Conditional Asset Pricing with Higher Moments

Diplomarbeit

zur Erlangung des Grades eines Diplom-Volkswirtes
an der Wirtschaftswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

vorgelegt von

Volker Ziemann
(Matrikel-Nr.: 157147)

Prüfer: Prof. Harald Uhlig, Ph.D.

Berlin, 14. Oktober 2004

Abstract

The Capital Asset Pricing Model (CAPM) is, despite its simplicity, still the most cited model in financial theory. However, empirically, it yields rather poor results in explaining cross-sectional returns of assets. The most successful extensions of the standard model often add factors that lack simple interpretations in terms of risk. In this study, instead of developing more factors, we will develop in detail the analysis (of dependencies of asset returns) with a single factor. We will concentrate on two main criticisms on the CAPM, questioning the hypothesis of normal distributed asset returns and the single-period character of the standard model. In the first step, we extend the model by taking higher moments into account, leading to risk premia for co-skewness and co-kurtosis. Second, we will allow these factors to vary over time in an autoregressive heteroscedastic context, using the new four-moment bivariate GARCH-in-mean model. This will lead us to the time-conditional four-moment Capital Asset Pricing Model.

Contents

Introduction

A substantial part of today's research in finance is directed towards explaining the cross-section of expected returns. Concerning the understanding of how investors value risky assets, it is commonly agreed that investors demand a higher expected return for investment in riskier projects. The Capital Asset Pricing Model (CAPM), developed by Sharpe (1964) and Lintner (1965) is still the most widely used approach to relative asset evaluation, in spite of its weak empirical success. In its simplest form this theory predicts that the expected returns on an asset above the risk free rate is proportional to the non-diversifiable risk, which is measured by the covariance of the asset return with a portfolio composed of all the existing assets, called the market portfolio. Based on the mean-variance description developed by Markowitz (1952) in the framework of his portfolio theory, the CAPM fundamentally relies on a Gaussian asset return distribution. However, it has long been recognized that financial asset returns are non-normal. Hence, it is appropriate to stress the importance of taking more moments into account. While normal distributions are entirely described by the first two moments (mean and variance), asymmetric and fat-tailed distribution are not. These two phenomena demand to consider moments of higher order. A number of studies have proven that the comparison of direct expected utility and utility approximated by the first two moments, leads to very small differences in terms of optimal portfolio allocation (Simaan (1993), Chamberlain (1983)). An explanation of the good performance of the mean-variance criterion in these papers may be that, although asset returns are non-normal they are at least elliptically distributed. Chamberlain (1983) showed that in this case the mean-variance approximation is equivalent to direct utility maximization.

However, recent research has shown that asset returns tend to deviate from elliptical distributions, stressing the importance to consider higher moments. Moreover, behavior studies proved that investors reveal a preference for skewness (third moment) and are averse towards kurtosis (fourth moment). This stylized fact coincides with the findings of the usual utility functions (Hwang and Satchell (1999)). Empirical results have proven that the corresponding third and fourth moment (skewness and kurtosis) of asset distributions significantly differ from those of the normal distribution (Kraus and Litzenberger (1976), Campbell and Siddique (1999), Hwang and Satchell (1999), Fang and Lai (1997)).

Consequently, this work will highlight the recent research in this direction and propose an extension of the traditional CAPM leading to a Four-Moment-Capital-Asset-Pricing-Model, referring to Fang and Lai (1997), Hwang and Satchell (1999), Rohan and Chaudhry (2000) and Kraus and Litzenberger (1976).

During the last decades, a second track has often been followed in order to improve the standard CAPM. The hypothesis that the risk associated with an asset does not vary over time, seems to be inappropriate. The traditional Capital Asset Pricing Model is built on the hypothesis that the investor lives only one period and optimizes her wealth at the end of this period. However, in the real world the investment horizon consists of many periods. Applying the CAPM in this context implies constant risk parameters which seems to be over-simplified. By now, it has long been recognized that there is no doubt about the empirical evidence for time-varying risk premia of financial assets. This stylized fact was first attributed to the time-varying behavior of conditional covariances (see Engle et al. (1987) or Bollerslev et al. (1988)). A common way in order to model time-conditional moments is to consider an autoregressive moving average relations. Engle (1982) first introduced the Autoregressive Conditional Heteroscedasticity (ARCH) model followed by Bollerslev (1986) with the Generalized ARCH (GARCH) aiming to parameterize the conditional mean and the conditional covariance of financial time series. These models are often built on the assumption that error terms are normally distributed, and the parameters are estimated using the conditional normal log-likelihood framework. However, as already mentioned, kurtosis and skewness play an essential role in asset distributions.

This work will attempt to incorporate the need of taking higher moments into account with the stated time-conditional context. Leon et al. (2002) and Campbell and Siddique (1999) contributed to this area of research through the consideration of univariate time-series. In both approaches, asymmetric and fat-tailed distributions have been applied in order to parameterize the conditional moments. We will extend these models to arrive at a bivariate model, which is composed of one asset and the market portfolio. We will introduce this model as the new Four-Moment Time-Conditional Capital Asset Pricing Model using the bivariate GARCH approach with time-conditional skewness and kurtosis (GARCHSK). We will use the Gram-Charlier expansions of the normal distribution in order to model the higher moments explicitly in the likelihood estimation. In doing so, we need to introduce tensor methods and hermite polynomials that are often applied in physics and mathematics, although these methods are yet to be widely used in financial or economic theory.

The traditional CAPM is an intuitive and rather simple theory, but yields rather poor empirical results. The most successful extensions in the sense of empirical results often add factors that lack simple interpretations in terms of risk (cf. the empirical nature of the Fama French factors). Instead of developing new factors, we will develop the "detail" in the analysis of dependencies between asset returns with a single factor, taking higher moments into account and letting the betas vary over time.

The rest of this study will be organized as follows: In the first chapter we give a brief review on asset pricing theory. We revisit the CAPM and its shortcomings as well as existing extensions. The derivation of the four-moment CAPM will be discussed in the second chapter. In the third chapter we will transform this model to a time-conditional one, using the bivariate GARCH-in-mean model extended with skewness and kurtosis.

Chapter 1

A Review of the CAPM

1.1 Mean-Variance Portfolio Choice

In his seminal work, Harry Markowitz (1952) proposed a model that combined the preference of high expected investment returns and the aversion against the risk. His model considers the portfolio as a whole and does not focus on securities on an individual basis. In order to measure the market risk of a portfolio, he used the variance (volatility) of the underlying portfolio returns. Markowitz defined the notion *efficient portfolio*. This approach can be understood in two different ways. First, a portfolio is considered as *efficient* when it minimizes the volatility for a given expected return. Second, a portfolio that maximizes the expected return for a fixed level of market risk (volatility) is called *efficient*. Considering the universe of all feasible portfolios these two definitions describe a curve in the expected-return-variance space, known as the *efficient frontier*. Due to its two objectives this approach is called *Mean-Variance-Analysis*.

We will now formalize this model. The first two moments of the portfolio can be written as follows:

$$
\begin{aligned}
\mathrm{E}(R_P) &= \sum_{i=1}^{N} \omega_i R_i \\
\mathrm{Var}(R_P) &= \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j \sigma_{ij}
\end{aligned}
$$

where R_P defines the portfolio return, R_i the return on asset i, σ_{ij} the covariance between asset i and asset j, ω_i the proportion of asset i held in the portfolio and N the number of disposable assets. As already stated above the standard assumption is that the agent prefers a higher expect return and a lower variance. Hence, this investor only takes those two parameters into account, i.e. the first two moments of the distribution of the portfolio return. In this approach the only income of investors is the outcome of their portfolio. These agents construct their portfolio in order to maximize their wealth at the end of the period. Thus, the underlying utility function is an increasing function of the expected return. In addition, this function is assumed to be concave since the investor is risk averse. This mean-variance analysis is fully consistent with expected utility maximization if the expected returns are normally distributed, an assumption that is respected over a short time horizon. This assumption may not hold over a longer period. In this case, one can consider the risk of loss (shortfall risk), which only considers negative deviations from a target return (e.g. semi-variance proposed by Markowitz (1959)).

As described above, the efficient frontier can be obtained by using either a maximization-problem (expected returns) or a minimization-problem (variance). The simplest way to determine this line can be expressed in the following matrix-form:

$$\min_{\omega} \frac{1}{2}\omega'\Sigma\omega \tag{1.1}$$

$$\begin{aligned} \text{s.t.}: \quad \mathbf{1}'\omega &= 1 \\ R'\omega &= \mu \end{aligned}$$

with Σ the variance-covariance matrix of returns, R the vector of expected returns and ω the vector of asset weights in the portfolio. As long as there are no further constraints (e.g. $\omega_i \geq 0$) this problem can easily be solved by using the Lagrange multiplier method (Merton (1973)). We obtain a hyperbolic relation between the expected return and the portfolio's variance. Only the upper part of this hyperbola defines the efficient frontier. This relation is shown in figure 1.1.

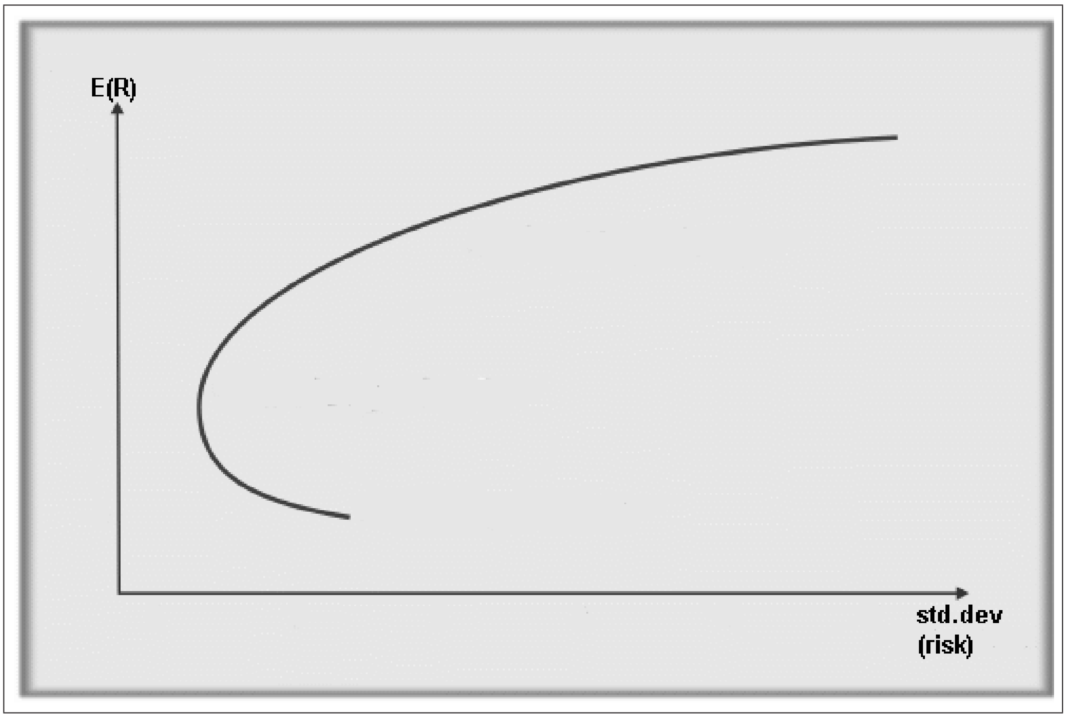

Figure 1.1: Efficient frontier hyperbola

We can define more elaborate models by adding linear constraints, in particular non-negative weights, which corresponds to restricted short sales. Elton and Gruber (1995) present the different cases in detail.

The choice of a particular point on the efficient frontier depends on the specific utility function of each investor.

We should mention that all the points on the efficient frontier can be obtained by linear combination of two efficient portfolios. This result is known as Black's theorem and has interesting applications in the domain of mutual funds: an agent who wants to hold a specific efficient portfolio can achieve this by combining two efficient mutual funds.

1.2 The Traditional Capital Asset Pricing Model

In the previous section, we introduced the Markowitz model, also known as the mean-variance-approach. We considered a portfolio from the point of view of one investor. However, this model cannot explain the observed asset allocations, since it is only a normative analysis. It postulates that the chosen portfolio should represent a point on the efficient frontier, but does not say which one.

The next step is to consider the universe of investors as a whole, which leads us to the deduction of the *Capital Asset Pricing Model*(CAPM). The idea is to derive a theory of asset valuation in an equilibrium situation, drawing together expected returns and market risk. This model is considered as the first to explain asset valuation by using the notion of risk. We will see that this risk can be split into a systematic risk, common to all assets in the same market segment and an unsystematic risk attributed to the specific asset.

The CAPM is built on the following main assumptions:

1. Investors are risk averse and maximize the expected utility of wealth at the end of the period
2. The asset returns are normally distributed or the investor only considers the first two moments of their return distribution (e.g. quadratic utility function)
3. Investors only consider one investment period which is the same for all investors
4. Investors have limitless access to financial markets and can borrow and lend at a risk-free rate R_F
5. Markets are complete (perfect information) and perfect (no taxes, no transaction costs)

In the Markowitz world we considered an investor who acted in isolation and only possessed risky assets. We will now introduce a risk-free asset and consider the market equilibrium.

1.2.1 Portfolio Choice with a risk-free Asset

The introduction of a risk-free asset, whose return is denoted R_F, enables the investor to spread her wealth between an efficient portfolio and this risk-free asset, which leads us to the following equation, where R_E denotes the return of the chosen efficient portfolio and R_P the return of the resulting portfolio composed of the risky and the risk-less asset:

$$\mathrm{E}(R_P) = xR_F + (1-x)\mathrm{E}(R_E) \tag{1.2}$$

with x the proportion of wealth invested in the risk-free asset. The portfolio's risk is than simply given by:

$$\sigma_P = (1 - x)\sigma_E \tag{1.3}$$

σ denoting the respective standard deviation of the portfolio returns. Combining (1.2) and (1.3) leads us to the following expression:

$$\mathrm{E}(R_P) = R_F + \left(\frac{\mathrm{E}(R_E) - R_F}{\sigma_E}\right)\sigma_P. \tag{1.4}$$

For each efficient portfolio we obtain a line representing all linear combinations of this efficient portfolio and the risk-free asset. Among this set of lines there is one that dominates all others.

