COMENIUS UNIVERSITY BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

ON APPLICATION OF RISK MEASURES IN PORTFOLIO SELECTION PROBLEMS

DISSERTATION THESIS

BRATISLAVA 2009 **RNDr. Martin Jandačka**

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

APLIKÁCIA RIZIKOVÝCH MIER NA PROBLÉMY VOĽBY PORTFÓLIA

DIZERTAČNÁ PRÁCA

BRATISLAVA 2009 RNDr. Martin Jandačka

On Application of Risk Measures in Portfolio Selection Problems

Dissertation Thesis

RNDr. Martin Jandačka

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

11-14-9 Applied Mathematics

Supervisor: Doc. RNDr. Daniel Ševčovič, CSc.

Bratislava 2009

Acknowledgement

I would like to express my gratitude to all those who gave me the possibility to complete this thesis.

First and foremost I offer my sincerest gratitude to my supervisor Daniel Ševčovič for his support and guidance throughout my thesis.

Deepest gratitude are also due to Thomas Breuer for providing a stimulating research environment and his support and collaboration on research field.

Martin Jandačka
Bratislava, September 2009

Abstrakt

RNDr. Martin Jandačka

Aplikácia rizikových mier na problémy voľby portfólia

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Školiteľ: Doc. RNDr. Daniel Ševčovič, CSc.

Bratislava 2009

Dizertačná práca sa zaoberá rizikom a zahrnutím rizika do optimálnej voľby portfólia. Cieľom práce je preskúmať vplyv rizika na správanie investora.

V prvej časti definujeme problém voľby portfólia s rizikovým ohraničením na infímum konečného majetku investora. Následne na numerickom príklade analyzujeme zmenu správania investora, ktorý je vystavený rizikovému ohraničeniu.

V druhej časti dizertačnej práce zahrnieme riziko do modelu na oceňovanie opcií. Zameriame sa na riziko vyplývajúce z nedokonalého zabezpečenia syntetického portfólia ako aj na riziko bankrotu obchodného partnera. V numerickej analýze modelu ukážeme, že model dokáže vysvetliť aj takzvaný "volatility smile", nekonštantný priebeh implikovanej volatility. Následne na tomto modeli ilustrujeme problém integrovanej a čiastočnej analýzy trhového a kreditného rizika a ukážeme za akých podmienok čiastočná analýza rizika nepodcení skutočné riziko.

V tretej časti dizertačnej práce analyzujeme agregáciu časových radov. Zameriame sa na agregáciu časových radov počas dlhšieho obdobia a na optimálnu voľbu frekvencie dát. Podrobnejšie sa zaoberáme agregáciou GARCH modelov a ich podmienenou varianciou a kurtosisom (špicatosťou). Odvodíme limitné vlastnosti podmienenej variance a kurtosisu, keď sa časový horizont blíži k nekonečnu.

Abstract

RNDr. Martin Jandačka

On application of risk measures in portfolio selection problems

Comenius University Bratislava
Faculty of mathematics, physics and informatics

Supervisor: Doc. RNDr. Daniel Ševčovič, CSc.

Bratislava 2009

The dissertation thesis deals with risk and integration of risk to the portfolio selection problem. The goal of the dissertation thesis is to explore the influence of risk on investor's behavior.

In first part we formalize the portfolio selection problem with risk constraint on the final infimum wealth of investor. We then analyse behavior of the investor under the presence of the risk constraint on a numerical example.

In second part of the dissertation we integrate the risk generated by imperfect portfolio hedging and from the possibility of default of the counterparty, into the option pricing model. We numerically analyse the model and show its ability to explain volatility smile. We illustrate the problem of integrated versus separated analysis of credit and market risk and show when the separated analysis will not underestimate the overall risk.

In the last part of the dissertation we focus on aggregation of time series models. We deal with the aggregation of time series over longer time horizon and with the choice of optimal data frequency. We analyse the aggregation of GARCH model and its conditional variance and kurtosis. We derive the limits behavior of conditional variance and kurtosis when the time horizon goes to infinity.

Foreword

Risk was always a key concept in portfolio selection problems. One of the goals of portfolio selection is to minimise the risk of the portfolio. In some models it is even possible to reduce the risk completely. Such a reduction is usually possible only with respect to one source of risk.

There can be many different source of risk in a portfolio: risk from volatile nature of the equities, credit risk due to default of the counterparty, imperfect hedging, missspecification of the time series models, ...

It is important to study and understand these risks, even if we will never be able to control or reduce these risks completely. The understanding of the risk can help us to better choose from different investment possibility, from different strategies or regulation conditions.

The goal of this thesis is to explore the influence of risk on investor's behavior. We will focus on longer time horizon as for the short time period and one period models is the influence of risk relatively well known.

Martin Jandačka, Fall 2009
Author

Contents

Chapter 1

Introduction

Risk was always a key concept in portfolio selection. In one period portfolio selection models (see e.g. [75], [76], [94], [98]) risk, measured as variance of return, has the same importance as return. In multi-period portfolio selection, however, risk mainly appeared only in the utility function encoding the risk behavior of the investor. Multi-period portfolio selection is a prominent topic in finance for almost forty years. The vast body of literature can be classified along several criteria:

- continuous time (see e.g. [3, 14, 30, 36, 57, 63, 74, 79, 80, 83, 97]) versus discrete time models (see e.g. [31, 48]),

- investment-consumption models (see e.g. [3, 27, 31, 36, 49, 57, 74, 79, 80, 83]) versus pure investment models (see e.g. [14, 30, 81, 97]),

- single asset (see e.g. [36, 48, 61, 74, 79, 80, 97]) versus multi asset models (see e.g. [3, 14, 27, 30, 31, 57, 63, 81, 83]),

- models without transaction costs (see e.g. [48, 49, 68, 79, 80]) versus models with transaction costs (see e.g. [3, 14, 27, 30, 31, 36, 57, 63, 74, 81, 83]) or brokerage fees (see e.g. [97]),

- asset prices following Brownian motion versus more general stochastic processes (see e.g. [34, 44, 64, 89]),

- complete versus incomplete market models (see e.g. [49, 61, 71, 72]).

In this multi-period literature the concept of risk plays a marginal role. There is, however, substantial work on one-period portfolio selection under risk constraints other than variance, see [47] and reference therein. Usually in the multi-period setting, risk enters via utility functions, which encode the risk attitude of the investor, or via short selling constraints, which restrict the portfolio value to be positive, or via margin requirements. In contrast to this literature we put central emphasis on risk constraints formulated as restrictions on economic capital.

Financial institutions usually have no specified utility function. Rather they have some economic capital at their disposal and try to conduct business so as to maximize profits making sure that their economic capital is sufficient for the business. How much economic capital is needed for some business is described by risk measures. Coherent one-period risk measures were introduced some time ago [6, 37, 45]. Recently the concept of coherent risk measure has been extended to a multiperiod setting [7, 28, 29, 100]. In such a framework it is possible to pose the portfolio selection problem faced in reality by many financial institutions: In markets with stochastic prices and transaction costs, choose a portfolio strategy which maximizes expected long term growth and ensures that economic capital is sufficient at all times.

In Chapter 3 we specify the stochastic control problem with the risk constraints. We introduce the limit control strategies and regions of no transaction for this problem and discuss a simplified numerical approximation method, which displays some of the key phenomena of stochastic control problems with dynamic risk constraints. The optimisation problem we consider is relevant not only for portfolio management but also for risk measurement. Integrating credit and market risk requires the choice of one common time horizon for credit and market risk. This usually will be the longer time horizon of credit risk, e.g. one year. When determining the profit-loss distribution of market risk on such a long time horizon, we cannot assume anymore that the trading book is largely the same at the end of the time horizon. This crucial assumption is usually made in calculation of market risk on short time horizons of a day or a week. Without this assumption the rebalancing behavior of the portfolio manager has to be taken into account when determining the portfolio distribution. Within a time horizon of one year the portfolio manager receives new information about the market and has the opportunity to sell and buy assets. For evaluation purposes not the actual rebalancing strategy but the optimal rebalancing strategy is relevant. To determine the optimal rebalancing strategy is exactly the topic of Chapter 3, parts of this chapter were published in [22].

The standard option pricing theory is derived by complete elimination of risk from hedged portfolio (see Black and Scholes [15, 16]), which is possible due to the non existing transaction costs in this model. In the past years, the Black–Scholes equation and its generalizations for pricing derivatives has attracted a lot of attention from both theoretical as well as practical point of view. According to the classical Black-Scholes theory [12, 15, 16, 51, 66, 92, 101] the present cost of an option equals to the initial value of a solution to the so called Black-Scholes equation. This theory is capable of valuing options and other derivative securities over moderate time intervals in which transaction costs and the risk from a volatile portfolio are negligible. On the other hand, if transaction costs like e.g. bid-ask spreads are taken into account the classical Black-Scholes theory is no longer valid. In order to maintain delta hedge one has to make frequent portfolio adjustments yielding thus substantial increase in transaction costs. On the other hand, rare portfolio adjustments leads to increase of the risk from a volatile (unprotected) portfolio.

One of the interesting problems in the modelling of pricing of financial derivatives is the question how to incorporate both transaction costs and risk arising from a volatile portfolio into the governing equation. In [65], M. Kratka derived a mathematical model for pricing derivative securities in the case when both transaction costs as well as the risk from a volatile portfolio are taken into account. The model is based on the Black-Scholes parabolic PDE in which transaction costs are described by Leland's approach (see e.g. [9, 51, 66, 70]) whereas the risk from a volatile portfolio is described by the average value of the variance of the synthetized portfolio.

In Chapter 4 we revisit Kratka's approach in order to derive a model which is mathematically well posed and scale invariant. We will extend the model by the possibility of default of counterparty and compare the integrated model with adding up the market and credit risk, which is often seen as conservative risk assessment. Recent author's paper (see [24]) shows that separating the market and credit risk can lead to significant underestimation of the integrated risk. We will show when the separated analysis of the market and credit risk will not underestimate the overall risk. We present qualitative analysis of the governing equation and we derive a robust numerical scheme. We perform extensive numerical testing of the model and compare the results to real option market data. We also introduce a concept of the so-called implied RAPM volatility and implied risk premium coefficient. These results for option model without default probability and the dividends were published in [53].

In any model of portfolio selection or option pricing the fundamental position have estimation of the model and time series parameters. With higher frequency data being increasingly available and attention focusing on longer time horizons we face the choice whether or not to use the higher frequency data available in long term analysis. At first sight it seems clear that it should be used. If we restricted ourselves to the low frequency data we either would have very few data points or use very old historical data for getting reliable parameter estimates. Neither is desirable. On the other hand, when we use the higher frequency data the time horizon of the forecast is several time steps ahead. A long term analysis then has to analyse the distribution arising from aggregating the high frequency model over several time steps. This motivates our analysis of aggregated distributions. We concentrate our analysis on GARCH model as introduced by Bollerslev [17]. The approach we take is to estimate a strong GARCH model for single time steps of suitable length and then aggregate over sufficiently many time steps to arrive at the desired time horizon. Drost and Nijman [39] in a landmark paper showed that the temporal aggregation of a strong GARCH process is in general not a strong GARCH. Therefore they introduced the larger classes of semi-strong and weak GARCH models. For semi-strong GARCH processes the mean and variance of innovations are determined, but other properties of the distribution of innovations are not determined. In particular, the innovations need not be independent or identically distributed. For weak GARCH processes not even the mean and variance are determined, we just have a linear predictors.

Weak GARCH processes have the advantage of aggregating to weak GARCH processes, but for purposes of risk management they do not convey much information. For mean and variance they only specify the best linear predictor, other properties of the distribution are not specified at all. In risk management we often need more information about the conditional distribution: quantiles, higher moments, and for risk measures like Expected Shortfall even the full distribution function in the tails. This information is not specified by semi-strong or weak but only by strong GARCH processes. For this reason we will focus on analysis of the aggregated distribution of strong GARCH processes accepting that this aggregated distribution is itself not a strong GARCH process, but we will derive the properties of higher moments of aggregated time series also for more general semi-strong GARCH process.

In Chapter 5 we do not deal with contemporaneous aggregation, as do for example Nijman and Sentana [84]. Drost and Werker [40] define continuous time GARCH processes which exhibit weak GARCH behavior at all discrete frequencies. They show that the discrete time GARCH processes arising from the observation of continuous time GARCH processes have excess kurtosis even if the continuous time process does not. In our paper we also describe a similar phenomenon: The processes arising from an aggregation of strong GARCH processes have excess kurtosis even if the basic process does not. But obviously the class of processes we consider do not satisfy the assumptions of continuous time GARCH. Meddahi and Renault [78] also investigate the temporal aggregation of volatility models. They consider a class of processes more general than weak GARCH, which works even if fourth moments are not finite. Our expressions for the variance of the aggregated conditional distribution (Theorems 5.1 and 5.2) do not assume either that fourth moments are finite, but the analysis of the kurtosis of the aggregated conditional distribution assumes innovations to be symmetric and have finite fourth moments. Baillie and Bollerslev [10] investigate conditional mean and variance (but not conditional kurtosis) of GARCH error distributions and specify all unconditional moments of the error distribution. However, they do not analyse conditional kurtosis. The chapter is part of working paper [23].

Chapter 2

Goals of the thesis

Risk was always a key concept in portfolio management. In portfolio selection model risk appears in the utility function encoding the risk behavior of the investor. In one period model risk constraint and their influence on the investor behavior is well explored. Recently the concept of coherent risk measure has been extended to a multiperiod setting. This allow us to explore the problem of dynamic multiperiod portfolio selection problem under risk constraint.

Risk also plays the crucial role in the pricing of the options. In the well-known Black-Scholes theory risk is completely eliminated from the portfolio by a continuous hedge. This theory is capable of valuing options and other derivative securities over moderate time intervals in which transaction costs and the risk from a volatile portfolio are negligible. On the other hand, over the longer time period, the transaction costs can not be ignored. Imperfect hedge of the synthetic portfolio will result in increase of the risk from a volatile portfolio. This risk must be taken into the account on the longer time intervals. Additionally the unprotected portfolio must not be the only source of the risk for the investor. The investor can be subject to the default of counterparty which can open previously well hedged positions of investor.

For dynamic portfolio the properties of the portfolio over short period are equally important as the properties of the portfolio over longer time horizon. The short period properties of time series are important for hedge of the portfolio. On the other hand for risk analysis of static portfolio we usually need to know the properties of win/loss distribution over longer time horizon. The question arise whether or not to use the higher frequency data available in long term analysis. For static portfolio the calculation of risk over longer time interval require the choice of time series model which capture the properties of the given variables on the end of the time horizon.

Our goal is to explore the influence of risk on investor's behavior. We will focus on the longer time horizon as for the short period, resp. one period, models is the influence of the risk relatively well known. We will mainly explore the problem of risk in multi period dynamic portfolio selection, the risk in synthetic portfolio for option pricing over longer time intervals and problem of time forecasting over a longer time horizon.

Most of the results of this dissertation thesis were already published in recent papers [22, 23, 24, 25, 53] co-authored by me.

Chapter 3

Portfolio selection with transaction costs under risk constraints

In this chapter we formalize the following portfolio selection problem: An investor subject to proportional transaction costs allocates funds to multiple stocks and a bank account, to maximize the expected growth rate of the portfolio value under a risk constraints.

The chapter is structured as follows. In Section 3.1 we specify the stochastic control problem with the risk constraints. Limit control strategies and regions of no transaction for the problem are introduced in Section 3.2. In Section 3.3 strongly simplified numerical example for risk in form of Expected Shortfall (ES) is discussed, which nevertheless displays some of the key phenomena of stochastic control problems with dynamic risk constraints. These results were published in [22].

3.1 The Stochastic Control Problem

We assume the investor operates on a market of one riskless bond ("bank") with constant interest rate r and m different stocks. The evolution of the riskless bond B is given by

$$dB(t) = Brdt. \tag{3.1}$$

The evolution of stock price $S(t)$ is described by an m-dimensional Wiener process $W(t)$ as

$$dS_i(t) = S_i(t)\left(\mu_i dt + \sum_{j=1}^{m} \sigma_{ij} dW_j(t)\right), \quad i = 1, ..., m. \tag{3.2}$$

Here σ is a $m \times m$ positive definite matrix representing the covariance structure and μ represents the drift. The covariance matrix is $\sigma'\sigma$, where σ' denotes the transpose of matrix σ.

The investor has initially x_0 Euros invested in the bank and $(x_1, ..., x_m)$ Euros invested in the stocks $1, ..., m$. He can control his portfolio composition by buying and selling arbitrarily large or small amounts of stock from his bank account at any time, exchanging directly one stock against the other is not allowed. His portfolio selection strategy π is described by control processes $Z(t)$, $U(t)$. Here the i-th component of $U(t)$ represents the cumulative amount of money obtained from selling stock i before incurring transaction costs. The i-th component of $Z(t)$ represents the cumulative amount of money used to buy stock i before incurring transaction costs.

Buying and selling stock incur proportional transaction costs. Let

$$\begin{aligned} C_b &= (C_{b1}, ..., C_{bm}) \geq 0, \\ C_s &= (C_{s1}, ..., C_{sm}) \geq 0 \end{aligned}$$

be vectors of proportional transaction costs for buying and selling. Buying one Euro worth stock i will cost $(1 + C_{bi})$ Euro in cash from the bank. Selling one Euro worth of stock i will result in $(1 - C_{si})$ Euro in cash that is added to the bank.

Given the portfolio strategy π in terms of buy and sell processes $Z(t)$, $U(t)$ the cotrolled evolution of values V_0^π of investment in bond and V_i^π of investment in stocks follows the stochastic differential equations

$$dV_0^\pi(t) = V_0^\pi(t) r dt - (\mathbf{1} + C_b) \cdot dZ(t) + (\mathbf{1} - C_s) \cdot dU(t) \tag{3.3}$$

$$dV_i^\pi(t) = V_i^\pi(t) \left(\mu_i dt + \sum_{j=1}^{m} \sigma_{ij} dW_j(t) \right) + dZ_i(t) - dU_i(t). \tag{3.4}$$

Here $\cdot$ denotes the standard dot product, e.g. $(\mathbf{1} + C_b) \cdot dZ(t) = \sum_{i=1}^m (1 + C_{bi}) dZ_i(t)$, and $\mathbf{1}$ denotes a vector of ones. Since the investor starts with x_0 Euros invested in the bank and $x := (x_1, ..., x_m)$ Euros invested in stocks $1, ..., m$ we have $V_0^\pi(0) = x_0$ and $V_i^\pi(0) = x_i$. We can rewrite equation (3.4) in vector notation $V^\pi = (V_1^\pi, ..., V_m^\pi)$,

$$dV^\pi(t) = V^\pi(t) * (\mu dt + \sigma dW(t)) + dZ(t) - dU(t).$$

Here $*$ denotes componentwise multiplication of vectors, $x * y := (x_1 y_1, ..., x_m y_m)$.

The total market value p of the portfolio $(V_0^\pi, ..., V_m^\pi)$ can be defined as sum:

$$\begin{aligned} p(V_0^\pi(t), V^\pi(t)) &= V_0^\pi(t) + \sum_{i=1}^{m} V_i^\pi(t) \\ &= x_0 + \sum_{i=1}^{m} x_i + \int_0^t \left(rV_0^\pi(s) + \sum_{i=1}^{m} \mu_i V_i^\pi(s) \right) ds \qquad (3.5) \\ &\quad + \int_0^t \sum_{i,j=1}^{m} V_i^\pi \sigma_{ij} dW_j(s) - \sum_{i=1}^{m} C_{bi} Z_i(t) - \sum_{i=1}^{m} C_{si} U_i(t), \end{aligned}$$

or in vector notation

$$\begin{aligned} p(V_0^\pi, V^\pi(t)) &= (V_0^\pi, V^\pi(t)) \cdot \mathbf{1} \\ &= (x_0, x) \cdot \mathbf{1} + \int_0^t (r, \mu) \cdot (V_0^\pi(s), V^\pi(s)) ds \qquad (3.6) \\ &\quad + \int_0^t V^\pi \cdot (\sigma dW(s)) - C_b \cdot Z(t) - C_s \cdot U(t). \end{aligned}$$

We define the net value of portfolio $(V_0^\pi, ..., V_m^\pi)$ as the total market value of portfolio after transfer of all stock wealth to the bond

$$P(V_0^\pi, V^\pi) := V_0^\pi + \sum_{i=1}^{m} \min[(1 - C_{si})V_i^\pi, (1 + C_{bi})V_i^\pi]. \qquad (3.7)$$

In our approach, the risk is measured by a dynamic risk measure ρ corresponding to a risk-adjusted value measure $v = -\rho$. A coherent risk measure (see i.e. [1, 6]) is defined by a set of coherency axioms as

Definition 3.1. *Consider a set V of real-valued random variables. A function $\rho : V \to \mathbb{R}$ is called risk measure if it is*
(i) monotonous:

$$X, Y \in V,\ Y \geq X \quad \Rightarrow \quad \rho(Y) \leq \rho(X),$$

(ii) sub-additive:

$$X, Y, X + Y \in V \quad \Rightarrow \quad \rho(X + Y) \leq \rho(X) + \rho(Y),$$

(iii) positively homogenous:

$$X \in V, h > 0, hX \in V \quad \Rightarrow \quad \rho(hX) = h\rho(X),$$

(iv) translation invariant:

$$X \in V, a \in \mathbb{R}, X + a \in V \quad \Rightarrow \quad \rho(X + a) = \rho(X) - a.$$

In a continuous time setting dynamic risk measures were introduced by Cheridito et al. [28, 29]. A strategy π is admissible for a starting point x if it is $\mathcal{F}$-adapted and the controlled process V^π to which this strategy leads has positive risk adjusted value, i.e. $v(p^\pi) \geq 0$, or negative risk, i.e. $\rho(p^\pi) \leq 0$.

We assume the dynamic risk measure ρ is derived from an one-period risk measure ρ_0 via

$$\rho(X) = \rho_0 \left(\inf_{t \in [0,\infty]} X_t \right). \tag{3.8}$$

This implies that if $\rho(X) \leq 0$ then $\rho_0(X_t) \leq 0$ for all t. Therefore the strategy π must be admissible for all times with respect to risk measure ρ_0. Then the static risk adjusted value of controlled wealth will always be positive. Associated in a one-to-one way to the static coherent risk measure ρ_0 is a norm-closed convex cone $C \subset L^\infty(\Omega, \mathcal{F}, P)$, see Delbaen [37], Theorem 2.3. C is the set of all real valued random variables X, representing profits and loses of portfolios, for which $\rho(X) \leq 0$. A strategy $\pi_t = (Z_t, U_t)$ is acceptable for a starting point $x \in C$ if and only if V_0^π, V^π remains in C for all times. Denote $\mathcal{A}_x^\rho$ the set of all strategies which satisfy the risk constraint $\rho(V^\pi) \leq 0$ for the starting point (x_0, x).

The objective of the investor is to maximize the long-term average expected growth of the portfolio value by using an optimal admissible strategy. The goal function is

$$\lim_{T \to \infty} E(\log(P(V_0^\pi, V^\pi)))/T. \tag{3.9}$$

The stochastic control problem is to find the admissible strategy which for a given starting point x maximizes the long term growth rate where the process under control π follows the dynamic (3.4).

3.2 Limit control strategies and regions of inaction

For portfolio selection problem under transaction costs optimal strategies can be usually identified with control limit strategies [3, 36, 95, 97]. The control limit strategy with the control limits $[A_i, B_i]$ looks as follows. If the proportion of the stock i is below the limit A_i the strategy is to buy the minimal amount of the stock necessary to bring the proportion of the stock i back to A_i. We are in the buy region $\mathcal{B}_i$ of the the stock i. If the proportion of the stock i is above the limit

B_i the strategy is to sell the minimal amount of the stock necessary to bring the proportion of the stock i back to B_i. We are in the sell region $\mathcal{S}_i$ of the stock i. If the proportion of the stock i is in the interval $[A_i, B_i]$ the stock i is neither bought nor sold. We are in the no transaction region NT_i of the stock i. If for all the stocks i the proportions are in $[A_i, B_i]$, we do not buy or sell any stock. This define the no transaction region NT. The regions of inaction for the stock i contain the optimal stock proportion in the absence of transaction costs.

These concepts carry over almost unchanged to the stochastic control problem with coherent risk constraints. The main difference is that the no transaction regions are not only defined by the trade-off between expected return to be gained and transaction costs to be paid, but also by the risk constraint. Intuitively, the no transaction region determined by considerations of expected return maximisation may but need not overlap with the region satisfying the risk constraint. If it does not overlap the investor is required to perform risk reducing transactions even if he would not make any transactions to improve return.

The optimal proportions are closely related to Merton's [80] optimal proportions $(\sigma\sigma')^{-1}(\mu - r\mathbf{1})/\gamma$ for an investor optimising expected utility $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ from consumption c. Here γ is the coefficient of relative risk aversion and $\mathbf{1}$ represents vectors with all elements equal to one. In the presence of transaction costs it is not feasible any more to trade continuously in order to maintain the optimal stock proportions. The limit control strategies with their transaction regions are a natural response to the imposition of transaction costs.

For dynamic portfolio optimisation problems in the presence of transaction costs there are some traditional main solution techniques:

- Approximate solution of temporally and/or spatially discretized versions of the stochastic control problem with scenario trees [2, 11, 19, 56, 73, 91]
- Martingale Techniques [32, 60, 86]
- Stochastic Duality Theory [33, 34, 35, 49, 50, 64, 89]
- Finite difference PDE solution methods of the corresponding HJB-equation [20, 21, 102]
- Markov chain approximations [82]

These approaches have been applied to various forms of portfolio selection problems, but not yet to the problem with coherent risk constraints, as introduced above. In the following section we will give a simple technique to approximate solution to the stochastic control problem with coherent risk constraints based on the approximation with scenario trees and we will discuss the first impression of

- how the risk constraints affect expected return and risk,
- how the possibility to rebalance the portfolio affects risk.

