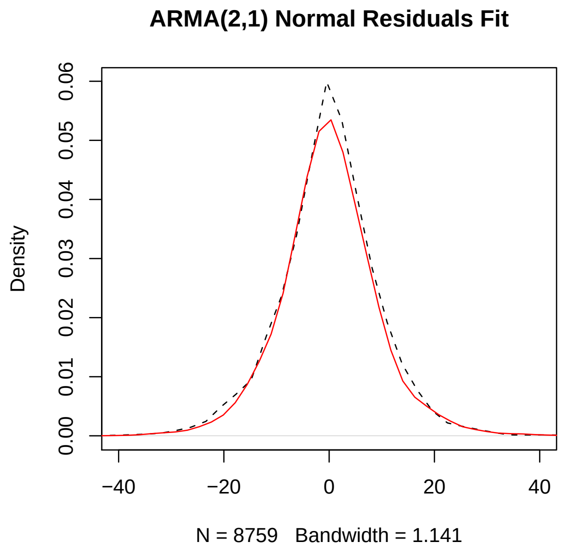

Figure 7.3: Normal ARMA(2,1) Residuals

7.2 Heteroskedastic Model

While there exist several popular model classes designed to parsimoniously and effectively fit return data, the GARCH class of models is arguably the most common.

An ARCH or GARCH model extends the mean equation, here the form (7.2), by assuming that

$$\varepsilon_t = c_t u_t,$$

where, in the normal case, $u_t \sim N(0, 1)$ and

$$c_t^2 = \omega + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_j c_{t-j}^2, \quad 0 < c_t. \tag{7.3}$$

A standard approach to detecting GARCH-dependencies in a time series y_t, is to compute the SACF of the squared series, y_t^2. Figure 7.4 and 7.5 show the SACF and PACF of the squared returns. The returns show substantial evidence of ARCH effects, as judged by the autocorrelations in Figure 7.4. The first order autocorrelation is significant, but they immediately decline. Standard Box-Jenkins methodology would suggest the need for a mixed model, i.e., one with r and s both greater than zero. As is common in economical GARCH modeling (see, for example, Bollerslev et al. [3]), it

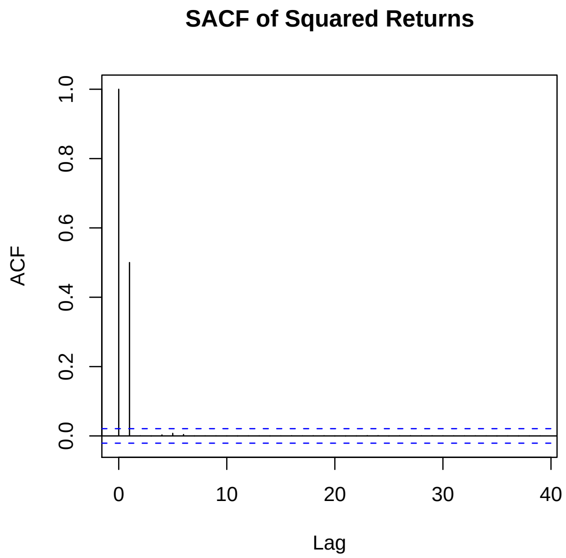

Figure 7.4: Sample Correlation Function of Squared Returns

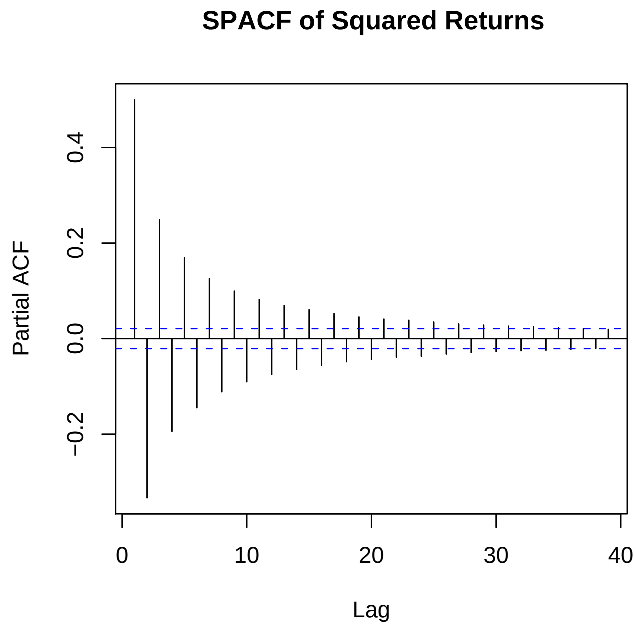

Figure 7.5: Sample Partial Correlation Function of Squared Returns

	$\hat{a}_1$	$\hat{a}_2$	$\hat{b}_1$
ESTIMATIONS	-0.7158821	-0.4048588	0.6157279
S.E.	0.01628052	0.01373888	0.01345973
	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}_1$
ESTIMATIONS	8.2386495	0.2340199	0.7139701
S.E.	0.66475579	0.01265286	0.01404709

Table 7.3: Parameter Estimates of ARMA(2,1)-GARCH(1,1)

was found that $r = s = 1$ is adequate in capturing the correlation structure for this squared series (absolute series as well, but is not shown).