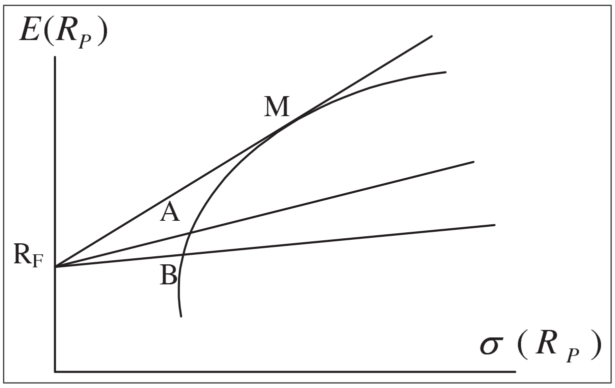

Figure 1.2: Efficient frontier in the presence of a risk-free asset

This line corresponds to the point M. Thus, we can consider the efficient frontier in a world with a risk-free asset as a straight line from the point R_F to M (see figure 1.2). This line is also called *capital market line*. As we can easily see in this figure, investors benefit from the risk-free opportunity, since for a given expected return the risk on this new efficient frontier is less than or equal to the risk of the corresponding portfolio consisting of risky assets. In other words, the efficient frontier with a risk-free asset dominates the efficient frontier without a risk-free asset.

As in the Markowitz Model the choice of a specific point on this line depends on the utility function and more precisely on the level of risk-aversion of the investor. If the investor has unlimited access to an efficient financial market, i.e. he can borrow money to a rate R_F, the efficient frontier to the right of point M corresponds to the extension of the line between R_F and M. If the investor is constrained to borrow to a rate $\widetilde{R_F} > R_F$ the efficient frontier is flatter to the right of point M. If the investor has no access at all to financial markets the efficient frontier with a risk-free asset corresponds to the efficient frontier without a risk-free asset for $\mathrm{E}(R_P) > \mathrm{E}(R_M)$. These results do not only depend on the respective financial market and its accessibility but also on the assumption that all assets are infinitely divisible.

The following derivation of the CAPM assumes an efficient financial market ($\widetilde{R_F} = R_F$) and infinite divisibility of assets. We call markets that fulfil these conditions as frictionless.

Figure 1.2 and equation (1.2) show that the portfolio decision problem can be divided into two parts: first, the investor chooses the optimal risky portfolio corresponding to one point on the efficient frontier without a risk-free asset and second, the choice of the split between the risk-free asset and the risky portfolio. This is known as Tobin's (1958) two-fund separation theorem.

1.2.2 Equilibrium Analysis

Until now, we have only considered one isolated investor. Subsequently, taking all investors into account, we move on to the market equilibrium. Since every investor is supposed to hold a mean-variance portfolio and we assume homogenous expectations of moments, all agents optimize the same program and derive the same efficient frontier, since the one passing through M dominates all the others. Depending on their level of risk aversion they invest a certain proportion in the risk-free asset and the rest in the portfolio M. In equilibrium, all assets are held and since the only traded risky portfolio is portfolio M it must contain all assets. Hence, this portfolio is the market portfolio and it holds all the assets in proportions of their market capitalization.

The Capital Asset Pricing Model aims to value each asset by considering an equilibrium situation. By applying the two-fund theorem we only have to consider the risky part of the portfolio in order to price each asset since the two decisions are independent. Similar to (1.2) and (1.3) we define:

$$\mathrm{E}(R_P) = xR_i + (1-x)\mathrm{E}(R_M) \tag{1.5}$$

$$\sigma_P = \left[x^2\sigma_i + (1-x)^2\sigma_M^2 + 2x(1-x)\sigma_{iM}\right]^{1/2}. \tag{1.6}$$

By varying x we can derive all possible efficient portfolios consisting of the risky asset i and the market portfolio in the $\mathrm{E}(R_P)-\sigma_P-$space (cf. figure 1.3).

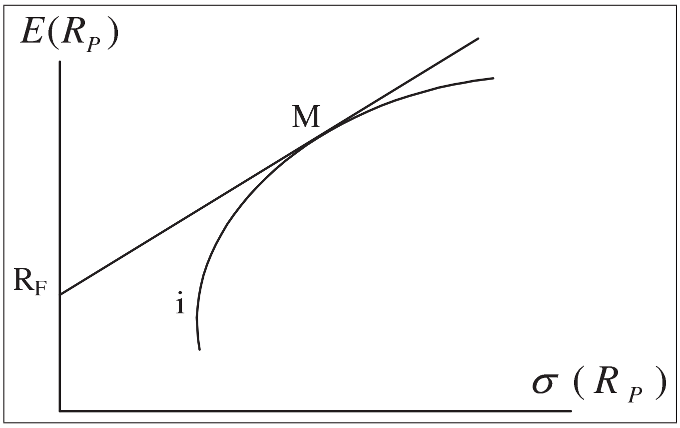

Figure 1.3: efficient frontier combining the market portfolio and a risky asset

In the optimum we can see that the leading coefficient of the tangent to this efficient frontier must be equal to the slope coefficient of the market line:

$$\frac{\partial \mathrm{E}(R_P)}{\partial \sigma_P} = \frac{\mathrm{E}(R_M) - R_F}{\sigma_M}. \tag{1.7}$$

Considering the functional relations above we have:

$$\frac{\partial \mathrm{E}(R_P)}{\partial \sigma_P} = \frac{\partial \mathrm{E}(R_P)}{\partial x}\frac{\partial x}{\partial \sigma_P}. \tag{1.8}$$

The two derivatives are:

$$\frac{\partial \mathrm{E}(R_P)}{\partial x} = \mathrm{E}(R_i) - \mathrm{E}(R_M) \tag{1.9}$$

$$\frac{\partial \sigma_P}{\partial x} = \frac{2x\sigma_i - 2(1-x)\sigma_M^2 + 2(1-2x)\sigma_{iM}}{2\sigma_P} \tag{1.10}$$

which gives us:

$$\frac{\partial \mathrm{E}(R_P)}{\partial \sigma_P} = \frac{(\mathrm{E}(R_i) - \mathrm{E}(R_M))\sigma_P}{x(\sigma_i^2 + \sigma_M^2 - 2\sigma_{iM}) + \sigma_{iM} - \sigma_M^2}. \tag{1.11}$$

In equilibrium the market portfolio contains all assets. The proportion x is therefore an excess in asset i in the portfolio P that must be zero at equilibrium (considering all investors). The Portfolio P is then the market portfolio ($\sigma_P = \sigma_M$) and we obtain for the point M:

$$\frac{\partial \mathrm{E}(R_P)}{\partial \sigma_P}(M) = \frac{(\mathrm{E}(R_i) - \mathrm{E}(R_M))\sigma_M}{\sigma_{iM} - \sigma_M^2} \tag{1.12}$$

From (1.7) we derive:

$$\frac{(\mathrm{E}(R_i) - \mathrm{E}(R_M))\sigma_M}{\sigma_{iM} - \sigma_M^2} = \frac{\mathrm{E}(R_M) - R_F}{\sigma_M}. \tag{1.13}$$

That last expression can also be written as

$$\mathrm{E}(R_i) = R_F + \frac{\mathrm{E}(R_M) - R_F}{\sigma_M^2}\sigma_{iM}. \tag{1.14}$$

Defining

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2} \tag{1.15}$$

we obtain the characteristic relationship of the CAPM:

$$\mathrm{E}(R_i) = R_F + \beta_i(\mathrm{E}(R_M) - R_F). \tag{1.16}$$

This equation can be understood as follows: The expected return of the risky asset i equals the return of the risk-free asset plus a risk premium. β_i is also called the systematic risk of asset i. By consequence, the risk-free asset has a beta of zero and the market portfolio a beta of one.

The CAPM establishes a theory for valuing individual securities and highlights the importance of taking risk into account. It states that there are two kinds of risk. First, we have the systematic risk, common to all assets, which is rewarded by the market (risk premium). Second, each asset has an individual non-rewarded risk which can also be called diversifiable risk, since it can be avoided by constructing diversified portfolios. We can also write:

$$R_i = R_F + \beta_i(\mathrm{E}(R_M) - R_F) + \varepsilon_i \tag{1.17}$$

with $\mathrm{E}(\varepsilon_i)=0$ and $\mathrm{Var}(\varepsilon)$ the unsystematic risk, according to the CAPM.

In the following section we will point out the shortcomings as well as some extensions of the traditional CAPM.

1.3 Shortcomings and Extensions of the CAPM

Empirical studies trying to justify the Capital Asset Pricing Model in its traditional form have shown that this model is inadequate to explain the observed cross-sectional returns. This is partly due to its strong assumptions. However, there exist extensions that are appropriate to eliminate some of these assumptions.

Black's zero-beta model

One example is the often criticized existence of a risk-free asset. Ingersoll (1987) showed that in the absence of a risk-free asset the Capital Asset Pricing Model still holds by introducing a zero-beta portfolio:

We assume that by combining risky assets it is possible to create portfolios that are uncorrelated with the respective market portfolio and therefore have a beta of zero. As a matter of fact, these portfolios have the same expected return $\mathrm{E}(R_Z)$ since the risk premium rewarded by the market is zero. Among all these zero-beta portfolios only one is located on the efficient frontier. It minimizes the risk subject to the expected return $\mathrm{E}(R_Z)$.

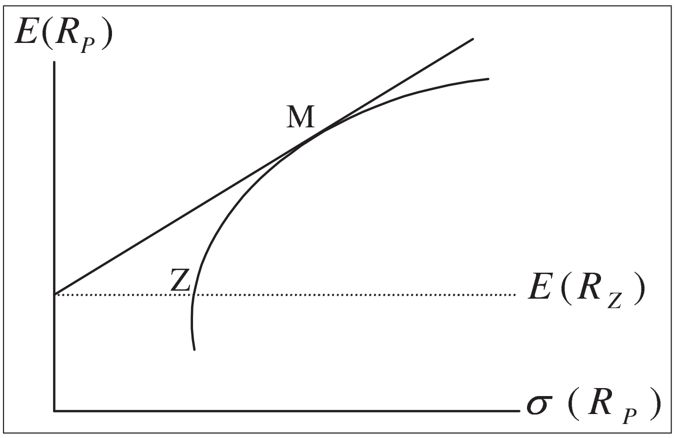

Figure 1.4: Efficient frontier in the absence of a risk-free asset

Ingersoll (1987) applied the same reasoning as in the basic case (section 1.2.2.) this time combining the zero-beta portfolio and the market portfolio

$$
\begin{aligned}
\mathrm{E}(R_P) &= x\mathrm{E}(R_Z) + (1-x)\mathrm{E}(R_M) && (1.18)\\
\sigma(R_P) &= \left[x^2\sigma_Z^2 + (1-x)^2\sigma_M^2\right]^{1/2}. && (1.19)
\end{aligned}
$$

They proved that the following equation holds:

$$
\mathrm{E}(R_i) = \mathrm{E}(R_Z) + \beta_i(\mathrm{E}(R_M) - \mathrm{E}(R_Z)). \tag{1.20}
$$

This is the Black-CAPM (1972) or two-factor-model and it eliminates one major criticism on the CAPM: the existence of a risk-free asset and the unlimited access to it.

Intertemporal Capital Asset Pricing Model (ICAPM)

Since Merton's (1973) seminal article on the intertemporal approach of the CAPM, many successors followed him and developed different models in the same line of criticism on the traditional CAPM: Unlike the basic model by Sharpe and Lintner, Merton assumed that the agent maximizes her expected utility of wealth over a time horizon longer then one period. As a consequence, he allowed state variables, like the fixed risk-free return R_F in the basic model, to vary randomly over time. In this case Merton showed that the investor's behavior can be better described by taking a third asset into account, that is perfectly negatively correlated with changes in the return on the risk-free asset. Thus it hedges against changes in the agent's investment opportunity set (here: the risk-free asset). In the same way, any other state variable can be taken into consideration. Our two-fund separation model now becomes a three-fund separation model, including the risk-free asset, the market portfolio and the so called hedge portfolio. Merton proposed a long term bond portfolio for this hedge portfolio. According to the standard logic of the CAPM the expected return is then increased by the risk premium associated with this last portfolio.

$$
\mathrm{E}(R_i) = R_F + \lambda_{iM}(\mathrm{E}(R_M) - R_F) + \lambda_{iH}(\mathrm{E}(R_H) - R_F) \tag{1.21}
$$

with:

$$
\lambda_{iM} = \frac{\sigma_i(\rho_{iM} - \rho_{iH}\rho_{MH})}{\sigma_M(1-\rho_{HM}^2)} \qquad \text{and} \qquad \lambda_{iM} = \frac{\sigma_i(\rho_{iH} - \rho_{iM}\rho_{MH})}{\sigma_H(1-\rho_{HM}^2)}. \tag{1.22}
$$

Hence, this approach was a first contribution to the explanation of the empirical "anomaly" that small cap firms that have the same market beta as comparable large cap firms reveal higher expected returns. According to Merton's intertemporal version of the CAPM this could be due to a higher sensitivity to interest rate fluctuations.

Furthermore, this model allows the expected return of a zero-beta portfolio to differ from the risk-free rate, unlike the traditional form of the CAPM. Merton argues that this third factor could be the expected return of a zero-beta portfolio and he obtains the characteristic form of his intertemporal version of the CAPM:

$$
\mathrm{E}(R_i) = R_F + \beta_i(\mathrm{E}(R_M) - R_F) + \gamma_i(\mathrm{E}(R_Z) - R_F) \tag{1.23}
$$

He uses the zero-beta portfolio as a proxy for the long term bond portfolio measuring the volatility of the risk-free rate, since empirical evidence showed that the correlation of a bond portfolio and the market portfolio is close to zero. Furthermore, Merton shows that the gamma is negatively correlated with the beta and thus, explains why empirically low-beta portfolios have higher expected returns than high-beta portfolios.

To conclude, Merton's multi-period version of the CAPM produces a multi-factor model, when the investment opportunity set is stochastic. As we will see in the following sections, this model can then be generalized by considering other risk-factors contributing through risk premiums to the expected return on a specific asset.

