The Dynamics of Pricing Kernels and Relative Risk Aversion

Master Thesis submitted to

Prof. Dr. Wolfgang Härdle
Institute for Statistics and Econometrics
CASE - Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin

by

Michael Handel
(189036)

in partial fulfillment of the requirements
for the degree of
Master of Science in Economics and Management Science

Berlin, December 8th 2005
ABSTRACT

Risk management and the thorough understanding of the relations between financial markets and the standard theory of macroeconomics have always been among the topics most addressed by researchers, both financial mathematicians and economists. This work aims at explaining investors’ behavior from a macroeconomic aspect (modeled by the investors’ pricing kernel and their relative risk aversion) using stocks and options data. Daily estimates of investors’ pricing kernel (PK) and relative risk aversion (RRA) are obtained and used to construct and analyze a three-year long time-series. The first four moments of these time-series as well as their values at the money are the starting point of a principal component analysis. The relation between changes in a major index level and implied volatility at the money and between the principal components of the changes in RRA is found to be linear. The relation of the same explanatory variables to the principal components of the changes in PK is found to be log-linear, although this relation is not significant for all of the examined maturities.

Declaration of Authorship

I hereby confirm that I have authored this master thesis independently and without use of others than the indicated sources. All passages which are literally or in general matter taken out of publications or other sources are marked as such.

Berlin, December 8th 2005

Michael Handel
## CONTENTS

1. Introduction .................................................. 11

2. The Pricing Kernel in Macroeconomic Asset-Pricing Models .... 14
   2.1 Defining the Pricing Kernel .................................. 17
   2.2 Defining the Investors’ Relative Risk Aversion .......... 18
   2.3 Investors with Power Utility .............................. 19

3. The Pricing of Options .......................................... 23
   3.1 Black and Scholes and Implied Volatility ................. 23
   3.2 The Pricing Kernel under Black and Scholes Assumptions .. 26

   4.1 Subjective Probability Estimation .......................... 28
   4.2 SPD Estimation ............................................ 30

5. A Static Model: Daily Estimation ............................... 34
   5.1 The Database ............................................. 34
   5.2 Subjective Density Estimation ............................ 35
   5.3 State-Price Density Estimation .......................... 38
   5.4 Deriving the Pricing Kernel and Risk Aversion ........... 41

6. A Dynamic Model: Time-Series Analysis ......................... 45
   6.1 Moments of the Pricing Kernel and Relative Risk Aversion . 45
LIST OF TABLES

6.1 Descriptive statistics: Moments of pricing kernel and relative risk aversion ........................................... 52
6.2 Descriptive statistics: Differences of the moments of pricing kernel and relative risk aversion ...................... 53
6.3 Descriptive statistics: Differences of the logarithms of the moments of pricing kernel and relative risk aversion 54
6.4 KPSS test for stationarity of the mean pricing kernel ................................................................................. 56
6.5 Conditions for stationarity of pricing kernel and relative risk aversion ...................................................... 57
6.6 Pricing kernel differences covariance matrix (τ=60 days) ......................................................................... 58
6.7 Principal components analysis of pricing kernel differences ........................................................................ 59
6.8 Principal components analysis of relative risk aversion differences ........................................................ 60
6.9 Descriptive Statistics: The PCs of the PK differences ................................................................................ 64
6.10 Descriptive Statistics: The PCs of the RRA differences ............................................................................ 64
6.11 Comparison between different time-series models for the PCs .............................................................. 70
LIST OF FIGURES

3.1 Butterfly spread ................................................... 25

5.1 Estimated subjective density for different maturities ........... 37
5.2 Estimated SPD for different maturities .......................... 40
5.3 Estimated pricing kernel for different maturities ............... 42
5.4 Estimated RRA for different maturities ........................... 43