3.3 A numerical example of portfolio selection under transaction costs and expected shortfall constraints

We consider an investor who choose only between the bank and one stock on a finite time horizon, with 10 discrete equally spaced points in time in which investor can hedge his portfolio by selling or buying stocks. The evolution of the stock value is approximed by a binomial tree. In the first setting, which we include for comparison, the investor cannot make any transactions at intermediate time and he will have the same amount of stocks and "bonds" through all time steps. In the other setting the investor can make transactions at these 10 points in time. In second setting the investor follows a portfolio selection strategy which aims at maximising the expectation value of the log-return of the portfolio, the only constraint being that he cannot sell short the bond or the stock. In the third setting the investor also tries to maximize the expected value of the log-return, but is subject to the constraint that from one time step to the next the Expected Shortfall (ES) be smaller then 3% of the current portfolio value.

Definition 3.2. Expected shortfall *with* $100(1-\alpha)\%$ *confidence level is defined as*

$$ES_{1-\alpha} = -E[X|x \leq x_\alpha], \tag{3.10}$$

where x_α *denotes* $100\alpha\%$*-percentile of the profit/loss distribution of* X.

In this setting the investor is given some economic capital and is allowed to make only transactions which require at most this amount of economic capital. To calculate the behavior of the investor in the third setting we go backward through the tree and in every node calculate the no transaction region.

The following Table 3.1 and 3.2 represent the resulting expected return and risk numbers. The proportional transaction costs for buying stock are equal to transaction costs for selling stocks and are equal to 0.5% and 0.1% resp., the risk free interest rate r over the time horizon is 2%, the stock has a log-normal distribution with $\mu = 0.05$ over the ten periods time horizon and $\sigma = 0.2$. The first two columns give the ratio of bonds and stocks in the initial portfolio. The expected return numbers are given in the last three columns of Table 3.1. The third column gives the expected returns when no transactions are possible. The fourth column gives expected returns of an investor subject only to a no-short-selling constraint, the last column gives expected returns for an investor subject to the ES constraint.

With transaction costs at 0.5% we observe that expected returns are lowest for the investor who is subject to the ES constraint, as long as the initial portfolio less than roughly 50% bonds. As the number of bonds in the initial portfolio increase the return diminishing effect of the ES constraints becomes smaller until at an initial position of roughly 50% bonds the expected return of the investor subject to the ES constraint is higher than for the investor not able to perform any transactions. The investor subject only to the no-short-selling constraint achieves highest expected returns. The expected returns of the investor subject only to the no-short-selling constraint are higher then those of the investor subject to the ES constraint by about 0.03 to 0.27 percentage points.

When transaction costs are only at 0.1% the expected returns improve in the two settings which allow intermediate transactions. Comparing the Tables 3.1a and 3.1b we see that for the investor subject only to the no-short-selling constraint expected returns are considerably higher at 0.1% than at 0.5% transaction costs when the initial portfolio consists primarily of bonds. This is because at low transaction costs the investor can redirect his investment into stocks with smaller loss given by transaction costs. When the initial portfolio consists primarily of stocks expected returns are not improved by lower transaction costs because it is not necessary to shift from bonds into stocks. For the investor subject to the ES constraint expected returns are consistently higher when transactions costs are lower, for all initial portfolios. With low transactions costs expected returns of the investor subject to the ES constraint hardly depend on the initial portfolio. This can be explained by the fact that the ES constraint and the goal to maximize expected log-returns force the investor quickly achieve a portfolio with roughly 50% stocks.

Table 3.2 gives the ES numbers of the three investors. The third column gives the total ES over the time horizon when no transaction are possible. The fourth column gives the total ES of an investor subject only to a no-short-selling constraint, the last column gives expected returns for an investor subject to the ES constraint of 3% at each time step. We observe that for initial portfolios consisting primarily of stock, ES is highest when no transactions are possible, and lowest when the ES constraint is in force. For initial portfolios risk consisting primarily of bonds, ES is lowest when no transactions are possible, and highest when only the no-short-selling constraint is in force. This is due to the fact that the optimal portfolio with only the no-short-selling constraint in force carries $20-30\%$ bonds. Without transaction possibilities and an initial position of more 40% bonds risk is lower, and so is expected return. Under the ES constraint, total ES does depend significantly on the initial composition of the portfolio. It is between one half

Table 3.1: Maximal expected returns achievable without transactions (column 3), with transactions subject only to a no-short-selling constraint (column 4), and with transactions subject to an ES constraint (column 5). Column 1 and 2 give the proportions of bonds and stocks in the initial portfolio. The risk free interest rate r is 2% over 10 periods, the stock has log-normal distribution with $\mu = 0.05$ and $\sigma = 0.2$.

(a) 0.5% transaction costs

initial value		expected return		
bonds	stocks	no trnsct	no short-sell.	ES cnstr
0.1	0.9	0.0281	0.0281	0.0254
0.2	0.8	0.0284	0.0284	0.0259
0.3	0.7	0.0284	0.0284	0.0264
0.4	0.6	0.0280	0.0280	0.0269
0.5	0.5	0.0273	0.0275	0.0272
0.6	0.4	0.0262	0.0270	0.0267
0.7	0.3	0.0248	0.0265	0.0262
0.8	0.2	0.0229	0.0260	0.0257
0.9	0.1	0.0208	0.0255	0.0252

(b) 0.1% transaction costs

initial value		expected return		
bonds	stocks	no trnsct	no short-sell.	ES cnstr
0.1	0.9	0.0281	0.0281	0.0270
0.2	0.8	0.0284	0.0284	0.0271
0.3	0.7	0.0284	0.0284	0.0272
0.4	0.6	0.0280	0.0283	0.0273
0.5	0.5	0.0273	0.0282	0.0274
0.6	0.4	0.0262	0.0281	0.0273
0.7	0.3	0.0248	0.0280	0.0272
0.8	0.2	0.0229	0.0279	0.0271
0.9	0.1	0.0208	0.0278	0.0270

and three quarters of the ES numbers when only the no-short-selling constraint is in force. This compares to expected returns lowered by $1 - 10\%$ when the ES constraint is introduced.

With lower transaction costs this picture does not change qualitatively, as Table 3.2(b) shows. Under the mere no-short-sell constraint ES numbers are higher when transaction costs are lower and the initial portfolio carries primarily bonds. This is due to the incentive to shift to stock more quickly when transactions costs are lower. Under the ES constraint, total ES is higher when transactions costs are lower, but again does depend significantly on the initial composition of the portfolio. The higher total ES is caused by increased incentive to shift to stock when transaction costs are low.

The temporal evolution of no transaction regions are shown in Table 3.1. The

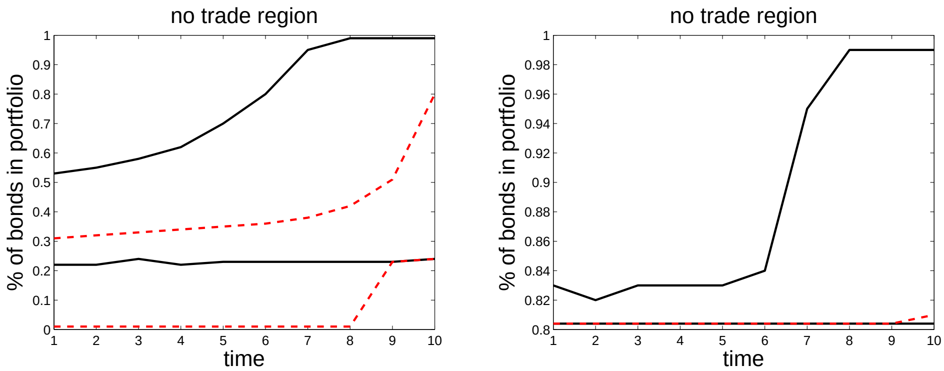
Table 3.2: ES at 0.5% and 0.1% transaction costs:

(a) 0.5% transaction costs

initial value		Expected Shortfall (95% conf.int.)		
bonds	stocks	no trnsct	no short-sell.	ES cnstr
0.1	0.9	0.263-0.279	0.257-0.273	0.141-0.149
0.2	0.8	0.232-0.246	0.227-0.241	0.143-0.153
0.3	0.7	0.200-0.213	0.196-0.208	0.144-0.154
0.4	0.6	0.169-0.180	0.170-0.180	0.138-0.147
0.5	0.5	0.138-0.147	0.170-0.181	0.136-0.145
0.6	0.4	0.107-0.114	0.172-0.183	0.137-0.146
0.7	0.3	0.075-0.080	0.173-0.183	0.137-0.145
0.8	0.2	0.044-0.048	0.174-0.184	0.136-0.144
0.9	0.1	0.013-0.015	0.174-0.185	0.140-0.148

(b) 0.1% transaction costs

initial value		Expected Shortfall (95% conf.int.)		
bonds	stocks	no trnsct	no short-sell.	ES cnstr
0.1	0.9	0.253-0.267	0.253-0.267	0.155-0.166
0.2	0.8	0.223-0.236	0.223-0.236	0.149-0.159
0.3	0.7	0.193-0.236	0.204-0.216	0.151-0.161
0.4	0.6	0.162-0.172	0.204-0.216	0.150-0.160
0.5	0.5	0.132-0.140	0.204-0.216	0.153-0.163
0.6	0.4	0.102-0.109	0.204-0.216	0.156-0.166
0.7	0.3	0.072-0.077	0.204-0.216	0.152-0.162
0.8	0.2	0.042-0.045	0.205-0.217	0.154-0.165
0.9	0.1	0.012-0.013	0.205-0.217	0.149-0.159



(a) With no-short-selling constraint (b) With ES constraint

Figure 3.1: Temporal evolution of no transaction regions at 0.5% transaction costs (black, solid line) and at 0.1% (red, dashed line). Figure 3.1(a) shows the no transaction region for the investor subject only to the no-short-selling constraint. Figure 3.1(b) shows the no transaction region for the investor subject to the ES constraint. When transaction costs are higher no transaction regions in general include portfolios with higher proportions of bonds. This is due to the lacking incentive to move into stock when transaction costs are high. We observe the well-known fact that towards the end of the investment it is optimal to be more conservative and increase the investment in bonds. For an investor subject only to the no-short-selling constraint, we see that at lower transaction costs it is better to be less conservative and invest more into stock. For an investor subject to the ES constraint at low 0.1% transaction costs the no transaction region in the early stages is very narrow, at roughly 48% bonds. At higher 0.5% transactions costs the no transaction region of the investor with the ES constraint is wider.

upper and lower boundary of the no transaction region are shown for 0.5% transaction costs (black, solid line) and at 0.1% (red, dashed line). Figure 3.1(a) shows the no transaction region for the investor subject only to the no-short-selling constraint. Figure 3.1(b) shows the no transaction region for the investor subject to the ES constraint. In general no transaction regions include portfolios with higher proportions of bonds when transaction costs are higher. This is due to the lacking incentive to move into stock when transaction costs are high. We observe the well-known fact that towards the end of the investment it is optimal to be more conservative and increase the investments in bonds. For an investor subject only to the no-short-selling constraint, we see that at lower transaction costs it is better to be less conservative and invest more into stock. For an investor subject to the ES constraint at low 0.1% transaction costs the no transaction region in the early stages is very narrow, at roughly 48% bonds. At higher 0.5% transactions costs the no transaction region of the investor with the ES constraint is wider.

Chapter 4

Risk adjusted pricing methodology

One of the interesting problems in the modeling of pricing of financial derivatives is the question how to incorporate both transaction costs and risk arising from a volatile portfolio into the governing equation. In [65], M. Kratka derived a mathematical model for pricing derivative securities in the case when both transaction costs as well as the risk from a volatile portfolio are taken into account. The model is based on the Black-Scholes parabolic PDE in which transaction costs are described by Leland's approach (see e.g. [9, 51, 66, 70] whereas the risk from a volatile portfolio is described by the average value of the variance of the synthetized portfolio. Transaction costs as well as the volatile portfolio risk depend on the time-lag between two consecutive transactions. Minimizing the total costs functional yields the optimal length of the hedge interval. It also gives us a new strategy for hedging derivative securities. This strategy is associated with a solution to a fully nonlinear parabolic equation with varying diffusion coefficient. In this chapter we revisit Kratka's approach in order to derive a model which is mathematically well posed and is scale invariant. These two important features were missing in the original Kratka's model. The key idea of our modification of Kratka's approach consists in a slightly different definition of the risk measure. Furthermore we will extend the model with the possibility of default of counterparty. We will show the importance of integration of default on comparison of integrated vs. separated market and credit risk. The resulting governing equation is scale invariant and it can be mathematically treated. We present qualitative analysis of the governing equation and we derive a robust numerical scheme. We perform extensive numerical testing of the model and compare the results to real option market data. We also introduce a concept of the so-called implied RAPM volatility and implied risk premium coefficient. Implied quantities are computed for large option data sets. We discuss how they can be used in qualitative analysis of option market data.

The chapter is organized as follows. In section 4.1 we derive a scale invariant risk adjusted model for pricing options on assets. We follow the original Leland's and modified Kratka's approach in order to incorporate both transaction costs as well as the risk value arising from a volatile portfolio and the risk from default of counterparty. Based on this model it turns out that prices of options are solutions to a fully nonlinear parabolic partial differential equation. We discuss optimal time interval between consecutive portfolio adjustments. We also show scale invariance of the model. In section 4.2 we analyze the resulting nonlinear partial differential equation. We focus our attention on qualitative aspects of a solution. We also show how to transform the governing fully nonlinear parabolic equation into a quasilinear parabolic equation for the Gamma factor. For such a system of equations we can construct an effective numerical discretization scheme allowing us to find an approximate solution. Qualitative properties of the full space-time discretization scheme are analyzed in section 4.3. Next section 4.4 contains results of numerical simulation and comparison of results based on the RAPM model to real market data. We also show how to calibrate the model. Implied RAPM volatility and implied risk premium are introduced. Finally, we present several numerical experiments comparing computational results to real option market data and take a detailed look on the separated vs. integrated analysis of market and credit risk. The RAPM model for the case when no default can occurs and no dividend are paid was published in author's paper [53] and the results on separated vs. integrated analysis of market and credit risk for loan portfolio were published in author's paper [24].

4.1 Derivation of a scale invariant RAPM model

Before describing the derivation of the RAPM model, we discuss first the basic assumptions we will be making. Throughout this chapter we assume that the asset price $S = S(t), t \geq 0$, follows a geometric Brownian motion with a drift μ and standard deviation $\sigma > 0$, i.e.

$$dS = \mu S dt + \sigma S dW$$

where dW denotes the differential of the standard Wiener process. Additionally the investor can invest to risk free zero coupon bonds B which evolution in time is given by equation $B(t) = B(0)e^{rt}$, where r denotes risk free interest rate. Respectively we can write this equation in differential form

$$dB = rBdt. \tag{4.1}$$

Similarly as in the derivation of the classical Black-Scholes equation we construct a synthetized portfolio Π consisting of one option with the price V, δ assets with the price S per one asset and α risk free zero coupon bonds with a price B

$$\Pi = V + \delta S + \alpha B. \tag{4.2}$$

We recall that the key idea in the Black-Scholes theory is to examine the differential of equation (4.2)

$$d\Pi = dV + Sd\delta + \delta dS + Bd\alpha + \alpha dB. \tag{4.3}$$

In a self financed strategy no additionally investment is used to rehedge the portfolio. To buy a stocks the investor must sell corresponing amount of the bonds. In world without transaction costs, without dividend and without a risk from nonperfect hedging the selffinancing strategy can be expressed as

$$Sd\delta + Bd\alpha = 0. \tag{4.4}$$

For the asset paying continuous dividends the investor earn additionaly amount $DSdt$ in time intervat dt for each stock in his portfolio. He can use these money to buy additional stocks or bonds. The condition for selffinancing strategy then becomes

$$Sd\delta + Bd\alpha = \delta DSdt. \tag{4.5}$$

The equation (4.5) does not include the transaction costs which must investor pay for each transaction. We will denote $r_{TC}Sdt$ the transaction costs which results from rehedging the portfolio on time interval dt. In presence of nontrivial transaction costs, continuous adjustment of portfolio may lead to infinite total transaction costs, while adjustment of portfolio in discrete times leads to additional risk in portfolio. The selffinanced strategy must include also the risk premium r_V. This risk premium can be interpretted either as the premium which must the investor get to willingly invest in a risky portfolio, reps. as the amount which is needed to buy an insurrence for this risk. The selffinancing strategy including the dividends, transaction costs and the portfolio volatility risk premium can be written as

$$d\delta S + d\alpha B = \delta DSdt - r_{TC}Sdt - r_V Sdt. \tag{4.6}$$

Applying this self financed strategy on equation (4.3) gives the differential of the portfolio value as

$$d\Pi = dV + \delta dS + \alpha dB \quad + \delta DSdt - (r_{TC} + r_V)Sdt. \tag{4.7}$$

The differential of risk free zero coupon bond dB is given by equation (4.1) as $rBdt$. From equation (4.2) we can express the term αB under zero investment strategy as

$$\alpha dB = \alpha rBdt = -r(V + \delta S)dt. \tag{4.8}$$

In the classical Black-Scholes theory the option price $V(S,t)$ from equation (4.3), resp. equation (4.7) in case of RAPM model, is differentiated by using Itô's formula (see e.g. chapter 2.4.2 in [66])

$$dV = \partial_t V dt + \frac{1}{2}\sigma^2 S^2 \partial_S^2 V \phi^2 dt + \partial_S V dS, \tag{4.9}$$

where ϕ is normally distributed with zero mean and unit variance. In the classical Black-Scholes theory no default is possible and the Itô's formula can be used. However in the presence of the possibility to default we can apply the Itô's formula only on interval where no default is generated. In our model we assume that the counterparty for the options can default. In case of default the value of options will change discontinuously, depending on the recovery rate. More precissly we will denote the probability of default on time interval dt as ρdt. Under the assumption that, the defaults are independent on history, the probability of default pd on time interval $[0,t]$ simplifies to

$$pd(t) = 1 - e^{-\rho t}. \tag{4.10}$$

We assume that in case of default, the investor will recover the losses only partially. We assume that the default will influence only the options. In case of default investor will recover $R_R V$ from his investment in option insted of V, which represent the options price in case of no default. Here R_R denotes recovery rate. In case of default the total investment in options will be discontinuous as

$$\lim_{dt \to 0+} (V_{t+dt} - V_t) = (R_R - 1)V_t. \tag{4.11}$$

However this loss will be generated only with probability ρdt and therefore its contribution to expected change of the option investment in case of default is

$$\lim_{dt \to 0+} \rho dt (V_{t+dt} - V_t) = -\rho dt(1 - R_R V_t). \tag{4.12}$$

Therefore for small time intervall dt the contribution to the expected change of options investment will be $-\rho(1 - R_R V)dt$. This contribution is achieved in case of default. The counterparty will not default with probability $1 - \rho dt$, in which case the contribution to the expected change of options investment is $(1 - \rho dt)dV$, in this case we can apply Itô's lema on dV. The conditional expected value of the change of options investment in the lowest approximation order becomes

$$E[dV|dS] = -\rho(1 - R_R)V dt + \partial_t V dt + \frac{1}{2}\sigma^2 S^2 \partial_S^2 V \phi^2 dt + \partial_S V dS. \tag{4.13}$$

Our next goal is to show how the transaction costs r_{TC} and the risk premium r_V depend on other quantities, like e.g. σ, S, V and derivatives of V.

Modeling transaction costs

In practice, we have to adjust our portfolio by frequent buying and selling of assets. In the presence of nontrivial transaction costs, the continuous portfolio adjustment may lead to infinite total transaction costs. A natural way how to consider transaction costs in the frame of the Black-Scholes theory is to follow well known Leland's approach [70]. In what follows we recall crucial lines of the derivation of Leland's model in order to show how to incorporate effect of transaction costs into the governing equation. More precisely, we will derive the coefficient of transaction costs r_{TC} occurring in (4.6).

Let us denote by C the round trip transaction cost per unit dollar of transaction. Then

$$C = (S_{ask} - S_{bid})/S \tag{4.14}$$

where S_{ask} and S_{bid} are the so-called Ask and Bid prices of the asset, i.e. the market price offers for selling, resp. for buying assets, which include all additionally costs for transaction. Here $S = (S_{ask} + S_{bid})/2$ denotes the mid value. It means that the transaction costs are given by the value $C|k|S/2$ where k is the number of sold assets ($k < 0$) or bought assets ($k > 0$). The number of stocks in portfolio is given by δ and the number of sold or bought assets are given by change of δ, i.e. $k = \Delta\delta$. Using the Itô's formula on δ as function of S and t we get in the lowest order approximation

$$\Delta\delta = \sigma S \partial_S \delta \Delta W. \tag{4.15}$$

If portfolio adjustments follow the so-called δ-hedging strategy then $\delta = -\partial_S V$ and so, in the lowest order approximation in Δt, we obtain

$$\Delta\delta = -\sigma S \partial_S^2 V \Delta W. \tag{4.16}$$

Since W is the Wiener process with $\Delta W \sim N(0, \Delta t)$ we have $E(|\Delta W|) = \sqrt{2/\pi}\sqrt{\Delta t}$. If Δt is small compared to $T - t$ Leland in [70] proved that we can take $|\Delta W| \approx E(|\Delta W|)$ and thus the coefficient r_{TC} of transaction costs is given by the formula:

$$r_{TC} = \frac{C\sigma S}{\sqrt{2\pi}} \left|\partial_S^2 V\right| \frac{1}{\sqrt{\Delta t}}\,. \tag{4.17}$$

Clearly, increasing the time-lag Δt between portfolio adjustments decreases transaction costs. Therefore, in order to minimize transaction costs we have to take a larger time-lag Δt. On the other hand, as it will be obvious from the next section, choosing a larger time-lag Δt could lead to a higher investor's exposition to the risk from an unprotected portfolio.

Modeling risk from a volatile portfolio

In this section we focus our attention to the question how to include the risk from a volatile portfolio into the model. In the case the portfolio consisting of options and assets is highly volatile, an investor usually asks for a price compensation.

The risk of a fluctuating portfolio can be measured by the variance in relative increments of the replicating portfolio $\Pi = V + \delta S + \alpha B$, i.e. by the term $Var((\Delta\Pi)/S)$. Therefore it is convenient to define the measure of the portfolio volatility risk r_{VP} as follows:

$$r_{VP} = R\frac{Var\left(\frac{\Delta\Pi}{S}\right)}{\Delta t}\,. \tag{4.18}$$

In other words, r_{VP} is proportional to the variance of the relative change of the portfolio per time interval Δt. The constant R is the so-called *risk premium coefficient*. It represents the marginal value of investor's risk exposition. Now applying Itô's formula to the differential $\Delta\Pi$ we obtain

$$\Delta\Pi = (\partial_S V + \delta)\,\sigma S \Delta W + \frac{1}{2}\sigma^2 S^2 \Gamma(\Delta W)^2 + \mathcal{G}$$

where $\Gamma = \partial_S^2 V$ and $\mathcal{G}$ is a deterministic term, i.e $E(\mathcal{G}) = \mathcal{G}$. Additionaly we assume that this risk premium cover only the risk of volatile portfolio when the counterparty does not default. Thus

$$\Delta\Pi - E(\Delta\Pi) = (\partial_S V + \delta)\,\sigma S \phi\sqrt{\Delta t} + \frac{1}{2}\sigma^2 S^2 \Gamma(\phi^2 - 1)\Delta t$$

where ϕ is a random variable with standard normal distribution with zero mean and unit variance such that $\Delta W = \phi\sqrt{\Delta t}$. Hence the variance of the change $\Delta\Pi$ in the portfolio Π can be computed as follows:

$$\begin{aligned} Var(\Delta\Pi) &= E\left[E\left[(\Delta\Pi - E(\Delta\Pi))^2\middle|\, dS\right]\right] \\ &= E\left[\left((\partial_S V + \delta)\,\sigma S\phi\sqrt{\Delta t} + \frac{1}{2}\sigma^2 S^2 \Gamma\left(\phi^2 - 1\right)\Delta t\right)^2\right] \\ &= \frac{1}{2}\sigma^4 S^4 \Gamma^2 \Delta t^2 + (\partial_S V + \delta)^2 \sigma^2 S^2 \Delta t\,. \end{aligned} \tag{4.19}$$

Substituting (4.19) into equation (4.18) we get the following form of risk premium

$$r_{VP} = \frac{1}{2}R\sigma^4 S^2 \Gamma^2 \Delta t + R(\partial_S V + \delta)^2 \sigma^2. \tag{4.20}$$

Similarly, as in the derivation of transaction costs measure r_{TC} we assume δ-hedging of the portfolio adjustment, i.e. we choose $\delta = -\partial_S V$. Note that the choice $\delta = -\partial_S V$ minimise the risk of portfolio and therefore also the risk premium, which will be reduced to

$$r_{VP} = \frac{1}{2} R\sigma^4 S^2 \Gamma^2 \Delta t\,. \tag{4.21}$$

It means that increase in the time-lag Δt between consecutive transactions leads to a linear increase of the risk from a volatile portfolio. In other words, larger time interval Δt means higher risk exposition for an investor.