Consumption-based Capital Asset Pricing Model (CCAPM)

Another way of implementing a multi-period model is to explain the asset return through the consumption growth rate, instead of explaining it through market return. Breeden (2003) was the first to come up with a continuous-time version of this so-called Consumption-based Asset Pricing Model (CCAPM). The expected return for the asset i in the equilibrium is given by:

$$\mathrm{E}(R_i) = \mathrm{E}(R_Z) + \gamma(\mathrm{Cov}(\Delta c_t, R_i - R_Z)) \tag{1.24}$$

where γ is the coefficient of relative risk aversion with respect to intraperiod utility flow. This was also a response to the main criticism by Roll (1977) that the market portfolio does not exist and is difficult to approximate.

Multi-Factor-Models

One major problem of the traditional CAPM is that it can not explain the empirical fact that small cap firms have significantly higher expected returns than big cap firms. Similarly, firms with a high book-value-to-price ratio (called "value" stocks) do better than firms with a low ratio (called "growth" stocks). Fama and French (1993) therefore introduced two additional factors to the traditional CAPM:

- SMB (small minus big): denotes the excess returns of a small-capitalization portfolio compared to a large-capitalization portfolio
- HML (high minus large): denotes the excess returns of value stocks compared to growth stocks

The model then becomes:

$$\mathrm{E}(R_i) = R_F + \beta_i(\mathrm{E}(R_M) - R_F) + b_{is}\mathrm{E}(SMB) + b_{iv}\mathrm{E}(HML) \tag{1.25}$$

Fama and French still see high returns as a reward for taking on high risks. Stocks with a high book-to-price ratio are supposed to have higher expected returns. Consequently, they must be more risky. The idea is that relatively low priced stocks are largely due to the investors' mentality that these stocks are riskier.

In the same way as Fama and French introduced the two additional factors other authors tried to take more factors into account. Carhart (1997), for example, added an index measuring the difference between the average of the highest and the average of the lowest return.

Summarizing the CAPM extensions

Each of these approaches called only one essential assumption of the traditional CAPM into question. They derived models for a world where a risk-less asset does not exist or the market portfolio is not observable. Furthermore, they added more factors in order to explain more and more of the observed deviations from the predictions by the traditional CAPM. As a matter of fact, taking more factors into account, leads to higher R^2 and a better description of the asset return. However, there is a trade off between the interpretability and the deviation from true values. For an empirical review one can refer to Cochrane (2001).

In the next chapter we will present a model that achieves better empirical results without introducing new factors, keeping the relative simplicity of the CAPM in comparison to other models.

Chapter 2

The Four-Moment-CAPM

In this chapter we will present a first approach in order to reinstate the traditional one-factor Capital Asset Pricing Model. We go back to the simple idea that the investor has a specific utility function and is willing to optimize her expected utility of wealth. In the mean-variance approach of Markowitz, upon which the CAPM is built, the risk was represented by the variance (or the standard deviation) of the portfolio returns. Consequently, the investor tried to maximize the expected portfolio return given a certain standard deviation or tried to minimize the standard deviation given a level of expected return of her portfolio. Depending on her level of risk aversion the investor chose a point on the efficient frontier representing an expected-utility-standard-deviation couple.

Considering only the first two moments of the portfolio return distribution is only an approximation of the real portfolio allocation game, except in two situations:

1. When the portfolio returns are normally distributed, and hence the distribution is perfectly determined by the first two moments.
2. When the representative agent (investor) has an utility function only depending on the first two moments (e.g. a quadratic utility function).

However, empirical results have proved that the corresponding third and fourth moment (skewness and kurtosis) significantly differ from those of the normal distribution (e.g. Kraus and Litzenberger (1976), Campbell and Siddique (1999), Hwang and Satchell (1999), Fang and Lai (1997)). Hence, an equilibrium analysis such as the CAPM, that is built on the expected utility - risk duality, should take these higher moments into account. In addition, this approach is justified by the fact that the most used utility functions yield existing derivatives of higher order different from zero. Instead of fixing one particular utility function, we will concentrate on a general method that is applicable to a large class of functions with the objective to demonstrate the need of consideration of higher moments.

2.1 Approximation of Utility Functions

We consider any arbitrary utility function. As in Markowitz (1952), the investor only tries to maximize her wealth stemming from her asset investment, since we assume a world without labor income. Hence, she will consider the utility of the investment return (R). The fourth-order Taylor expansion gives us:

$$\mathrm{U}(R)=\sum_{i=0}^{4}\left[\frac{\mathrm{U}^{(i)}(\mathrm{E}(R))}{i!}(R-\mathrm{E}(R))^{i}\right]+o\left[(R-\mathrm{E}(R))^{4}\right] \tag{2.1}$$

where $U^{(n)}$ denotes the nth derivative of the function U. Taking the expectation on both sides yields:

$$\mathrm{E}[\mathrm{U}(R)]=\mathrm{U}(\mathrm{E}(R))+\frac{\mathrm{U}^{(2)}(\mathrm{E}(R))}{2}\sigma^{2}+\frac{\mathrm{U}^{(3)}(\mathrm{E}(R))}{6}s+\frac{\mathrm{U}^{(4)}(\mathrm{E}(R))}{24}\kappa \tag{2.2}$$

with s the non-standardized skewness and κ the non-standardized kurtosis of the portfolio return distribution. Note that the usual definitions of skewness and kurtosis are normalized:

$$S=\frac{\mathrm{E}(R-\mathrm{E}(R))^{3}}{\sigma^{3}} \quad \text{and} \quad K=\frac{\mathrm{E}(R-\mathrm{E}(R))^{4}}{\sigma^{4}}. \tag{2.3}$$

If we neglect the influence of the last two components, (2.2) coincides with the mean-variance approach, since the second derivative is negative[1]. Maximizing expected utility is equivalent to the portfolio trade-off between mean and variance and depends on the level of risk aversion. Similarly, if the utility function only depends on the first two moments (i.e. $\mathrm{U}\equiv\mathrm{U}(\mu,\sigma^{2})$), the third and fourth derivatives are zero and the last two terms equal consequently zero. This is the underlying quadratic utility function in the Markowitz approach.

But what happens in a world where investors are sensitive to skewness and kurtosis and portfolio returns are not normally distributed? First, we consider special utility functions with desirable features[2] in order to determine the sign of the third and the fourth moment.

	$\mathrm{U}(W)$	$\mathrm{U}^{(1)}(W)$	$\mathrm{U}^{(2)}(W)$	$\mathrm{U}^{(3)}(W)$	$\mathrm{U}^{(4)}(W)$
CRRA	$\frac{W^{1-\gamma}}{1-\gamma}$	$W^{-\gamma}$	$-\gamma W^{-(\gamma+1)}$	$\gamma(\gamma+1)W^{-(\gamma+2)}$	$-\gamma(\gamma+1)(\gamma+2)W^{-(\gamma+3)}$
CARA	$-\frac{\exp(-\lambda W)}{\lambda}$	$\exp -\lambda W$	$-\lambda\exp(-\lambda W)$	$\lambda^{2}\exp(-\lambda W)$	$-\lambda^{3}\exp(-\lambda W)$
DARA	$\ln(W)$	$\frac{1}{W}$	$-\frac{1}{W^{2}}$	$\frac{2}{W^{3}}$	$-\frac{3}{W^{4}}$

λ and γ are respectively positive, denoting the absolute and the relative risk aversion. For each utility function we can state that the third derivative is positive and the fourth negative. Analyzing (2.2), we can therefore assume that investors have preference for a higher skewness and an aversion towards kurtosis which had already been found by Horvath and Scott (1980). Looking at the distribution this is rather easy to understand. A positive skewness means a higher probability for higher values of wealth relative to lower values:

$$\mathbb{P}[x>\mathrm{E}(R)+c]>\mathbb{P}[x<\mathrm{E}(R)-c] \tag{2.4}$$

[1] We still assume that investors are risk averse.

[2] Constant Relative Risk Aversion (CRRA), Constant Absolute Risk Aversion (CARA) and Decreasing Absolute Risk Aversion (DARA)

Concerning the fourth moment, a high kurtosis reflects so-called "fat tails", i.e. a higher probability for extreme values than in the case of a normal distribution. Following the prospect theory, the negative value that is attributed to the chance of highly negative returns excesses the positive value that is attributed to the opportunity of higher returns.

Jondeau and Rockinger (2004) showed that the fourth-order Taylor expansion of a CARA-Utility function leads to excellent approximations of the underlying function in the framework of optimal portfolio allocation even under large departure from normally distributed portfolio returns. On the contrary, the mean-variance approach yields large deviation from the optimal portfolio constructed by direct expected utility maximization.

As a result, we consider the approximation of utility function by a fourth-order Taylor expansion as satisfactory. Beside using the variance as a risk and uncertainty measure, this approach also incorporates skewness and kurtosis. Since the traditional CAPM does not consider the latter, we will correct the model by introducing two factors.

2.2 Higher Moment Risk Premia

As explained in the previous section, we can approximate any utility function as a function of the expected return, the standard deviation, the skewness and the kurtosis of the portfolio return distribution function. The maximization problem yields therefore:

$$\max_{\omega_P, \omega_{0P}} \Phi(\mu_P, \sigma_P^2, s_P, \kappa_p) \qquad (2.5)$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \omega_{iP} = 1 - \omega_{0P}$$

with:

$$\begin{aligned}
\mu_P &= \omega_{0P} R_0 + \mathrm{E}(\omega_P' R) = \omega_{0P} R_0 + \omega_P' \mu \\
\sigma_P^2 &= \omega_P' \mathrm{E}[(R-\mu)(R-\mu)']\omega_P = \omega_P' \Sigma \omega_P \\
s_P &= \omega_P' \mathrm{E}[(R-\mu)(R-\mu)' \otimes (R-\mu)'](\omega_P \otimes \omega_P) = \omega_P' \Omega_{\omega_P} \\
\kappa_P &= \omega_P' \mathrm{E}[(R-\mu)(R-\mu)' \otimes (R-\mu)' \otimes (R-\mu)'](\omega_P \otimes \omega_P \otimes \omega_P) = \omega_P' \Psi_{\omega_P}.
\end{aligned}$$

$\otimes$ denotes the Kronecker-product, $R = (R_1, \ldots, R_N)'$ the vector of asset returns, $\omega_P = (\omega_{1P}, \omega_{2P}, \ldots, \omega_{NP})'$ the vector of portfolio proportions invested in these assets, ω_{0P} the part invested in the risk-free asset (returning R_0) and $\mu = (\mathrm{E}(R_1), \mathrm{E}(R_2), \ldots, \mathrm{E}(R_N))'$. Ω_{ω_P} is the vector of co-skewnesses for the weighting vector ω_P and Ψ_{ω_P} the vector of co-kurtosises respectively. We took this matrix representation from Jondeau and Rockinger (2004) who used it in a different approach (Asset allocation problem). Writing the Lagrangian of this problem gives:

$$\mathcal{L}(\omega_P, \lambda) = \Phi(\mu_P, \sigma_P^2, s_P, \kappa_p) - \lambda(\omega_P' \mathbf{1} + \omega_{0P} - 1). \qquad (2.6)$$

The first order conditions yields:

$$
\begin{aligned}
\frac{\partial \mathcal{L}(\omega_P, \lambda)}{\partial \omega_P'} &= \Phi_1 \mu + 2\Phi_2 \Sigma \omega_P + 3\Phi_3 \Omega_{\omega_P} + 4\Phi_4 \Psi_{\omega_P} - \lambda \overset{!}{=} 0 && (2.7)\\
\frac{\partial \mathcal{L}(\omega_P, \lambda)}{\partial \omega_{0P}} &= \Phi_1 R_0 - \lambda \overset{!}{=} 0 && (2.8)\\
\frac{\partial \mathcal{L}(\omega_P, \lambda)}{\partial \lambda} &= \omega_P' \mathbf{1} + \omega_{0P} - 1 \overset{!}{=} 0 && (2.9)
\end{aligned}
$$

where Φ_i is the partial derivative of the ith argument. In the equilibrium every investor behaves optimal: From (2.7) and (2.8) we derive the condition for an optimum:

$$
\mu - R_0 = \alpha_1 \Sigma \omega_P + \alpha_2 \Omega_{\omega_P} + \alpha_3 \Psi_{\omega_P} \tag{2.10}
$$

with:

$$
\alpha_1 = -\frac{2\Phi_2}{\Phi_1}, \quad \alpha_2 = -\frac{3\Phi_3}{\Phi_1} \quad \text{and} \quad \alpha_3 = -\frac{4\Phi_4}{\Phi_1} \tag{2.11}
$$

In order to move from optimal conditions for individuals to the resulting market equilibrium, we refer to Cass and Stiglitz (1970). Under identical agent's probability beliefs, a necessary and sufficient condition to apply a two-fund separation theorem is that all agents have a hyperbolic absolute risk aversion utility function (HARA), i.e. that each investor's risk tolerance is a linear function of his wealth $(-U_i'/U_i'' = a_i + b_i W)$ with the same parameter b_i. In this case the portfolio weights of each portfolio are the same. Summing up all these portfolios lets us conclude that the condition (2.10) must also hold for the market portfolio. Defining:

$$
\begin{aligned}
\beta &= \frac{\Sigma \omega_M}{\sigma_M^2} && (2.12)\\
\gamma &= \frac{\Omega_{\omega_M}}{s_M} && (2.13)\\
\vartheta &= \frac{\Psi_{\omega_M}}{\kappa_M} && (2.14)
\end{aligned}
$$

we obtain the four-moment Capital Asset Pricing Model:

$$
\mu - R_0 = b_1 \beta + b_2 \gamma + b_3 \vartheta \tag{2.15}
$$

The variables b_i can be understood as the corresponding risk premia associated with the respective risk. Moving to a single asset we obtain a more intuitive and comprehensible version, resembling the usual form of the CAPM:

$$
\mathrm{E}(R_i) - R_0 = b_1 \beta_i + b_2 \gamma_i + b_3 \vartheta_i \tag{2.16}
$$

Referring to the traditional CAPM β_i denotes the systematic beta, γ_i the systematic skewness and ϑ_i the systematic kurtosis of asset i. It is easy to verify that:

$$\beta_i = \frac{\mathrm{Cov}(R_i, R_M)}{\sigma_M^2}\,, \quad \gamma_i = \frac{\mathrm{CoS}(R_i, R_M)}{s_M} \quad \text{and} \quad \vartheta_i = \frac{\mathrm{CoK}(R_i, R_M)}{\kappa_M} \tag{2.17}$$

with:

$$\mathrm{CoS}(X,Y) = \mathrm{E}\left[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))^2\right] \tag{2.18}$$
$$\mathrm{CoK}(X,Y) = \mathrm{E}\left[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))^3\right] \tag{2.19}$$

the corresponding co-moments and

$$s_M = \mathrm{E}\left[R_M - \mathrm{E}(R_M)\right]^3 \quad \text{and} \quad \kappa_M = \mathrm{E}\left[R_M - \mathrm{E}(R_M)\right]^4. \tag{2.20}$$

We have thus derived a multi-factor model. However, the three factors go back to the same root. Indeed, they measure the relation of the asset with the market portfolio concerning the respective risk. Hence, we need only one appropriated index in contrast to "real" multi-factor models. We should stress the important problem, that indices are mostly not predefined and often hard to estimate, which goes back to critique of Roll (1977). We have to admit that the stated market portfolio is probably the index that is the least observable.