The conditional mean was modeled with the same ARMA parameterization as used in the homoskedastic case. The following conditional heteroskedastic model for the return series was specified:

- ARMA(2,1)-GARCH(1,1)

The parameters in the ARMA and GARCH equation of the model were jointly estimated via conditional ML, where it is assumed that the scaled innovations, $u_t = \varepsilon_t / c_t$, are i.i.d. normal and c_t satisfies GARCH recursion (7.3). The parameter estimates are reported in Table 7.3. The three coefficients in the variance equation (7.3) are listed as $\hat{\omega}$, the intercept; $\hat{\alpha}_1$, the first lag of the squared return and $\hat{\beta}_1$, the first lag of the conditional variance.

One can observe that the estimates of ARMA parameters from already estimated ARMA-GARCH model, which correspond now to the scaled innovations, u_t, are larger in absolute value than those in Table 7.1. All the estimated parameters are statistically significant, as it follows from the p-values, computed by the "R" software.

Figure 7.6 and 7.7 show the SACFs corresponding to both the ARMA(2,1)-GARCH(1,1) residuals themselves and their squares. One can see that the parsimoniously parameterized ARMA(2,1)-GARCH(1,1) model is capable of extracting the majority of the outstanding serial correlation exhibited by both the mean and variance of the returns. Clearly, the autocorrelations are reduced from that observed in the portfolio returns themselves.

The associated SPACFs were qualitatively similar and are not shown.

The GARCH(1,1) model can be generalized to a GARCH(p,q) model–that is, a model with additional lag terms. Such higher-order models are often useful when a long span of data is used, like several decades of daily data or a year of hourly data (the case of this returns). With additional lags, such models allow both fast and slow decay of information. A particular