Gamma hedging strategy based on the RAPM model

The total risk premium $r_R = r_{TC} + r_{VP}$ consists of two parts: transaction costs premium r_{TC} and the risk from a volatile portfolio premium r_{VP} defined as in (4.17) and (4.21), resp. An investor usually seeks for a minimal value of the total risk premium r_R. To this end, an investor has to choose an optimal time-lag Δt between consecutive portfolio adjustments. As both r_{TC} as well as r_{VP} depend on the time-lag Δt so does the total risk premium r_R. In order to find an optimal value of Δt we have to minimize the function

$$\Delta t \mapsto r_R = r_{TC} + r_{VP} = \frac{C|\Gamma|\sigma S}{\sqrt{2\pi}} \frac{1}{\sqrt{\Delta t}} + \frac{1}{2} R\sigma^4 S^2 \Gamma^2 \Delta t\,.$$

A graph of the function $\Delta t \mapsto r_R$ is depicted in Fig. 4.1. The minimum of the function $\Delta t \mapsto r_R$ is attained at the time-lag

$$\Delta t_{opt} = \frac{K^2}{\sigma^2 |S\Gamma|^{\frac{2}{3}}} \quad \text{where } K = \left(\frac{C}{R}\frac{1}{\sqrt{2\pi}}\right)^{\frac{1}{3}}\,. \tag{4.22}$$

For the minimal value of the function $\Delta t \mapsto r_R(\Delta t)$ we have

$$r_R(\Delta t_{opt}) = \frac{3}{2}\left(\frac{C^2 R}{2\pi}\right)^{\frac{1}{3}} \sigma^2 |S\Gamma|^{\frac{4}{3}}\,. \tag{4.23}$$

Since S follows the geometric Brownian motion we have, in the lowest order approximation w.r.t. to Δt

$$E(|\Delta S|/S) = \sigma E(|\Delta W|) = \sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta t}\,.$$

As a consequence from minimizing of the total risk premium r_R we can conclude:

Corollary 4.1. The optimal hedging strategy.

If $|\Delta S|/S \approx K\sqrt{\frac{2}{\pi}}|S\Gamma|^{-\frac{1}{3}}$ in the sense of expected values then adjustment of the portfolio is required. The portfolio is adjusted according to δ-hedging.

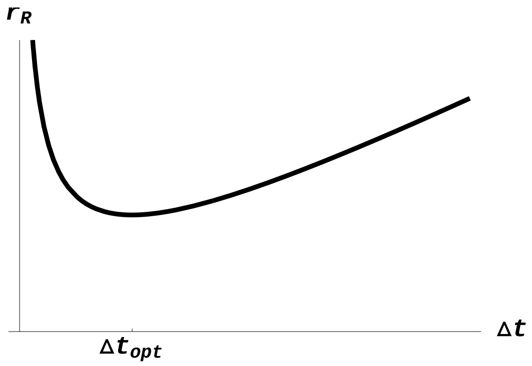


Figure 4.1: The total risk premium $r_R = r_{TC} + r_{VP}$ as a function of the time-lag Δt between two consecutive portfolio adjustments.

Risk adjusted Black-Scholes equation

In previous sections we have shown that taking into account both transaction costs as well as risk from a volatile portfolio into the equation for the change $\Delta\Pi$ of the portfolio Π we obtain

$$\Delta\Pi = \Delta V + \delta\Delta S + (\delta D S - rV - r\delta S - r_R S)\Delta t$$

where r_R represents the total risk premium, $r_R = r_{TC} + r_{VP}$. On the other hand, the investor create the portfolio with zero investment and also zero investment is needed for rehedging the portfolio. By the no-arbitrage principle the expected value of change in the portfolio $\Delta\Pi$ must be equal to zero. Additionally with equation (4.13) we finally obtain following generalization of the Black-Scholes equation for valuing options:

$$\partial_t V + \frac{\sigma^2}{2} S^2 \partial_S^2 V = (r + \rho(1 - R_R))V + (D - r)S\partial_S V + r_R S.$$

Taking the optimal value of the total risk coefficient r_R derived in (4.23) we obtain that the option price is a solution to the following nonlinear parabolic equation:

(*Risk adjusted Black-Scholes equation*)

$$\partial_t V + \frac{\sigma^2}{2} S^2 \Gamma \left(1 - \upsilon (S\Gamma)^{\frac{1}{3}}\right) = \left(r + \rho(1 - R_R)\right) V + (D - r)S\partial_S V \quad (4.24)$$

where

$$\Gamma = \partial_S^2 V \qquad \text{and} \qquad \upsilon = 3\left(\frac{C^2 R}{2\pi}\right)^{\frac{1}{3}}. \quad (4.25)$$

Here and after we will denote by $x^{\frac{1}{3}}$ the signed power function, defined as: $x^p = |x|^{p-1}x = |x|^p \mathrm{sign}(x)$ for all $x \in R$, $p > 0$. In the case there are either no transaction costs ($C = 0$) or no risk from the volatile portfolio ($R = 0$) we have $\upsilon = 0$ if additionally there is no default possibility ($\rho = 0$) or we have full recovery in case of default ($R_R = 1$) then the equation (4.24) reduces to the original Black-Scholes linear parabolic equation

$$\partial_t V + \frac{\sigma^2}{2} S^2 \Gamma = rV + (D - r)S\partial_S V\,. \tag{4.26}$$

Behavior near the exercise time

Our next goal is to analyze the behavior of the option price $V = V(S,t)$ near the exercise time T, i.e. when $T - t$ is small. Recall that we have followed Leland's methodology in modeling transaction costs. In this approach one has to assume that the time-lag Δt between consecutive portfolio adjustments is small compared to $T - t$ (see [51, 66, 70]). A natural way how to satisfy the condition $\Delta t_{opt} \ll T - t$ is to disallow portfolio adjustments when the time t is close to the exercise time T. Hence it is convenient to assume that $V = V(S,t)$ is a solution to the classical Black-Scholes equation (4.26) for times t close to T. The same kind of approximation is used in [46]. More precisely, we will assume that $V = V(S,t)$ is a solution to the Black-Scholes equation (4.26) on some small time interval $(T - \tau_\upsilon, T)$ whereas $V(S,t)$ solves the Risk adjusted Black-Scholes equation (4.24) on $(0, T - \tau_\upsilon)$ where $0 < \tau_\upsilon \ll T$. There is another, purely mathematical, justification for such an assumption. Equation (4.24) is a parabolic PDE if and only if the function

$$\beta(H) = \frac{\sigma^2}{2}(1 - \upsilon H^{\frac{1}{3}})H \tag{4.27}$$

is an increasing function in the variable $H := S\Gamma = S\partial_S^2 V$. Hence, in order to verify parabolicity of (4.24), we have to assume

$$S\Gamma < \left(\frac{3}{4\upsilon}\right)^3 . \tag{4.28}$$

However, if we consider either Call or Put options on assets paying no dividends then the term $S\Gamma = S\partial_S^2 V(S,t)$ becomes infinite at $S = X$ for $t \to T^-$ and the above condition is violated. For both Call and Put options we have the same expression for the term $S\Gamma$,

$$S\Gamma = \frac{N'(d)}{\sigma\sqrt{T-t}}\,, \quad d = \frac{\ln(S/X) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \tag{4.29}$$

where $N'(d) = \frac{1}{\sqrt{2\pi}} e^{-d^2/2}$ is the density function of the standard normal distribution.

As it is usual in the theory of nonlinear diffusion (see e.g. [58, 59]) the idea how to overcome this difficulty is to modify the function $\beta(H) = (1 - \upsilon H^{\frac{1}{3}})H$ for large values of $H = S\Gamma$ and then to prove *á priori* bounds of a solution enabling us to conclude that the solution of the modified equation satisfies the original equation on $(0, T - \tau_\upsilon)$. We first modify the function $m(H) = \upsilon H^{\frac{1}{3}}$ for large values of H as follows:

$$m_\varepsilon(H) = \begin{cases} \upsilon H^{\frac{1}{3}} & \text{if } H < \kappa_\upsilon = \left(\frac{3(1-\varepsilon)}{4\upsilon}\right)^3 ; \\ (1-\varepsilon)\left(1 - \frac{\kappa_\upsilon}{4H}\right) & \text{otherwise,} \end{cases} \tag{4.30}$$

where $0 < \varepsilon \ll 1$ is a small regularization parameter. With this regularization the function

$$\beta_\varepsilon(H) = \frac{\sigma^2}{2}(1 - m_\varepsilon(H))H \tag{4.31}$$

is C^1 continuous, and, moreover

$$\beta_\varepsilon(H) = \frac{\sigma^2}{2}\left(1 - \upsilon H^{\frac{1}{3}}\right) H \quad \text{for } H < \kappa_\upsilon\,. \tag{4.32}$$

and

$$\beta'_\varepsilon(H) = \frac{\sigma^2}{2}\varepsilon > 0 \quad \text{for } H \geq \kappa_\upsilon\,. \tag{4.33}$$

Recall that we have to assume that $V(S,t)$ is a solution to the Black-Scholes equation (4.26) for $t \in (T - \tau_\upsilon, T)$ where $0 < \tau_\upsilon \ll 1$ is small. Hence, for both Call and Put options, it follows from (4.29) that

$$\lim_{t \to T - \tau_\upsilon^+} \max_{S>0} S\partial_S^2 V(S,t) = \lim_{t \to T - \tau_\upsilon^+} 1/\sqrt{2\pi\sigma^2(T-t)} = \kappa_\upsilon$$

provided that the switching time $0 < \tau_\upsilon < T$ is defined as follows:

$$\tau_\upsilon = \frac{1}{2\pi\sigma^2\kappa_\upsilon^2}\,. \tag{4.34}$$

Remark 4.1. *We also remind ourselves that the terminal pay-off for a Call option at $t = T$ is given by $V(S,T) = \max(S - X, 0)$ whereas $V(S,T) = \max(X - S, 0)$ for a Put option. Here and after X denotes the exercise price and T stands for the exercise time. Furthermore, a Call option price $V(S,t)$ is subject to boundary conditions $V(0,t) = 0, V(S,t) \to S$ as $S \to \infty$, $t \in (0,T)$, and, Put option price satisfies: $V(0,t) = Xe^{-r(T-t)}, V(S,t) \to 0$ as $S \to \infty$.*

Having modified the function β we are in a position to introduce a notion of a solution to the Risk adjusted Black-Scholes equation.

Definition 4.1. *By a solution to the Risk adjusted Black-Scholes equation we mean a continuous function $V = V(S,t), S \in (0,\infty), t \in (0,T)$, satisfying boundary conditions, the terminal payoff condition at $t = T$, and such that*

a) $V(S,t)$ *is a classical (smooth) solution to the Black-Scholes equation*

$$\partial_t V + \frac{\sigma^2}{2} S^2 \Gamma = rV + (D-r)S\partial_S V\,, \quad S > 0\,,$$

on the time interval $(T - \tau_v, T)$.

b) $V(S,t)$ *is a classical (smooth) solution to the equation*

$$\partial_t V + S\beta_\varepsilon(S\Gamma) = (r+\rho(1-R_R))V + (D-r)S\partial_S V\,, \quad S > 0\,, \tag{4.35}$$

on the time interval $t \in (0, T - \tau_v)$.

Scale invariance property

The governing equation (4.24) as well as (4.35) have the scale invariance property. Indeed, let us multiply the asset and option prices by the same scaling factor $\kappa > 0$, i.e. we take $\tilde{S} = \kappa S, \tilde{V} = \kappa V$. Then $\tilde{S}\tilde{\Gamma} = \tilde{S}\partial^2_{\tilde{S}}\tilde{V} = S\partial^2_S V = S\Gamma$. i.e. the term $S\Gamma$ remains unchanged after scaling of S and V by the factor $\kappa > 0$. Therefore the scaled option price $\tilde{V}$ satisfies the same governing equation (4.24) in which we change the variable S to $\tilde{S}$. This is a very important property of the governing equation which was missing in original Kratka's approach based on a different definition of the risk coefficient r_{VP} measuring volatility of the portfolio. More precisely, in [65] the risk measure was defined as follows:

$$r_{VP} = R\frac{Var(\Delta\Pi)}{\Delta t}\,.$$

The equation for valuing the price of an option then reads as:

$$\partial_t V + \frac{\sigma^2}{2} S^2 \left(1 - \upsilon\Gamma^{\frac{1}{3}}\right)\Gamma = r\left(V - S\partial_S V\right)\,.$$

However, this equation is not scale invariant with respect to the scaling: $V \leftrightarrow \kappa V, S \leftrightarrow \kappa S$.

4.2 Analysis of the RAPM model

The idea how to analyze and solve equation (4.35) is based on transformation technique. As it is usual in similar circumstances (see e.g. [51, 66]) we consider the change of independent variables:

$$x := \ln(S/X)\,,\ \ x \in \mathbb{R}\,, \qquad \tau := T - t\,,\ \ \tau \in (0, T)\,.$$

As equation (4.35) contains the term $S\Gamma = S\partial_S^2 V$ it is convenient to introduce the following transformation:

$$H(x,\tau) := S\Gamma = S\partial_S^2 V(S,t)\,.$$

Recall that the option price $V(S, T-\tau)$ for $0 < \tau \le \tau_v$ can be valuated by an explicit formula for both Call and Put options, resp. (see e.g. [51, 66]). More precisely, the valuation formulae for pricing European Call and Put options read as follows:

$$\begin{aligned} V_{ec}(S, T-\tau) &= SN(d_1) - Xe^{-r\tau}N(d_2)\,, \\ V_{ep}(S, T-\tau) &= Xe^{-r\tau}N(-d_2) - SN(-d_1)\,, \end{aligned}$$

where $d_1 = (\ln(S/X) + (r + \sigma^2/2)\tau)/(\sigma\sqrt{\tau}), d_2 = d_1 - \sigma\sqrt{\tau}$.
Therefore the $H(x,\tau)$ for $0 < \tau \le \tau_v$ becomes

$$H(x,\tau) = \frac{N'(d_1)}{\sigma\sqrt{(\tau)}}, \qquad d_1 = \frac{x + (r+\sigma^2/2)\tau}{\sigma\sqrt{(\tau)}}. \tag{4.36}$$

Valuation formula for option price

Suppose for a moment that the function $S\Gamma$ is already known. Setting

$$H(x,\tau) = S\Gamma(S,t) \tag{4.37}$$

$$U(x,\tau) = e^{(r+\rho(1-R_r))(T-t)}V(S,t), \tag{4.38}$$

with $\tau = T - t$ and $x = \ln\frac{S}{X} - (D-r)\tau$ we get from (4.24) differential equation for $U(x,\tau)$ in a simple form

$$\partial_\tau U = Xe^{x+(D+\rho(1-R_R))\tau}\beta(H(x,\tau)). \tag{4.39}$$

Value of $U(x,\tau_v)$ is given by the initial condition on V in time $T-\tau_v$. With the initial condition $U(x,\tau)$ one can integrate (4.39) as

$$U(x,\tau) = U(x,\tau_v) + \int_{\tau_v}^{\tau} Xe^{x+(D+\rho(1-R_R))\theta}\beta(H(x,\theta))d\theta. \tag{4.40}$$

Applying the inverse transform to (4.38) finally we get the option price.

Γ equation

Next we derive an equation for the function H on time interval (τ_v, T). It turns out that the function $H(x,\tau)$ is a solution to a nonlinear parabolic equation subject to the initial and boundary condition. More precisely, by taking the second derivative of (4.24) with respect to x, we obtain, after some calculation, that $H = H(x,\tau)$ is solution to the quasilinear parabolic equation

$$\partial_\tau H = \partial_x^2 \beta_\epsilon(H) + \partial_x \beta_\epsilon(H) - (D + \rho(1 - R_R))H + (r - D)\partial_x H. \quad (4.41)$$

Henceforth, we will refer to (4.41) as Γ equation. A solution H to (4.41) is subject to the initial condition at $\tau = \tau_v$:

$$H(x, \tau_v) = \tilde{H}(x), \quad x \in \mathbb{R}, \quad (4.42)$$

where $\tilde{H}(x) = N'(d_1)/(\sigma\sqrt{\tau_v})$ (see (4.36)). In the case of Call or Put options, the function is subject to boundary conditions at $x = \pm\infty$,

$$H(-\infty, \tau) = H(\infty, \tau) = 0, \quad \tau \in (0, T). \quad (4.43)$$

Next we show useful bounds for a solution H to the Γ equation (4.41). Notice that for any constant c the function

$$H_c(x,\tau) = ce^{-(D+\rho(1-R_R))(\tau-\tau_v)} \quad (4.44)$$

is a solution to (4.41), where the term $D + \rho(1 - R_R)$ is by its definition non-negative. For this solution of the Γ equation we get

$$H_0(x, \tau_v) \le \tilde{H}(x) \le H_{\kappa_v}(x, \tau_v), \quad \text{for any } x \in \mathbb{R}. \quad (4.45)$$

From the classical maximum principle for parabolic equation (see [87]) follows that a solution $H(x,\tau)$ to the initial-boundary problem (4.41)-(4.43) satisfies the estimate

$$0 \le H(x,\tau) \le H_{\kappa_v}(x,\tau) \le \kappa_v, \quad \text{for any } x \in \mathbb{R}, \tau \in (\tau_v, T) \quad (4.46)$$

The above estimate enable us to conclude that a solution $V(S,t)$ to the risk-adjusted Black-Scholes equation (see Definition 4.1) is indeed a solution to (4.24) on time interval $t \in (0, T - \tau_v)$.

4.3 Numerical scheme for full space-time discretization

In this section we describe a full space-time discretization scheme for solving (4.41) and (4.40). The idea of construction of a numerical approximation to (4.41) is based on the finite-volume method (see, e.g., [43]).

Discretization of the Γ equation

In order to find a numeric solution to (4.41), we have to restrict ourselves to a finite spatial interval $x \in (-L, L)$ where $L > 0$ is sufficiently large. Since $S = Xe^x$, we have restricted the interval of asset values to $S \in (Xe^{-L}, Xe^{L})$. From practical point of view, it is therefore sufficient to take $L \approx 1.5$ in order to include important values of S. Subsequently, we have also to modify boundary conditions (4.43). Instead of (4.43), we will consider Dirichlet boundary conditions at $x = \pm L$, that is,

$$H(-L, \tau) = H(L, \tau) = 0, \quad \tau \in (\tau_v, T). \tag{4.47}$$

We take a uniform division of the time interval $[0, T]$ with a time step $k = T/m$ and a uniform division $x_i = ih$, $i = -n, ..., n$, of the interval $[-L, L]$ with a step $h = L/n$. To construct numerical approximation of a solution H to (4.41), we derive a system of difference equations corresponding to (4.41) to be solved at every discrete time step. Difference equations involve discrete values of $H_i^j \approx H(ih, jk)$ where $j = p, ..., m$. Here the index p corresponds to the initial time τ_v, that is $\tau_v \approx pk$. We choose the time step k less then Δt_{opt} (see (4.22)).

Our numerical algorithm is semi-implicit in time. It means that all nonlinear terms in equations are treated from the previous time step whereas linear terms are solved at the current time level. In order to guarantee stability of the scheme, we assume the CLF condition (see [52]) for the time step k and spatial step h: $(k/h^2)\lambda_+ < 1/2$. Such a discretization leads to a solution of linear systems of equations at every discrete time level. Now, by replacing the time derivative by the time difference, approximating H in nodal points by the average value of neighboring segments, collecting all linear terms at the new time level j, and taking all the remaining terms from the previous time level $j-1$, we obtain a tridiagonal system subject to homogeneous Dirichlet boundary conditions imposed on new discrete values of H^j:

$$a_i^j H_{i-1}^j + b_i^j H_i^j + c_i^j H_{i+1}^j = d_i^j, \quad H_{-n}^j = 0, \; H_n^j = 0, \tag{4.48}$$

for $i = -n+1, ..., n-1$, and $j = p+1, ..., m$, where $H_i^p = \tilde{H}(x_i)$ and

$$a_i^j = -\frac{k}{h^2}\beta'_\epsilon(H_{i-1}^{j-1}) + \frac{k}{h}(r - D), \tag{4.49}$$

$$b_i^j = 1 - (a_i^j + c_i^j), \tag{4.50}$$

$$c_i^j = -\frac{k}{h^2}\beta'_\epsilon(H_i^{j-1}), \tag{4.51}$$

$$d_i^j = (1 - k(D + \rho(1 - R_R)))H_i^{j-1} + \frac{k}{h}(\beta_\epsilon(H_i^{j-1}) - \beta_\epsilon(H_{i-1}^{j-1})). \tag{4.52}$$

Since triagonal systems admit a simple LU-matrix decomposition, we can solve the above tridiagonal system in every time step in a fast and effective way.

Computation of option prices

Equation (4.40) is simple updating formula once a numerical approximation of a solution $H(x,\tau)$ to Γ equation is known. We can use a simple trapezoidal rule for numerical integration of equation (4.40), when the value of a function H at a spatial point $x \in [x_i, x_{i+1}]$ is computed by a linear approximation of H using the neighboring values H_i, H_{i+1}.

4.4 Computational results

The purpose of this section is to discuss application of the Risk adjusted pricing methodology to the real market option price data. We introduce a concept of the so-called implied RAPM volatility σ_{RAPM} and implied risk premium coefficient R. Furthermore, we discuss the volatility smile phenomenon and explanation of this paradox within the frame of RAPM model. In the first part of this section we will focus on the RAPM model with zero default probability and in the second part we will explore the influence of the presence of default probability to the options prices.

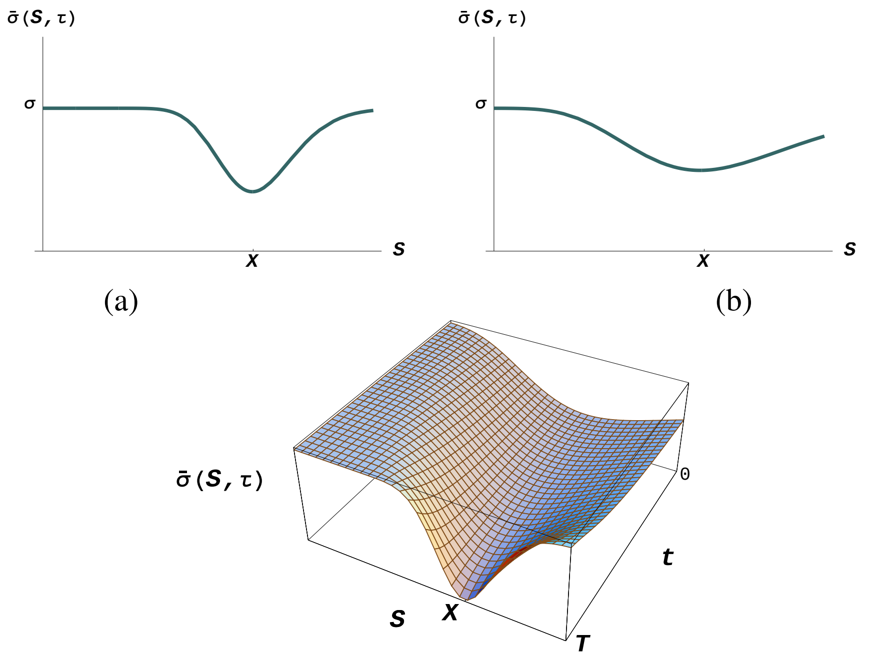
Volatility smile explained

One of the most striking phenomena in the Black-Scholes theory is the so-called *volatility smile* phenomenon. Notice that derivation of the classical Black-Scholes equation (4.26) relies on the assumption on a constant value of the volatility parameter σ. On the other hand, as it was documented by many examples obtained from real market data (see e.g. [8, 42, 55, 99]) this assumption is often violated. More precisely, the implied volatility σ_{impl} is no longer constant and it may depend on the asset price ratio S/X as well as the time t. Typically, the dependence $S \mapsto \sigma_{impl}$ has a convex shape. Clearly, the concept of the implied volatility is one of the weakest points in the Black-Scholes theory because the implied volatility is being computed from a model which is based on the constant volatility assumption.

In the Risk adjusted pricing methodology approach we are yet able to explain volatility smile without breaking the RAPM model. The risk adjusted Black-Scholes equation (4.24) can be viewed as an equation with variable volatility coefficient, i.e.

$$\partial_t V + \frac{\bar{\sigma}^2(S,t)}{2} S^2 \Gamma = \bar{r} V - (\bar{r} - \bar{D}) S \partial_S V$$

where $\Gamma = \partial_S^2 V$, $\bar{r} = r + \rho(1 - R_R)$, $\bar{D} = D + \rho(1 - R_R)$ and the volatility



(a) (b)

(c)

Figure 4.2: Explanation of the volatility smile. The dependence of $\bar{\sigma}(S,t)$ on S is depicted in (a) for t close to the expiration T and in (b) for time $0 < t << T$. The function $(S,t) \mapsto \bar{\sigma}(S,t)$ is shown in (c).

$\bar{\sigma}^2(S,t)$ depends itself on a solution $V = V(S,t)$ as follows:

$$\bar{\sigma}^2(S,t) = \sigma^2 \left(1 - \upsilon(S\Gamma)^{1/3}\right) .$$

In Fig. 4.2 we show the dependence of the function $\bar{\sigma}(S,t)$ on the asset price S and time t. It should be obvious that the function $S \mapsto \bar{\sigma}(S,t)$ has a convex shape near the exercise price X. We have used the RAPM model in order to compute values of $\Gamma = \partial_S^2 V$. We choose $\mu = 0.2, \sigma = 0.3, r = 0.011, T = 0.5, D = 0$ and $\rho = 0$.