As a consequence of the four-moment CAPM, the agent will be rewarded a risk premium not only for the volatility (variance) of the market portfolio but also premia for the market skewness and the market kurtosis provided the asset is correlated with the market portfolio in the sense of the specific order (i.e. the co-moments are significant).

The expected excess return of the market portfolio becomes now:

$$\mathrm{E}(R_M) - R_0 = b_1 + b_2 + b_3. \tag{2.21}$$

We will now have a look on these premia: As developed in the previous section, investors have a positive preference for expected returns and skewness, on the contrary they have an aversion towards high variance (standard deviation) and high kurtosis. As a consequence, we can state:

$$\Phi_1 > 0\,, \quad \Phi_2 < 0\,, \quad \Phi_3 > 0\,, \quad \Phi_4 < 0. \tag{2.22}$$

Regarding the risk premia we obtain:

$$b_1 = -\frac{2\Phi_2}{\Phi_1}\sigma_M^2 > 0 \tag{2.23}$$
$$b_2 = -\frac{3\Phi_3}{\Phi_1}s_M \gtreqless 0 \tag{2.24}$$
$$b_3 = -\frac{4\Phi_4}{\Phi_1}\kappa_M > 0 \tag{2.25}$$

For the systematic beta we get the same sign as already determined in the framework of the traditional CAPM. The risk premium that is rewarded for a beta reduction is assumed to be positive ($b_1 > 0$). Higher market risk results in a higher risk premium.

For the systematic skewness the result concerning the sign is ambiguous. b_2 will take the opposite sign of the market skewness. Since agents have a preference for high skewness, a negative market skewness is considered as a risk and will be rewarded with a positive skewness-risk premium. For the Kurtosis we can apply the same argument as for the second moment: high kurtosis ("fat tails") is a negative investment incentive and the corresponding kurtosis-risk premium will be positive.

Hwang and Satchell (1999) showed that the four-moment CAPM yields better results in terms of explanation of cross-sectional returns than the standard mean-variance approach. This is especially true in the case of emerging markets or hedge funds, since skewness and kurtosis are particulary significant in these contexts. However, the model, as presented in this chapter, applies a stationary analysis to an evolutionary problem. Especially in emerging markets the assumption of stationarity of risk premia does not hold. The next chapter will challenge this assumption by allowing the corresponding moments and co-moments to vary over the investment horizon.

Chapter 3

The Time-Conditional CAPM

This chapter will highlight another substantial part of today's research efforts in finance in order to improve the standard Capital Asset Pricing Model. One rather strong assumption of the latter is that the investment horizon is one period and the investor optimizes her expected return (wealth) at the end of this period. In the real world, however, investors live many periods. Hence, applying the CAPM on the real world, we make the assumption that the specific risk (beta) of an asset remains constant over time. Considering the well-accepted fact that the relative risk of a firm's cash flow is varying over, time this assumption seems to be misleading. During a recession, for example, firms' need for financial leverage depends on their cash-flow situation. Firms that actually have problems may be obliged to borrow more money than firms that are more substantial. This makes them more dependent on the market and increases their covariance with it. In addition, the positioning of firms in the different markets is likely to vary, inducing fluctuations in the co-moments, too. Consequently, this chapter will offer a model which allows systematic risk parameters to vary over time, in contrast to the standard CAPM.

3.1 The One-Factor Model

First, due to its simplicity, we will derive the one factor conditional CAPM. R_{it} and R_{Mt} will denote the excess returns of the Portfolio i and the market portfolio respectively at period t. The conditional version is given as:

$$\mathrm{E}_t(R_{it}|\mathcal{F}_{t-1}) = \beta_{it}\mathrm{E}_t(R_{Mt}|\mathcal{F}_{t-1}) \tag{3.1}$$

with

$$\beta_{it} = \frac{\mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1})}{\mathrm{Var}_t(R_M|\mathcal{F}_{t-1})} \quad \text{and} \quad R_{Mt} = \omega_t' R_t \tag{3.2}$$

$\omega = (\omega_1, \omega_2, \ldots, \omega_N)'$ being the vector of asset proportions in the market portfolio, $R = (R_1, R_2, \ldots, R_N)'$ the vector of excess asset returns and $\mathcal{F}_t$ the information set available at period t. As a result of the consumption based Capital Asset Pricing, it is often argued that the risk premium should be constant over the time. As derived by Jensen (1972) and Campbell and Viceira (2001), we have:

$$\mathrm{E}_t(R_{it}|\mathcal{F}_{t-1}) = \gamma \mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1}) \tag{3.3}$$

with

$$\gamma = \frac{\mathrm{E}_t(R_{Mt}|\mathcal{F}_{t-1})}{\mathrm{Var}_t(R_M|\mathcal{F}_{t-1})}. \tag{3.4}$$

Following the cited theory, γ can be understood as a measure of relative risk aversion. The decomposition of the market portfolio yields:

$$\begin{aligned}
\mathrm{E}_t(R_{it}|\mathcal{F}_{t-1}) &= \gamma \mathrm{Cov}_t(R_i, \omega'_{t-1}R|\mathcal{F}_{t-1}) && (3.5)\\
&= \gamma \sum_{k=1}^{N} \omega_{k,t-1} \mathrm{Cov}_t\left(R_i, R_k|\mathcal{F}_{t-1}\right). && (3.6)
\end{aligned}$$

For the whole vector of conditional expected excess returns we obtain:

$$\mu_t = \gamma \mathbf{H}_t \omega_{t-1} \tag{3.7}$$

where $\mu_t = \mathrm{E}_t(R_t|\mathcal{F}_{t-1})$ denotes conditional mean vector and $\mathbf{H}_t = \mathrm{Cov}_t(R, R|\mathcal{F}_{t-1})$ the conditional covariance matrix of the excess returns given the information available at time $t-1$. Hence, we can derive the conditional moments of the market excess return as:

$$\sigma^2_{Mt} = \omega'_{t-1}\mathbf{H}\omega_{t-1} \quad \text{and} \quad \mu_{Mt} = \omega'_{t-1}\mu_t. \tag{3.8}$$

Substituting (3.4) in (3.7) and defining $\beta_t = \mathbf{H}_t\omega_{t-1}/\sigma^2_{Mt}$ lead us back to our starting model:

$$\mu_t = \beta_t \mu_{Mt}. \tag{3.9}$$

The introductory ideas at the beginning of this chapter and former empirical work (e.g. Jagannathan and Wang (1996)) lead to the hypothesis that the covariances of asset returns and thus $\mathbf{H_t}$ vary over time. Hence, the definition of beta and (3.9) show that the betas as well as the conditional means are time-dependent and likely to vary over the investment horizon.

We will now prove that under the stated conditions the conditional CAPM deviates from the unconditional CAPM. For this purpose, we remind the decomposition formula of the variance:

$$\mathrm{Var}(X) = \mathrm{Var}[\mathrm{E}(X|Y=y)] + \mathrm{E}[\mathrm{Var}(X|Y=y)]. \tag{3.10}$$

This formula is also known under the name of analysis equation of variance. Applying (3.10) to our model yields:

$$\Sigma_t = \mathrm{Var}(\mu_t) + \mathrm{E}(\mathbf{H}_t) \tag{3.11}$$

where $\Sigma_t = \mathrm{Cov}(R, R)$ denotes the unconditional covariance matrix of excess asset returns. Another formula stemming from standard probability courses shows:

$$\mathrm{E}[\mathrm{E}(X|Y=y)] = E(X) \tag{3.12}$$

which gives us:

$$\mathrm{E}(\mu_t) = \mathrm{E}[\mathrm{E}(R_t|\mathcal{F}_{t-1})] = \mathrm{E}(R_t). \tag{3.13}$$

Taking the expectation in t on both side of (3.7) leads us to:

$$\mathrm{E}(R_t) = \mathrm{E}(\mu_t) = \gamma \mathrm{E}(\mathbf{H}_t \omega_{t-1}). \tag{3.14}$$

In the special case where the market portfolio weights are fixed to $\omega_t \equiv \omega$ we can multiply (3.11) by ω from the right hand side. By rearranging terms we obtain:

$$\mathrm{E}(R_t) = \gamma(\Sigma_t \omega - \mathrm{Var}(\mu_t)\omega). \tag{3.15}$$

Using (3.7) finally yields:

$$\begin{aligned} \mathrm{E}(R_t) &= \gamma(\Sigma_t \omega - \gamma^2 \mathrm{Var}(\mathbf{H}_t \omega)\omega) && (3.16) \\ &= \gamma \Sigma_t \omega - \gamma^3 \mathrm{Var}(\mathbf{H}_t \omega)\omega. && (3.17) \end{aligned}$$

The latter equation shows that the unconditional moments satisfy the same CAPM relation as the conditional moments (see (3.7)) only if $\mathrm{Var}(\mathbf{H}_t \omega) = 0$. Since $\mathbf{H}_t$ is supposed to vary over time the conditional version of the CAPM deviates from the unconditional CAPM. Hence, we need a model that incorporates this time-conditional context.

3.2 A Model with Autoregressive and Heteroscedastic Covariances

We have derived a conditional version of the Capital Asset Pricing Model without determining the set of available information. It is possible to estimate this model for example by the generalized methods of moments using instrumental variables (cf. Hansen (1982)). For an application in a similar framework one can refer to Narasimhan and Pradhan (2002). The advantage of this method is its relative simplicity due to a fixed number of variables and their realizations. However, if the set of chosen variables is not convenient, the econometrician will fail to estimate the true time-dependent moments. A different approach is to determine the covariance as well as the higher co-moments (in a second step) matrices using a generalized autoregressive conditional heteroscedastic (GARCH) process.

The GARCH approach assumes that agents learn from the past, i.e. that they update their estimates of the means and covariances each period using the errors in last period's expectations. Thus, agents adapt their expectations only from information on the excess asset returns.

From (3.7) we derive the standard expression of the multidimensional GARCH(1,1)-in-mean model (GARCH(1,1)-M) (see Bollerslev et al. (1988)):

$$
\begin{aligned}
R_t &= \alpha + \gamma \mathbf{H}_t \omega_{t-1} + \varepsilon_t \\
\mathrm{vech}(\mathbf{H}_t) &= C + \mathbf{A}\mathrm{vech}(\varepsilon_{t-1}\varepsilon'_{t-1}) + \mathbf{B}\mathrm{vech}(\mathbf{H}_{t-1}) \\
\varepsilon_t|\mathcal{F}_{t-1} &\sim N(0, \mathbf{H}_t).
\end{aligned} \tag{3.18}
$$

where vech($\cdot$) denotes the column stacking operator of the lower portion of a symmetric matrix. According to the Capital Asset Pricing Model, the $N \times 1$ vector α should be zero. The $N(N+1)/2 \times N(N+1)/2$ matrices $\mathbf{A}$ and $\mathbf{B}$ determine the first order autoregressive and moving average characteristics of the excess asset price time series.

The CAPM states furthermore that the evaluation of the asset price can be obtained by its relation to the market portfolio. Thus, from (3.18) we derive the bivariate model:

$$
\begin{pmatrix} R_{it} \\ R_{Mt} \end{pmatrix} = \begin{pmatrix} \alpha_i \\ \alpha_M \end{pmatrix} + \gamma \begin{pmatrix} \mathrm{Var}_t(R_i|\mathcal{F}_{t-1}) & \mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1}) \\ \mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1}) & \mathrm{Var}_t(R_M|\mathcal{F}_{t-1}) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \varepsilon_{it} \\ \varepsilon_{Mt} \end{pmatrix} \tag{3.19}
$$

$$
\begin{pmatrix} \mathrm{Var}_t(R_i|\mathcal{F}_{t-1}) \\ \mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1}) \\ \mathrm{Var}_t(R_M|\mathcal{F}_{t-1}) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \mathbf{A} \begin{pmatrix} \varepsilon^2_{i,t-1} \\ \varepsilon_{i,t-1}\varepsilon_{M,t-1} \\ \varepsilon^2_{M,t-1} \end{pmatrix} + \mathbf{B} \begin{pmatrix} \mathrm{Var}_{t-1}(R_i|\mathcal{F}_{t-2}) \\ \mathrm{Cov}_{t-1}(R_i, R_M|\mathcal{F}_{t-2}) \\ \mathrm{Var}_{t-1}(R_M|\mathcal{F}_{t-2}) \end{pmatrix}
$$

Hence, we obtain the characteristic equation of the conditional one-factor CAPM:

$$
R_{it} = \alpha_i + \gamma \mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1}) + \varepsilon_{it} \tag{3.20}
$$

with

$$
\mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1}) = c_2 + \mathbf{A}_{2\cdot} \begin{pmatrix} \varepsilon^2_{i,t-1} \\ \varepsilon_{i,t-1}\varepsilon_{M,t-1} \\ \varepsilon^2_{M,t-1} \end{pmatrix} + \mathbf{B}_{2\cdot} \begin{pmatrix} \mathrm{Var}_{t-1}(R_i|\mathcal{F}_{t-2}) \\ \mathrm{Cov}_{t-1}(R_i, R_M|\mathcal{F}_{t-2}) \\ \mathrm{Var}_{t-1}(R_M|\mathcal{F}_{t-2}) \end{pmatrix} \tag{3.21}
$$

where $\mathbf{A}_{2\cdot}$ and $\mathbf{B}_{2\cdot}$ denote, respectively, the second rows of the matrices.