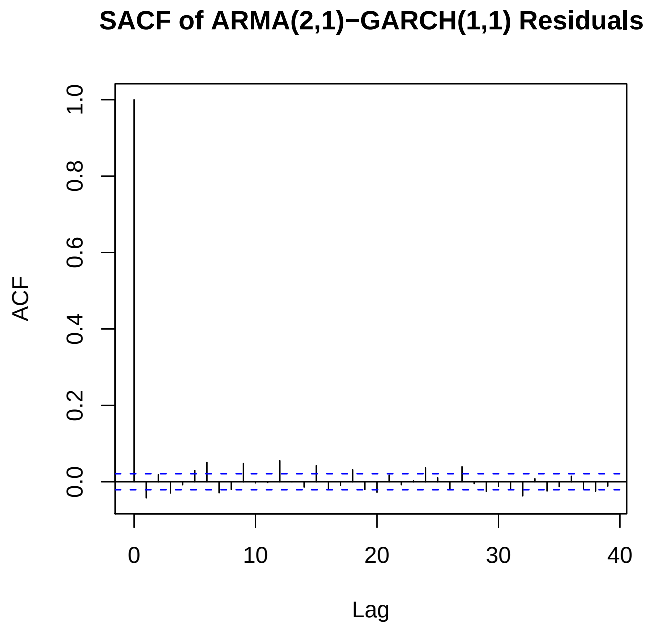

Figure 7.6: SACF of Conditional Heteroskedastic Residuals

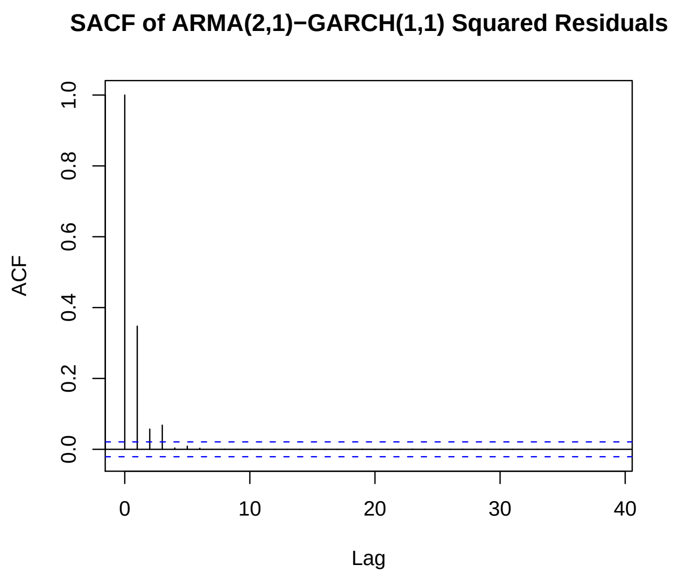

Figure 7.7: SACF of Conditional Heteroskedastic Squared Residuals

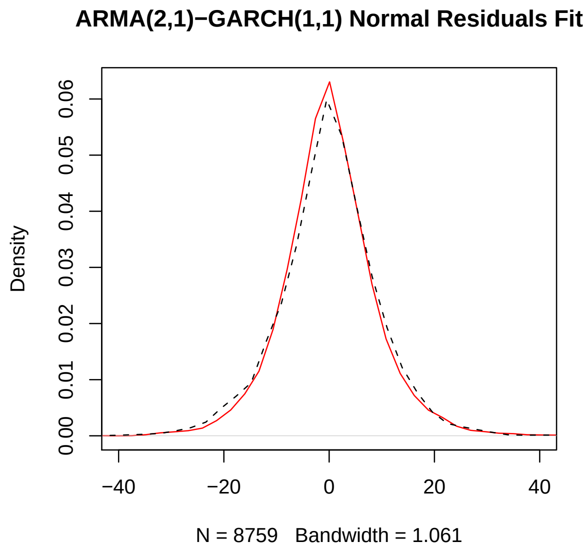

Figure 7.8: Fitted Normal ARMA(2,1)-GARCH(1,1) Densities

$\hat{\alpha}$ (index)	$\hat{\beta}$ (skewness)	$\hat{c}$ (scale)	$\hat{\delta}$ (location)
2	0	5.5347405100	0.0003108482

Table 7.4: ML Estimates of Conditional Heteroskedastic Normal Distribution

specification of the GARCH(2,2), sometimes called the "component model", is a useful starting point to this approach. Therefore, it was tried to "over"-parameterize ("under"- as well) the chosen ARMA(2,1)-GARCH(1,1) model, but there were not better results (for example, neither in AIC nor in ML value).

The parameter estimates of the fitted conditional distribution are reported in Table 7.4. Comparing the results to those of the fitted unconditional *normal* distributions - Table 6.1, one can see a significant difference in the scale parameter $\hat{c}$ and in location parameter $\hat{\delta}$ again. One can observe considerable changes against the results in Table 7.2 too.

Similar to the Figure 6.1 and Figure 7.3, Figure 7.8 shows the kernel density, and as well fitted density of the residuals corresponding to the normal conditional heteroskedastic model. The result demonstrates the dominance of combined ARMA-GARCH fit over the more simple ARMA fit, when comparing with the Figure 7.3.

Chapter 8

Conditional Stable Fitting

8.1 Simulation of Stable Distributed Innovations

In this section the practical matters of simulation, according to Klebanov et al. [21] and Devroye [11], are discussed.

A random variate with a given non-uniform distribution can often be generated in one assignment statement if an uniform source and some simple functions are available. In next text, such one-line methods for stable distributions will be reviewed. The simplest end of the spectrum — the generators that can be implemented in the one line of code, will be explored and surveyed. Random variate generators that are conceptually simple and quick to program become invariably popular, even if they are not as efficient as some more complicated methods.

The standard way of generating sequences of random variables given by a distribution function F is the inversion method (see, e.g., Devroye [11]), consisting of generating uniform $[0, 1]$ variates U_i and returning $F^{-1}(U_i)$, where F^{-1} is the inverse of F. The inversion method is based upon the property that $F^{-1}(U_i)$ has distribution function F if U_i is uniformly distributed on $[0, 1]$. It leads to one-liners only if F is explicitly invertible in terms of functions that are in $\mathcal{F}$ — one may think of a one-liner as an expression tree in which the leaves are uniform $[0, 1]$ random variables or constants, and the internal nodes are the operators or functions in the accepted class of operators, which will be called $\mathcal{F}$.

This method can be applied to geometric stable (GS) laws with $\alpha = 2$, since the skew Laplace distribution functions are available in closed forms.

Here is an algorithm based on the representation in terms of two i.i.d. uniform variables.