Modeling Bid - Ask spreads of option values

In real market quotes data sets there are listed two different option prices $V_{bid} < V_{ask}$ called Bid and Ask price representing thus offers for buying and selling options, respectively (see [85]). In our approach of derivation of RAPM model the asset transaction costs as well as the risk from unprotected portfolio were on the side of a holder of an option because he/she has to keep a fixed amount of options

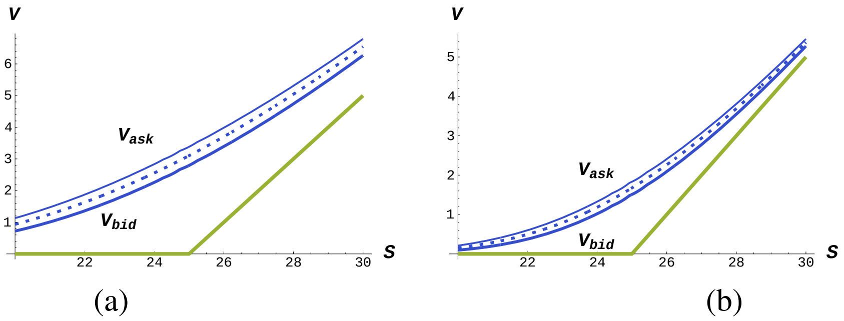


Figure 4.3: A comparison of bid and ask option prices computed by means of the RAPM model. The dotted line in the middle is options price computed from the Black-Scholes equation. We choose (a) $\sigma = 0.3$, $\upsilon = 0.2$, $r = 0.011$, $X = 25$, $D = 0$, $\rho = 0$ and $T = 1$ and (b) $T = 0.3$.

and has to adjust portfolio by buying or selling assets. Having assumed such a long option position the solution to the RAPM model (4.24) corresponds to the Bid option price V_{bid}. If we switch to the short positioned option we transfered both transaction costs and the risk from unprotected portfolio to the buyer of an option. In this setting we just changed the governing equation slightly - the coefficient υ has a reverse sign. It means that the RAPM equation modeling higher Ask option prices reads as follows:

$$\partial_t V + \frac{\sigma^2}{2} S^2 \left(1 + \upsilon (S\Gamma)^{\frac{1}{3}}\right) \Gamma = \bar{r} V + \bar{D} S \partial_S V \,. \tag{4.53}$$

The above PDE can be numerically computed exactly in the same way as the RAPM equation (4.24) for the bid option price. In fact, we only change the sign of the coefficient υ in our numerical scheme. Let us denote $V(S,t;\sigma,\upsilon)$ the value of a solution to (4.24). In order to calibrate RAPM model we seek for the pair (σ_{RAPM}, R) such that $V_{bid} = V(S,t;\sigma,\upsilon)$ and $V_{ask} = V(S,t;\sigma,-\upsilon)$. It leads us to the following definition of implied RAPM volatility and risk premium coefficient.

Definition 4.2. *Let V_{bid}, V_{ask} denote the market option data for the Bid and Ask option price. By the implied RAPM volatility σ_{RAPM} and implied RAPM risk premium coefficient R we mean the unique values of σ and R such that $V_{bid} = V(S,t;\sigma_{RAPM},\upsilon)$ and $V_{ask} = V(S,t;\sigma_{RAPM},-\upsilon)$ where $\upsilon = 3(C^2R/(2\pi))^{1/3}$ and $C > 0$ is the asset transaction cost rate.*

In figure 4.3 we show a comparison of V_{bid} and V_{ask} options prices to the Call option payoff diagram. We also show the solution to the classical Black-Scholes equation (4.26) lying on between V_{bid} and V_{ask} prices. Notice that a solution σ

to the equation $V_{mid} = V(S, t; \sigma, t_\upsilon, 0)$, where $V_{mid} = (V_{bid} + V_{ask})/2$ is just the usual implied volatility σ_{impl} (see [66]).

Remark 4.2. *In modeling bid-ask spreads, we have unambiguously associated a long positioned option with a lower bid price, and a short positioned option with a higher ask price. In a real market, it need not be so easy to switch costs and the risk to the other side of the contract. A consistent way how to calibrate the RAPM model should be to work with either one of V_{bid} or V_{ask} and stick to it. It turned out from the calibration of implied pairs (σ_{RAPM}, R) that σ_{RAPM} is very close to the Black-Scholes implied volatility σ_{impl}, their relative difference being less than 5.10^{-3} (see section Example of calibration of the RAPM model). Of course they need not coincide as the governing parabolic equation is nonlinear and so $\frac{1}{2}(V(S, t; \sigma, t_\upsilon, -\upsilon) + V(S, t; \sigma, t_\upsilon, \upsilon)) \neq V(S, t; \sigma, t_\upsilon, 0)$, in general. Nevertheless, from practical point of view, we may take $\sigma_{RAPM} \approx \sigma_{impl}$ leading to calibration of the remaining parameter υ (and subsequently R) from the single equation $V_{bid} = V(S, t; \sigma, t_\upsilon, \upsilon)$ only.*

In order to find a pair (σ_{RAPM}, R) of the implied volatility and risk premium R, we have to solve the following system of nonlinear equations $F(\sigma, \upsilon) = (V_{bid}, V_{ask})$, where the mapping $F : \mathbb{R}^2 \to \mathbb{R}^2$ is defined as

$$F(\sigma, R) = (V(s, t; \sigma, t_\upsilon, \upsilon), V(s, t; \sigma, t_\upsilon, -\upsilon)). \tag{4.54}$$

To find a solution to (4.54), we make use of the iterative Newton-Kantorovich method (see [4]),

$$y^{n+1} = y^n - [F'(y^n)]^{-1} F(y^n), \quad n = 0, 1, ..., \tag{4.55}$$

where $y^n = (\sigma^n, R^n)$ and F' is the derivative of F. Taking a good initial approximation (σ^0, R^0) of an implied pair, the Newton-Kantorovich sequence $y^n = (\sigma^n, ^n)$ defined as in (4.55) converges to a solution (σ, R) of (4.54). In practice, we replace partial derivatives in the Jacobe matrix F' by their central difference approximations. Notice that the overall complexity of a single Newton-Kantorovich step is therefore 10 times the complexity of computation of a particular RAPM option price $V(S, t; \sigma, t_\upsilon, \upsilon)$. In our experiments to follow, we needed (in average) $5 - 15$ steps in the Newton-Kantorovich scheme in order to find a solution to (4.54) with accuracy less than 0.1% of the option price.

Example of calibration of the RAPM model

In this section, we summarize results of several numerical experiments and comparison of results to market option datasets. We focus on calibration of the RAPM model. The main goal is to analyze time series of option prices and to compute the implied RAPM volatility σ_{RAPM} and risk premium coefficient R. The analyzed datasets consisted of several hundreds of option prices for different exercise prices X and exercise times T. These results were part of the author's paper [53].

As an example we considered sample datasets for Microsoft Corporation. In all studied cases, we computed the implied RAPM volatilities and risk premium coefficients. We considered a flat interest rate $r = 0.02$ and a constant transaction cost coefficient $C = 0.01$. We also compared implied RAPM volatilities to standard implied volatilities σ_{impl} computed by means of the classical Black-Scholes equation (4.26). It turned out that time series of σ_{RAPM} and σ_{impl} are almost perfectly correlated with correlation higher than 0.99. On the other hand, in all studied cases we have $\sigma_{RAPM} > \sigma_{impl}$ with the relative difference $(\sigma_{RAPM} - \sigma_{impl})/\sigma_{RAPM}$ less then 0.005. Notice that we have considered only Call option price records in which $V_{bid} > S_{bid} - X$.

In Figures 4.4a and 4.4b, we present the behavior of the mid value asset price $S = (S_{bid} + S_{ask})/2$ during April 4, 2003. We choose three Call options with the same expiration date T=April 19, 2003, and different expiration prices $X = 23$, $X = 25$, $X = 30$. The behavior of the implied volatility σ_{RAPM} and implied risk premium R is depicted Figures 4.4c and 4.4d. For Call options with expiration prices $X = 25$ and $X = 30$, implied risk coefficients are almost constant during the day except for the initial shock for the option with $X = 30$. On the other hand, implied risk coefficients for Call option with expiration price $X = 23$ as well as all implied volatility σ_{RAPM} are highly volatile during this day. The lowest risk (measured by R) is achieved by holding the Call option on $X = 25$. These results could indicate that holding $X = 25$ Call option is less risky compared to other analyzed call options.

In Figures 4.5a and 4.5b, we present analogous results for Microsoft stocks and Call options having a longer expiration date T=January 22, 2005. Again we choose three Call options with different expiration prices $X = 20$, $X = 25$ and $X = 30$. The behavior of the implied volatility σ_{RAPM} and implied risk premium R is depicted Figures 4.5c and 4.5d. Similarly as in Figure 4.4, Call options with expiration prices $X = 25$ and $X = 30$ have almost constant implied risk coefficients. Interestingly enough, in the first half of the day, the implied risk coefficient R for the Call option with $X = 20$ is much higher compared to those corresponding to $X = 25$ and $X = 30$, respectively. During the second part of the day, it is jumping up and down between them. It could give some indication to an investor that the portfolio consisting of $X = 25$ Call options is less risky.

Finally, in Figure 4.6 we present one-week behavior of implied volatilities and risk premium coefficients for the Microsoft Call option on $X = 25$ expiring at T=April 19, 2003. In the beginning of the investigated period, the risk premium coefficient R was rather high and fluctuating. On the other hand, it tends to a flat value of $R \approx 5$ at the end of the week.

Default probability in the RAPM model

In classical Black-Scholes model the derivative depends only on the stock properties and are not specific to a counterparty. Integration of the default possibility to the pricing model made the derivative dependend also on the investors specific paramaters, default intensity ρ and recovery rate R_R.

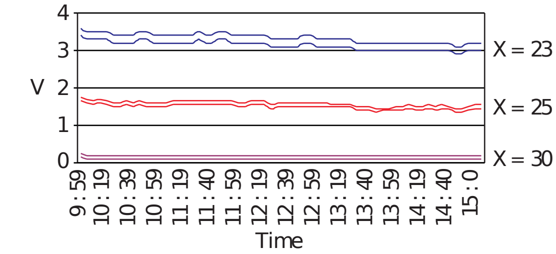
In the case $\upsilon = 0$ the RAPM will reduce to Black-Scholes model with integrated default possibility. In this case we can express the options price as

$$V(S,\tau) = e^{-\rho(1-R_R)\tau} BS(S,\tau), \tag{4.56}$$

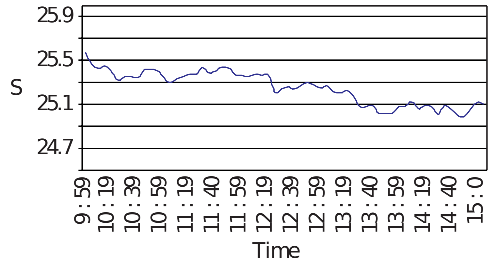
where $BS(S,\tau)$ represents the solution to the classical Black-Scholes model. We should remind that equation (4.56) holds only if $\upsilon = 0$, otherwise the non-linearity of $RAPM$ model make this kind of transformation impossible. The price of such a options is decreased by the possibility of default. The higher the default probability (resp. the smaller recovery rate) the lower the price of option.

In Figure 4.7 we present the comparison of option price calculated with Black-Scholes equation and option price calculated with equation (4.56). We see that the absolute difference between the option prices is increasing with stock price S, while the relative difference is constant. As we calculate the European type of call options the option price will be under the payoff diagram for high value of stock S, resp. for high default intensity ρ.

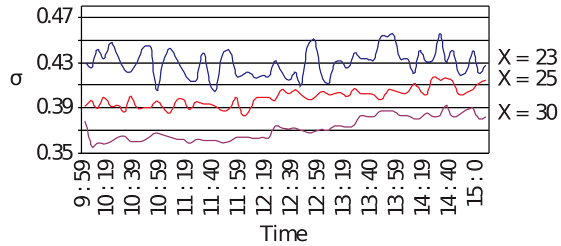
The highest influence have the integration of default possibility on nearly perfectly hedged portfolios (see author's paper [24, 25] for case of loan portfolio and integrated versus separated market and credit risk analysis). Such a portfolio is displayed on the Figure 4.8. The portfolio consist of a long position of call option and a short position of call option. We should remind that the impact of default probability is not symmetric on short and long position. For the long position a default will generate a loss of the option price. However for the short position a default of counterparty will not change our payment obligation. We see that the portfolio is perfectly hedged only for $\rho = 0$. In such case a change of stock price S will not change the portfolio value. Similar for a low value of stock price S any reasonable change of ρ will not influence the portfolio value. These two kind of analysis correpsond to separated credit and market risk calculation. On the other hand we can see that a joint move of stock price S and default intensity ρ can influence a high loss of portfolio. On the Figure 4.8 we can see that this joint move of risk factors (integrated risk analysis) can have much higher impact on portfolio value than a sum of portfolio changes in stock and default intensity directions. The equality between separated approximation of credit and market



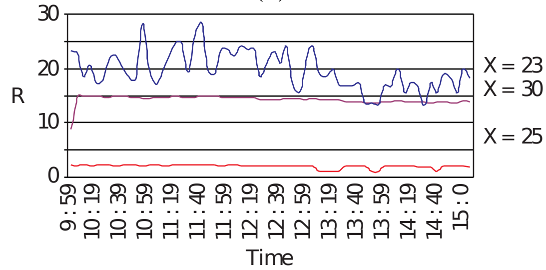
(a)



(b)



(c)



(d)

Figure 4.4: Intraday behavior of Microsoft stocks (April 4, 2003) and shortly expiring Call options with expiry date April 19, 2003, with computed implied volatilities σ_{RAPM} and risk premium coefficients R.

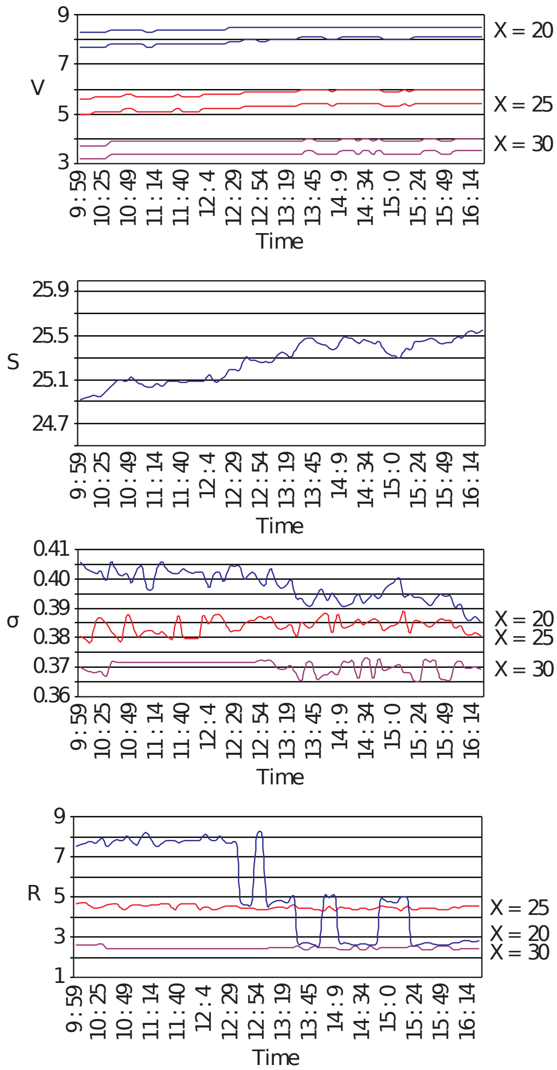
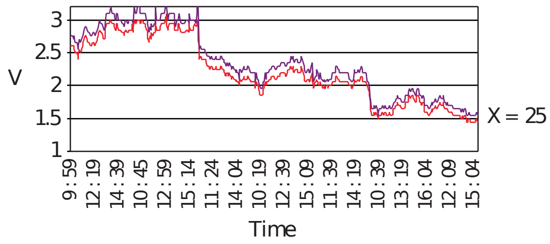
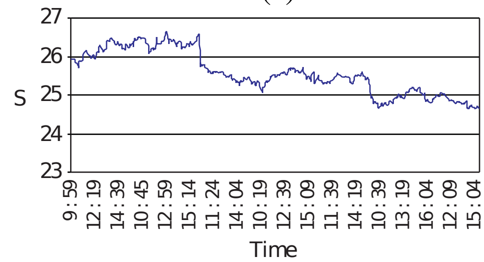


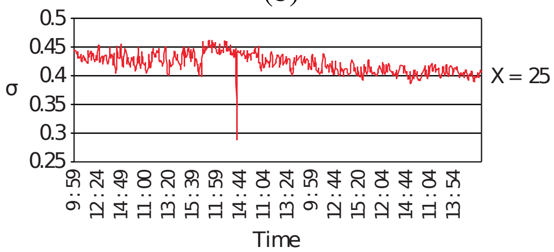
Figure 4.5: Intraday behavior of Microsoft stocks (April 17, 2003) and Call options with long expiration date January 22, 2005, with computed implied volatilities σ_{RAPM} and risk premium coefficients R.



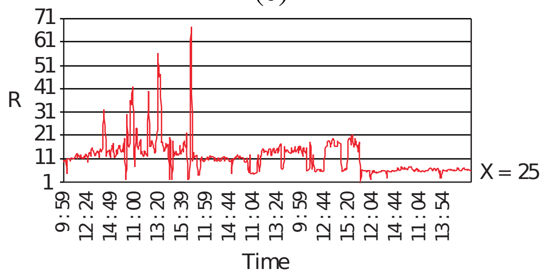
(a)



(b)



(c)



(d)

Figure 4.6: One week behavior of Microsoft stocks (March 20-27, 2003) and Call options with expiry date April 19, 2003, with computed implied volatilities σ_{RAPM} and risk premium coefficients R.

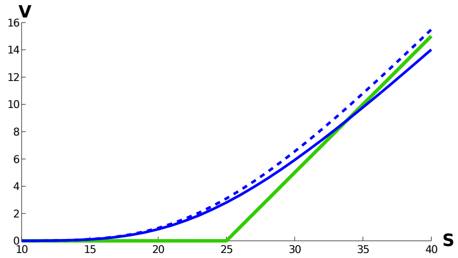


Figure 4.7: A comparison of bid option price for RAPM model with default probability and the option price for classical Black-Scholes model. The dotted line is option price computed with Black-Scholes equation. The solid blue line is computed with RAPM model with $\sigma = 0.3$, $\mu = 0$, $r = 0.011$, $X = 25$, $D = 0$, $\rho = 0.1$, $R_R = 0$ and $T = 1$.

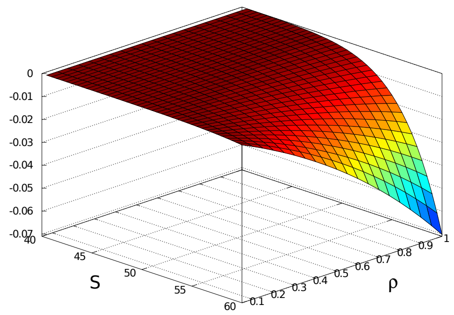


Figure 4.8: Graph shows the value of portfolio consisting from one short and one long position. The change of portfolio value is more significant when both risk factor can move, what represents the integrated risk, while move in one direction, corresponding with market risk (move in S direction) and credit risk (move in ρ direction), are neglible.

risk and integrated risk can be assured if and only if the portfolio value can be separated as $V(S, \rho) = f(S) + g(\rho)$. And only in this case the subaditivity of coherent risk measures with respect to credit and market risk can be assured (see author's paper [25]). Therefore is necessary to have the possibility to model the price of derivative also for counterparty which can default.

The following proposition summarize when we can assure that the integrated risk is not underestimated by separated risk analysis. Although the proofs of the following results can be found in author's paper [25] for less general case, we provide them for readers convenience. In these proposition function f represents the profit/loss distribution of portfolio and X represent the scenarios. In our case of market vs. credit risk analysis the X would consists of stock price S and default density ρ. For sake of simplicity we will use following short notations

$$\begin{aligned}
X_n^i &= (x_1^i, ..., x_n^i), \\
f^n(i_1...i_n) &= f^n(x_1^{i_1}, x_2^{i_2}, ..., x_n^{i_n}), \;\; i_j \in \{0,1\}, \;\; j = 1, ..., n, \\
f_j^n &= f^n(i_1...i_n), \;\; i_k = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}, \\
f_{jk}^n &= f^n(i_1...i_n), \;\; i_l = \begin{cases} 1 & l \in \{j,k\} \\ 0 & l \notin \{j,k\} \end{cases}, \\
f_\Omega^n &= f^n(X_n^1), \\
f_\varnothing^n &= f^n(X_n^0).
\end{aligned}$$

Proposition 4.1. *Assume we have groups of risk factors I_k for $k = 1, ..., s$ and each risk factor is exactly in one group. If the function $f^n : \mathbb{R}^n \to \mathbb{R}$ have continuous second order derivatives, then the sub-additivity with respect to the groups of risk factors I_k*

$$\rho\left(f_\Omega^n - f_\varnothing^n\right) \leq \sum_{k=1}^{s} \rho\left(f_{I_k}^n - f_\varnothing^n\right) \tag{4.57}$$

holds for all coherent risk measures ρ and for all X_n^0 if and only if the function f^n is separable with respect to the groups of risk factors I_k

$$f_\Omega^n(X) = \sum_{k=1}^{s} g_k(X_{I_k}). \tag{4.58}$$

Following propositions will help us to prove proposition 4.1.

Proposition 4.2. *Let $f^n : \mathbb{R}^n \to \mathbb{R}$, with scenarios $X_n^0, X_n^1 \in \mathbb{R}^n$, have continuous second order derivatives, then the value of function f^n in scenario X_n^1 can be calculated as*

$$f^n(1...1) = \sum_{i=1}^{n} f_i^n - (n-1)f^n(0...0) + \sum_{1 \leq i < j \leq n} I_{ij}^n, \tag{4.59}$$

where

$$I_{ij}^n = \int\limits_{x_i^0}^{x_i^1} \int\limits_{x_j^0}^{x_j^1} \frac{\partial^2 f^n}{\partial x_i \partial x_j}(x_1^0, x_2^0, ..., u_i, x_{i+1}^0, ..., u_j, x_{j+1}^1, ..., x_n^1) du_j du_i, \tag{4.60}$$

for $1 \leq i < j \leq n$.

Proof of proposition 4.2. For $n = 1$ the equation (4.59) reduce to

$$f^1(1) = f^1(1). \tag{4.61}$$

While the case $n = 1$ is sufficient basis for mathematical induction, in the inductive step we will use also the equation (4.59) for $n = 2$, therefore we need to proof it separately from inductive step. For $n = 2$ we get

$$\begin{aligned} f^2(x_1^1, x_2^1) &= f^2(x_1^0, x_2^1) + \int\limits_{x_1^0}^{x_1^1} \frac{\partial f^2}{\partial x_1}(u_1, x_2^1) du_1 \\ &= f^2(x_1^0, x_2^1) + \int\limits_{x_1^0}^{x_1^1} \frac{\partial f^2}{\partial x_1}(u_1, x_2^0) + \int\limits_{x_2^0}^{x_2^1} \frac{\partial^2 f^2}{\partial x_2 \partial x_1}(u_1, u_2) du_2 du_1 \\ &= f_1^2 + f_2^2 - f^2(00) + I_{12}^2, \end{aligned} \tag{4.62}$$

what proves the equation (4.59) for $n = 2$.
With equation (4.59) proved for $n \in \{1, 2\}$ we can continue with the induction step. Assume that the proposition holds for n, then the proposition holds also for function $h^n : \mathbb{R}^n \to \mathbb{R}$ defined as

$$h^n(x_1, ..., x_n) = f^{n+1}(x_1, ..., x_n, x_{n+1}^1)$$

and by equation (4.59) for function h^n we get

$$f^{n+1}(1...11) = \sum_{i=1}^{n} f_{i(n+1)}^{n+1} - (n-1) f^{n+1}(0...01) + \sum_{1 \le i < j \le n} I_{ij}^{n+1}. \tag{4.63}$$

Define function $g_{i(n+1)}^{n+1} : \mathbb{R}^2 \to \mathbb{R}$ as

$$g_{i(n+1)}^{n+1}(x_i^1, x_{n+1}^1) = f_{i(n+1)}^{n+1},$$

applying equation (4.62) we get

$$f_{i(n+1)}^{n+1} = f_i^{n+1} + f_{n+1}^{n+1} - f^{n+1}(0...0) + I_{i(n+1)}^{n+1} \tag{4.64}$$

Substituting equation (4.64) into the equation (4.63) we get

$$\begin{aligned} f^{n+1}(1...1) &= \sum_{i=1}^{n} \left(f_i^{n+1} + f_{n+1}^{n+1} - f^{n+1}(0...0) + I_{i(n+1)}^{n+1} \right) \\ &\quad -(n-1) f_{n+1}^{n+1} + \sum_{1 \le i < j \le n} I_{ij}^{n+1} \\ &= \sum_{i=1}^{n+1} f_i^{n+1} - n f^{n+1}(0...0) + \sum_{1 \le i < j \le n+1} I_{ij}^{n+1}, \end{aligned} \tag{4.65}$$

what is equation (4.59) for $n + 1$ and together with equation (4.61) and (4.62) proofs the proposition 4.2. $\square$

The proposition 4.2 gives a simple approximation of the change of function f^n between two point $X_n^0, X_n^1 \in \mathbb{R}^n$ in form

$$f^n(1...1) - f^n(0...0) \approx \sum_{i=1}^{n} \left(f_i^n - f^n(0...0)\right). \tag{4.66}$$

The error which we make using this approximation can be calculated as

$$\epsilon = \sum_{1 \leq i < j \leq n} I_{ij}^n. \tag{4.67}$$

In case of portfolio risk analysis and function f^n representing value of portfolio, left side of equation (4.66) represent portfolio profit/loss when moving from scenario X_n^0 to scenario X_n^1, while the right side is the sum of contribution of individual risk factors. The error term ϵ gives then the interaction between the risk factors and it can be by definition approximated as

$$|\epsilon| \leq K \frac{(n-1)n}{2} ||X_n^1 - X_n^0|| \tag{4.68}$$

if the absolute value of second order mixed derivatives are bounded by constant K. As a consequence of this approximation of error terms, the approximation (4.66) is exact when the second ordered mixed derivatives are zero, resp. when the cross-dependence of the risk factors is vanishing. The following proposition summarize these results.