3.3 The Four-Moment Extension of the GARCH

We now return to the four-moment CAPM proposed in the second chapter and transform it into a conditional model following the procedure stated for the two-moment model in the previous section. Introducing the third and the fourth moment into the conditional form, (3.19) becomes:

$$
\begin{aligned}
\begin{pmatrix} R_{it} \\ R_{Mt} \end{pmatrix} = \begin{pmatrix} \alpha_i \\ \alpha_M \end{pmatrix} &+ \gamma_1 \begin{pmatrix} \mathrm{Var}_t(R_i|\mathcal{F}_{t-1}) & \mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1}) \\ \mathrm{Cov}_t(R_i, R_M|\mathcal{F}_{t-1}) & \mathrm{Var}_t(R_M|\mathcal{F}_{t-1}) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&+ \gamma_2 \begin{pmatrix} s^\star_{it} & \mathrm{CoS}_t(R_i, R_M|\mathcal{F}_{t-1}) \\ \mathrm{CoS}_t(R_M, R_i|\mathcal{F}_{t-1}) & s^\star_{Mt} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&+ \gamma_3 \begin{pmatrix} \kappa^\star_{it} & \mathrm{CoK}_t(R_i, R_M|\mathcal{F}_{t-1}) \\ \mathrm{CoK}_t(R_M, R_i|\mathcal{F}_{t-1}) & \kappa^\star_{Mt} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \varepsilon_{it} \\ \varepsilon_{Mt} \end{pmatrix}
\end{aligned} \tag{3.22}
$$

with the common notation from the second chapter and $s^{\star}_{jt} = \mathrm{E}\left[(R_j - \mathrm{E}(R_j))^3 | \mathcal{F}_{t-1}\right]$ and $\kappa^{\star}_{jt} = \mathrm{E}\left[(R_j - \mathrm{E}(R_j))^4 | \mathcal{F}_{t-1}\right]$ the corresponding conditional third and fourth moments. As the second chapter, we find the three components of the risk premium associated with the covariance (γ_1), the co-skewness (γ_2) and the co-kurtosis (γ_3). The γ_i denote the market prices for the corresponding risks and are supposed to be constant over the investment horizon.

(3.22) can be rewritten as:

$$\begin{aligned}\begin{pmatrix} R_{it} \\ R_{Mt} \end{pmatrix} = \begin{pmatrix} \alpha_i \\ \alpha_M \end{pmatrix} &+ \gamma_1 \begin{pmatrix} \mathrm{Cov}_t(R_i, R_M | \mathcal{F}_{t-1}) \\ \mathrm{Var}_t(R_M | \mathcal{F}_{t-1}) \end{pmatrix} + \gamma_2 \begin{pmatrix} \mathrm{CoS}_t(R_i, R_M | \mathcal{F}_{t-1}) \\ s^{\star}_{Mt} \end{pmatrix} \\ &+ \gamma_3 \begin{pmatrix} \mathrm{CoK}_t(R_i, R_M | \mathcal{F}_{t-1}) \\ \kappa^{\star}_{Mt} \end{pmatrix} + \begin{pmatrix} \varepsilon_{it} \\ \varepsilon_{Mt} \end{pmatrix}. \end{aligned} \tag{3.23}$$

In the previous section, we introduced the conditional covariance matrix ($\mathbf{H}$) to allow heteroscedasticity and time-moving covariances. Thus, we dealt with one of the main restrictions in standard CAPM, due to its one-period horizon. However, we still assumed a normal distribution of the asset returns. As we introduced the co-skewness and the co-kurtosis, we should allow the third and the fourth moment of the asset distribution not only to deviate from the normal distribution, but also to vary over time. Hence, the error terms no longer follow a normal distribution. We will come back to this point after having defined the diverse conditional moments in (3.22). As in the previous section, we will use the vech-operator:

$$\begin{pmatrix} \mathrm{Var}_t(R_i | \mathcal{F}_{t-1}) \\ \mathrm{Cov}_t(R_i, R_M | \mathcal{F}_{t-1}) \\ \mathrm{Var}_t(R_M | \mathcal{F}_{t-1}) \\ \mathrm{CoS}_t(R_i, R_M | \mathcal{F}_{t-1}) \\ s^{\star}_{Mt} \\ \mathrm{CoK}_t(R_i, R_M | \mathcal{F}_{t-1}) \\ \kappa^{\star}_{Mt} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \end{pmatrix} + \mathbf{A} \begin{pmatrix} \varepsilon^2_{i,t-1} \\ \varepsilon_{i,t-1}\varepsilon_{M,t-1} \\ \varepsilon^2_{M,t-1} \\ \varepsilon_{i,t-1}\varepsilon^2_{M,t-1} \\ \varepsilon^3_{M,t-1} \\ \varepsilon_{i,t-1}\varepsilon^3_{M,t-1} \\ \varepsilon^4_{M,t-1} \end{pmatrix} + \mathbf{B} \begin{pmatrix} \mathrm{Var}_{t-1}(R_i | \mathcal{F}_{t-2}) \\ \mathrm{Cov}_{t-1}(R_i, R_M | \mathcal{F}_{t-2}) \\ \mathrm{Var}_{t-1}(R_M | \mathcal{F}_{t-2}) \\ \mathrm{CoS}_{t-1}(R_i, R_M | \mathcal{F}_{t-2}) \\ s^{\star}_{M,t-1} \\ \mathrm{CoK}_{t-1}(R_i, R_M | \mathcal{F}_{t-2}) \\ \kappa^{\star}_{M,t-1} \end{pmatrix} \tag{3.24}$$

In this case, the matrices $\mathbf{A}$ and $\mathbf{B}$ are 7×7 dimensional. As one can see, this model reveals the problem of a large number of parameters. As a consequence, our simple bivariate model has already 110 parameters to estimate. Thus, we will impose certain constraints. One method often used in general GARCH models is to consider the matrices $\mathbf{A}$ and $\mathbf{B}$ to be diagonal. By doing so, we reduce the number of parameters to estimate to 26. A common way to estimate these parameters is the maximum likelihood method.

The GARCH specification assumes that returns come from a conditionally normal distribution. However, stock market returns reveal thicker tails (excess kurtosis). Oelker (2004) showed that the GARCH process yields higher kurtosis than a normal distribution, even if the underlying error terms are normally distributed. Furthermore, the model needs to accommodate time-varying conditional skewness, which stresses the need of an asymmetric distribution function. In this context an underlying conditional t-Distribution would be appropriate. However, since we want to model the co-moments of the second, third and fourth order, we need a distribution function that includes these moments explicitly. In the case of small deviations from the normal distribution Blinnikov and Moessner (1998) showed that the Gram-Charlier expansion of the normal density offers such a solution.

We define $\mathbf{S}$ and $\mathbf{K}$ as the matrices of conditional co-moments according to (3.22) which yields:

$$\begin{pmatrix} R_{it} \\ R_{Mt} \end{pmatrix} = \begin{pmatrix} \alpha_i \\ \alpha_M \end{pmatrix} + \gamma_1 \mathbf{H}_t \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \gamma_2 \mathbf{S}_t \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \gamma_3 \mathbf{K}_t \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \varepsilon_{it} \\ \varepsilon_{Mt} \end{pmatrix}. \tag{3.25}$$

For the distribution of the asset returns we have:

$$\varepsilon_t | \mathcal{F}_{t-1} = \begin{pmatrix} \varepsilon_{it} | \mathcal{F}_{t-1} \\ \varepsilon_{Mt} | \mathcal{F}_{t-1} \end{pmatrix} \sim (0, \mathbf{H}_t, \mathbf{S}_t, \mathbf{K}_t). \tag{3.26}$$

In order to facilitate the following multivariate Gram-Charlier expansion, we will standardize the obtained error terms using

$$\mathbf{H}_t = \begin{pmatrix} \mathrm{Var}_t(R_i | \mathcal{F}_{t-1}) & \mathrm{Cov}_t(R_i, R_M | \mathcal{F}_{t-1}) \\ \mathrm{Cov}_t(R_i, R_M | \mathcal{F}_{t-1}) & \mathrm{Var}_t(R_M | \mathcal{F}_{t-1}) \end{pmatrix}. \tag{3.27}$$

Defining $\eta_t = \mathbf{H}_t^{-1/2} \varepsilon_t$ we have:

$$\eta_t | \mathcal{F}_{t-1} = \begin{pmatrix} \eta_{it} | \mathcal{F}_{t-1} \\ \eta_{Mt} | \mathcal{F}_{t-1} \end{pmatrix} \sim (0, I, \tilde{\mathbf{S}}_t, \tilde{\mathbf{K}}_t) \tag{3.28}$$

with I the two-dimensional unit matrix, $\tilde{\mathbf{S}}$ and $\tilde{\mathbf{K}}$ the matrices of standardized conditional co-moments of third and fourth order. The bivariate Gram-Charlier expansion of the two-dimensional standard normal density is given by (see McCullagh (1987)):

$$\Psi(\eta_t | \mathcal{F}_{t-1}) = \frac{1}{2\pi} e^{-\frac{\eta_t' \eta_t}{2}} \underbrace{\left(1 + \frac{1}{3!} \sum_{ijk} d_{ijk} h^{ijk}(\eta_t) + \frac{1}{4!} \sum_{ijkl} d_{ijkl} h^{ijkl}(\eta_t) \right)}_{\stackrel{\mathrm{def}}{=} \mathbf{Z}_t} \tag{3.29}$$

where the d's are coefficients denoting the difference between the cumulants of the underlying conditional error terms and the cumulants of the standard gaussian:

$$d_{1\ldots n} = \kappa^{\Psi}_{1\ldots n} - \kappa^{N}_{1\ldots n} \quad \text{and} \quad \kappa^{\Psi}_{1\ldots n} = \mathrm{cum}(x_1, \ldots, x_n). \tag{3.30}$$

In our case the two functions (Ψ and the gaussian) have zero mean vectors and unit variance matrices. Hence, the multivariate cumulants are defined by:

$$\mathrm{cum}(x_i, x_j, x_k) = \mathrm{E}(x_i x_j x_k) \tag{3.31}$$

$$\mathrm{cum}(x_i, x_j, x_k, x_l) = \mathrm{E}(x_i x_j x_k x_l) - \mathrm{E}(x_i x_j)\mathrm{E}(x_k x_l)[3]. \tag{3.32}$$

The bracket notation means that the corresponding term appears for all the relevant index permutations (the number in the brackets indicates the number of these permutations).

The functions denoted with h in (3.29) correspond to the multidimensional hermite polynomials (see McCullagh (1987)):

$$\begin{aligned} h^{ijk}(x) &= x_i x_j x_k - x_i \delta_{jk}[3] && (3.33)\\ h^{ijkl}(x) &= x_i x_j x_k x_l - x_i x_j \delta_{kl}[6] + \delta_{ij}\delta_{kl}[3] && (3.34)\end{aligned}$$

where δ_{ij} is the Kronecker symbol with:

$$\delta_{ij} = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j. \end{cases} \qquad (3.35)$$

For our bivariate model the mentioned indices take either the value i or the value M denoting respectively the asset and the market portfolio. In the case of only two different indices one can easily see that only the number of the occurring indices determines the value of the cumulants and the Hermite polynomials. For the third moment we have 8 terms and for the fourth moment we obtain 16 terms. However, for symmetric reasons we only have 4 different terms for the third and 5 different terms for the fourth moment. For the Hermite polynomials we obtain:

$$\begin{aligned} h^{iii}(\eta_t) &= \eta_{it}^3 - 3\eta_{it}\\ h^{iiM}(\eta_t) = h^{iMi}(\eta_t) = h^{Mii}(\eta_t) &= \eta_{it}^2\eta_{Mt} - \eta_{Mt}\\ h^{MMi}(\eta_t) = h^{MiM}(\eta_t) = h^{iMM}(\eta_t) &= \eta_{it}\eta_{Mt}^2 - \eta_{it}\\ h^{MMM}(\eta_t) &= \eta_{Mt}^3 - 3\eta_{Mt}\\ h^{iiii}(\eta_t) &= \eta_{it}^4 - 6\eta_{it}^2 + 3 \qquad (3.36)\\ h^{MMMi}(\eta_t) = h^{MMiM}(\eta_t) = h^{MiMM}(\eta_t) = h^{iMMM}(\eta_t) &= \eta_{it}\eta_{Mt}^3 - 3\eta_{it}\eta_{Mt}\\ h^{iiiM}(\eta_t) = h^{iiMi}(\eta_t) = h^{iMii}(\eta_t) = h^{Miii}(\eta_t) &= \eta_{it}^3\eta_{Mt} - 3\eta_{it}\eta_{Mt}\\ h^{MMii}(\eta_t) = h^{MiMi}(\eta_t) = h^{MiiM}(\eta_t) &\\ = h^{iiMM}(\eta_t) = h^{iMiM}(\eta_t) = h^{iMMi}(\eta_t) &= \eta_{it}^2\eta_{Mt}^2 - \eta_{it}^2 - \eta_{Mt}^2 + 1\\ h^{MMMM}(\eta_t) &= \eta_{Mt}^4 - 6\eta_{Mt}^2 + 3.\end{aligned}$$