6.1 Time-series: ATM pricing kernel ................................. 47
6.2 Time-series: Mean of pricing kernel ............................... 47
6.3 Time-series: STD of pricing kernel ................................. 48
6.4 Time-series: skewness of pricing kernel ......................... 48
6.5 Time-series: kurtosis of pricing kernel ........................... 49
6.6 Time-series: ATM relative risk aversion ......................... 49
6.7 Time-series: Mean of relative risk aversion ..................... 50
6.8 Time-series: STD of relative risk aversion ....................... 50
6.9 Time-series: skewness of relative risk aversion ............... 51
6.10 Time-series: kurtosis of relative risk aversion ............... 51
6.11 Time-series of pricing kernel and relative risk aversion principal components ........................................... 61
6.12 Correlation of PCs with ΔPK and ΔRA ......................... 63
6.13 ACF and PACF of the PCs of PK differences .................. 67
6.14 ACF and PACF of the PCs of RRA differences ................ 68
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1</td>
<td>Time-series: Differences of ATM pricing kernel</td>
<td>82</td>
</tr>
<tr>
<td>A.2</td>
<td>Time-series: Differences of mean pricing kernel</td>
<td>83</td>
</tr>
<tr>
<td>A.3</td>
<td>Time-series: Differences of STD of pricing kernel</td>
<td>83</td>
</tr>
<tr>
<td>A.4</td>
<td>Time-series: Differences of skewness of pricing kernel</td>
<td>84</td>
</tr>
<tr>
<td>A.5</td>
<td>Time-series: Differences of kurtosis of pricing kernel</td>
<td>84</td>
</tr>
<tr>
<td>A.6</td>
<td>Time-series: Differences of ATM relative risk aversion</td>
<td>85</td>
</tr>
<tr>
<td>A.7</td>
<td>Time-series: Differences of mean relative risk aversion</td>
<td>85</td>
</tr>
<tr>
<td>A.8</td>
<td>Time-series: Differences of STD of relative risk aversion</td>
<td>86</td>
</tr>
<tr>
<td>A.9</td>
<td>Time-series: Differences of skewness of relative risk aversion</td>
<td>86</td>
</tr>
<tr>
<td>A.10</td>
<td>Time-series: Differences of kurtosis of relative risk aversion</td>
<td>87</td>
</tr>
<tr>
<td>A.11</td>
<td>Time-series: Log differences of ATM pricing kernel</td>
<td>87</td>
</tr>
<tr>
<td>A.12</td>
<td>Time-series: Log differences of mean pricing kernel</td>
<td>88</td>
</tr>
<tr>
<td>A.13</td>
<td>Time-series: Log differences of STD of pricing kernel</td>
<td>88</td>
</tr>
<tr>
<td>A.14</td>
<td>Time-series: Log differences of skewness of pricing kernel</td>
<td>89</td>
</tr>
<tr>
<td>A.15</td>
<td>Time-series: Log differences of kurtosis of pricing kernel</td>
<td>89</td>
</tr>
<tr>
<td>A.16</td>
<td>Time-series: Log differences of mean relative risk aversion</td>
<td>90</td>
</tr>
<tr>
<td>A.17</td>
<td>Time-series: Log differences of STD of relative risk aversion</td>
<td>90</td>
</tr>
<tr>
<td>A.18</td>
<td>Time-series: Log differences of skewness of relative risk aversion</td>
<td>91</td>
</tr>
<tr>
<td>A.19</td>
<td>Time-series: Log differences of kurtosis of relative risk aversion</td>
<td>91</td>
</tr>
</tbody>
</table>
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation or Symbol</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>def</td>
<td>is defined as</td>
</tr>
<tr>
<td>,</td>
<td>derivative, ( u'(x) ) is the derivative of ( u ) with respect to ( x )</td>
</tr>
<tr>
<td>ACF</td>
<td>autocorrelation function</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike information criterion</td>
</tr>
<tr>
<td>AR</td>
<td>autoregressive process</td>
</tr>
<tr>
<td>ARMA</td>
<td>autoregressive moving average process</td>
</tr>
<tr>
<td>ATM</td>
<td>at the money</td>
</tr>
<tr>
<td>B&amp;S</td>
<td>Black &amp; Scholes (1973)</td>
</tr>
<tr>
<td>( c_t )</td>
<td>consumption at time ( t )</td>
</tr>
<tr>
<td>( C )</td>
<td>price of a call option</td>
</tr>
<tr>
<td>( Corr(X,Y) )</td>
<td>correlation of two random variables ( X ) and ( Y )</td>
</tr>
<tr>
<td>( Cov(X,Y) )</td>
<td>covariance of two random variables ( X ) and ( Y )</td>
</tr>
<tr>
<td>( \delta )</td>
<td>dividend yield</td>
</tr>
<tr>
<td>( E[X] )</td>
<td>expected value of a random variable ( X )</td>
</tr>
<tr>
<td>( E^Q[X] )</td>
<td>expected value of ( X ) under the risk-neutral probability measure</td>
</tr>
<tr>
<td>( E_t[X_{t+1}] )</td>
<td>expected value of ( X ) at time ( t+1 ) based on information at time ( t )</td>
</tr>
<tr>
<td>( F_t )</td>
<td>future price of an asset at time ( t )</td>
</tr>
<tr>
<td>GARCH</td>
<td>General Autoregressive Conditional Heteroscedastic</td>
</tr>
<tr>
<td>GLS</td>
<td>general least squares</td>
</tr>
<tr>
<td>( I_n )</td>
<td>( n \times n ) unity matrix</td>
</tr>
<tr>
<td>IV</td>
<td>implied volatility</td>
</tr>
<tr>
<td>IVS</td>
<td>implied volatility surface</td>
</tr>
<tr>
<td>( K )</td>
<td>option’s strike, exercise price</td>
</tr>
<tr>
<td>( K(\cdot) )</td>
<td>kernel function: continuous, bounded and symmetric real function satisfying: ( \int K(u)du = 1 )</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>future moneyness: ( \kappa \overset{\text{def}}{=} K/F_t )</td>
</tr>
</tbody>
</table>
\( Kurt \) kurtosis of a random variable
\( \log, \ln \) natural logarithm
\( \mu \) mean of a random variable
\( M_t(\cdot) \) pricing kernel at time \( t \)
\( \text{MA} \) moving average process
\( \text{MRS} \) marginal rate of substitution
\( \text{N.A.} \) Non applicable
\( N_p(\mu_p, \Sigma) \) multinormal distribution with \( p \times 1 \) mean vector \( \mu_p \) and \( p \times p \) covariance matrix \( \Sigma \)
\( \text{OLS} \) ordinary least squares
\( P_t \) price of an asset at time \( t \)
\( p(S_T|S_t), p_t(S_T) \) subjective density of an asset at maturity based on information at time \( t \)
\( \text{PACF} \) partial autocorrelation function
\( \text{PC} \) principal component
\( \text{PCA} \) principal component analysis
\( \text{PK} \) pricing kernel
\( q(S_T|S_t), q_t(S_T) \) state-price density of an asset at maturity based on information at time \( t \)
\( r \) interest rate
\( \rho_t(\cdot) \) relative risk aversion at time \( t \)
\( \text{RRA} \) relative risk aversion
\( s_t \) state of the world at time \( t \)
\( S_t \) price of an asset at time \( t \)
\( \sigma \) standard deviation of a random variable
\( \Sigma \) covariance matrix
\( \text{SIC} \) Schwarz information criterion
\( \text{STD} \) standard deviation
\( \text{Skew} \) skewness of a random variable
\( \text{SPD} \) state-price density
\( t \) time
\( T \) expiry date of a financial contract
\( T \) transpose, \( X^\top \) is the transpose of \( X \)
\( \tau \) time to maturity of a financial contract \( \tau \stackrel{\text{def}}{=} T - t \)
\( u(\cdot) \) one-period utility function
\( W_t \) Brownian motion
\( \mathbf{W}_t \) Brownian motion under the risk-neutral measure
\( \text{Var}(X) \) variance of a random variable \( X \)
\( \psi(\cdot) \) payoff function of an asset
Dedicated to Mali Lehobye, my better half
1. INTRODUCTION