Proposition 4.3. *Let $f^n : \mathbb{R}^n \to \mathbb{R}$ have continuous second order derivatives, then*

$$f^n(1...1) - f^n(0...0) = \sum_{i=1}^{n} \left(f_i^n - f^n(0...0)\right) \tag{4.69}$$

holds for all pairs $X_n^0, X_n^1 \in \mathbb{R}^n$ if and only if the function f^n can be written as

$$f^n(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} g_i(x_i). \tag{4.70}$$

Lema 4.1. *Let $f^n : \mathbb{R}^n \to \mathbb{R}$ have continuous second order derivatives, then*

$$f^n(1...1) - f^n(0...0) = \sum_{i=1}^{n} \left(f_i^n - f^n(0...0)\right) \tag{4.71}$$

holds $\forall X_n^0, X_n^1 \in \mathbb{R}^n$ *if and only if* $\forall X \in \mathbb{R}$ *and* $\forall (i,j) \in J$*, where* $J = \{(i,j) : 1 \le i < j \le n\}$

$$\frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) = 0. \tag{4.72}$$

Proof of lema 4.1. The if part of the lema follows directly from definition of I_{ij}^n and proposition 4.2. If $\forall X \in \mathbb{R}, \forall (i,j) \in J : \frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) = 0$, then by definition of I_{ij}^n, equation (4.60), we have $I_{ij}^n = 0$ for $(i,j) \in J$ and equation (4.59) from proposition (4.2) is reduced to equation (4.71).
The proof of the only if part follows. If $\exists X \in \mathbb{R} : \frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) > 0\ (< 0)$, then from the continuity of $\frac{\partial^2 f^n}{\partial x_i \partial x_j}$ we can construct such a neighborhood $O(X)$ of scenario X that $\forall Y \in O(X) : \frac{\partial^2 f^n}{\partial x_i \partial x_j}(Y) > 0\ (< 0)$. Taking two scenarios from this neighborhood $X_n^0, X_n^1 \in O(X)$ such, that $x_k^0 < x_k^1$ for $k \in \{i,j\}$ and $x_k^0 = x_k^1$ for $k \notin \{i,j\}$, then from definition of I_{ij}^n we get $I_{ij}^n < 0 (> 0)$ and $I_{kl}^n = 0$ for $(k,l) \neq (i,j)$. From proposition 4.2, equation (4.59), we get

$$f^n(1...1) - f^n(0...0) < \sum_{i=1}^{n} \left(f_i^n - f^n(0...0)\right),$$

when $\frac{\partial^2 f^n}{\partial x_i \partial x_j} > 0$, resp.

$$f^n(1...1) - f^n(0...0) > \sum_{i=1}^{n} \left(f_i^n - f^n(0...0)\right),$$

when $\frac{\partial^2 f^n}{\partial x_i \partial x_j} < 0$. $\square$

Note that choosing $x_i^0 > x_i^1$ instead of $x_i^0 < x_i^1$, resp. switching X_n^0 and X_n^1 will result in opposite inequality. As a consequence if $\exists X \in \mathbb{R} : \frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) \neq 0$ then there exists scenarios $X_n^0, X_n^1 \in \mathbb{R}^n$ such that

$$f^n(1...1) - f^n(0...0) < \sum_{i=1}^{n} \left(f_i^n - f^n(0...0)\right)$$

and there exists scenarios $\tilde{X}_n^0, \tilde{X}_n^1 \in \mathbb{R}^n$ such that

$$f^n(\tilde{1}...\tilde{1}) - f^n(\tilde{0}...\tilde{0}) > \sum_{i=1}^{n} \left(f_{\tilde{i}}^n - f^n(\tilde{0}...\tilde{0})\right)$$

and neither inequality can be assured.

Proof of proposition 4.3. Given the lema 4.1, we need to prove that $\forall X \in \mathbb{R}^n$, $i \neq j : \frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) = 0$ if and only if the function f^n can be separated as

$$f^n(x_1, ..., x_n) = \sum_{i=1}^{n} g_i(x_i). \tag{4.73}$$

The if part is quite trivial. If the function f^n is separable as defined in equation 4.73 then $\frac{\partial f^n}{\partial x_i}(X) = g_i'(x_i)$, where $g_i(x_i)$ is constant with respect to x_j for $i \neq j$ and therefore $\forall X \in \mathbb{R}^n, i \neq j : \frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) = 0$.
The only if part becomes the interesting part of the proposition and we will prove it with mathematical induction. Assume that

$$\forall X \in \mathbb{R}^n, i \neq j : \frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) = 0 \tag{4.74}$$

then for $n = 2$ from equation (4.74) we get $\frac{\partial f^2}{\partial x_1}(X) = h(x_1)$, resp. $f^2(X) = H(x_1) + g_2(x_2)$. Choosing $g_1(x_1) = H(x_1)$ we finaly get the equation (4.73) for $n = 2$

$$f^2(X) = \sum_{i=1}^{2} g_i(x_i), \tag{4.75}$$

which is sufficient basis for mathematical induction. As next we will continue with the induction step.
Assume that the proposition holds for n. We will first show that the function $f^{n+1}(X)$ can be written as

$$f^{n+1}(X) = u_j(x_1, x_j, ..., x_{n+1}) + v_j(x_2, ..., x_{n+1}), \tag{4.76}$$

for $j = 2, ..., n + 1$.
For $j = 2$ we can take $u_2 = f^{n+1}$ and $v_2 = 0$.
Now assume that the separation is possible till j, then we have

$$f^{n+1}(X) = u_j(x_1, x_j, ..., x_{n+1}) + v_j(x_2, ..., x_{n+1}). \tag{4.77}$$

As the function v_j does not depend on the variable x_1 we get we get

$$\frac{\partial^2 f^{n+1}}{\partial x_1 \partial x_j} = \frac{\partial^2 u_j}{\partial x_1 \partial x_j} \tag{4.78}$$

what equals to zero, based on equation (4.74). Therefore we can apply equation (4.75) on function $\tilde{v}_j(x_1, x_j)$ defined as $\tilde{v}_j(x_1, x_j) = u_j(x_1, x_j, ..., x_{n+1})$ and we get

$$u_j(x_1, x_j, ..., x_{n+1}) = u_{j+1}(x_1, x_{j+1}, ..., x_{n+1}) + h_j(x_j, ..., x_{n+1}). \tag{4.79}$$

Denoting $v_{j+1} = v_j + h_j$ we get

$$f^{n+1}(X) = u_{j+1}(x_1, x_{j+1}, ..., x_{n+1}) + v_{j+1}(x_2, ..., x_{n+1}). \tag{4.80}$$

For $j = n + 1$ this gives

$$f^{n+1}(X) = u_{n+1}(x_1, x_{n+1}) + v_{n+1}(x_2, ..., x_{n+1}). \tag{4.81}$$

As v_{n+1} does not depend on variable x_1, we get from equation (4.74) that

$$\frac{\partial^2 u_n}{\partial x_1 \partial x_{n+1}} = \frac{\partial^2 f^{n+1}}{\partial x_1 \partial x_{n+1}} = 0. \tag{4.82}$$

Again we can apply the equation (4.75) on the function u_{n+1}

$$u_{n+1}(x_1, x_{n+1}) = g_1(x_1) + h(x_{n+1}). \tag{4.83}$$

Finally we get

$$f^{n+1}(X) = g_1(x_1) + v_{n+2}(x_2, ..., x_{n+1}). \tag{4.84}$$

The function v_{n+2} is function of n variables with all mixed second orders derivatives equals to zero and therefore we can apply the assumption of the induction step what implies

$$f^{n+1}(X) = \sum_{i=1}^{n+1} g_i(x_i). \tag{4.85}$$

□

We define maximum loss contribution (MLC) as

$$MLC(i) := \frac{f_i^n - f^n(0...0)}{f^n(1...1) - f^n(0...0)}, \tag{4.86}$$

for $f^n(0...0) \neq f^n(1...1)$. As a consequency of proposition 4.3

$$\sum_{i=1}^{n} MLC(i) = 1 \tag{4.87}$$

holds for all X_n^0 and all plausibility domains if and only if the function f^n can be separated as

$$f^n(x_1, ..., x_n) = \sum_{i=1}^{n} g_i(x_i), \tag{4.88}$$

resp. if and only if the second order derivatives equals zero.

In risk analysis a class of risk is usually not function of one variable, reps. risk factors but is often calculated as risk achieved by change of a group of risk factors. As example the market risk will be the risk given by change of market risk factors (i.e. stock prices, exchange rates,...). Following proposition generalise the proposition 4.2 to case of groups of risk factors.

Proposition 4.4. *Assume we have groups of risk factors I_k for $k = 1...s$, where each risk factor is exactly in one group. Then if function $f : \mathbb{R}^n \to \mathbb{R}$ has continuous second order derivatives then for any two scenarios $X_n^0, X_n^1 \in \mathbb{R}$ we get*

$$f_\Omega^n - f_\varnothing^n = \sum_{k=1}^{s} \left(f_{I_k}^n - f_\varnothing^n\right) + \sum_{1 \le k < l \le s} \tilde{I}_{kl}^s, \tag{4.89}$$

where

$$f_S^n = f^n(i_1...i_n),\ i_k = \begin{cases} 1 & k \in S \\ 0 & k \notin S \end{cases}, \tag{4.90}$$

$$\tilde{I}_{kl}^s = \int_0^1 \int_0^1 \sum_{k \in I_k, l \in I_l} (x_k^1 - x_k^0)(x_l^1 - x_l^0) \frac{\partial^2 f^n}{\partial x_k \partial x_l}(y_1, ..., y_n) du dv, \tag{4.91}$$

where

$$y_i = \begin{cases} x_i^0 + u(x_i^1 - x_i^0) & i \in I_k, \\ x_i^0 + v(x_i^1 + x_i^0) & i \in I_l, \\ x_i^1 & i > max(I_l), \\ x_i^0 & otherwise. \end{cases} \tag{4.92}$$

Proof of proposition 4.4. Define a function $\tilde{f} : \mathbb{R}^s \to \mathbb{R}$ as

$$\tilde{f}(Y) = f(g_1(y_1), ..., g_s(y_s)), \tag{4.93}$$

where functions g_k represents paths between scenarios $X_{I_k}^0$ and $X_{I_k}^1$, resp. the functions g_k are two times differentiable with $g_k(0) = X_{I_k}^0$ and $g_k(1) = X_{I_k}^1$. The proposition 4.2 for function $\tilde{f}$ becomes

$$f_\Omega^n - f_\varnothing^n = \sum_{k=1}^{s} \left(f_{I_k}^n - f_\varnothing^n\right) + \sum_{1 \le k < l \le s} \tilde{I}_{kl}^s, \tag{4.94}$$

where $\tilde{I}_{kl}^s$ depends on the paths, which are defined by the functions g_k. Choosing the shortest and linear path between $X_{I_k}^0$ and $X_{I_k}^1$ as

$$g_k(t) = X_{I_k}^0 + t\left(X_{I_k}^1 - X_{I_k}^0\right) \tag{4.95}$$

we get the $\tilde{I}^s_{kl}$ for $1 \le k < l \le s$ in form

$$\tilde{I}^s_{kl} = \int_0^1 \int_0^1 \sum_{i \in I_k, j \in I_l} (x_i^1 - x_i^0)(x_j^1 - x_j^0) \frac{\partial^2 f^n}{\partial x_i \partial x_j}(y_1, ..., y_n) du dv. \tag{4.96}$$

$\square$

Similar as proposition 4.2 gives approximation of change of the function f^n by the change given by individual risk factors, the proposition 4.4 gives the approximation by the changes given by groups of risk factors.

Proposition 4.5. *Assume we have groups of risk factors I_k for $k = 1...s$, where each risk factor is exactly in one group. Then if function $f : \mathbb{R}^n \to \mathbb{R}$ has continuous second order derivatives then*

$$f^n_\Omega - f^n_\varnothing = \sum_{k=1}^{s} \left(f^n_{I_k} - f^n_\varnothing \right) \tag{4.97}$$

hold $\forall X_n^0, X_n^1 \in \mathbb{R}^n$ if and only if function f^n can be separated as

$$f^n_\Omega(X) = \sum_{k=1}^{s} g_k(X_{I_k}), \tag{4.98}$$

resp. if and only if $\forall X \in \mathbb{R}^n$

$$\frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) = 0, \tag{4.99}$$

for each $i \in I_k$ and each $j \in I_l$, with $k \neq l$.

Proof of proposition 4.5 is identical to proof of proposition 4.3, when we use the paths between the scenarios as in the proof of proposition 4.4.

In propositions 4.2-4.5 we analyzed the approximation of function by changes of individual risk factors, reps. by changes of groups of risk factors. Given these propositions we can investigate the subadditivity with respect to a groups of risk factors for coherent risk measures. The proof of proposition 4.1 follows.

Proof of proposition 4.1. Based on proposition 4.5 the equation (4.58) imply that

$$f^n_\Omega - f^n_\varnothing = \sum_{k=1}^{s} \left(f^n_{I_k} - f^n_\varnothing \right). \tag{4.100}$$

Assuming coherent risk measure ρ imply the subadditivity of this risk measure in sense

$$\rho(X+Y) \leq \rho(X) + \rho(Y). \tag{4.101}$$

Applying this $s-1$ times on equation (4.100) we get

$$\rho\left(f^n_\Omega - f^n_\varnothing\right) \leq \sum_{i=1}^{s} \rho\left(f^n_{I_k} - f^n_\varnothing\right). \tag{4.102}$$

Assume that the function f^n is not separable, then based on proposition 4.5 there $\exists X, k, l, i, j \,:\, k \neq l,\ i \in I_k,\ j \in I_l$, that $\frac{\partial^2 f^n}{\partial x_i \partial x_j}(X) \neq 0$. Based on proof of lema 4.1 there exists such X^0_n and X^1_n that

$$f^n_\Omega - f^n_\varnothing > \sum_{k=1}^{s} \left(f^n_{I_k} - f^n_\varnothing\right). \tag{4.103}$$

Taking $\rho_{X^1_n}$ as maximum loss with plausibility domain equals to points, namely X^1_n we get

$$\rho_{X^1_n}\left(f^n_\Omega - f^n_\varnothing\right) \geq \sum_{i=1}^{s} \rho_{X^1_n}\left(f^n_{I_k} - f^n_\varnothing\right), \tag{4.104}$$

for coherent risk measure $\rho_{X^1_n}$. □

Chapter 5

Temporal aggregation of GARCH models

In models of portfolio selection or option pricing the fundamental position has estimation of the model and time series parameters. The choice of time series models and calibration of these models influence also the results of portfolio selection problem discussed in Chapter 3 and the results of option pricing model discussed in Chapter 4. With higher frequency data being increasingly available and attention focusing on longer time horizons we face the problem whether or not to use the higher frequency data available in the long term analysis. At first sight it seems clear that it should be used. If we restricted ourselves to the low frequency data we either would have very few data points or use very old historical data for getting reliable parameter estimates. Neither is desirable. On the other hand, when we use the high frequency data the time horizon of the forecast is several time steps ahead. The long term analysis then has to calculate the distribution arising from aggregating the high frequency model over several time steps. With the high frequency model even the estimation error will be aggregated and can neglect the advantage of having many historic data for estimation. This motivates our analysis of aggregated distributions. We concentrate our analysis on GARCH model as introduced by Bollerslev [17]. The approach we take is to estimate a strong GARCH model for single time steps of suitable length and then aggregate over sufficiently many time steps to arrive at the desired time horizon. Drost and Nijman [39] in a landmark paper showed that the temporal aggregate of a strong GARCH process is in general not a strong GARCH. Therefore they introduced the larger classes of semi-strong and weak GARCH models. For semi-strong GARCH processes the mean and variance of innovations are determined, but other properties of the distribution of innovations are not determined. In particular, the innovations need not be independent or identically distributed. For weak GARCH processes not even the mean and variance are determined, we just have a linear predictor. Weak

GARCH processes have the advantage of aggregating to weak GARCH processes, but for purposes of risk management they do not convey much information. For mean and variance they only specify the best linear predictor, other properties of the distribution are not specified at all. In risk management we often need more information about the conditional distribution: quantiles, higher moments, and for risk measures like Expected Shortfall even the full distribution function in the tails. To calculate the distribution of tails of aggregated random variable one need the full distribution of non-aggregated model. This information is not specified by semi-strong or weak but only by strong GARCH processes. For this reason we will focus on analysis of the aggregated distribution of strong GARCH processes accepting that this aggregated distribution is itself not a strong GARCH process, but we will derive some of the properties of higher moments of aggregated time series also for more general semi-strong GARCH process. These results are part of the working paper [23]. As first we recall the definition of strong, semi-strong, and weak GARCH processes as introduced in [39].

Definition 5.1. *Let $\{h_t, t \in \mathbb{Z}\}$ be defined as the stationary solution of*

$$h_t = \psi + \alpha\epsilon_{t-1}^2 + \beta h_{t-1}.$$

A time series $\{\epsilon_t, t \in \mathbb{Z}\}$ is said to be generated by a strong GARCH(1,1) process if ψ, α, β can be chosen in such a way that

$$\xi_t \;:=\; \epsilon_t/\sqrt{h_t} \sim D(0,1) \qquad i.i.d., \tag{5.1}$$

where $D(0,1)$ is some fixed distribution of errors with zero mean and unit variance. The series $\{\epsilon_t, t \in \mathbb{Z}\}$ is said to be generated by a semi-strong GARCH(1,1) process if

$$\begin{aligned} E[\epsilon_t|I_{t-1}] &= 0, \\ E[\epsilon_t^2|I_{t-1}] &= \psi + \alpha\epsilon_{t-1}^2 + \beta h_{t-1} = h_t, \end{aligned} \tag{5.2}$$

where the information set $I_{t-1} := \{\epsilon_{t-1}, \epsilon_{t-2}, \ldots\}$ describes the information available at time $t-1$. A time series $\{\epsilon_t, t \in \mathbb{Z}\}$ is said to be generated by a weak GARCH(1,1) process if

$$\begin{aligned} P[\epsilon_t|\epsilon_{t-1}, \epsilon_{t-2}, ...] &= 0, \\ P[\epsilon_t^2|\epsilon_{t-1}, \epsilon_{t-2}, ...] &= \psi + \alpha\epsilon_{t-1}^2 + \beta h_{t-1} = h_t, \end{aligned} \tag{5.3}$$

where $P[x_t|\epsilon_{t-1}, \epsilon_{t-2}, ...]$ denotes the best linear prediction of x_t in terms of 1, ϵ_{t-1}, ϵ_{t-2} , ... ,$\epsilon_{t-1}^2, \epsilon_{t-2}^2, ...$ Furthermore we assume $\psi, \alpha, \beta > 0$ for volatility to be positive and $\alpha + \beta < 1$ in order to ensure stationarity of the process.

We will consider both stock random variable and flow random variable for the aggregation. The difference between the stock and flow random variable is in the way how they aggregate. An example of stock random variables are the stock prices or the wealth of a person. For stock random variable only the last observed value within the aggregation period is of interest. We will denote the aggregation of stock random variable ϵ_t over m time periods as $\epsilon_{(m)t}$, which we define as

$$\epsilon_{(m)t} := \epsilon_{t+m}. \tag{5.4}$$

An example of flow random variable are the log returns of stock prices or the income of a person. For flow random variable the aggregation is given by the sum of all observed values within the aggregation period. We will denote the aggregation of flow random varialbe ϵ_t over m time periods as $\epsilon_{[m]t}$, which is defined as

$$\epsilon_{[m]t} := \sum_{i=1}^{m} \epsilon_{t+i}. \tag{5.5}$$

Unlike the definition of aggregation addopted by Drost and Nijman we use slightly different definition for aggregated stock variable, which covers also the Drost and Nijman definition for $t := (\tilde{t} - 1)m$. Similar our definition of aggregated flow variable covers the Drost and Nijman definition for $t := m\tilde{t}$. This alternative definition simplify the proof as it allows for simpler recurrent calculation of aggregated variance and kurtosis. This results from the possibility of overlapping aggregation.

The densities of the conditional aggregated distribution can be calculated explicitly for both flow and stock random variables. For a flow variable the probability density function of m-period aggregation under the assumption, that the whole non aggregated history is known, is

$$\begin{aligned} f_{\epsilon_{[m]t}|I_t}(y) &= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} d\epsilon_{t+1} \ldots d\epsilon_{t+m-1} \\ & \quad f_{\epsilon_{t+m}|I_{t+m-1}}\left(y - \sum_{i=1}^{m-1} \epsilon_{t+i}\right) \prod_{i=1}^{m-1} f_{\epsilon_{t+i}|I_{t+i-1}}(\epsilon_{t+i}). \end{aligned} \tag{5.6}$$

Here I_t represents the information set available in time t. Information set I_{t+i} inherit all the information from information in I_{t+i-1} and additionally the value of the process in time $t+i$, which is ϵ_{t+i}.

For a stock variable the probability density function of m-period aggregation under the assumption, that the whole non aggregated history is known $(\epsilon_{(m)t}|I_t$

corresponds to m-step ahead forecast), is

$$f_{\epsilon_{(m)t}|I_t}(y) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} d\epsilon_{t+1} \ldots d\epsilon_{t+m-1} \qquad (5.7)$$
$$f_{\epsilon_{t+m}|I_{t+m-1}}(y) \prod_{i=1}^{m-1} f_{\epsilon_{t+i}|I_{t+i-1}}(\epsilon_{t+i}).$$

This equation can also be found in Andersen et al. [5, eq. (3.26)]. The aggregated density functions of equations (5.6), (5.7) are the basis of risk measurement at the aggregated time level. Apart from realizations not being serially independent the density functions do not necessarily have the same form as the density of one period returns. These are the reasons why the aggregated realizations of strong GARCH are not in generally strong GARCH.

In Section 5.1 we study the kurtosis and variance of the conditional aggregated distributions of equations (5.6), (5.7). We also analyse the limit behavior of conditional variance and kurtosis when aggregating over sufficiently many time steps in the Section 5.2. It turns out that in the limit of infinitely many aggregation steps (corresponding to an infinite time horizon) the conditional aggregated kurtosis approaches three (resp. a different constant, for stock variables) or infinity depending on whether or not a simple inequality in term of the GARCH parameters ψ, α, β (and additionally κ, for flow variables) is satisfied. In Section 5.3 we deal with the optimal data frequency for strong GARCH processes.

5.1 Conditional variance and kurtosis

Now we can specify the conditional variance and conditional kurtosis of the aggregated GARCH (1,1) processes for stock and flow variables.

Theorem 5.1. *Assume ϵ_t is a stock variable following a semi-strong GARCH process. Then the conditional variance of the m-step distribution $\epsilon_{(m)t} := \epsilon_{t+m}$ is given by*

$$\mathrm{Var}(\epsilon_{(m)t}|I_t) = \sigma_u^2 + (\alpha+\beta)^{m-1}\left(h_{t+1} - \sigma_u^2\right), \qquad (5.8)$$

where σ_u^2 is the unconditional variance of one period returns, $\sigma_u^2 = \psi/(1-\alpha-\beta)$. If innovations ξ_t are independent, have symmetric distribution and have finite fourth moments κ, then the conditional fourth moments of the m-step distribution $\epsilon_{(m)t}$ can be written recursively as

$$E(\epsilon_{(m)t}^4|I_t) = \kappa\psi^2 + \gamma E(\epsilon_{(m-1)t}^4|I_t) \qquad (5.9)$$
$$+2\psi\kappa(\alpha+\beta)\mathrm{Var}(\epsilon_{(m-1)t}|I_t),$$

where κ is the unconditional kurtosis of innovations ξ_t (which equals the conditional kurtosis of ϵ_t) and

$$\gamma := \alpha^2\kappa + \beta^2 + 2\alpha\beta.$$

Non-recursively the fourth moment of the aggregated distribution is given by

$$\begin{aligned} E(\epsilon^4_{(m)t}|I_t) &= \gamma^{m-1}\kappa h^2_{t+1} + \frac{\kappa\psi^2 + 2\psi\kappa(\alpha+\beta)\sigma^2_u}{1-\gamma}(1-\gamma^{m-1}) \\ &\quad +2\psi\kappa(\alpha+\beta)\frac{h_{t+1}-\sigma^2_u}{\alpha+\beta-\gamma}\left((\alpha+\beta)^{m-1}-\gamma^{m-1}\right), \end{aligned} \tag{5.10}$$

when $\gamma \notin \{1, \alpha+\beta\}$.