8.1.1 $AL^*(\kappa, \sigma)$

1. Generate a standard uniform variate U_1
2. Generate a standard uniform variate U_2, independent on U_1
3. Set $Y \leftarrow \frac{\sigma}{\sqrt{2}} \log \frac{U_1^\kappa}{U_2^{1/\kappa}}$
4. Return Y

The inversion method is not directly applicable to the general GS case, since there are no analytic expressions for the relevant distribution functions and their inverses. The solution to the problem of computer simulation of GS random variables can be obtained through the representation formula in the following Theorem:

Theorem 8.1 *Let $E_0, E_1, \ldots, E_n, \ldots$ be a sequence of i.i.d. standard exponential variables and let $R_1, \ldots, R_n, \ldots$ be a sequence of i.i.d. random variables, independent of the sequence $\{E_j\}$. If the series*

$$\sum_{k=1}^{\infty} \left(\frac{E_0}{E_1 + \cdots + E_k} \right)^{1/\alpha} R_k \tag{8.1}$$

converges a.s., then it converges to a strictly geometric stable random variable.

PROOF: See Klebanov et al. [21] Q.E.D.

Bellow, in the next subsections, several algorithms taken from Kozubowski [23] are presented.

8.1.2 A Strictly $GS_\alpha(\lambda, \tau)$ Generator

Here is a generator of a strictly GS random variable Y, given by the ch.f. (2.12).

1. Set $p \leftarrow \frac{(1+\tau)}{2}$
2. Generate a standard exponential variate Z
3. Generate uniform $[0, 1]$ variate U_1 independent of Z
4. IF $U_1 \leq p$
   THEN set $\rho \leftarrow \alpha p$ and $J = 1$
   ELSE set $\rho \leftarrow \alpha(1-p)$ and $J = -1$

5. IF $\rho = 1$
   THEN set $W \leftarrow 1$

   ELSE {
   generate uniform $[0, 1]$ variate U_2, independent of Z and U_1;
   set $W \leftarrow \sin(\pi\rho)\cot(\pi\rho U_2) - \cos(\pi\rho)$ }

6. Set $Y \leftarrow J \cdot Z \cdot (\lambda \cdot W)^{1/\alpha}$

7. Return Y

8.1.3 General $GS_\alpha(c, \beta, \delta)$ Generator

1. Generate a standard exponential variate Z

2. Generate standard stable variate $X \sim S_\alpha(1, \beta, 0)$, independent of Z (see,e.g., Section 8.1.4 or 8.1.5)

3. IF $\alpha \neq 1$
   THEN set $Y \leftarrow \delta Z + Z^{1/\alpha} cX$
   ELSE set $Y \leftarrow \delta Z + ZcX + cZ\beta(2/\pi)\log(Zc)$

4. Return Y

8.1.4 A Standard Stable $S_\alpha(1, \beta, 0)$ Generator

To generate the stable $S_\alpha(1, \beta, 0)$ r.v. X one can use the stable generator discussed in Weron [39].

1. Generate a standard exponential variate W

2. Generate uniform $(-\pi/2, \pi/2)$ variate V, independent of W

3. IF $\alpha = 1$
   THEN set $X \leftarrow \frac{2}{\pi}\left[(\frac{\pi}{2} + \beta V)\tan V - \beta\log(\frac{W\cos V}{\pi/2+\beta V})\right]$
   ELSE {
   set $b \leftarrow \frac{\arctan(\beta\tan(\pi\alpha/2))}{\alpha}$

   set $s \leftarrow [1 + \beta^2\tan^2(\frac{\pi\alpha}{2})]^{1/2\alpha}$

   set $X \leftarrow s \cdot \frac{\sin(\alpha(V+b))}{(\cos V)^{1/\alpha}} \cdot \left(\frac{\cos(V-\alpha(V+b))}{W}\right)^{\frac{1-\alpha}{\alpha}}$
   }

4. Return X

8.1.5 A Standard Stable $S_\alpha(c, \beta, \delta)$ Generator

Chambers et al. [9] suggested this algorithm:

1. Set $K(\alpha) \leftarrow \alpha - 2I_{1<\alpha}$
2. Set $\theta \leftarrow \frac{\beta K(\alpha)}{\alpha}$
3. Set $B_\alpha(z) \leftarrow \frac{\sin\left(\frac{\pi\alpha(z+\theta)}{2}\right)}{\cos\left(\frac{\pi((\alpha-1)z+\alpha\theta)}{2}\right)} \left(\frac{\cos\left(\frac{\pi((\alpha-1)z+\alpha\theta)}{2}\right)}{\cos(\frac{\pi z}{2})}\right)^{1/\alpha}$
4. Generate a standard exponential variate Z
5. Generate uniform $[0,1]$ variate U independent of Z
6. For $\beta \in [-1,1]$, IF $\alpha \neq 1$
   THEN set $X \leftarrow B_\alpha(U - 1/2)Z^{1-1/\alpha}$

   ELSE ($\alpha = 1$) set $X \leftarrow B_1(U - 1/2) - \frac{2\beta}{\pi} \log Z$,

   where $B_1(z) = \frac{2\beta}{\pi} \log\left(\frac{1+\beta z}{\cos(\frac{\pi z}{2})}\right) + (1 + \beta z)\tan(\frac{\pi z}{2})$
7. Return X

8.2 Homoskedastic Model

Now, the ARMA(2,1) model from the Section 7.1 will be driven by stable distributed innovations. Two case studies will be discussed: one for a standard α-stable distribution and one for a GS distribution.