Since Ψ is standardized it has the same first-order and second-order cumulants as the gaussian. That is why (3.29) only yields terms of the third and the fourth order (the coefficients d are zero for the first two moments). As we know the cumulants of higher order of the standard normal density, we obtain:

$$\begin{aligned} d_{iii}(\eta_t) &= \mathrm{E}(\eta_{it}^3|\mathcal{F}_{t-1})\\ d_{iiM}(\eta_t) = d_{iMi}(\eta_t) = d_{Mii}(\eta_t) &= \mathrm{E}(\eta_{it}^2\eta_{Mt}|\mathcal{F}_{t-1})\\ d_{MMi}(\eta_t) = d_{MiM}(\eta_t) = d_{iMM}(\eta_t) &= \mathrm{E}(\eta_{it}\eta_{Mt}^2|\mathcal{F}_{t-1})\\ d_{MMM}(\eta_t) &= \mathrm{E}(\eta_{Mt}^3|\mathcal{F}_{t-1})\\ d_{iiii}(\eta_t) &= \mathrm{E}(\eta_{it}^4|\mathcal{F}_{t-1}) - 3 \qquad (3.37)\\ d_{MMMi}(\eta_t) = d_{MMiM}(\eta_t) = d_{MiMM}(\eta_t) = d_{iMMM}(\eta_t) &= \mathrm{E}(\eta_{it}\eta_{it}^3|\mathcal{F}_{t-1})\\ d_{iiiM}(\eta_t) = d_{iiMi}(\eta_t) = d_{iMii}(\eta_t) = d_{Miii}(\eta_t) &= \mathrm{E}(\eta_{it}^3\eta_{Mt}|\mathcal{F}_{t-1})\\ d_{MMii}(\eta_t) = d_{MiMi}(\eta_t) = d_{MiiM}(\eta_t) &\\ = d_{iiMM}(\eta_t) = d_{iMiM}(\eta_t) = d_{iMMi}(\eta_t) &= \mathrm{E}(\eta_{it}^2\eta_{Mt}^2|\mathcal{F}_{t-1}) - 1\\ d_{MMMM}(\eta_t) &= \mathrm{E}(\eta_{Mt}^4|\mathcal{F}_{t-1}) - 3.\end{aligned}$$

The last difficulty is now to proceed from the observed co-moments of ε and its expected conditional co-moments in (3.24) to the respective expressions for the standardized error terms η. For the covariance matrix this is rather simple:

$$\begin{pmatrix} \eta_i^2 & \eta_i\eta_M \\ \eta_i\eta_M & \eta_M^2 \end{pmatrix} = \mathbf{H}^{-1/2} \begin{pmatrix} \varepsilon_i^2 & \varepsilon_i\eta_M \\ \varepsilon_i\varepsilon_M & \varepsilon_M^2 \end{pmatrix} \mathbf{H}^{-1/2}. \tag{3.38}$$

Second, we will consider the third-order moments. Using the Kronecker product yields:

$$\begin{pmatrix} \eta_i^2 & \eta_i\eta_M \\ \eta_i\eta_M & \eta_M^2 \end{pmatrix} \otimes (\eta_i \ \eta_M) = \begin{pmatrix} \eta_i^3 & \eta_i^2\eta_M & \eta_i^2\eta_M & \eta_i\eta_M^2 \\ \eta_i^2\eta_M & \eta_i\eta_M^2 & \eta_i\eta_M^2 & \eta_M^3 \end{pmatrix}. \tag{3.39}$$

Thus, we can derive the needed co-moments using the standardization formula $\eta = \mathbf{H}^{-1/2}\varepsilon$:

$$\begin{pmatrix} \eta_i^3 & \eta_i^2\eta_M & \eta_i^2\eta_M & \eta_i\eta_M^2 \\ \eta_i^2\eta_M & \eta_i\eta_M^2 & \eta_i\eta_M^2 & \eta_M^3 \end{pmatrix} = \mathbf{H}^{-1/2} \begin{pmatrix} \varepsilon_i^2 & \varepsilon_i\varepsilon_M \\ \varepsilon_i\varepsilon_M & \varepsilon_M^2 \end{pmatrix} \mathbf{H}^{-1/2} \otimes (\varepsilon_i \ \varepsilon_M)\mathbf{H}^{-1/2}. \tag{3.40}$$

We proceed in the same manner for the fourth-order moments:

$$\begin{pmatrix} \eta_i^2 & \eta_i\eta_M \\ \eta_i\eta_M & \eta_M^2 \end{pmatrix} \otimes (\eta_i \ \eta_M) \otimes (\eta_i \ \eta_M) = \begin{pmatrix} \eta_i^4 & \eta_i^3\eta_M & \eta_i^3\eta_M & \eta_i^2\eta_M^2 & \eta_i^3\eta_M & \eta_i^2\eta_M^2 & \eta_i^2\eta_M^2 & \eta_i\eta_M^3 \\ \eta_i^3\eta_M & \eta_i^2\eta_M^2 & \eta_i^2\eta_M^2 & \eta_i\eta_M^3 & \eta_i^2\eta_M^2 & \eta_i\eta_M^3 & \eta_i\eta_M^3 & \eta_M^4 \end{pmatrix} \tag{3.41}$$

and finally:

$$\begin{aligned} &\begin{pmatrix} \eta_i^4 & \eta_i^3\eta_M & \eta_i^3\eta_M & \eta_i^2\eta_M^2 & \eta_i^3\eta_M & \eta_i^2\eta_M^2 & \eta_i^2\eta_M^2 & \eta_i\eta_M^3 \\ \eta_i^3\eta_M & \eta_i^2\eta_M^2 & \eta_i^2\eta_M^2 & \eta_i\eta_M^3 & \eta_i^2\eta_M^2 & \eta_i\eta_M^3 & \eta_i\eta_M^3 & \eta_M^4 \end{pmatrix} \\ &= \mathbf{H}^{-1/2} \begin{pmatrix} \varepsilon_i^2 & \varepsilon_i\varepsilon_M \\ \varepsilon_i\varepsilon_M & \varepsilon_M^2 \end{pmatrix} \mathbf{H}^{-1/2} \otimes (\varepsilon_i \ \varepsilon_M)\mathbf{H}^{-1/2} \otimes (\varepsilon_i \ \varepsilon_M)\mathbf{H}^{-1/2}. \end{aligned} \tag{3.42}$$

(3.38), (3.40) and (3.42) give us the elements of the hermite polynomials in(3.36). For the conditional standardized co-moments defining the coefficients in (3.37) we impose the same autoregressive moving average relations as for the non-standardized error terms:

$$\begin{pmatrix} \mathrm{E}_t(\eta_{it}^3|\mathcal{F}_{t-1}) \\ \mathrm{E}_t(\eta_{it}^2\eta_{Mt}|\mathcal{F}_{t-1}) \\ \mathrm{E}_t(\eta_{it}\eta_{Mt}^2|\mathcal{F}_{t-1}) \\ \mathrm{E}_t(\eta_{Mt}^3|\mathcal{F}_{t-1}) \\ \mathrm{E}_t(\eta_{it}^4|\mathcal{F}_{t-1}) \\ \mathrm{E}_t(\eta_{it}^3\eta_{Mt}|\mathcal{F}_{t-1}) \\ \mathrm{E}_t(\eta_{it}^2\eta_{Mt}^2|\mathcal{F}_{t-1}) \\ \mathrm{E}_t(\eta_{it}\eta_{Mt}^3|\mathcal{F}_{t-1}) \\ \mathrm{E}_t(\eta_{Mt}^4|\mathcal{F}_{t-1}) \end{pmatrix} = \begin{pmatrix} c_8 \\ c_9 \\ c_{10} \\ c_{11} \\ c_{12} \\ c_{13} \\ c_{14} \\ c_{15} \\ c_{16} \end{pmatrix} + \tilde{\mathbf{A}} \begin{pmatrix} \eta_{i,t-1}^3 \\ \eta_{i,t-1}^2\eta_{M,t-1} \\ \eta_{i,t-1}\eta_{M,t-1}^2 \\ \eta_{M,t-1}^3 \\ \eta_{i,t-1}^4 \\ \eta_{i,t-1}^3\eta_{M,t-1} \\ \eta_{i,t-1}^2\eta_{M,t-1}^2 \\ \eta_{i,t-1}\eta_{M,t-1}^3 \\ \eta_{M,t-1}^4 \end{pmatrix} + \tilde{\mathbf{B}} \begin{pmatrix} \mathrm{E}_{t-1}(\eta_i^3|\mathcal{F}_{t-2}) \\ \mathrm{E}_{t-1}(\eta_i^2\eta_M|\mathcal{F}_{t-2}) \\ \mathrm{E}_{t-1}(\eta_i\eta_M^2|\mathcal{F}_{t-2}) \\ \mathrm{E}_{t-1}(\eta_M^3|\mathcal{F}_{t-2}) \\ \mathrm{E}_{t-1}(\eta_i^4|\mathcal{F}_{t-2}) \\ \mathrm{E}_{t-1}(\eta_i^3\eta_M|\mathcal{F}_{t-2}) \\ \mathrm{E}_{t-1}(\eta_i^2\eta_M^2|\mathcal{F}_{t-2}) \\ \mathrm{E}_{t-1}(\eta_i\eta_M^3|\mathcal{F}_{t-2}) \\ \mathrm{E}_{t-1}(\eta_M^4|\mathcal{F}_{t-2}) \end{pmatrix} \tag{3.43}$$

Having defined all the elements of (3.29), we can pass to the likelihood function. First, we denote θ the vector containing all the parameters to estimate:

$$\theta' = (\alpha_i, \alpha_M, \gamma_1, \gamma_2, \gamma_3, c', a', b') \tag{3.44}$$

where c is the 16×1 vector containing the constants in (3.24) and (3.43). a denotes the 16×1 vector of the diagonal elements of the matrices $\mathbf{A}$ $(a_1 \dots a_7)$ and $\tilde{\mathbf{A}}$ $(a_8 \dots a_{16})$. b contains the respective values of the diagonal matrices $\mathbf{B}$ and $\tilde{\mathbf{B}}$. Hence, we have 53 parameters to estimate.

From (3.29) we obtain the log-likelihood function:

$$L_t(\theta, \eta) = -\ln 2\pi - \frac{\eta_t' \eta_t}{2} + \ln \mathbf{Z}_t. \tag{3.45}$$

Thus the log-likelihood function for the sample is given by:

$$\begin{aligned} L(\theta) &= \sum_{t=1}^{T} L_t(\theta) && (3.46) \\ &= -T \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \eta_t' \eta_t + \sum_{t=1}^{T} \ln \mathbf{Z}_t. && (3.47) \end{aligned}$$

Some parameter constraints concerning the autoregressive moving average process are necessary. The stationarity condition for these processes implies:

$$a_i + b_i < 1 \qquad \forall \quad \mathrm{i} = 1 \dots 16. \tag{3.48}$$

Thus, by maximizing (3.47) with respect to (3.48) the parameters can be estimated.

We should stress the importance of using data sets with a sufficiently high number of observations. Otherwise, one can get to misleading results. Regarding the number of parameters (53), even sophisticated statistical software could eventually find local optima instead of the global solution. Hence, a simulation, using different starting parameters, could be appropriate.

The presented model can be tested in different stages in order to analyze the contribution of the corresponding terms. Information criteria (Akaike's (AIC) or Schwarz' (SC)) should be used in order to determine the significant factors (moments) in the model.

Conclusions

In this study we considered the Capital Asset Pricing Model, first developed by Sharpe (1964) and Lintner (1965), as the benchmark model in asset pricing theory. Despite its simplicity and its relatively weak empirical results in explaining cross-sectional returns, it is still the most widely used model in this field of financial theory, e.g. in MBA courses. We first stated the standard CAPM and pointed out the different hypotheses upon which it is founded. Furthermore, we discussed existing extensions that relax some of these assumptions, yielding better empirical results to further explain the cross-sectional returns. The most successful models, derived from the standard CAPM, achieve their goal in adding explanatory factors to the model (e.g. Fama and French (1993)). However, the problem of economic interpretability arises and many of these factors reveal a pure empirical nature instead of being derived from an underlying economic theory.

This work aimed to reinstate the main idea of the standard Capital Asset Pricing Model: The agent maximizes her expected return while minimizing the corresponding risk. Accordingly, this study tried to concentrate on the question how the risk can be determined. The standard CAPM has been built on a mean-variance approach by Markowitz (1952) defining the first two moments of a portfolio distribution as target variables. However, especially in the long run, asset returns often deviate from normality. This is particularly true for alternative investments such as hedge funds. We further conclude that the distribution of portfolio returns can not be properly described by expected returns and historic variances. Moreover, the mean-variance approach assumes (in the case of non-normal distributions) that the agent is not interested in higher moments of the portfolio return. We challenged this assumption and developed a hypothesis that any utility function can be approximated by the fourth-order Taylor-expansion. We cited recent research works that proved that in asset allocation this approximation yields results that are very similar to direct utility maximization. Thus, we defined the risk in a more complex way, taking skewness and kurtosis and the corresponding risk premia into account, leading to the four-moment-CAPM. We proved that, theoretically, agents prefer skewness and are averse towards kurtosis, thus, we expect higher risk premia for lower skewness and higher risk premia for higher kurtosis.

Next, we challenged the assumption of a one period model and considered a longer investment horizon. Accordingly, the factors should be allowed to vary over time, since the hypothesis of stationarity of the explanatory factors would be over-simplified. The literature offers two main directions in the modelling of time-varying factors. First, instruments that yield better predictability characteristics can be used to enhance the conditional CAPM. However, the choice of instruments is always arbitrary and the subset of used instruments does not necessarily capture the time-varying characteristics of the factors. Second, the GARCH model is often used to model time-conditional factors. It assumes autoregressive and heteroscedastic relations for the covariances and variances of time series. We took this model and extended it by the co-moments of third and fourth order. In order to model the higher moments we applied the Gram-Charlier expansion of the normal distribution. Whilst this approach is often used in physics and mathematics, economic and financial modelling has not yet taken advantage of it. The method enables us to explicitly enter all four moments and their co-moments in the distribution function. Hence, as in the one factor GARCH specification, the factors appear in their time-conditional form in the likelihood function. In the bivariate case (asset and market portfolio), the application of tensor methods (cf. McCullagh (1987)) is necessary. As a result we arrived at a four-moment time-conditional Capital Asset Pricing Model.

Instead of adding factors, we added detail in the analysis of dependencies between assets and the market portfolio. However, the latter is not observable (Roll's criticism) and the choice of an appropriate measure for it is ambiguous. Second, we assumed that risk premia for co-variance, co-skewness and co-kurtosis remain constant over time. This convention is a often applied in CAPM models, however, future research could challenge this assumption. An appropriate extension could, for example follow the approach of Hafner and Herwatz (1999) and parameterize the market risk premia for the given factors.

Empirically, our model could find different ways of application. First, it should be tested and compared to existing models (CAPM and its extensions) by analyzing cross sectional returns of Fama-French-portfolios. In other words, when our model holds, futures research must prove that the risk premia associated with co-skewness and co-kurtosis are significant. Second, it could be used as a performance measure of assets and funds, especially in the field of emerging markets and alternative investments, where skewness and kurtosis are highly significant. Provided that the explanation power of our model is higher than those of other models[1], the constant of the regression (alpha) yields a better measure of the unexplained performance than in former models. Recently, these alphas are receiving more attention and are increasingly applied as additional performance measures to the conventionally used Sharpe-ratios and other instruments. Especially in the world of hedge funds, our model can reveal interesting changes of rankings, since skewness and kurtosis are highly significant in the distributions of alternative investments.