Proof of Theorem 5.1. By the law of iterated expectations we can express conditional variance as

$$\operatorname{Var}(\epsilon_{(m)t}|I_t) = E(E(\epsilon^2_{t+m}|I_{t+m-1})|I_t) = E(h_{t+m}|I_t). \tag{5.11}$$

Therefore

$$\begin{aligned} \operatorname{Var}(\epsilon_{(m)t}|I_t) &= E(\psi + \alpha\epsilon^2_{t+m-1} + \beta h_{t+m-1}|I_t) \\ &= \psi + \alpha E(\epsilon^2_{t+m-1}|I_t) + \beta E(h_{t+m-1}|I_t). \end{aligned}$$

Applying equation (5.11) to $m-1$ we get

$$\begin{aligned} \operatorname{Var}(\epsilon_{(m)t}|I_t) &= \psi + (\alpha+\beta)\operatorname{Var}(\epsilon_{t+m-1}|I_t) \\ &= \psi\sum_{i=0}^{m-2}(\alpha+\beta)^i + (\alpha+\beta)^{m-1}h_{t+1} \\ &= \psi\frac{1-(\alpha+\beta)^{m-1}}{1-(\alpha+\beta)} + (\alpha+\beta)^{m-1}h_{t+1} \\ &= \sigma^2_u + (\alpha+\beta)^{m-1}\left(h_{t+1}-\sigma^2_u\right) \end{aligned}$$

which proves equation (5.8). Now consider the fourth moments. By the definition of $\epsilon_{(m)t}$ for stock variables we have

$$E(\epsilon^4_{(m)t}|I_{t-1}) = E(\epsilon^4_{t+m}|I_t) = E(\xi^4_{t+m}h^2_{t+m}|I_t).$$

Furthermore, since by the assumption of the theorem the ξ_{t+m} is independent on the h_{t+m}, as the h_{t+m} is generated by previous realization of the innovations, and the fourth moment of innovation ξ_{t+m} is equal to κ, we get

$$E(\epsilon^4_{(m)t}|I_t) = E(\xi^4_{t+m}|I_t)E(h^2_{t+m}|I_t) = \kappa E(h^2_{t+m}|I_t).$$

Reasoning as in equation (5.11), and using the definition of h_t we get

$$\begin{aligned}
E(\epsilon^4_{(m)t}|I_t) &= \kappa E((\psi + \alpha\epsilon^2_{t+m-1} + \beta h_{t+m-1})^2|I_t) \\
&= \kappa\psi^2 + \kappa 2\alpha\beta E(\epsilon^2_{t+m-1}h_{t+m-1}|I_t) \\
&\quad +\kappa\alpha^2 E(\epsilon^4_{t+m-1}|I_t) + \kappa\beta^2 E(h^2_{t+m-1}|I_t) \\
&\quad +2\kappa\psi\left(\alpha E(\epsilon^2_{t+m-1}|I_t) + \beta E(h_{t+m-1}|I_t)\right) \\
&= \kappa\psi^2 + 2\alpha\beta E(\epsilon^4_{(m-1)t}|I_t) \\
&\quad +\kappa\alpha^2 E(\epsilon^4_{(m-1)t}|I_t) + \beta^2 E(\epsilon^4_{(m-1)t}|I_t) \\
&\quad +2\kappa\psi(\alpha+\beta)\mathrm{Var}(\epsilon_{(m-1)t}|I_t) \\
&= \kappa\psi^2 + (\kappa\alpha^2 + \beta^2 + 2\alpha\beta)E(\epsilon^4_{(m-1)t}|I_t) \\
&\quad +2\kappa\psi(\alpha+\beta)\mathrm{Var}(\epsilon_{(m-1)t}|I_t)
\end{aligned}$$

which proves equation (5.9). □

Theorem 5.2. *Assume ϵ_t is a flow variable following a semi-strong GARCH process. Then the conditional variance of the aggregated flow variable $\epsilon_{[m]t} := \sum_{i=1}^m \epsilon_{t+i}$ is given by*

$$\mathrm{Var}(\epsilon_{[m]t}|I_t) = m\sigma^2_u + \frac{1-(\alpha+\beta)^m}{1-(\alpha+\beta)}(h_{t+1} - \sigma^2_u), \tag{5.12}$$

where σ^2_u is the unconditional variance of one period returns ϵ_t, $\sigma^2_u = \psi/(1-\alpha-\beta)$. If innovations ξ_t are independent, have symmetric distribution and have finite fourth moment equal to κ, then the conditional fourth moments of the aggregated variable $\epsilon_{[m]t}$ can be written as

$$E(\epsilon^4_{[m]t}|I_t) = B\sum_{i=0}^{m-1} A^i b, \tag{5.13}$$

where $B := (1, h_{t+1}, h^2_{t+1}, 0, 0)$, $b' := (0, 0, \kappa, 0, 6)$, and

$$A = \begin{pmatrix} 1 & \psi & \psi^2 & 0 & 0 \\ 0 & \alpha+\beta & 2\psi(\alpha+\beta) & 1 & \psi \\ 0 & 0 & \kappa\alpha^2 + 2\alpha\beta + \beta^2 & 0 & \kappa\alpha+\beta \\ 0 & 0 & 0 & 1 & \psi \\ 0 & 0 & 0 & 0 & \alpha+\beta \end{pmatrix}.$$

Proof of Theorem 5.2. To prove equation (5.12) one calculates

$$\begin{aligned} E(\epsilon^2_{[m]t}|I_t) &= E\left(\left(\epsilon_{t+m} + \sum_{i=1}^{m-1}\epsilon_{t+i}\right)^2 \middle| I_t\right) \\ &= E\left((\sum_{i=1}^{m-1}\epsilon_{t+i})^2 + 2\epsilon_{t+m}(\sum_{i=1}^{m-1}\epsilon_{t+i})^2 + \epsilon^2_{t+m}\middle| I_t\right) \\ &= E(\epsilon^2_{[m-1]t}|I_t) + E(\epsilon^2_{t+m}|I_t). \end{aligned}$$

The first term of the equation represent the variance of aggregated flow variable over $m-1$ periods, while the second term of the equation represent the aggregation of stock variable over m periods, which is given in equation (5.8). We get

$$\begin{aligned} E(\epsilon^2_{[m]t}|I_t) &= E(\epsilon^2_{[m-1]t}|I_t) + \sigma^2_u + (\alpha+\beta)^{m-1}(h_{t+1} - \sigma^2_u) \\ &= \sum_{i=0}^{m-1}\left(\sigma^2_u + (\alpha+\beta)^i(h_{t+1} - \sigma^2_u)\right) \\ &= m\sigma^2_u + \frac{1-(\alpha+\beta)^m}{1-(\alpha+\beta)}(h_{t+1} - \sigma^2_u). \end{aligned}$$

This proves equation (5.12).

Instead of proving equation (5.13) we show the equivalent

$$E(\epsilon^4_{[m]t}|I_{t-1}) = BY_m, \tag{5.14}$$

with

$$Y_{i+1} = AY_i + b, \quad \text{for } i \geq 0 \tag{5.15}$$

$$Y_0' = (0,0,0,0,0) \tag{5.16}$$

and $b' = (0,0,\kappa,0,6)$, $B = (1, h_{t+1}, h^2_{t+1}, 0, 0)$.
We will denote the components of Y_i as $Y_i = (Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4}, Y_{i5})'$. By mathematical induction with respect to i we will show that

$$\begin{aligned} E(\epsilon^4_{[m]t}|I_t) &= E\Bigg(Y_{i1} + Y_{i2}h_{t+1+m-i} + Y_{i5}h_{t+1+m-i}(\sum_{j=1}^{m-i}\epsilon_{t+j})^2 \\ &\quad + Y_{i3}h^2_{t+1+m-i} + Y_{i4}(\sum_{j=1}^{m-i}\epsilon_{t+j})^2 + (\sum_{j=1}^{m-i}\epsilon_{t+j})^4 \Bigg| I_t\Bigg) \end{aligned} \tag{5.17}$$

holds for $i = 0, 1, ..., m$. For $i = 0$ we have $Y_0' = (0, 0, 0, 0, 0)$ and therefore we get

$$E(\epsilon_{[m]t}^4 | I_t) = E\left((\sum_{j=1}^{m} \epsilon_{t+j})^4 \middle| I_t \right),$$

which follows directly from the definition of the aggregated flow variable $\epsilon_{[m]t}$.

The induction step is established by applying the definition of h_t. As in proof of theorem 5.1 we get

$$E[h_{t+1+m-i} | I_t] = E[\varphi + (\alpha + \beta) h_{t+1+m-(i+1)} | I_t] \tag{5.18}$$

for $i \le m$. This imply that the second column of matrix A, which corresponds to h_t, will have ψ on first row, which corresponds to constant, and value $\alpha + \beta$ in second column, which corresponds to h_t. Furthermore we split all the sums by removing the last element

$$\sum_{j=1}^{m-i} \epsilon_{t+j} = \epsilon_{t+m-i} + \sum_{j=1}^{m-(i+1)} \epsilon_{t+j}. \tag{5.19}$$

Reorganizing the components of the expectation and using the definition of innovations we get also the other columns of matrix A and vector b, therefore the equation holds also for $i + 1$.
Finally for $i = m$ equation (5.17) reads

$$\begin{aligned} E(\epsilon_{[m]t}^4 | I_{t-1}) &= E\left(Y_{m1} + Y_{m2} h_t + Y_{m3} h_t^2 | I_t\right), \\ &= Y_{m1} + Y_{m2} h_t + Y_{m3} h_t^2 = BY_m. \end{aligned}$$

which is equation (5.14). □

Remark 5.1. *Drost and Nijman [39, p. 916, eq. (14)] give the unconditional kurtosis of the aggregated distribution. Our results specify the conditional kurtosis, which is relevant for purposes of risk management.*

5.2 The limiting behavior of the conditional aggregated kurtosis

Let us now consider the long-term behavior of the conditional aggregated kurtosis, still under the assumption that the one period process is a semi-strong GARCH process with symmetric innovations, which have finite fourth moment and are independent. We let the number of aggregation steps increase. With fixed length of the basic period, this amounts to a proportional increase of the time horizon.

Diebold [38] has shown that a version of the central limit theorem implies that conditional heteroscedasticity disappears with increasing sampling intervals. Moreover, it is generally accepted that for most financial time series return innovations tend to normality as the sampling interval increases. If return innovations indeed approach normality with increasing sampling intervals, the kurtosis of the conditional aggregate should approach the value of three as $m \to \infty$.

Corollary 5.1. *For a stock variable, under the assumptions of Theorem 5.1 the kurtosis of the conditional aggregated distribution can be defined as*

$$\kappa_{(m)} := \frac{E(\epsilon^4_{(m)t}|I_t)}{\Big(E(\epsilon^2_{(m)t}|I_t)\Big)^2}. \tag{5.20}$$

Aggregated kurtosis $\kappa_{(m)}$ *goes to infinity as* $m \to \infty$ *in the case* $\gamma := \alpha^2\kappa + 2\alpha\beta + \beta^2 \geq 1$. *In the case* $\gamma < 1$ *the kurtosis goes to the following finite value*

$$\lim_{m\to\infty} \kappa_{(m)} = \kappa\psi \frac{\psi + 2(\alpha+\beta)\sigma_u^2}{(1-\gamma)\sigma_u^4}. \tag{5.21}$$

Remark 5.2. *While limit of kurtosis for a stock variable is finite as* m *goes to infinity, this finite value can not be bounded without the knowledge of the process parameters* κ, ψ, α *and* β. *Choosing* γ *near one, but smaller than one, can lead to the limit of any size.*

Proof of Corollary 5.1. Because of Theorem 5.1 the limit of the variance of an aggregated stock variable is equal to

$$\sigma_u^2 = \frac{\psi}{1-(\alpha+\beta)}. \tag{5.22}$$

So the limit of the aggregated kurtosis defined by equation (5.20) equals

$$\lim_{m\to\infty} \kappa_{(m)} = \frac{1}{\sigma_u^4} \lim_{m\to\infty} E(\epsilon^4_{(m)t}|I_t). \tag{5.23}$$

If $\gamma < 1$ the limit of the fourth moments given by equation (5.10) is

$$\lim_{m\to\infty} E(\epsilon^4_{(m)t}|I_t) = \kappa\psi \frac{\psi + 2(\alpha+\beta)\sigma_u^2}{1-\gamma},$$

which together with equation (5.23) proves the first part of the corollary.
If $\gamma \geq 1$ equation (5.22) implies that the fourth moment goes to infinity as $m \to \infty$ because from equation (5.9) we get

$$E(\epsilon^4_{(m)t}|I_t) \geq \kappa\psi^2 + E(\epsilon^4_{(m-1)t}|I_t) \geq m\kappa\psi^2.$$

Additionally by equation (5.8) we have $\lim_{m\to\infty} \mathrm{Var}(\epsilon_{(m)t}|I_t) = \sigma_u^2$, which is by assumption positive as $\psi, \alpha, \beta > 0$. Accordingly the aggregated kurtosis goes to infinity. □

For flow variables the limit behavior of the conditional aggregated kurtosis is given by following theorem.

Theorem 5.3. *Assume ϵ_t is a flow variable following a semi-strong GARCH process and innovations ξ_t are independent with symmetric distribution and with a finite fourth moments κ. Then kurtosis of the conditional aggregated distribution defined as*

$$\kappa_{[m]} := \frac{E(\epsilon^4_{[m]t}|I_t)}{\Big(E(\epsilon^2_{[m]t}|I_t)\Big)^2}. \tag{5.24}$$

has the following limit

$$\lim_{m\to\infty} \kappa_{[m]} = \begin{cases} 3 & \text{if } \gamma < 1 \\ \infty & \text{if } \gamma > 1 \\ 3+d & \text{if } \gamma = 1 \end{cases}, \tag{5.25}$$

where

$$\begin{aligned} \gamma &= (\alpha^2\kappa + 2\alpha\beta + \beta^2), \\ d &= \big((\psi\kappa)/(2\sigma_u^2) + 3(\kappa\alpha + \beta)\big)\big(\psi/(\sigma_u^2) + 2(\alpha+\beta)\big). \end{aligned}$$

Proof of Theorem 5.3. From the definition of aggregated flow variable in equation (5.24) we get

$$\lim_{m\to\infty} k_{[m]} = \lim_{m\to\infty} \frac{E\left(\epsilon^4_{[m]t}|I_t\right)}{\left(E\left(\epsilon^2_{[m]t}|I_t\right)\right)^2}. \tag{5.26}$$

Because of Theorem 5.2 and its equation (5.12) for fourth moment and equation (5.13) for variance we can write

$$\begin{aligned}
\lim_{m\to\infty} k_{[m]} &= \lim_{m\to\infty} \frac{B\sum_{i=0}^{m-1} A^i b}{\left(m\sigma_u^2 + \frac{1-(\alpha+\beta)^m}{1-(\alpha+\beta)}(h_{t+1}-\sigma_u^2)\right)^2} && (5.27)\\
&= \lim_{m\to\infty} \frac{\frac{1}{m^2}B\sum_{i=0}^{m-1} A^i b}{\left(\sigma_u^2 + \frac{1}{m}\frac{1-(\alpha+\beta)^m}{1-(\alpha+\beta)}(h_{t+1}-\sigma_u^2)\right)^2} && (5.28)\\
&= \frac{1}{\sigma_u^2} B\left(\lim_{m\to\infty}\frac{1}{m^2}\sum_{i=0}^{m-1} A^i\right) b. && (5.29)
\end{aligned}$$

To express the limit of the kurtosis we need to calculate A^i. While calculating A^i directly is complex, we can search for similar matrix J for which J^i is simple to calculate. For similar matrix holds $A = S^{-1}JS$ and therefore we have $A^i = S^{-1}J^iS$. An example of matrix for which J^i is simple to calculate are diagonal matrices or the Jordan matrices[1]. In the proof we will not use complete Jordan decomposition[2], we will write the matrices in a form which gives a little bit simpler matrices S and J and still allow for simple calculation of J^i.
The Jordan decomposition of the matrix A depends on the value of $\gamma := \kappa\alpha^2 + 2\alpha\beta + \beta^2$. More precisely, we need to distinguish three different cases. The first case is $\gamma = \alpha+\beta$, the second case is $\gamma = 1$, and in the third case γ is not equal to one and not equal to $\alpha+\beta$. The three cases differ only in the form of the decomposition. All other steps of the argument are the same. We will give the argument for the third case.
Since γ differs from one and $\alpha+\beta$, the matrix J of the decomposition is

$$J = \begin{pmatrix} 1 & 0 & 0 & \sigma_u^2 & 0 \\ 0 & \alpha+\beta & 0 & 0 & d_3 \\ 0 & 0 & \gamma & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \alpha+\beta \end{pmatrix}$$

and

$$S = \begin{pmatrix} 1 & \sigma_u^2 & \frac{d_2}{1-\gamma} & 0 & -\frac{d_4}{\alpha+\beta-1} \\ 0 & 1 & \frac{2\psi(\alpha+\beta)}{\alpha+\beta-\gamma} & \frac{1}{\alpha+\beta-1} & \frac{-\sigma_u^4}{\psi} \\ 0 & 0 & 1 & 0 & -\frac{\kappa\alpha+\beta}{\alpha+\beta-\gamma} \\ 0 & 0 & 0 & 1 & \sigma_u^2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

[1] For details on Jordan decompisition see e.g. [41]

[2] To get the full Jordan decomposition one would need to follow with swap of lines and rows and a simple division to get the element over diagonal equal to one

with $d_2 = \psi^2 + 2\psi(\alpha+\beta)\sigma_u^2$, $d_3 = \psi - \sigma_u^2 + \frac{2\psi(\alpha+\beta)(\kappa\alpha+\beta)}{\alpha+\beta-\gamma}$ and $d_4 = \sigma_u^2(\psi - \sigma_u^2) + \frac{d_2(\kappa\alpha+\beta)}{1-\gamma}$.
From J we can compute J^i as

$$J^i = \begin{pmatrix} 1 & 0 & 0 & i\sigma_u^2 & 0 \\ 0 & (\alpha+\beta)^i & 0 & 0 & id_3(\alpha+\beta)^{i-1} \\ 0 & 0 & \delta^i & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & (\alpha+\beta)^i \end{pmatrix}$$

and also the sum

$$\sum_{i=0}^{m-1} J^i = \begin{pmatrix} m & 0 & 0 & \frac{1}{2}\sigma_u^2 m(m-1) & 0 \\ 0 & \frac{1-(\alpha+\beta)^m}{1-(\alpha+\beta)} & 0 & 0 & d_4(m) \\ 0 & 0 & \frac{1-\gamma^m}{1-\gamma} & 0 & 0 \\ 0 & 0 & 0 & m & 0 \\ 0 & 0 & 0 & 0 & \frac{1-(\alpha+\beta)^m}{1-(\alpha+\beta)} \end{pmatrix},$$

where $d_4(m) = \frac{d_3}{1-(\alpha+\beta)}\left(\frac{1-(\alpha+\beta)^{m-1}}{1-(\alpha+\beta)} - (m-1)(\alpha+\beta)^{m-1}\right)$.
Knowing the $\sum_{i=0}^{m-1} J^i$ we can calculate the limit as

$$\lim_{m\to\infty} k_{[m]} = \frac{1}{\sigma_u^2} B \left(\lim_{m\to\infty} \frac{1}{m^2} \sum_{i=0}^{m-1} A^i\right) b \tag{5.30}$$

$$= \frac{1}{\sigma_u^2} BS \left(\lim_{m\to\infty} \frac{1}{m^2} \sum_{i=0}^{m-1} J^i\right) S^{-1} b \tag{5.31}$$

$$= \frac{1}{\sigma_u^2} BS \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \eta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} S^{-1} b, \tag{5.32}$$

where

$$\eta = \lim_{m\to\infty} \sum_{i=0}^{m-1} \frac{\gamma^i}{m^2} = \begin{cases} 0 & \text{if } \gamma \le 1, \\ \infty & \text{if } \gamma > 1. \end{cases}$$

Multiplying the matrices in equation (5.32) we get

$$\lim_{m\to\infty} \kappa_{[m]} = 3, \quad \text{for } \gamma < 1. \tag{5.33}$$

For $\gamma > 1$ the η is always multiplied by a positive constant and therefore

$$\lim_{m\to\infty} \kappa_{[m]} = \infty, \quad \text{for } \gamma > 1. \tag{5.34}$$

In case $\gamma = 1$ or $\gamma = \alpha + \beta$ the proof follows essentially the same lines, but with a slightly different matrix J and matrix S, which leads to the additional term d in equation (5.25).

For $\gamma = \alpha + \beta$ we have

$$J_{\alpha+\beta} = \begin{pmatrix} 1 & 0 & 0 & \sigma_u^2 & 0 \\ 0 & \alpha+\beta & 2\psi(\alpha+\beta) & 0 & 0 \\ 0 & 0 & \alpha+\beta & 0 & \kappa\alpha+\beta \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \alpha+\beta \end{pmatrix}, \tag{5.35}$$

$$S_{\alpha+\beta} = \begin{pmatrix} 1 & \sigma_u^2 & \frac{d_2}{1-(\alpha+\beta)} & 0 & \frac{d_3}{1-(\alpha+\beta)} \\ 0 & 1 & 0 & -\sigma_u^2/\psi & -\sigma_u^4/\psi \\ 0 & 0 & 1 & 0 & \frac{\psi-\sigma_u^2}{2\psi(\alpha+\beta)} \\ 0 & 0 & 0 & 1 & \sigma_u^2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.36}$$

with $d_2 = \psi^2 + 2\psi(\alpha+\beta)\sigma_u^2$ and $d_3 = \sigma_u^2(\psi - \sigma_u^2) + \frac{d_2}{1-(\alpha+\beta)}(\kappa\alpha+\beta)$.

For $\gamma = 1$ we get

$$J_1 = \begin{pmatrix} 1 & 0 & d_2 & 0 & 0 \\ 0 & \alpha+\beta & 0 & 0 & d_3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \alpha+\beta \end{pmatrix}, \tag{5.37}$$

$$S_1 = \begin{pmatrix} 1 & \sigma_u^2 & 0 & 0 & \frac{d_4}{1-(\alpha+\beta)} \\ 0 & 1 & \frac{2\psi(\alpha+\beta)}{\alpha+\beta-1} & -\sigma_u^2/\psi & -\sigma_u^4/\psi \\ 0 & 0 & 1 & \frac{\sigma_u^2}{d_2} & -\frac{\kappa\alpha+\beta}{\alpha+\beta-1} + \frac{\sigma_u^4}{d_2} \\ 0 & 0 & 0 & 1 & \sigma_u^2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.38}$$

with $d_2 = \psi^2 + 2\psi(\alpha+\beta)\sigma_u^2$, $d_3 = \psi - \sigma_u^2 + \frac{2\psi(\alpha+\beta)(\kappa\alpha+\beta)}{\alpha+\beta-1}$ and $d_4 = \sigma_u^2(\psi - \sigma_u^2) + d_2\frac{\kappa\alpha+\beta}{\alpha+\beta-1}$. $\square$

To analyse the aggregated conditional kurtosis in the long term limit we estimated the GARCH parameters $\psi, \alpha, \beta, \kappa$ of some time series at daily frequency for real data. For these series we used daily data from Bloomberg starting at the dates indicated below and ending 28 February, 2004. The source of the series Germany Euro-deposits 6 months is Datastream. This analysis is part of author's paper [23].

Equity Indices	
Dow Jones Industrial Average	03-Jan-1970
DAX	03-Jan-1970
Nikkei 225	06-Jan-1970
Austrian Traded Index	09-Jan-1986
FTSE 100	04-Jan-1984
Swiss Market Index	02-Jul-1988
Interest Rates	
US Govt. 3 months	02-Jun-1983
US Govt. 6 months	02-Jun-1983
US Govt. 2 years	01-Feb-1977
US Govt. 5 years	03-Jan-1970
US Govt. 10 years	03-Jan-1970
US Govt. 30 years	02-Dec-1980
Germany Euro-deposits 6 months	03-Jan-1975
Germany Govt. 10 years	04-Jan-1989
Japan Govt. 10 years	23-Oct-1987
Exchange Rates	
EUR/USD	05-Jan-1971
EUR/GBP	05-Jan-1971
EUR/CHF	05-Jan-1971
EUR/JPY	05-Jan-1971

The values of parameters α, β and κ estimated by a QMLE[3] and the resulting γ are shown in Table 5.1. Value of γ is an indicator of the long time limit of aggregate conditional kurtosis. For time series with $\gamma > 1$ the aggregated conditional kurtosis goes to infinity according Theorem 5.3. The second last column gives the values of γ together with its 80% confidence levels. These confidence levels were approximately determined by Monte Carlo simulations[4]. The error distribution of model parameters $\psi, \alpha, \beta, \kappa$ is asymptotically normal for MLE estimates. We drew 10.000 values of these parameters and from them calculated γ. The last column indicates the limit behavior of aggregate conditional kurtosis: c stands for

[3]For details on quasi-maximum likelihood estimator for GARCH processes see e.g. [69]

[4]For details on approximation of confidence intervals with Monte Carlo methods see e.g. [26]

Table 5.1: The GARCH parameters α, β, κ and the indicator γ of the long time limit of conditional kurtosis. For time series with $\gamma > 1$ the aggregated conditional kurtosis goes to infinity according Theorem 5.3. The second last column gives the values of γ together with the 80% confidence levels. The last column indicates the limit behavior of aggregate conditional kurtosis: c stands for convergence to the value of 3, d stands for divergence, u stands for undecided which is used if the confidence interval contains the value of 1. We observe that for 6 time series the conditional aggregate kurtosis diverges, for 5 time series it converges to three, and for 6 time series the limit behavior is undecided by this analysis.

	α	β	κ	γ	limit
DJI	0.05	0.94	8.1	1.0011 ± 0.0000	d
DAX	0.11	0.87	11.4	1.0168 ± 0.0033	d
Nikkei 225	0.10	0.89	5.3	1.0158 ± 0.0232	u
ATX	0.07	0.92	8.2	1.0872 ± 0.0851	d
FTSE 100	0.08	0.90	3.8	0.9904 ± 0.0014	c
Swiss Market Index	0.14	0.81	15.8	1.0089 ± 0.0329	u
USD/EUR via DEM	0.10	0.00	4.3	0.9903 ± 0.0125	u
GBP/EUR via DEM	0.09	0.87	4.7	0.9625 ± 0.0000	c
CHF/EUR via DEM	0.07	0.91	7.3	1.0140 ± 0.0045	d
JPY/EUR via DEM	0.07	0.92	4.4	0.9981 ± 0.0000	c
US Govt 3m	0.18	0.82	15.8	1.0574 ± 0.3083	u
US Govt 6m	0.08	0.92	11.6	0.9513 ± 0.0863	u
US Govt 2y	0.05	0.95	5.2	0.9776 ± 0.0344	u
US Govt 5y	0.05	0.95	4.4	1.0068 ± 0.0017	d
US Govt 10y	0.05	0.94	4.1	0.9983 ± 0.0000	c
US Govt 30y	0.03	0.96	5.1	0.9911 ± 0.0000	c
DEM Govt 10y	0.05	0.95	4.7	1.0105 ± 0.0000	d
JPY Govt 10y	0.10	0.90	6.7	1.0192 ± 0.0250	d

convergence to the value of 3, d stands for divergence, u stands for undecided which is used if the confidence interval contains the value of 1. We observe that for 7 time series the conditional aggregate kurtosis diverges, for 5 time series it converges to three, and for 6 time series the limit behavior is undecided by this analysis.