As the first thing, standard α-stable random variables were simulated. To obtain such innovations, one moves according to the Section 8.1.5. Then, GS variables according to the Section 8.1.3 were simulated.

Using the form (7.2), the fitted values of the ARMA(2,1) model driven by this already generated α-stable and GS innovations were counted. Also, the coefficients of the model are the same as for the normal case (compare Table 7.1), only the innovations have been changed. With this method, the residuals for the α-stable conditional homoskedastic case and as well for the GS conditional homoskedastic case were obtained.

Similar to the Figure 7.3, Figure 8.1, or 8.2 respectively, show the kernel density (dashed line) and fitted densities of the residuals corresponding to the normal (red line) and to the α-stable, or GS respectively, (steel blue line) conditional homoskedastic model. One is clearly able to compare both cases:

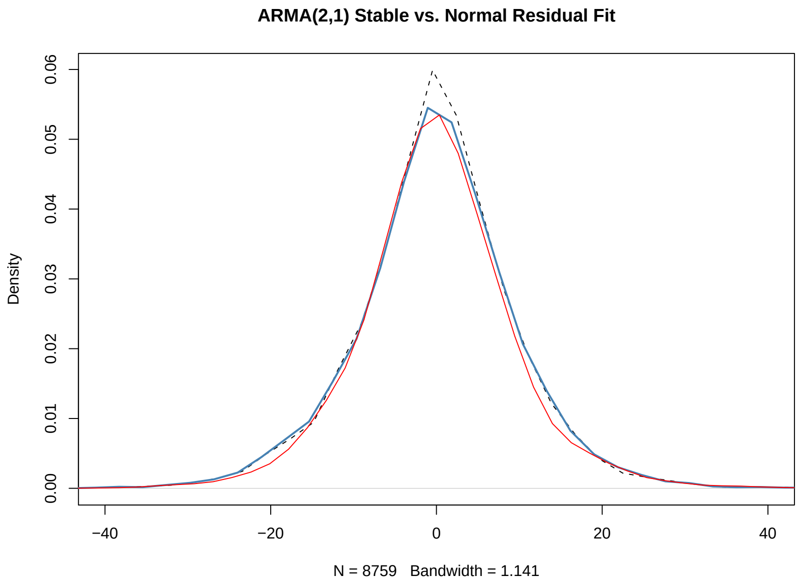

Figure 8.1: ARMA(2,1) Driven by α-Stable (steel blue line) and Normal (red line) Innovations

the normal vs. the α-stable conditional homoskedastic model and the normal vs. the GS homoskedastic model. One can see that in both figures the stable model overwhelms the normal one. Comparing α-stable and GS case, one can observe that the GS one renders the real data distribution better.

Again, conditional ML estimation was used to estimate both the fitted α-stable and GS distributions of the ARMA(2,1) model, whereby one took the first 2 values of the return series to be fixed. The parameter estimates of fitted conditional distributions are reported in Table 8.1 for both cases. One can observe that there are considerable discrepancies between them. Comparing the results to those of the *normal* fitted conditional distributions in Table 7.2, one can see a significant reduction in the scale parameters $\hat{c}$ and a significant increase in the location parameters $\hat{\delta}$. The shape (index) parameter $\hat{\alpha}$, of the α-stable distribution is slightly bellow 2, the normal case. For the GS conditional distribution, one can observe that the shape parameter is lower than in the α-stable case, therefore the excess of the kurtosis is covered better by the GS case, as one can see in Figure 8.2.

Note that the estimation of all stable models is *approximate* in the sense that the stable density function, $S_\alpha(c, \beta, \delta)$, is approximated via FFT of the α-stable characteristic function (2.2), or (2.12) respectively.