Another possible extension of our model could be the analysis of structural breaks in the investment horizon. The introduction of states of nature in the context of Regime-Switching (Markov) could ameliorate the time-varying description of the factors and reduce the occurring error terms. The autoregressive and heteroscedastic process in every state of nature would probably be a better approximation of the "real" relation. However, a potential problem of an overwhelming number of parameters could arise and constrain future research in this direction.

[1]measured by Akaike's Information Criterion (AIC) or by Schwarz's Criterion (SC)

Appendix A

Pricing of Hedge Funds

According to the U.S. *Securities and Exchange Commission* (SEC) a Hedge fund is a general, non-legal term that was originally used to describe a type of private and unregistered investment pool that employed sophisticated hedging and arbitrage techniques to trade in the corporate equity markets. Hedge funds have traditionally been limited to sophisticated, wealthy investors. Over time, the activities of hedge funds have broadened into other financial instruments and activities. Today, the term "hedge fund" refers not so much to hedging techniques, which hedge funds may or may not employ, but to their status as private and unregistered investment pools.

Hedge funds are similar to mutual funds in that they both are pooled investment vehicles that accept investors' money and generally invest it on a collective basis. Hedge funds differ significantly from mutual funds, however, because hedge funds are not as strictly regulated under the federal securities laws.

Hedge funds also are not subjected to the numerous regulations that apply to mutual funds for the protection of investors such as regulations requiring a certain degree of liquidity, regulations requiring that mutual fund shares be redeemable at any time, regulations protecting against conflicts of interest, regulations to assure fairness in the pricing of fund shares, disclosure regulations, regulations limiting the use of leverage, and more. This freedom from regulation permits hedge funds to engage in leverage and other sophisticated investment techniques to a much greater extent than mutual funds.

The primary aim of most hedge funds is to reduce volatility and risk while attempting to preserve capital and deliver positive returns under all market conditions.

A.1 The different Strategies

The existing hedge fund strategies can be classified in three broad categories:

Arbitrage Strategies

Arbitrage is the exploitation of an observable price inefficiency and, as such, pure arbitrage is considered riskless. For example, *Convertible Arbitrage* entails buying a corporate convertible bond, which can be converted into common shares, while simultaneously selling short the underlying stock of the same company that issued the bond. This strategy tries to exploit the relative prices of the convertible bond and the stock: the arbitrageur of this strategy would think the bond is a little cheap and the stock is a little expensive.

The idea is to make money from the bond's yield if the stock goes up but also make money from the short sale if the stock goes down.

Event Driven Strategies

Event-driven strategies take advantage of transaction announcements and other one-time events. One example is *Merger Arbitrage*, which is used in the event of an acquisition announcement and involves buying the stock of the target company and hedging the purchase by selling short the stock of the acquiring company.

Another example of Event Driven Strategies are *Distressed Securities* funds, that invest in securities (equity and/or debt) of a company either already in bankruptcy or facing it. These securities are purchased by the investor inexpensively. It is hoped that as the company emerges from bankruptcy the securities will appreciate.

Directional or Tactical Strategies

The largest group of hedge funds uses directional or tactical strategies. One example is the *Global Macro* fund that aim to profit from major economic trends and events in the global economy, such as large currency and interest shifts. Macro managers employ a "top-down" global approach, and may invest in any markets using any instruments to participate in expected market movements. These movements may result from forecasted shifts in world economies, political fortunes or global supply and demand for resources, both physical and financial. Exchange-traded and over-the-counter derivatives are often used to magnify these price movements.

Long/short strategies combine purchases (long positions) with short sales. For example, a long/short manager might purchase a portfolio of core stocks that occupy the S&P 500 and hedge by selling (shorting) S&P 500 Index futures. If the S&P 500 goes down, the short position will offset the losses in the core portfolio, limiting overall losses.

Market neutral strategies are a specific type of long/short whose goal is to negate the impact and risk of general market movements, trying to isolate the pure returns of individual stocks. This type of strategy is a good example of how hedge funds can aim for positive, absolute returns even in a bear market.

Emerging Markets Fund invests in securities of companies in developing, or, emerging countries. The strategy consists of purchasing sovereign or corporate debt and/or equity in such countries.

Fixed Income Arbitrage is a market neutral hedging strategy that seeks to profit by exploiting pricing inefficiencies between related fixed income securities while neutralizing exposure to interest rate risk. Fixed Income Arbitrage is a generic description of a variety of strategies involving investment in fixed income instruments, and weighted in an attempt to eliminate or reduce exposure to changes in the yield curve. Managers attempt to exploit relative mispricing between related sets of fixed income securities. The generic types of fixed income hedging trades include: yield-curve arbitrage, corporate versus Treasury yield spreads, municipal bond versus Treasury yield spreads and cash versus futures.

Emerging Markets funds invest in securities of companies or the sovereign debt of developing or "emerging" countries. Investments are primarily long. "Emerging Markets" include countries in Latin America, Eastern Europe, the former Soviet Union, Africa and parts of Asia. Emerging Markets - Global funds will shift their weightings among these regions according to market conditions and manager perspectives. In addition, some managers invest solely in individual regions.

A.2 Data

In order to analyze the different strategies presented in the previous section we use the *CSFB/Tremont* indices for each strategy respectively. The table (A.1) shows the first four moments of the empirical distribution of the corresponding indices for the period of January 1994 to December 2003.

	CONVARB	DISTRESSED	EMERGEMA	FIXINCOME	GLMACRO	LONGSHORT	MANEUTRAL	MERGARB
Mean	0,0084	0,0107	0,0071	0,0056	0,0119	0,0101	0,0085	0,0068
Std.Dev.	0,0138	0,0202	0,0513	0,0114	0,0350	0,0318	0,0089	0,0129
Skewness	-1,5532	-2,7147	-0,5713	-3,2135	-0,0372	0,2154	0,2100	-1,3206
Kurtosis	6,8448	18,4340	6,5098	18,8755	4,8483	6,1637	3,1791	8,9544
Jarque-Bera	122,16	1338,43	68,12	1466,69	17,11	50,97	1,04	212,16
Prob.	0	0	0	0	0,000193	0	0,593863	0

Table A.1: Distribution statistics of strategy indices

This table underlines the already mentioned departure from normality which is characteristic for hedge funds. The Jarque-Bera test for normality indicates that with exception of the *Market Neutral* fund all hedge funds distributions differ significantly from the normal distribution. This stresses the need to take higher moments into account and reinforces our model.

Appendix B

MATLAB Source Codes

The next section provides MATLAB source codes for the following model stages:

1. bivariate GARCH-in-mean (main1, model1, core1)
2. bivariate GARCHS[1]-in-mean (main1, model2, core2)
3. bivariate GARCHSK[2]-in-mean (main3, model3, core3)

We applied these source codes on hedge funds and tried to parameterize the models. The database CISDM provided the return series of the various hedge funds. Unfortunately, in the framework of this study, we cannot present reliable results. In the first model (bivariate GARCH) the matrix **H** yields values that are 10 times higher than the values of the true covariance matrix. This can either be due to a bug in the source code or to the fact that the disposable 150 observations are not sufficient to test the model. Hence, we decided not to provide the estimated parameters. Future research should correct this problem. For instance, it can be thought of simulation procedures.