For four selected time series (DAX, USD/EUR, GBP/EUR, US Govt.10y) the long time behavior of 80% confidence intervals of the aggregated conditional kurtosis is illustrated in Figure 5.1. For the confidence intervals of Table 5.1 the finite time behavior of aggregated kurtosis was calculated from equation (5.13). While for e.g. the DAX $\gamma \in [1.0135, 1.0198]$ the 90% confidence interval are above 1, the limiting kurtosis will go to infinity. For US Govt 10y $\gamma \in [0.9983, 0.9983]$

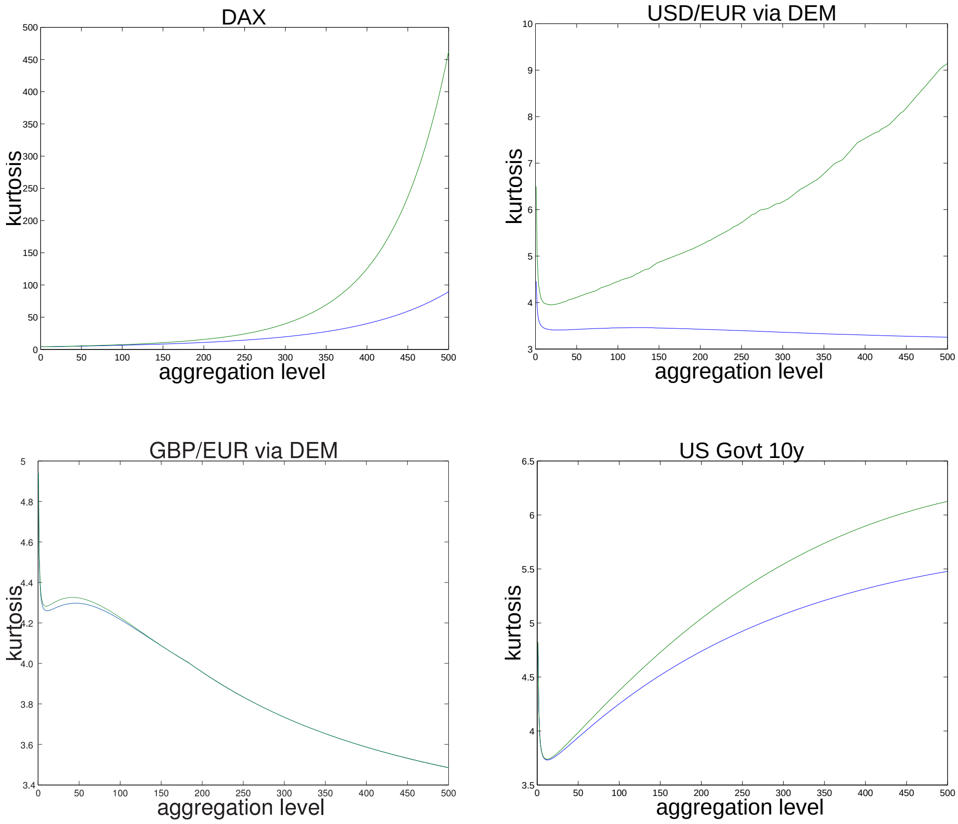


Figure 5.1: Finite long term behavior of 80% confidence intervals of the aggregated conditional kurtosis for four selected time series. For the γ confidence intervals of Table 5.1 the aggregated kurtosis for the upper and lower ends of the interval was calculated from equation 5.13.

and GBP/EUR $\gamma \in [0.9625, 0.9625]$ the aggregated conditional kurtosis will go to 3. For USD/EUR γ is in the interval$[0.9780, 1.0061]$ containing 1. The confidence level is widening as the number of aggregation steps increases, indicating that for this time series the long time behavior of conditional aggregated kurtosis is undecided on the basis of our analysis.

Figure 5.2 compares the conditional kurtosis as calculated with equation (5.13) from the γ values in Table 5.1 to the empirical kurtosis estimated directly from the time series, using Matlab's kurtosis function. While our time series are too short to estimate empirical kurtosis over longer horizons, the conditional kurtosis, calculated with the equations from Theorem 5.2, can be scaled to any aggregation level as shown in Figure 5.1.

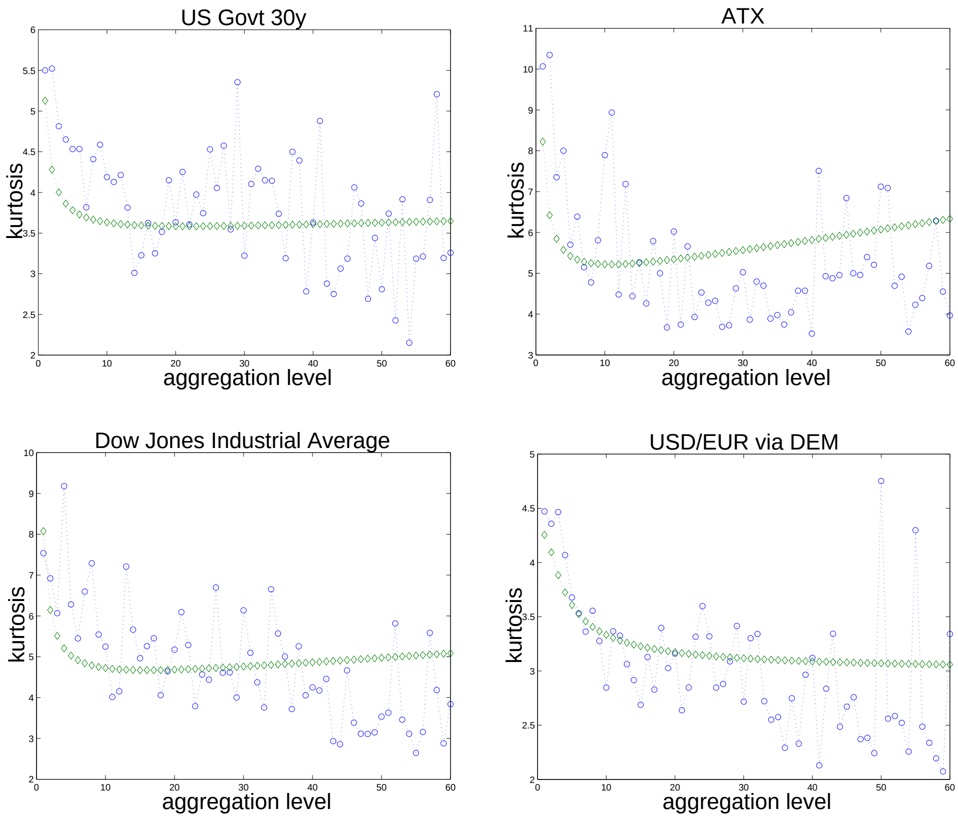


Figure 5.2: Conditional aggregated kurtosis in dependence of time horizon. Green diamonds represent the values of the conditional aggregated kurtosis, calculated with equation (5.13) from the estimated parameter values of the daily time series given in Table 5.1. Blue circles represent empirical kurtosis estimated directly from the time series.

5.3 The optimal frequency for strong GARCH models

Since strong GARCH processes do not aggregate to strong GARCH processes two questions arise: (1) at which frequency is the assumption of strong GARCH processes best justified?, and (2) When making forecasts over a long time horizon, should we use higher frequency data if available? In this Section we first give a Quasi Maximum Likelihood estimation procedure for the optimal frequency. Then we evaluate how well models of various frequencies predict densities over a time horizon of three months.

Estimation of the optimal frequency

In this subsection we will give a Quasi Maximum Likelihood procedure to estimate the optimal frequency of strong GARCH models.

For the estimation of GARCH parameters one often makes the restrictive assumption of having a strong GARCH process with some known type of error distribution and uses the Maximum Likelihood method. When comparing different data frequencies we cannot make the assumption that we have a strong GARCH process at all frequencies. We either have weak GARCH processes arising from the aggregation of some basic frequency strong GARCH process, as in Drost and Nijman [39], or some more general processes, as in Meddahi and Renault [78], or a continuous time GARCH process exhibiting GARCH behavior at all sampling frequencies, as in Drost and Werker [40]. In any case, if we do not have a strong GARCH process estimating the GARCH parameters with a Quasi Maximum Likelihood method introduces some errors. For the question whether or not these errors are neglible we refer to [39, 62, 78].

The basic idea of our Quasi Maximum Likelihood procedure to estimate the optimal data frequency is as follows. The estimated parameter values of a GARCH model depend on the data frequency, represented by the number m of aggregation steps from some given basic frequency. Therefore we have to estimate both, the GARCH parameters and the m which fits the data best. The optimal parameters $\theta = (m, \psi, \alpha, \beta)$ are those which maximize the log-likelihood function

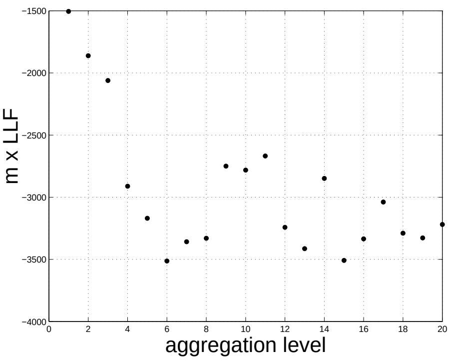
$$\begin{aligned} \theta^* &= \underset{\theta}{\operatorname{argmax}} \;\; m \cdot LLF(\theta) && (5.39) \\ &= \underset{m>0\ ,\psi,\alpha,\beta>0}{\operatorname{argmax}} m \sum_{j=1}^{n} \ln f_{\psi,\alpha,\beta}\left(\epsilon_t | I_{t-1}\right). && (5.40) \end{aligned}$$

The factor m corrects for the fact that for a lower frequency, where we aggregate m time steps of the basic frequency, we have m times less observations and thus fewer contributions to the LLF. We denote N to be total number of observation for the basic frequency and $n := \lfloor N/m \rfloor$ denotes the number of observation available at the aggregated frequency.

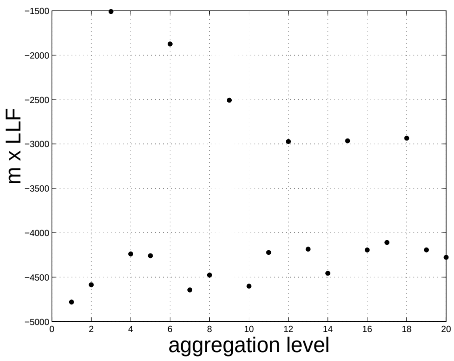
It is impossible to calculate the LLF for all possible values of m. We restrict ourselves to natural number m between 1 and 30. The m for which the maximum of m·LLF is achieved is our estimation of the optimal frequency:

$$\hat{m^*} = \arg\max_{1\leq m\leq 30} m \max_{\psi,\alpha,\beta>0} \ln \prod_{j=1}^{n} f_{\psi,\alpha,\beta}\left(\epsilon_t | I_{t-1}\right). \tag{5.41}$$

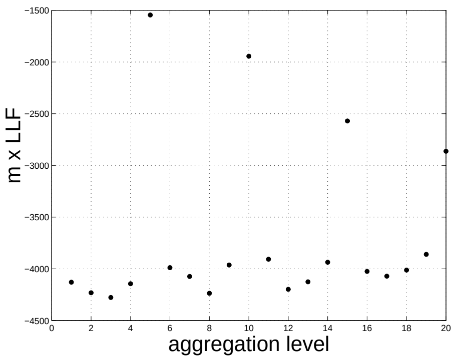
To examine the reliability of this estimation procedure we simulated three time series from a strong GARCH process with basic periods equal to one, three, and



(a) Basic period equals one day



(b) Basic period equals three days



(c) Basic period equals five days

Figure 5.3: Example of m·LLF as function of aggregation level for simulated GARCH processes, with true basic period equal to 1, 3, and 5 times the data period. We see that in all cases the estimated optimal period $\hat{m}^*$ is the true period.

five days. On the simulated time series we estimated the optimal frequency following the procedure above. The results are given in Fig. 5.3. We see that for all three time series the optimal LLF is achieved for the m^* corresponding to the true frequencies $m = 1, 3, 5$, which is reassuring. Still, we cannot exclude that the true basic frequency is different from $\hat{m}^*$, which is the one element of the set $\{1, 2, 3, \ldots, 30\}$ maximising the LLF. What happens if the true basic period is not a natural number? For example, if we have only weekly data (5 days), but the true frequency is equal to 3 days, this method cannot give the right result.

Typically the least common multiple of the sample frequency and the true frequency have the highest m·LLF value. The m·LLF of larger common multiples is somewhat smaller than the m·LLF of the least common multiple but larger than the m·LLF for other values of m. This is illustrated in two pictures of Figure 5.4, where we tested the estimation method on simulated data of a true frequency of 3 resp. 5 periods.

We can examine also the m·LLF for GARCH processes, where the true fre-

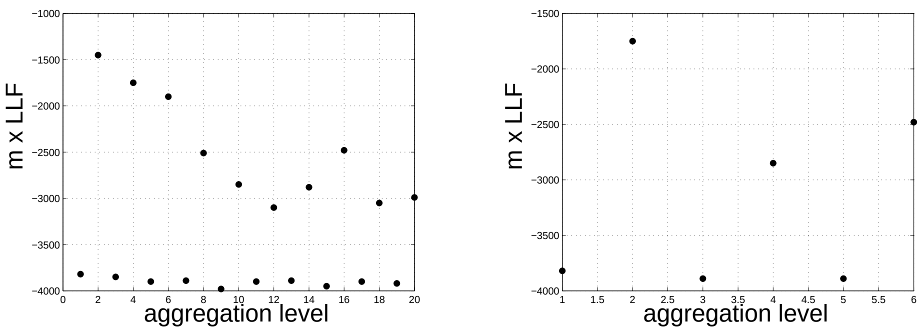


Figure 5.4: The m with the highest value of the m·LLF is not necessarily the true frequency. Left: The time series follows a GARCH process with true frequency equal to 2 data periods. The highest level of the m·LLF is achieved for $m = 2$, which is indeed the true frequency. Right: The time series follows a GARCH process with true frequency equal to 2/3 of a data period. The highest level of the m·LLF is achieved for $m = 2$, which is the least common multiple of the true frequency $2/3$ and the data frequency 1.

quency is higher than the data frequency. Assume we want to calculate the m·LLF for $m = 2/3$. The least common denominator of $m = 2/3$ and the data frequency 1 is 2. Aggregating the data at level 2 will yield the same series as aggregating the true process at level 3. From (5.6) we know the density function of returns aggregated at level 3. Entering this density for f_θ into (5.39) we can calculate the m·LLF of $m = 2/3$.

A serious disadvantage of this method to identify the optimal frequency is its computational complexity. Actually we can find the aggregated density function by numerically evaluating the integral (5.6). This has to be done n times, where n is the number of m-values we consider—e.g. 30 in equation (5.41)—times the number of iterations which the m·LLF maximisation algorithm needs.

Figure 5.5 shows (for $m = 1, 2, \ldots, 30$) the maximum over ψ, α, β of m·LLF for four time series: ATX, USD/EUR, DJI, US Govt 30y. We observe that for all four time series the optimal frequency determined by our Quasi Maximum Likelihood procedure is the basic frequency, $m^* = 1$. This suggests that the assumption of strong GARCH is better satisfied for daily data than for lower frequency data.

Statistical tests of long term density forecasts

A related albeit distinct question is the decision which frequency, when aggregated, leads to the best forecasts of long term return distributions. On the one hand, high frequency data allows for more reliable parameter estimates and the assumption of strong GARCH seems to be better justified for high frequency data. On the other hand, the aggregation procedure magnifies estimation and model errors. In this section we will perform statistical tests of the 60 day density fore-

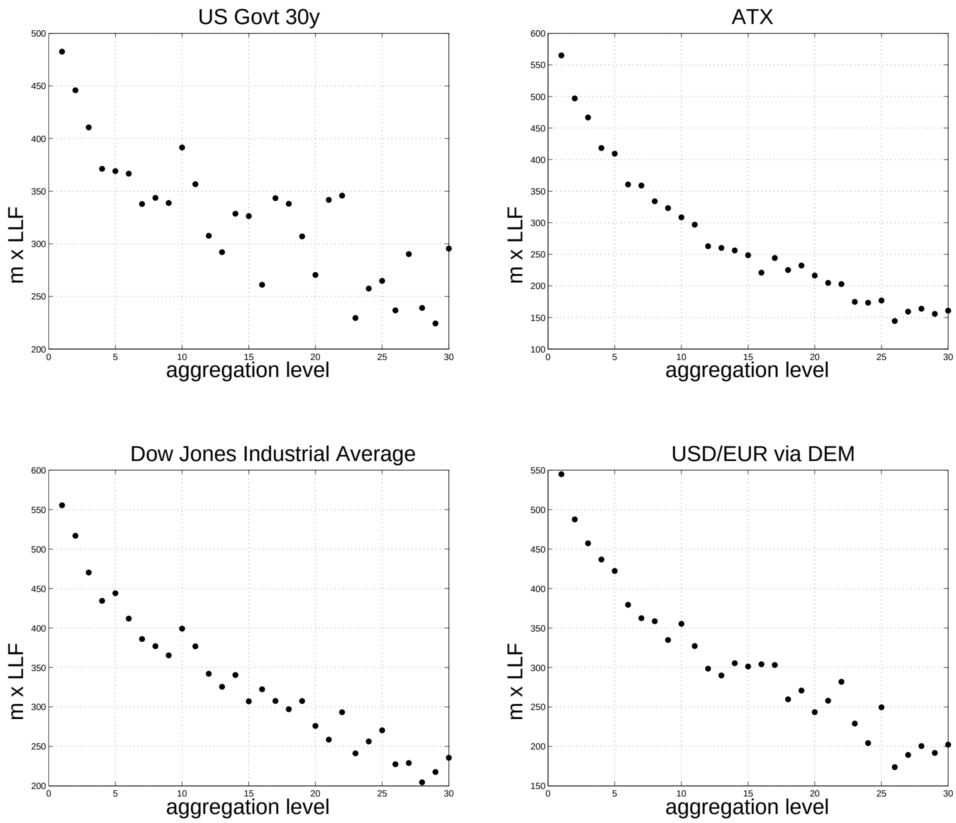


Figure 5.5: Estimating the optimal frequency of GARCH description of four market time series: USGovt30y, ATX, DJI, USD/EUR. The plots show $m{\cdot}$LLF as function of aggregation level m. The highest values of $m{\cdot}$LLF are always achieved for $m = 1$. This suggests that the assumption of strong GARCH is better satisfied for daily data than for lower frequency data.

casts produced by aggregating strong GARCH models of various basic periods with various error distributions.

As possible frequencies of strong GARCH models we consider 1 day, 5 days, 10 days, 20 days, 30 days, and 60 days. Aggregating these models over 60, 12, 6, 3, 2, resp. 1 period we get distributions of 60 day returns. In order to test the 60 days distribution forecasts produced by the various models, it is not enough to assess whether the means, variances, or some quantiles of the distributions were correctly predicted. For many applications, in particular in risk management, the overall distributional properties are important, not just the means or variances. Therefore, based on [88], we test for the adequacy of the density forecasts.

At the above frequencies we compare the following models. In order to account for possible excess kurtosis of the residuals $D(0, 1)$ of the strong GARCH models, we consider several possible residual distributions: normal, Student, and

EVT (extreme value theory). The EVT-distribution results from modelling the body of the distribution by historic simulation and the left and right tails by a Generalized Pareto distribution. For the left tails we took the lowest 10%, for the right tail the highest 10%. The details of the procedure are described in McNeil and Frey [77].

Consider a time series of returns r_t ($t = 1, \ldots, n$) generated from some true conditional densities $f_t(.)$ ($t = 1, \ldots, n$). Now some model produces a series of 60 days conditional density forecasts $p_t(.)$ ($t = 1, \ldots, n$). The task is to evaluate whether the true conditional densities $f_t(.)$ agree with the predicted conditional densities $p_t(.)$. Applying the Rosenblatt transformation [90] to the observed returns r_t,

$$r_t \mapsto z_t := \int_{-\infty}^{r_t} p_t(u)du \tag{5.42}$$

we get a transformed series z_t which should be i.i.d. U(0,1) if the predicted conditional densities $p_t(.)$ agree with the true conditional densities $f_t(.)$. Applying the inverse of the normal distribution function

$$z_t \mapsto n_t := N^{-1}(z_t), \tag{5.43}$$

produces a series n_t which is standard normally i.i.d. if the original returns r_t are distributed according to the predicted densities p_t (see [13]). There are myriad of tests for normality which could be applied to the $\{z_t\}$, see [18] and references therein.

Berkowitz [13] applied a likelihood-ratio test to the z_t against the first order autoregressive alternative $n_t - \mu = \rho_1(n_{t-1} - \mu) + \epsilon_t$ to test for i.i.d. N(0,1). Instead, we can perform a Kolmogorov-Smirnov test for the simple hypothesis that the n_t are sampled from a standard normal distribution. This is our Test 1. A model is accepted if the p-value is higher than 5%. It will turn out (see Table 5.2) that the Kolmogorv-Smirnov test is no very selective and accepted for many models all time series.

In order to test additionally whether the variance of the n_t is constant and equal to one, de Raaij and Raunig [88] consider the regressions

$$\begin{aligned} n_t &= \beta_0 + \beta_1 n_{t-1} + u_t && (5.44)\\ n_t^2 &= \gamma_0 + \gamma_1 n_{t-1}^2 + v_t && (5.45) \end{aligned}$$

where u_t and v_t are non-autocorrelated with zero expectation conditional on their own past values. In case the n_t have zero mean and are uncorrelated we have $\beta_0 = 0$ and $\beta_1 = 0$. In case the n_t have constant conditional unit variance we have $\gamma_0 = 1$ and $\gamma_1 = 0$. To test whether these restrictions are satisfied, de Raaij and Raunig [88] propose a joint Wald test of the four equalities $\beta_0 = 0$, $\beta_1 = 0$,

$\gamma_0 = 1$, and $\gamma_1 = 0$. This will be our Test 2. Additionally, they use the Jarque-Bera [54] test to see whether the n_t have skewness zero and kurtosis equal to three. This will be our Test 3. (The Jarque-Bera test by itself is not very powerful since it does not test for mean and variance.)

The fourth and fifth tests are variants of Pearson's χ^2-test. In Test 4, we split the interval $[0, 1]$ on which the random variable z_t is defined on five mutually distinct interval, while in Test 5 we use two adjacent observation of n_t to get bivariate observation and we apply the Pearson's chi-square test, with event described by the four quadrants. While Test 4 checks for the rough shape of the density, Test 5 is more sensitive on the dependence between two adjacent observation. The test statistic is defined as

$$\chi^2 = \sum_{i=1}^{m} \frac{(E_i - O_i)^2}{E_i}, \tag{5.46}$$

where O_i are observed frequencies corresponding to the $i-$th intervals and E_i are expected frequencies (based on the theoretical distribution).

The five tests outlined above were applied to the 19 market time series described in the previous section. In every test a model was accepted if the p-value of the given test was higher then 5%. Table 5.2 summarizes the test results. For each model, this table shows for how many of the 19 time series the model was accepted in Tests 1 to 5.

Summarizing results of this chapter

- We observe that results improve as the length of the basic period increases and the number of aggregation steps decreases. Aggregating models for high frequency data in general leads to worse results than discarding the high frequency data and estimating the models for 60 days returns only from 60 days data. This is in contrast to statements in the literature (e.g. [39, p. 922]) that in general it is preferable to estimate high frequency models if data are available and then aggregate to get long term models.

- We also observe that modelling the residual with Student t-distributions causes models to fail Tests 2, 3, and 5. This observation adds to the scepticism against the formerly popular use of Student t-distributions as residuals in strong GARCH processes. Allowing for fat-tailed residuals is not a good motivation. Leptokurtic distributions also result from the continuous time GARCH models of Drost and Werker [40]. Additionally, our results from Section 5.1 show that fat tails arise for the aggregated distribution even if the one-period model is not leptokurtic. Aggregation of some perhaps hidden high frequency strong GARCH model may account for the fat tails in the time series we observe.

Table 5.2: Summary of test results. For each model, this table shows for how many of the 19 time series the model was accepted in Test 1 to Test 5. In the model description N represents normal distribution of residual, while t denotes student distribution of residual and EVT represents residuals with EVT distribution. The first part of model description represents the basic time period (e.g. 5d_G_N represent GARCH model with normal residuals with 5 days sample frequency).