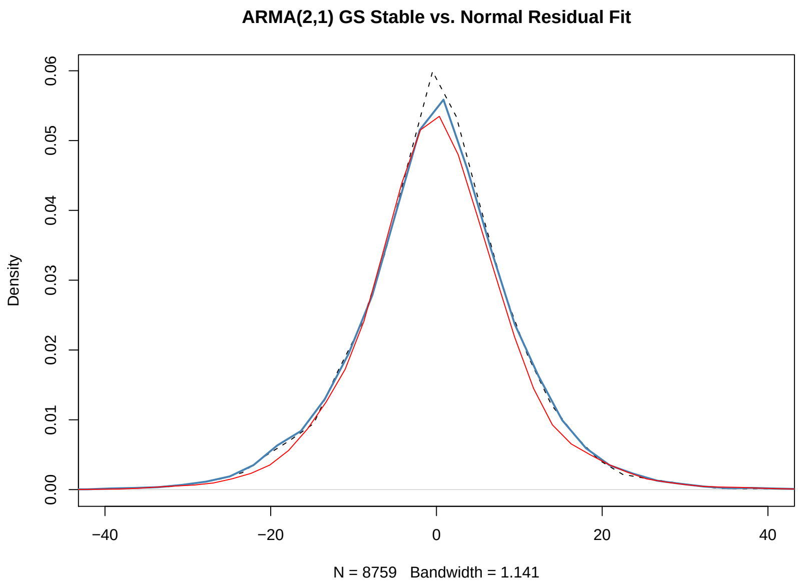

Figure 8.2: ARMA(2,1) Driven by GS (steel blue line) and Normal (red line) Innovations

	$\hat{\alpha}$ (index)	$\hat{\beta}$ (skewness)	$\hat{c}$ (scale)	$\hat{\delta}$ (location)
α-Stable case	1.786640295	-0.009850426	1.124063689	0.010957325
GS case	1.575808071	0.006353435	0.957653547	0.008025484

Table 8.1: ML Estimates of Conditional Homoskedastic Stable Distributions

8.3 Heteroskedastic Model

It turns out that ARCH-type models driven by normally distributed innovations ("building blocks") imply unconditional distributions, which themselves possess heavier tails (see Section 4.1). However, many studies have shown that GARCH-filtered residuals are themselves heavy-tailed, so that stable distributed innovations would be a reasonable distributed assumption. Thus, a combination of fat-tailed innovations and a GARCH structure appears necessary to successfully account for the excess kurtosis in this time-series research.

One uses the α-stable and the GS innovations, which were simulated in the previous section – Section 8.2.

This innovations were fitted to the ARMA(2,1)-GARCH(1,1) model, obtained in the Section 7.2. The coefficients of the model are the same as for the normal case (compare Table 7.3) again, only the innovations have been changed.

The graphical results of the ARMA(2,1)-GARCH(1,1) model driven by the α-stable and GS innovations are optically almost identical. Therefore, one provides only the α-stable case for graphical comparison. Also, similar to the Figure 8.1 and 8.2, Figure 8.3 shows the kernel density (dashed line) and fitted densities of the residuals corresponding to normal (red line) and α-stable (steel blue line) conditional ARMA-GARCH model. One is clearly able to compare the normal vs. the α-stable conditional heteroskedastic model. However, one can see that the α-stable model outperforms the normal one.

Once again, conditional ML estimation was used to estimate both the fitted α-stable and the GS distributions of the ARMA(2,1)-GARCH(1,1) model. The parameter estimates of the fitted conditional distributions are reported in Table 8.2 for both cases. One can observe that there are not considerable discrepancies between them. Comparing the results to those of the *normal* fitted conditional distributions in Table 7.4, one can see a significant reduction in the scale parameters $\hat{c}$ and a significant increase in the location parameters $\hat{\delta}$. The shape (index) parameters $\hat{\alpha}$ are slightly bellow 2, the normal case. One can see that the estimates of stable index α, which correspond now to the scaled innovations, u_t, are larger than those for the distributions in Table 8.1 and 6.1. This is what one expects, as ARCH/GARCH components absorb a portion of the excess of the kurtosis of the unconditional distribution. The skewness parameter $\hat{\beta}$ has a notable increase in his absolute value. However, one should keep in mind that as $\hat{\alpha}$ increases towards 2, the effect of the skewness parameter diminishes. For the $\hat{\alpha}$ near 1.828, even a skewness component of -0.415 is very mild, so that the large change is somewhat illusory.