Risk management has developed in the recent decades to be one of the most fundamental issues in quantitative finance. Various models are being developed and applied by researchers as well as financial institutions. By modeling price fluctuations of assets in a portfolio, the loss can be estimated using statistical methods. Different measures of risk, such as standard deviation of returns or confidence interval Value at Risk, have been suggested. These measures are based on the probability distributions of assets’ returns extracted from the data-generating process of the asset.

However, an actual one dollar loss is not always valued in practice as a one dollar loss. Purely statistical estimation of loss has the disadvantage of ignoring the circumstances of the loss. Hence the notion of an investor’s utility has been introduced. Arrow (1964) and Debreu (1959) were the first to introduce elementary securities to formalize economics of uncertainty. The so-called Arrow-Debreu securities are the starting point of all modern financial asset pricing theories. Arrow-Debreu securities entitle their holder to a payoff of 1$ in one specific state of the world, and 0 in all other states of the world. The price of such a security is determined by the market, on which it is tradable, and is subsequent to a supply and demand equilibrium. Moreover, these prices contain information about investors’ preferences due to their dependence on the conditional probabilities of the state of the world at maturity and the imposition of market-clearing and general equilibrium conditions. The prices reflect investors’ beliefs about the future, and the fact that they are priced differently in different states of the world implies, that a one-dollar gain is not always worth the same, in fact its value is exactly the price of the security.