Model	Test 1 (KS) # accept.	Test 2 (JB) # accept.	Test 3 (W) # accept.	Test 4 (χ^2-1d) # accept.	Test 5 (χ^2-2d) # accept.
1d_G_N	19	0	0	0	0
1d_G_t	0	0	0	0	0
1d_G_EVT	19	0	0	0	0
2d_G_N	19	0	0	0	0
2d_G_t	0	0	0	0	0
2d_G_EVT	19	0	0	0	0
5d_G_N	19	0	0	0	0
5d_G_t	0	0	0	0	0
5d_G_EVT	19	0	0	0	0
10d_G_N	19	0	0	0	0
10d_G_t	0	0	0	0	0
10d_G_EVT	19	0	0	0	0
20d_G_N	19	0	0	0	17
20d_G_t	19	0	1	0	17
20d_G_EVT	19	0	0	0	16
30d_G_N	19	19	19	19	19
30d_G_t	19	0	0	19	0
30d_G_EVT	19	19	19	19	19
60d_G_N	19	19	19	19	19
60d_G_t	19	0	0	7	0
60d_G_EVT	19	19	8	16	19

- It is important to interpret our test results carefully. They do not imply that the assumptions of strong GARCH are best satisfied for the low frequency 60d models—which would contradict the results of the Maximum Likelihood estimates of the optimal frequency in the previous Subsection 5.3. Rather they show that the low frequency models which require fewer aggregation steps produce better 60d density forecasts. When it comes to producing long term density forecasts it seems that for high frequency models the advantage of having more data points available is outweighed by the disadvantage of estimation errors being magnified by the aggregation.

Chapter 6

Conclusions

In the first part, we formalized the portfolio optimisation problem under risk constraint on the infimum of wealth process. Our comparative analysis of expected return and expected shortfall (ES) with and without the ES constraint suggests the following conclusions: First, expected returns are reduced by less than one tenth when the ES constraint is introduced. In comparison, economic capital as measured by ES, is reduced to amounts between one half and three quarters when the ES constraint is introduced. Second, the dependence of the expected return and ES on the initial portfolio, in particular when transaction costs are high, is largely removed by introduction of the ES constraint. We analyzed how does the possibility to perform intermediate transaction affect risk measured over long time horizon. Our analysis shows that both expected return and risk as measure by ES are brought to some intermediate level when intermediate transaction are made possible subject to an ES constraint. Without the ES constraint, intermediate transaction aiming at maximising expected log-return lead to higher returns and higher risk.

In the second part (Chapter 4), we introduced a new model for pricing derivative securities in the presence of both transaction costs as well as the risk from unprotected portfolio. The risk which we introduced consists of two parts, the risk from not perfectly hedged portfolio and the risk of default of counterparty, in which case investor loss part of his investment. The derivation of the Risk adjusted pricing methodology (RAPM) model is a modification of original Kratka's approach. The option prices can be computed from a solution to a highly nonlinear parabolic PDE. The governing equation extends the classical Black-Scholes equation and Leland's equation to the case when the risk from unprotected portfolio is taken into account. We showed how this equation can be approximated by a stable numerical scheme. We performed extensive numerical testing of the model and compared the results to real option market data. We also introduced the concept of a so-called implied RAPM volatility and implied risk premium coefficient. We computed these implied quantities for large option data sets and we showed how they can be used in qualitative analysis of option market data. We introduced also the default possibility into the pricing models and analyzed when the inclusion of default possibility is required to achieve correct risk analysis.

Finally, in the third part of this thesis, we analyzed the problem of optimal aggregation frequency. For models with different basic frequency and with different residual distributions we perform out of sample tests of three months density forecasts on the basis of daily market prices. It turns out that low frequency models with longer basic periods and fewer aggregation steps perform fare better than high frequency models. This seems to imply that for high frequency models the advantage of having more data available for estimation is outweighed by the disadvantage of aggregation magnifying estimation errors. Contrary to some statements in the literature (see e.g. Drost and Nijman [39, p. 922]) producing long term forecasts from aggregating higher frequency models need not be better than using only low frequency data. We derived explicit expressions for the conditional volatility and kurtosis of the aggregated GARCH distribution and we derived the limit behavior of these conditional moments, when time horizon gets longer. Given that the aggregation of a strong GARCH process is not any more a strong GARCH process, the question arises for which data frequency a description by a strong GARCH process fits the data best. We proposed a quasi maximum likelihood method to determine the optimal data frequency for a GARCH description.

Our main results can be summarized as follows:

1. We proposed and formalized portfolio selection problem with risk constraint and made its numerical analysis.

2. We derived a option pricing model which contain the transaction costs, the risk from unprotected portfolio and the risk from default of a counterparty. We proposed a numerical scheme for this model and shown that it can explain the so called volatility smile.

3. We analyzed the problem of integrated vs. separated market and credit risk. We have also shown when the separated risk analysis does not underestimate the overall risk for coherent risk measure.

4. For models with different basic frequency and with different residual distributions we performed out of sample tests of three months density forecasts on the basis of daily market prices. and we derived the conditional volatility and kurtosis and its behavior when time horizon goes to infinity.

Bibliography

[1] Acerbi, C. (2002), *Spectral measures of risk: a coherent representation of subjective risk aversion*, Journal of Banking and Finance, Volume 26, Number 7, July 2002 , pp. 1505-1518(14)

[2] Ait-Sahalia, Y., Brandt, M. (2001), *Variable selection for porttfolio choice*, J Finance 56, 1297-1351

[3] Akian, M., Menaldi, J.L., Sulem, A. (1996), *On an investment-consumption model with transaction costs*, SIAM J Control Optim 34, 329-364

[4] Algower, E. L. and Georg, K. (1990), *Numerical continuation methods*, New York, Heidelberg, Berlin: Springer-Verlag.

[5] Andersen, T.G., Bollerslev, T., Christoffersen, P.F., Diebold, F.X. (2005), *Volatility forecasting*. Working paper 11188, National Bureau of Economic Research, Cambridge, MA, also available at:
http://www.nber.org/papers/w11188

[6] Artzner, P., Delbean, F., Ebner, J.M., Heath, D. (1999), *Coherent measures of risk*, Mathematical Finance 9, 203-228, Also available as:
http://www.math.ethz.ch/∼delbaen/ftp/preprints/CoherentMF.pdf

[7] Artzner, P., Delbaen, F., Eber, J.M., Heath, D., Ku, H. (2002), *Coherent multiperiod risk measurement*, Technical report, ETH Zurich, Also available as:
http://www.math.ethz.ch/ delbaen/ftp/preprints/MULTIPERIOD-3-02.pdf

[8] Avellaneda, M. and Paras, A. (1994), *Dynamic Hedging Portfolios for Derivative Securities in the Presence of Large Transaction Costs*, Applied Mathematical Finance 1, 165-193.

[9] Avellaneda, M. and Paras, A. (1995), *Pricing and hedging derivative securities in markets with uncertain volatilities*, Applied Mathematical Finance 2, 73-88.

[10] Baillie, R.T., Bollerslev, T. (1992), *Prediction in dynamic models with time-dependent conditional variance*, Journal of Econometrics 52, 91-113

[11] Barberis, N. (2000), *Investing for the long run when returns are predictable*, J Finance 55, 225-264

[12] Baxter, M. and Rennie, A. (1996) *Financial Calculus: An Introduction to Derivative Pricing*, Cambridge University Press.

[13] Berkowitz. J. (2001), *Testing density forecasts, with application to risk management*, Journal of Business And Economic Statistics 19, 465-474

[14] Bielecki, T.R., Pliska, S.R. (2000), *Risk sensitive asset management with transaction costs*, Finance Stochastics 4, 1-33

[15] Black, F. and Scholes, M. (1972), *The valuation of options contracts and a test of marker efficiency* , Journal of Finance 27, 399-417.

[16] Black, F. and Scholes, M. (1973), *The pricing of options and corporate liabilities*, J. Political Economy 81, 637-659.

[17] Bollerslev, T. (1986), *Generalized autoregressive conditional heteroscedasticity*, Journal of Econometrics 31, 307-327

[18] Bontemps, C., Meddahi, N. (2005), *Testing normality: A GMM approach*, Journal of Econometrics 124, 149-186

[19] Brandt, M., Goyal, A., Santa-Clara, P. (2005), *A simulation approach to dynamic portfolio choice with an application to learning predictability*, Rev Financ Stud 18, 831-873

[20] Brennan, M., Xia, Y. (2002), *Dynamic asset allocation under inflation*, J Finance 57, 1201-1238

[21] Brennan, M., Schwartz, E., Lagnado, R. (1997), *Strategic asset allocation*, J Econ Dyn Control 21, 1377-1403

[22] Breuer T., Jandačka M. (2008), *Portfolio Selection with Transaction Costs under Expected Shortfall Constraints*, Computational Management Science 5, 305-316

[23] Breuer, T., Jandačka, M. (Working paper), *Temporal Aggregation of GARCH Models: Conditional Kurtosis and Optimal Frequency*, Available at SSRN: http://ssrn.com/abstract=967824

[24] Breuer, T., Jandačka, M., Rheinberger, K., Summer, M. (2008), *Compounding effects between market and credit risk: The case of variable rate loans*, 371-384 in: A. Resti (ed.): The Second Pillar in Basel II and the Challenge of Economic Capital, Risk Books

[25] Breuer, T., Jandačka, M., Rheinberger, K., Summer, M. (2008), *Regulatory capital for market and credit risk integration: is current regulation always conservative?*, Discussion Paper, Series 2: Banking and financial Studies No 14/2008

[26] Buckland, S.T. (1984), *Monte Carlo Confidence Inervals*, Biometrics 40, 811-817

[27] Cadenillas, A. (2000), *Consumption-investment problems with transaction costs: survey and open problems*, Math Methods Oper Res 51, 43-68

[28] Cheridito, P., Delbaen, F., Kupper, M. (2004), *Coherent and convex risk measures for bounded cádlág processes*, Stochastic Proc Appl 112, 1-22

[29] Cheridito, P., Delbaen, F., Kupper, M. (2005), *Coherent and convex monetary risk measures for unbounded cádlág processes*, Finance Stochastic 9, 369-287

[30] Coling, A.S., Pliska, P.W. (1997), *Portfolio management with transaction costs*, Proc R Soc A 453, 551-562

[31] Constantinides, G.M. (1979), *Multiperiod consumption and investment behavior with convex transaction costs*, Manag Sci 25, 1127-1137

[32] Cox, J., Huang, C.F. (1989), *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, J Econ Theory 49, 33-83

[33] Cuoco, D. (1997), *Optimal consumption and equilibrium prices with portfolio constraints and stochastic income*, J Econ Theory 72, 33-83

[34] Cuoco, D., Liu, H. (2000), *A martingale characterization of consumption choice and hedging costs with margin requirements*, Math Finance 10, 355-385

[35] Cvitanic, J., Karatzas, I. (1992), *Convex Duality in constrained portfolio optimization*, Ann Appl Prob 2, 767-818

[36] Davis, M.H.A., Norman A.R. (1990), *Portfolio selection with transaction costs*, Math Oper Res 15, 676-713

[37] Delbaen, F. (2003), *Coherent risk measures on general probability spaces*. In: Sandmann, K., Schonbucher, P.J. (eds) Advences in stochastics and finance: essays in honour of Dieter Sondermann. Springer, Heidelberg, 1-37, also available as:
http://www.math.ethz.ch/ delbaen/ftp/preprints/RiskMeasuresGeneralSpaces.pdf

[38] Diebold, F.X. (1988), *Empirical Modelling of Exchange Rates*, Springer, New York

[39] Drost, F.C., Nijman, T.E. (1993), *Temporal aggregation of GARCH processes*, Econometrica 61, 909-927

[40] Drost, F.C., Werker, B.J.M (1996), *Closing the GARCH gap: Continuous time GARCH modeling*, Journal of Econometrics 74, 31-57

[41] Dunford, N., Schwartz, J.T. (1958), *Linear Operators, Part I: General Theory*, Interscience,

[42] Duque J. and Paxson D. (1994), *Implied Volatility and Dynamic Hedging*, The Review of Futures Markets 13, 381-421.

[43] Eymard, R., Gutnic, M. and Hilhorst, D. (1998), *The finite volume method for an elliptic-parabolic equation*, Acta Mathematica Univ. Comenianae 67, 181-195.

[44] Fleming, W.H., Sheu, S.J. (1999), *Optimal long term growth rate of expected utility of wealth*, Ann Appl Prob 9, 871-903

[45] Föllmer, H., Schied, A. (2004), *Stochastic finance: an introduction in discrete time*, 2nd edn. vol 27 of de Gruyter Studies in Mathematics, Walter de Gruyter

[46] Fouque, J.P., Papanicolaou, G., Sircar, K.R. (2000), *Derivatives in Markets with Stochastic Volatility*, Cambridge University Press 2000

[47] de Giorgi, E. (2002), *A note on portfolio selection under various risk measures*, Technical report, University of Lugano, also available as:
http://ssrn.com/abstract=762104

[48] Hakkansson, N.H. (1970), *Optimal investment and consumption strategies under risk for a class of utility functions*, Econometrica 38, 585-607

[49] Haugh, M.B., Kogan, L., Wang, J. (2006), *Evaluating portfolio policies: a duality approach*, Oper Res 54, 405-418

[50] He, H., Pearson, N. (1991), *Consumption and portfolio policies with incomplete markets and short sale constraints: the infinite-dimensional case*, J Econ Theory 52, 259-304

[51] Hull, J. (1989), *Options, Futures and Other Derivative Securities*, New York: Prentice Hall.

[52] Hundsdorfer, W., Verwer, J.H. (2003), *Numerical solution of time-dependent advection-diffusion-reaction equations*, Springer-Verag Berlin-Heidelberg-New York

[53] Jandačka, M., Ševčovič, D. (2005), *On the risk adjusted pricing methodology based valuation of vanilla options and explanation of the volatility smile*, J. Appl. Math. 3, 235-258

[54] Jarque, C.M., Bera, A.K. (1980), *Efficient tests for normality, homoscedasticity and serial independence of regression residuals*, Economics Letters 6, 255-259

[55] Johnson H. and Shanno D. (1987), *Option Pricing when the Variance is Changing*, Journal of Financial and Quantitative Analysis 22, 143-151.

[56] Judd, K. (1996), *Approximation, peturbation, and projection methods in economic analysis*, In: Ammann, H.M., Kendrick, D.A., Rust, J., (eds) Handbook of computational economics. vol 1, Elsevier Science, Amsterdam, 509-585

[57] Kabanov, Y., Klüppelberg, C. (2004), *A geometric approach to portfolio optimization in models with transaction costs*, Finance Stochastics 8, 207-227

[58] Kacur, J. (1985), *Method of Rothe in evolution equations*, Teubner-Texte zur Mathematik, B. G. Teubner Verlagsgesellschaft.

[59] Kacur, J. and Mikula, K. (1995), *Solution of nonlinear diffusion appearing in image smoothing and edge detection*, Appl. Numer. Math. 17, 47-59.

[60] Karatzas, I., Lehoczky, J., Sethi, S., Shreve, S.E. (1986), *Explicit solution of a general consumption/investment problem*, Math Oper Res 11, 261-294

[61] Kim, T.S., Omberg, E. (1996), *Dynamic nonmyopic portfolio behavior*. Rev Financ Stud 9, 141-161

[62] Komunjer, I. (2001), *Consistent estimation for aggregated GARCH processes*, Working Paper Series 2001-08, University of California San Diego, also available at:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=276672

[63] Korn, R. (1998), *Optimal Portfolios: Stochastic models for optimal investment and risk management in continuous time*, World Scientific, Singapore

[64] Kramkov, D., Schachermayer, W. (1999), *The asymptotic elasticity of utility functions and optimal investment in incomplete markets*, Ann Appl Prob 9, 904-950

[65] Kratka, M. (1998), *No Mystery Behind the Smile*, Risk 9, 67-71.

[66] Kwok, Y. K. (1998), *Mathematical Models of Financial Derivatives*, New York, Heidelberg, Berlin: Springer Verlag.

[67] Ladyzhenskaya, O.A., Solonnikov, V.A. and Ural'ceva, N. N. (1968), *Linear and Quasilinear Equations of Parabolic Type*, American Mathematical Society. Providence, Rhode Island.

[68] Lakner, P., Nygren, L.M. (2006), *Portfolio optimization with downside risk constraints*, Math Finance 16, 283-299

[69] Lee, S. W., Bruce, E.H. (1994), *Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator*, Econometric Theory 10, 29-52

[70] Leland, H. E. (1985), *Option pricing and replication with transaction costs*, Journal of Finance 40, 1283-1301.

[71] Liu, J. (2006), *Portfolio choice in stochastic environments*, Technical report, University of California at San Diego, also available at: http://management.ucsd.edu/pdf/portfolio.pdf

[72] Liu, J., Longstaff, F. (2004), *Losing money arbitrages: Optimal dynamic portfolio choice in markets with arbitrage opporunities*, Rev Finance Stud 17, 611-641

[73] Lynch, A. (2001), *Portfolio choice and equity characteristics: Characterizing the hedging demands induced by return predictability*, J Finance Econ 62, 67-130

[74] Magill, M.J.P., Constantinides G.M. (1976), *Portfolio selection with transaction costs*, Journal of Econ Theory 13, 245-263

[75] Markowitz, H.M. (1952), *Portfolio selection*, Journal of Finance 7, 77-91

[76] Markowitz, H.M. (1959), *Portfolio selection*, 1991 edn. Blackwell, Oxford

[77] McNeil, A.J., Frey, R. (2000), *Estimation of tail related risk measures for heteroscedastic financial time series: an extreme value approach*, Journal of Empirical Finance 7, 271-300, also available at: http://math.uni-leipzig.de/%tEfrey/evt-garch.pdf

[78] Meddahi, N., Renault, E. (2004), *Temporal aggregation of volatility models*, Journal of Econometrics 119, 355-379

[79] Merton, R.C. (1969), *Life time portfolio selection under uncertainty: the continuous-time model*, Rev Econ Stat LI, 247-257

[80] Merton, R.C. (1971), *Optimal consumption and portfolio rules in a continuous-time model*, J Econ Theory 3, 373-413 (Erratum in Journal of Economic Theory 6, 213-214, 1973)

[81] Morton, A.J., Pliska, S.A. (1995), *Optimal portfolio management with fixed transaction costs*, Math Finance 5, 337-356

[82] Munk, C. (2000), *Optimal consumption/investment policies with unciversifiable income risk and liquidity constraints*, J Econ Dyn Control 24, 1315-1343

[83] Muthuraman, K., Kumar, S. (2006), *Multidimensional portfolio optimization with proportional transaction costs*, Math Finance 16, 301-335

[84] Nijman, T.E., Sentana, E. (1996), *Marginalization and contemporaneous aggregation in multivariate garch processes*, Journal of Econometrics 71, 71-87

[85] Pena I., Rubio G. and Serna G. (2001), *Smiles, Bid-Ask Spreads and Option Pricing*, European Financial Management, forthcoming.

[86] Pliska, S.A. (1986), *A stochastic calculus model of continuous trading: optimal portfolios*, Math Oper Res 11, 371-382

[87] Protter, M. and Weinberger, H. F. (1984), *Maximum principles in differential equations*, New York, Heidelberg, Berlin: Springer-Verlag.

[88] Raaij, G., Raunig. B. (2002), *Evaluating density forecasts from models of stock market return*, Working Paper 59, Oesterreichische Nationalbank, also available at: http://www.oenb.at/de/img/wp59_tcm14-6147.pdf

[89] Rogers, L.C.G. (2003), *Duality in constrained optimal investment and consumption problems: a synthesis*, In: Bank, P., Baudoin, F., Föllmer, H., Rogers, L.C.G., Soner, M., Touzi, N., (eds) Paris-Princeton Lectures on

Mathematical Finance 2002. vol 1814 of Lecture Notes in Mathematics. Berlin, 95-131

[90] Rosenblatt, M. (1952), *Remarks on a multivariate transformation*. Annals of Mathematical Statistics 23, 470-472

[91] Rust, J. (1996), *Numerical dynamic programming in economics*, In: Ammann, H.M., Hendrick, D.A., Rust, J., (eds) Handbook of computational economics, vol 1, Elsevier, Amsterdam, 619-729

[92] Salopek, D. M. (1997), *American Put Options*, Pitman Monographs and Surveys in Pure and Applied Mathematics 84, Addison Wesley Longman Inc.

[93] Ševčovič, D. (2001), *Analysis of the free boundary for the pricing of an American call option*, Euro. Journal on Applied Mathematics 12, 25-37.

[94] Sharpe, W.F. (1964), *Capital asset prices: a theory of market equilibrium under conditions of risk*, Journal of Finance 19, 425-442

[95] Shreve, S.E., Soner, H.M. (1994), *Optimal investment and consumption with transaction costs*, Ann Appl Prob 4, 909-962

[96] Stamicar, R., Ševčovič, D. and Chadam, J. (1999), *The early exercise boundary for the American put near expiry: numerical approximation*, Canad. Appl. Math. Quarterly 7, 427-444.

[97] Taksar, M., Klass, M.J., Assaf, D. (1988), *A diffusion model for optimal portfolio selection in the presence of brokage fees*, Math Oper Res 13, 277-294

[98] Tobin, J. (1958), *Liquidity preference as behavior towards risk*, Rev Econ Stud 25, 65-86

[99] Tompkins, R. G. (2001), *Implied volatility surfaces: uncovering regularities for options on financial futures*, The European Journal of Finance 7, 198-230.

[100] Weber, S. (2006), *Distribution-invariant risk measures, information, and dynamic consistency*, Math Finance 16, 419-442

[101] Wilmott, P., Dewynne, J. and Howison, S. (1995), *Option Pricing: Mathematical Models and Computation*, UK: Oxford Financial Press.

[102] Xia, Y. (2001), *Learning about predictability: The effect of parameter uncertainty on optimal dynamic asset allocation*, J Finance 56, 205-247

List of symbols

The following symbols are used in Chapter **Portfolio selection with transaction costs under risk constraints**

r	risk free interest rate
B	risk free zero coupon bonds
	$dB(t) = rBdt$
m	number of stocks
μ	mean of log-normal process, which follow the stock prices
$\sigma'\sigma$	covariance of log-normal process, which follow the stock prices
W	m-dimensional Wiener process
$S_i(t)$	evolution of $i-$th stock price
	$dS_i(t) = S_i(t)\left(\mu_i dt + \sum_{j=1}^{m} \sigma_{ij} dW_j(t)\right)$
x_i	initial investment to bond ($i = 0$) and stocks ($i = 1, ..., m$)
π	portfolio selection strategy
$Z_i(t)$	cumulative amount of money used to buy stock i
$U_i(t)$	cumulative amount of money obtained from selling stock i
C_{bi}	proportional transaction costs for buying $i-$th stock
C_{si}	proportional transaction costs for selling $i-$th stock
$V_i^{\pi}(t)$	value invested in bond ($i = 0$) and stocks ($i = 1, ..., m$) in portfolio
p	total market value of portfolio
P	total market value of portfolio after transfer of all stocks to bond
ρ_0	one period risk measure
ρ	multiperiod risk measure
	$\rho(X) = \rho_0\left(\inf_{t\in[0,\infty]} X_t\right)$
$[A_i, B_i]$	no transaction region for stock i

The following symbols are used in Chapter **Risk adjusted pricing methdology**

r	risk free interest rate
B	risk free zero coupon bonds $dB = rBdt$
μ	drift of stock price
σ	volatility of stock price
W	Wiener process
$S(t)$	stock price $dS = \mu Sdt + \sigma SdW$
Π	synthetic portfolio
α	number of bonds in synthetic portfolio
δ	number of stocks in synthetic portfolio
V	price of option
T	expiration time
X	expiration price
r_{TC}	transaction premium for proportional transaction costs (RAPM)
r_{VP}	risk premium from volatile portfolio
r_R	total risk premium: $r_R = r_{TC} + r_V$
D	continuous dividend
ϕ	$\phi \sim N(0,1)$, $\phi\sqrt{dt} = dW$
ρ	default intensity
R_R	recovery rate
C	proportional transaction costs for buying or selling stock
R	risk aversion coefficient
Γ	$\frac{\partial^2 V}{\partial S^2}$
$\triangle t_{opt}$	optimal hedging time
t_υ	swithing time
τ	$\tau = T - t$
υ	$\upsilon = 3\left(\frac{C^2 R}{2\pi}\right)^{\frac{1}{3}}$
H	$H = S\Gamma$
$\beta(H)$	$\beta(H) = \frac{1}{2}\sigma^2(1 - H^{1/3})H$
$N(x)$	cumulative distribution function of normal distribution

The following symbols are used in Chapter **Temporal aggregation of GARCH models**

ϵ_t	a time serie
$\epsilon_{(m)t}$	$\epsilon_{(m)t} = \epsilon_{t+m}$ aggregation of a stock random variable
$\epsilon_{[m]t}$	$\epsilon_{[m]t} = \sum_{i=1}^{m} \epsilon_{t+i}$ aggregation of a flow random variable
m	time horizon of aggregation
ψ, α, β	parameters of GARCH process
h_t	the volatility process $h_t = \psi + \alpha\epsilon_t^2 + \beta h_{t-1}$, where $\alpha + \beta < 1$ and $\psi, \alpha, \beta > 0$.
ξ_t	innovations
I_t	information available at time t
κ	conditional kurtosis of the innovations
σ_u^2	unconditional variance of the one period returns $\sigma_u^2 := \psi/(1 - \alpha - \beta)$
γ	$\gamma = \alpha^2\kappa + \beta^2 + 2\alpha\beta$
$N(x)$	cumulative distribution function of normal distribution
A	$A = \begin{pmatrix} 1 & \psi & \psi^2 & 0 & 0 \\ 0 & \alpha+\beta & 2\psi(\alpha+\beta) & 1 & \psi \\ 0 & 0 & \kappa\alpha^2 + 2\alpha\beta + \beta^2 & 0 & \kappa\alpha+\beta \\ 0 & 0 & 0 & 1 & \psi \\ 0 & 0 & 0 & 0 & \alpha+\beta \end{pmatrix}.$
B	$B = (1, h_{t+1}, h_{t+1}^2, 0, 0)$
b	$b' = (0, 0, \kappa, 0, 6)$