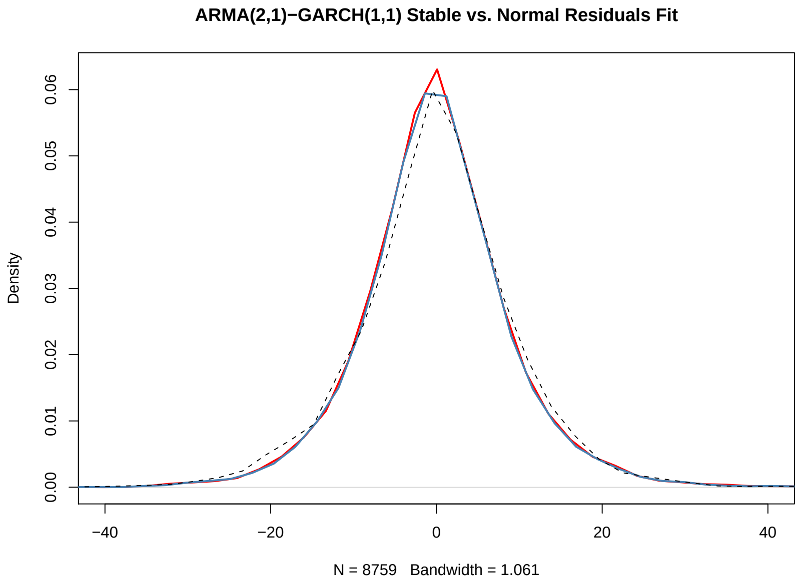

Figure 8.3: ARMA(2,1)-GARCH(1,1) Driven by α-Stable (steel blue line) and Normal (red line) Innovations

	$\hat{\alpha}$ (index)	$\hat{\beta}$ (skewness)	$\hat{c}$ (scale)	$\hat{\delta}$ (location)
α-Stable case	1.8285234	-0.4153421	2.3169318	0.2186468
GS case	1.8284702	-0.4153296	2.3228284	0.2192745

Table 8.2: ML Estimates of Conditional Heteroskedastic Stable Distributions

Chapter 9

Conclusions

Many researchers in stable distribution area argue that the Gaussian modeling of the real data is not sufficient describing all aspects of reality. They try to cover possible skewness and heavy tails by building more sophisticated models. Rendering such structures should be much more effective.

This paper investigates a couple of the stable laws (α-stable and GS) and compares it to the Gaussian one on a sample of real data from energetic field of economy. In successive steps, an unconditional, a conditional homoskedastic and a conditional heteroskedastic model is demonstrated. The stable and the normal approach to the data is confronted in all three cases. The obtained results clearly confirm that the stable models are more forcible describing the data than the normal one. However, the GS distributional assumption was found to be superior comparing to the classic Paretian α-stable, at least for the ARMA model.

Furthermore, the usual assumption, and that which is implemented in popular software packages, is that the driving innovations are either normally or Student's t distributed. The former is the "standard" assumption in financial and even most econometric or statistical models, but fails demonstrably in empirical applications (see, e.g., Gouriéroux [16]). Indeed, normality is a special, limiting case of the stable Paretian distribution, which, otherwise, allows for fatter-than-normal tails and skewness, these being precisely two of the typical "stylized facts" associated with the real returns data.

The problem is that despite of the former facts, there is almost none complex statistical software for ARMA-GARCH processes driven by stable innovations (such as SAS, SPSS, R or S-plus) until these days. There exists only a couple of small questionable commercial softwares for stable GARCH fitting and these are designed only for this purpose. (Maybe, there exists one exception and that is a "stable" library package for S-plus, but which is also extra paid.) Therefore, one can see future development of the stable laws, among the other things, in building of useful statistical utilities for the next academic research and applied mathematicians.

Further discussion along these lines and a test for the summability property in the context of ARMA-GARCH models were proposed in Rachev [33, Chapter 9], Paolella [32] and further applied in Mittnik et al. [28].

Bibliography

[1] Akaike, H. Information theory and an extension of the maximum likelihood principle. In Petrov, B. and Csaki, F.: 2nd international symposium on information theory, 267-281, (1973).

[2] Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(1986), 307–327.

[3] Bollerslev, T., Chou, R. Y., and Kroner, K. F. ARCH modeling in finance. A review of the theory and empirical evidence. *Journal of Econometrics 52*(1992), 5–59.

[4] Bougerol, P., and Picard, N. Stationarity of GARCH processes and some nonnegative time series. *Journal of Econometrics 52*(1992), 115–127.

[5] Bougerol, P., and Picard, N. Strict stationarity of generalized autoregressive processes. *Annals of Probability 20*, 4(1992), 1714–1730.

[6] Box, G. E., and Jenkins, G. M. *Time series analysis: Forecasting and control*, 2nd ed. San Francisco, CA: Holden-Day, (1976).

[7] Brockwell, P. J., and Davis, R. A. *Time series: theory and methods*, 2nd ed. Berlin, Germany: Springer, (1991).

[8] Campbell, J. Y., Lo, A. W., and MacKinlay, A. *The econometrics of financial markets.* Princeton, NJ: Princeton University Press., (1997).

[9] Chambers, J. M., Mallows, C. L., and Stuck, B. W. A method for simulating stable random variables. *Journal of the American Statistical Association 71*(1976), 340–344.

[10] Cipra, T. *Analýza časových řad s aplikacemi v ekonomii.* Praha, Czech Republic: Státní nakladatelství technické literatury, (1986).