[1]GARCH with skewness

[2]GARCH with skewness and kurtosis

```
function [parameters,LogL,Ht]=main1(data)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%         'data' is a bivariate time series
%         simulates different starting values for the 2-moment-model
%         calling 'model1', which calls 'core1'
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

C=cov(data);                %unconditional covariance matrix
g=mean(data(:,2))/C(2,1);   %gamma proxy
for i=1:10                  %simulate 10 different starting value sets
    gamma1=-g;
    par1=gamrnd(2,2,[3 1]);
    par1=par1/sum(par1);
    par2=gamrnd(2,2,[3 1]);
    par2=par2/sum(par2);
    par3=gamrnd(2,2,[3 1]);
    par3=par3/sum(par3);
    par=[par1(1);par2(1);par3(1);par1(2);par2(2);par3(2);par1(3);par2(3);par3(3)];
    alpha=gamrnd(2,2,[2 1]);
    for j=1:10
        starting=[par; alpha; gamma1];
        [para1(10*(i-1)+j,:),LLF1(10*(i-1)+j),H(:,:,10*(i-1)+j),...
        e(:,:,10*(i-1)+j),eta(:,:,10*(i-1)+j)]=model1(data,starting);
        gamma1=gamma1+0.4*g;
        loop=10*(i-1)+j
    end
end
[LogL,ind]=min(LLF1);
parameters=para1(ind,:)';
Ht=ones(size(data,1)+1,3);
Ht(:,:)=H(:,:,ind);
eopt(:,:)=e(:,:,ind);
etaopt(:,:)=eta(:,:,ind);
```

```
function [parameters,LogL,H,e,eta]=model1(data,starting)

options = optimset('fmincon');
options = optimset(options, 'Display', 'iter');
options = optimset(options, 'Diagnostics', 'off');
options = optimset(options, 'LevenbergMarquardt', 'off');
options = optimset(options, 'LargeScale', 'off');
options = optimset(options, 'MaxFunEvals', 1000);
options = optimset(options, 'MaxIter', 20);
options = optimset(options, 'TolX', 13e-06);
options = optimset(options, 'TolFun', 13e-06);
options = optimset(options, 'TolCon', 13e-06);


A=[eye(3) eye(3) zeros(3,6); -eye(6) zeros(6,6)];
B=[ones(3,1);zeros(6,1)];

parameters=fmincon('core1',starting,A,B,[],[],[],[],[],options,data);
[LogL,H,e,eta]=core1(parameters,data);
```

```
function [logL,H,e,eta]=core1(starting,data)

[t,k]=size(data);
a=starting(1:3);
b=starting(4:6);
c=starting(7:9);
alpha1=starting(10);
alpha2=starting(11);
gamma1=starting(12);
Cov=cov(data);
H=zeros(t+1,3);
H(1,1)=Cov(1,1);
H(1,2)=Cov(1,2);
H(1,3)=Cov(2,2);
e=zeros(t,2);
eta=zeros(t,2);
logL=0;
likelihoods=zeros(t,1);
for i=1:t
    e(i,1)=data(i,1)-alpha1-gamma1*H(i,2);
    e(i,2)=data(i,2)-alpha2-gamma1*H(i,3);
    he=e;
    he(1,:)=[0 0];
    c=[0;0;0];
    H(i+1,1:3)=c'+a'.*[he(i,1)^2 he(i,1)*he(i,2) he(i,2)^2]+b'.*H(i,1:3);
    if (isinf(H(i,1:3))==[0 0 0]) & (isnan(H(i,1:3))== [0 0 0])
    Cov=[H(i,1) H(i,2); H(i,2) H(i,3)];
    eta(i,:)=he(i,:)*Cov^-0.5;
    likelihoods(i)=log(2*pi)+eta(i,:)*eta(i,:)'/2;
    logL=logL+likelihoods(i);
    else logL=10000000;
    end
end
```

```
function [parameters,LogL,Ht]=main2(data)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%         'data' is a bivariate time series
%         simulates different starting values for the 3-moment-model
%         calling 'model2', which calls 'core2'
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

C=cov(data);                %unconditional covariance matrix
g=mean(data(:,2))/C(2,1);   %gamma proxy
for i=1:10                  %simulate 10 different starting value sets
    gamma1=-g;
    par1=gamrnd(2,2,[3 1]);
    par1=par1/sum(par1);
    par2=gamrnd(2,2,[3 1]);
    par2=par2/sum(par2);
    par3=gamrnd(2,2,[3 1]);
    par3=par3/sum(par3);
    par4=gamrnd(2,2,[3 1]);
    par4=par4/sum(par4);
    par5=gamrnd(2,2,[3 1]);
    par5=par5/sum(par5);
    par6=gamrnd(2,2,[3 1]);
    par6=par6/sum(par6);
    par7=gamrnd(2,2,[3 1]);
    par7=par7/sum(par7);
    par8=gamrnd(2,2,[3 1]);
    par8=par8/sum(par8);
    par9=gamrnd(2,2,[3 1]);
    par9=par9/sum(par9);
    par=[par1(1);par2(1);par3(1);par4(1); par5(1); par6(1);par7(1);par8(1);par9(1)];
    par=[par;par1(2);par2(2);par3(2);par4(2); par5(2); par6(2);par7(2);par8(2);par9(2)];
    par=[par;par1(3);par2(3);par3(3);par4(3); par5(3); par6(3);par7(3);par8(3);par9(3)];
    alpha=gamrnd(2,2,[2 1]);
    gamma2=0;
    for j=1:10
        starting=[par; alpha; gamma1; gamma2];
        [para2(10*(i-1)+j,:),LLF2(10*(i-1)+j),H(:,:,10*(i-1)+j),e(:,:,10*(i-1)+j),...
        eta(:,:,10*(i-1)+j)]=model2(data,starting);
        gamma1=gamma1+0.4*g;
        loop=10*(i-1)+j
    end
end [LogL,ind]=min(LLF2);
parameters=para2(ind,:)';
Ht=ones(size(data,1)+1,3);
Ht(:,:)=H(:,:,ind);
eopt(:,:)=e(:,:,ind);
etaopt(:,:)=eta(:,:,ind);
```

```
function [parameters,LogL,H,e,eta]=model2(data,starting)

options = optimset('fmincon');
options = optimset(options, 'Display', 'iter');
options = optimset(options, 'Diagnostics', 'off');
options = optimset(options, 'LevenbergMarquardt', 'off');
options = optimset(options, 'LargeScale', 'off');
options = optimset(options, 'MaxFunEvals', 1000);
options = optimset(options, 'MaxIter', 20);
options = optimset(options, 'TolX', 13e-06);
options = optimset(options, 'TolFun', 13e-06);
options = optimset(options,'TolCon', 13e-06);

A=[eye(9) eye(9) zeros(9,13); -eye(18) zeros(18,13)];
B=[ones(9,1);zeros(18,1)];

parameters=fmincon('core2',starting,A,B,[],[],[],[],[],options,data);
[LogL,H,e,eta]=core2(parameters,data);
```

```
function [parameters,LogL,Ht]=main3(data)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%         'data' is a bivariate time series
%         simulates different starting values for the 4-moment-model
%         calling 'model3', which calls 'core3'
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

C=cov(data);                %unconditional covariance matrix
g=mean(data(:,2))/C(2,1);   %gamma proxy
for i=1:10                  %simulate 10 different starting value sets
    gamma1=-g;
    par1=gamrnd(2,2,[3 1]);
    par1=par1/sum(par1);
    par2=gamrnd(2,2,[3 1]);
    par2=par2/sum(par2);
    par3=gamrnd(2,2,[3 1]);
    par3=par3/sum(par3);
    par4=gamrnd(2,2,[3 1]);
    par4=par4/sum(par4);
    par5=gamrnd(2,2,[3 1]);
    par5=par5/sum(par5);
    par6=gamrnd(2,2,[3 1]);
    par6=par6/sum(par6);
    par7=gamrnd(2,2,[3 1]);
    par7=par7/sum(par7);
    par8=gamrnd(2,2,[3 1]);
    par8=par8/sum(par8);
    par9=gamrnd(2,2,[3 1]);
    par9=par9/sum(par9);
    par10=gamrnd(2,2,[3 1]);
    par10=par10/sum(par10);
    par11=gamrnd(2,2,[3 1]);
    par11=par11/sum(par11);
    par12=gamrnd(2,2,[3 1]);
    par12=par12/sum(par12);
    par13=gamrnd(2,2,[3 1]);
    par13=par13/sum(par13);
    par14=gamrnd(2,2,[3 1]);
    par14=par14/sum(par14);
    par15=gamrnd(2,2,[3 1]);
    par15=par15/sum(par15);
    par16=gamrnd(2,2,[3 1]);
    par16=par16/sum(par16);
    par=[par1(1);par2(1);par3(1);par4(1); par5(1); par6(1);par7(1);par8(1);...
    par9(1);par10(1);par11(1);par12(1);par13(1);par14(1); par15(1); par16(1)];
    par=[par;par1(2);par2(2);par3(2);par4(2); par5(2); par6(2);par7(2);par8(2);...
    par9(2);par10(2);par11(2);par12(2);par13(2);par14(2); par15(2); par16(2)];
    par=[par;par1(3);par2(3);par3(3);par4(3); par5(3); par6(3);par7(3);par8(3);...
```

```
    par9(3);par10(3);par11(3);par12(3);par13(3);par14(3); par15(3); par16(3)];
    alpha=gamrnd(2,2,[2 1]);
    gamma2=0;
    gamma3=0;
    for j=1:10
        starting=[par; alpha; gamma1; gamma2; gamma3];
        [para3(10*(i-1)+j,:),LLF3(10*(i-1)+j),H(:,:,10*(i-1)+j),e(:,:,10*(i-1)+j),...
        eta(:,:,10*(i-1)+j)]=model3(data,starting);
        gamma1=gamma1+0.4*g;
        loop=10*(i-1)+j
    end
end [LogL,ind]=min(LLF3);
parameters=para3(ind,:)';
Ht=ones(size(data,1)+1,3);
Ht(:,:)=H(:,:,ind);
eopt(:,:)=e(:,:,ind);
etaopt(:,:)=eta(:,:,ind);
```

```
function [parameters,LogL,H,e,eta]=model3(data,starting)

options = optimset('fmincon');
options = optimset(options, 'Display', 'iter');
options = optimset(options, 'Diagnostics', 'off');
options = optimset(options, 'LevenbergMarquardt', 'off');
options = optimset(options, 'LargeScale', 'off');
options = optimset(options, 'MaxFunEvals', 1000);
options = optimset(options, 'MaxIter', 20);
options = optimset(options, 'TolX', 13e-06);
options = optimset(options ,'TolFun', 13e-06);
options = optimset(options ,'TolCon', 13e-06);

A=[eye(16) eye(16) zeros(16,21); -eye(32) zeros(32,21)];
B=[ones(16,1);zeros(32,1)];

parameters=fmincon('core3',starting,A,B,[],[],[],[],[],options,data);
[LogL,H,e,eta]=core3(parameters,data);
```

```
function [logL,H,e,eta]=core3(starting,data)

[t,k]=size(data);
a=starting(1:16);
b=starting(17:32);
c=starting(33:48);
alpha1=starting(49);
alpha2=starting(50);
gamma1=starting(51);
gamma2=starting(52);
gamma3=starting(53);

H=zeros(t+1,16);

est1=(data(:,1)-mean(data(:,1)));
est2=(data(:,2)-mean(data(:,2)));
Cov=cov(data);
H(1,1)=Cov(1,1);
H(1,2)=Cov(1,2);
H(1,3)=Cov(2,2);
H(1,4)=mean(est1.*(est2.^2));
H(1,5)=mean(est2.^3);
H(1,6)=mean(est1.*(est2.^3));
H(1,7)=mean(est2.^4);
H(1,12)=3;
H(1,14)=1;
H(1,16)=3;

e=zeros(t,2);
eta=zeros(t,2);
her=zeros(9,1);

logL=0;
likelihoods=zeros(t,1);
for i=1:t
    e(i,1)=data(i,1)-alpha1-gamma1*H(i,1)-gamma2*H(i,4)-gamma3*H(i,6);
    e(i,2)=data(i,2)-alpha2-gamma1*H(i,2)-gamma2*H(i,5)-gamma3*H(i,7);
    H(i+1,1:7)=c(1:7)'+a(1:7)'.*[e(i,1)^2 e(i,1)*e(i,2) e(i,2)^2 e(i,1)*e(i,2)^2...
                            e(i,2)^3 e(i,1)*e(i,2)^3 e(i,2)^4]+b(1:7)'.*H(i,1:7);
    if (isinf(H(i,1:7))==[0 0 0 0 0 0 0]) & (isnan(H(i,1:7))== [0 0 0 0 0 0 0])
        Cov=[H(i,1) H(i,2); H(i,2) H(i,3)];
        eta(i,:)=e(i,:)*Cov^-0.5;
        eta2=eta(i,:)'*eta(i,:);
        eta3=kron(eta2,eta(i,:));
        eta4=kron(eta3,eta(i,:));
        H(i+1,8:16)=c(8:16)'+a(8:16)'.*[eta3(:,1)' eta3(:,4)' eta4(:,1)' eta4(1,4)...
                                                  eta4(:,8)']+b(8:16)'.*H(i,8:16);
        d=H(i,8:16);
```

```
        d(5)=d(5)-3;
        d(7)=d(7)-1;
        d(9)=d(9)-3;
        her(1)=eta3(1,1)-3*eta(1);
        her(2)=eta3(2,1)-eta(2);
        her(3)=eta3(1,4)-eta(1);
        her(4)=eta3(2,4)-3*eta(2);
        likelihoods(i)=t*log(2*pi)+eta(i,:)*eta(i,:)'/2+log(1+d*her);
        logL=logL+likelihoods(i);
    else logL=10000000;
    end
end
```

Bibliography

Amenc, N. and V. Le Sourd (2003). *Portfolio Theory and Performance Analysis.* Wiley.

Black (1972). "Capital Market Equilibrium with restricted borrowing". *Journal of Business 45*, 444–455.

Blinnikov, S. and R. Moessner (1998). "Expansions for nearly Gaussian distributions". *Astronomy and Astrophysics Supplement Series* (130), 193–205.

Bollerslev, T. (1986). "Generalised Autoregressive Conditional Heteroscedasticity". *Journal of Econometrics 31*, 307–327.

Bollerslev, T., R. Engle, and J. Wooldridge (1988). "A Capital Asset Pricing Model with Time-varying Covariances". *Journal of Political Economics 96*(1), 116–131.

Breeden, D. (2003). "An Intertemporal Asset Pricing Model with Stochastic Concumption and Investment Opportunities". *Journal of Financial Economics 7*(3), 265–297.

Brown, C. and D. Robinson (2002). "Skewness and Kurtosis implied by Option Prices: A Correction". *Journal of Financial Research 25*(2), 279–282.

Campbell, H., J. Liechty, W. Liechty, and P. Müller (2004). "Portfolio Selection with Higher Moments".

Campbell, H. and A. Siddique (1999). "Autoregressive Conditional Skewness". *Journal of Financial and Quantitative Analysis.*

Campbell, J. and J. Cochrane (2000). "Explaining the Poor Performance of Consumption-based Asset Pricing Models". *Journal of Finance 55*(6).

Campbell, J., A. Lo, and C. MacKinlay (1997). *The Econometrics of Financial Markets.* Princeton University Press.

Campbell, J. and L. Viceira (2001). "Strategic Asset Allocation: Portfolio Choice for Long-Term Investors".

Cass, D. and J. Stiglitz (1970). "The Structure of Investor Preferences and Asset Returns, and Separability in Portfolio Allocation: A Contribution to the pure theory of Mutual Funds". *Journal of Political Economy.*

Chamberlain, G. (1983). "A Characterisation of the Distributions that imply Mean-Variance Utility Functions". *Journal of Economic Theory 29*(1), 185–201.

Chiang, K., M.-L. Lee, and C. Wisen (2004). "Another Look at the Asymmetric REIT-Beta Puzzle". *JRER 26.*

Cochrane, J. (2001). *"Asset Pricing"*. Princeton University Press.

Cvitanic, J., A. Lazrak, L. Martellini, and F. Zapatero (2003). "Optimal Allocation to Hedge Funds: an Empirical Analysis". *Quantitative Finance 3*, 1–12.

Cvitanic, J., A. Lazrak, L. Martellini, and F. Zapatero (2004). "Dynamic Portfolio Choice with Parameter Uncertainty and the Economic Value of Analysts'".

Davies, R., H. Kat, and S. Lu (2003). "Fund of Hedge Funds Portfolio Selection: A Multiple Objective Approach".

Elton, E. and M. Gruber (1995). *"Modern Portfolio Theory"*.

Engle, R. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation". *Econometrica 50*, 987–1007.

Engle, R., D. Lilien, and R. Robins (1987). "Estimating time varying risk premia in the term structure: The Arch-M model". *Econometrica 55*, 391–407.

Fama, E. and K. French (1993). "Common Risk Factors in the Returns on Stocks and Bonds". *Journal of Financial Economics 33*, 3–56.

Fang, H. and T.-Y. Lai (1997). "Co-Kurtosis and Capital Asset Pricing". *Financial Review 32*(2), 293–307.

Guidolin, M. and A. Timmermann (2002). "Size and Value Anomalies under Regime Shifts".

Gupta, B., B. Cerrahoglu, and A. Daglioglu (2003). "Hedge Fun Strategy Performance: Using Conditional Approaches".

Hafner, C. and H. Herwatz (1999). "Time-Varying Market Price of Risk in the CAPM-Approaches, Emprirical Evidence and Implications".

Hansen, L. (1982). "Large Sample Properties of Generalized Method of Moments Estimators". *Econometrica 50*, 1029–1054.

Horvath, P. and R. Scott (1980). "On the Direction of Preference for Moments of higher order than the Variance". *Journal of Finance 35*, 915–919.

Hwang, S. and S. Satchell (1999). "Modelling Emerging Risk Premia Using Higher Moments".

Ingersoll, J. (1987). *"Theory in Financial Decision Making"*. Rowman and Littlefield.

Jagannathan, R. and Z. Wang (1996). "The Conditional CAPM and the Cross-Section of Expected Returns". *Journal of Finance*.

Jensen, M. (1972). "Capital Market: Theory and Evidence". *The Bell Journal of Economics and Management Science 3*(2), 357–398.

Jiang, G. and J. Knight (2001). "Efficient Estimation of Markov Models where the Transition Density is unknown".

Jondeau, E. and M. Rockinger (2004). "Optimal Portfolio Allocation Under Higher Moments".

Kraus, A. and R. Litzenberger (1976). "Skewness Preference and the valuation of Risk Assets". *Journal of Finance 31*(4).

Krouwenberg, R. "Do Hedge Funds Add Value to A Passive Portfolio?".

Leon, A., G. Rubio, and G. Serna (1996). "Skewness and Kurtosis in SP 500 index returns implied by Option Prices". *Journal of Financial Research 19*(2), 175–192.

Leon, A., G. Rubio, and G. Serna (2002). "Autoregressive Conditional Volatility, Skewness and Kurtosis".

Lintner, J. (1965). "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolio and Capital Budgets". *Review of Economics and Statistics 47*, 13–37.

Malevergne, Y. and D. Sornette (2002). "Multi-Moments for Portfolio Management: Generalized Capital Asset Pricing Model in Homogeneous and Heterogeneous Markets".

Markowitz, H. (1952). "Portfolio Selection". *Journal of Finance 7*, 77–91.

McCullagh, P. (1987). *"Tensor methods in Statistics"*. Chapman and Hall.

Merton, R. (1973). "An Intertemporal Capital Asset Pricing Model".

Narasimhan, L. and H. Pradhan (2002). "Conditional CAPM and Cross-Sectional Returns - A Study of Indian Securities Market".

Oelker, J. (2004). "Modelle mit generalisierter Heteroskedastie und Anwendungen in der Kapitalmarkttheorie".

Premaratne, G. and A. Bera (2000). "Modeling Asymmetry and Excess Kurtosis in Stock Return Data".

Ranaldo, A. and L. Favre (2003). "How to price Hedge Funds: From Two- to Four-Moment CAPM".

Rohan, C.-D. and M. Chaudhry (2000). "Coskewness and Cokurtosis in Future Markets". *Journal of Financial Economics*.

Roll, R. (1977). "A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory". *Journal of Financial Economics*.

Scleicher, M. (2000). "Time-varying risk in the German stock market". *The European Journal of Finance 6*, 70–91.

Sharpe, W. (1964). "Capital Asset Prices: A theory of Market Equilibrium under conditions of risk". *Journal of Finance 19*, 425–442.

Simaan, Y. (1993). "Portfolio Selection and Asset Pricing - Three Parameter Framework". *Management Science 39*(5), 568–577.

Hiermit erkläre ich, dass ich die vorliegende Arbeit allein und nur unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Volker Ziemann,

Berlin, 14.10.2004