[11] Devroye, L. *Random variate generation in one line of code.* New York, NY: IEEE Press, pp. 265-272, (1996).

[12] Engle, R. F. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica 50*(1982), 987–1007.

[13] Engle, R. F., and Bollerslev, T. Modelling the persistence of conditional variances. *Econometric Reviews 5*(1986), 1–50.

[14] Feller, W. *An introduction to probability theory and its applications.*, 2nd ed. New York, NY: John Wiley & Sons, Inc., (1957).

[15] Feller, W. *An introduction to probability theory and its applications*, 2nd ed., vol. II. New York, NY: John Wiley & Sons, Inc., (1971).

[16] Gouriéroux, C. *ARCH models and financial applications.* New York, NY: Springer, (1997).

[17] Johnston, J. *Econometric methods*, 3rd ed. Singapore: Mc-graw Hill Book Company, (1984).

[18] Kesten, H., and Spitzer, F. Convergence in distribution of products of random matrices. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 67*(1984), 363–386.

[19] Klebanov, L. B., Manija, G. M., and Melamed, I. A. A problem of Zolotarev and analogs of infinitely divisible and stable distributions in a scheme for summing a random number of random variables. *Theory of Probability and its Applications 29*(1984), 791–794.

[20] Klebanov, L. B., Manija, G. M., and Melamed, I. A. Non-strictly stable laws and estimation of their parameters. *Stability Problems for Stochastic Models*(1986), 23–31.

[21] Klebanov, L. B., Kozubowski, T. J., and Rachev, S. T. *Ill-posed problems in Probability and stability of random sums.* New York, NY: Nova Science Publishers, Inc., (2006).

[22] Kokoszka, P. S., and Taqqu, M. S. Infinite variance stable ARMA processes. *Journal of Time Series Analysis 15*, 2(1994), 203–220.

[23] Kozubowski, T. J. Computer simulation of geometric stable distributions. *Journal of Computational and Applied Mathematics 116*, 2(2000), 221–229.

[24] Liu, S. M., and Brorsen, B. W. Maximum likelihood estimation of a GARCH-stable model. *Journal of Applied Econometrics 10*(1995), 273–285.

[25] Mandelbrot, B. The pareto-lévy law and the distribution of income. *International Economic Review 1*(1960), 79–106.

[26] Mandelbrot, B. The variation of certain speculative prices. *Journal of Business 36*(1963), 394–419.

[27] Mandelbrot, B. New methods in statistical economics. *Bulletin Institute of International Statistics 40*(1964), 699–721.

[28] Mittnik, Stefan, Paolella, M. S., and Rachev, S. T. Stationarity of stable power-GARCH processes. *Journal of Econometrics 106*, 1(2002), 97–107.

[29] Mittnik, S., and Rachev, S. T. Modeling asset returns with alternative stable distributions. *Econometric Reviews 12*(1993), 347–389.

[30] Nelson, D. B., and Foster, D. P. Asymptotic filtering theory for univariate ARCH models. *Econometrica 62*, 1(1994), 1–41.

[31] Panorska, A. K., Mittnik, S., and Rachev, S. T. Stable GARCH models for financial time series. *Applied Mathematics Letters 8*, 5(1995), 33–37.

[32] Paolella, M. S. Testing the stable Paretian assumption. *Mathematical and Computer Modelling 34*, 9-11(2001), 1095–1112.

[33] Rachev, S. T., Ed. *Handbook of heavy tailed distributions in finance.*, 1st ed. Amsterdam, The Netherlands: Elsevier Science B.V., (2003).

[34] Rachev, S. T., and Mittnik, S. *Stable Paretian models in finance.* Chichester, Baffins Lane, England: John Wiley & Sons Ltd. , (2000).

[35] Samorodnitsky, G., and Taqqu, M. S. *Stable non-Gaussian random processes: stochastic models with infinite variance.* New York, NY: Chapman & Hall., (1994).

[36] Schwert, W. G. Why does stock market volatility change over time? *Journal of Finance 44*(1989), 1115–1153.

[37] Taylor, S. *Modelling financial time series.* Chichester, Baffins Lane, England: John Wiley & Sons Ltd. , (1986).

[38] Wei, W. W. *Time series analysis. Univariate and multivariate methods.* Redwood City, CA: Addison-Wesley Publishing Company, (1990).

[39] Weron, R. On the Chambers-Mallows-Stuck method for simulating skewed stable random variables. *Statistics & Probability Letters 28*, 2(1996), 165–171.

[40] Zolotarev, V. *One-dimensional stable distributions.*, vol. 65. Translations of Mathematical Monographs. Providence, R.I.: American Mathematical Society (AMS), (1986). Translation from the original – 1983 Russian edition by H. H. McFaden, ed. by Ben Silver.