A very simple security that demonstrates the concept of Arrow-Debreu securities is a European option. An option is a security with a payoff that depends on the performance of an underlying asset. According to Definition 1.3 in Franke et al. (2004), "A European call option is an agreement which gives the holder the right to buy the underlying asset at a specified date $T > t$, (expiration date or maturity), for a specified price $K$, (strike price or
exercise price). If the holder does not exercise, the option expires worthless. A European put option is an agreement which gives the holder the right to sell the underlying asset at a specified date $T$ for a specified price $K^n$. The payoff function of a call option at maturity $T$ is therefore

$$\psi(S_T) = (S_T - K)^+ \overset{\text{def}}{=} \max(S_T - K, 0)$$

(1.1)

where $S_T$ is the asset's price at maturity. The payoff of a put option is respectively

$$\psi(S_T) = (K - S_T)^+$$

(1.2)

Since an option is a state-dependent contingent claim, it can be valued using the concept of Arrow-Debreu securities. Bearing in mind, that Arrow-Debreu prices can be perceived as a distribution (when the interest rate is 0, they are non negative and sum up to one), the option price is the discounted expectation of random payoffs received at maturity. Since the payoff equals the value of the claim at maturity time (to eliminate arbitrage opportunities), the value process is by definition a martingale. Introducing a new probability measure $Q$, such that the discounted value process is a martingale, we can write

$$C_t = e^{-r(T-t)} E^Q_t[\psi(S_T)] = e^{-r(T-t)} \sum_s q_s \psi_s(S_T)$$

(1.3)

where $r$ is the interest rate and $q_s$ is the price of an Arrow-Debreu security if $r = 0$, paying 1$ in state $s$ and nothing in any other state. The superscript $Q$ denotes the expectation based on the risk neutral probability measure, the subscript $t$ means that the expectation is conditioned on the information known at time $t$. The continuous counterpart of the Arrow-Debreu state contingent claims will be defined in the next chapter as the risk-neutral density or in its more commonly used name, the State Price Density (SPD).

Based on the relations between the actual data generating process of a major stock index and its risk-neutral probability measure, we can derive measures that help us learn a lot about investors’ beliefs and get an idea of the forces which drive them. This work aims at investigating the dynamics of investors’ beliefs. The second chapter reviews the classic theory of macroeconomic asset-pricing models and defines the two measures that will be applied and investigated later on in this work: the pricing kernel (PK) and the relative risk aversion (RRA). The third chapter reviews the classic theory of option pricing and demonstrates the calculation of the PK and RRA under the classic option pricing assumptions. The fourth chapter deviates from the classical world and reviews various methods for pricing kernel estimation based on options data, as discussed in the scientific literature. In
chapters 5 and 6 the chosen statistical model and the results are described in detail. Chapter 5 describes the database and the static daily estimation model, whereas Chapter 6 deals with the dynamic time-series, created by estimating the parameters every day for three years of options data. Main conclusions and final statements are given at the end of the work.

The appendices include a graphical representation of the various time-series and an implementation guide to the quantlets written in XploRe. This guide has to be thoroughly read and followed before any quantlet can be executed. Although the XploRe files are linked and can be read, they can not be directly executed, as they require the data files to be available, the paths to be properly specified and the quantlets themselves to be executed in a specific sequence.
2. THE PRICING KERNEL IN MACROECONOMIC ASSET-PRICING MODELS

The distinction between the actual data generating process of an asset and the market valuations is the essence of macroeconomic dynamic equilibrium asset-pricing models, in which market forces and investors’ beliefs are key factors to value an asset with uncertain payoffs.

A standard dynamic exchange economy as discussed by Lucas (1978), Rubinstein (1976) and many others, imposes that securities markets are complete, that they consist of one consumption good and that the investors, which have no exogenous income other than from trading the goods, seek to maximize their state-dependent utility function. There is one risky stock $S_t$ in the economy, corresponding to the market portfolio in a total normalized supply. In addition, the economy is endowed by a riskless bond $B_t$ with a continuously compounded rate of return $r$. The stock price follows the stochastic process

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{2.1}$$

where $\mu$ denotes the drift, $\sigma$ is the volatility and $W_t$ is a standard Brownian motion. The drift and volatility can be functions of the asset price, time and many other factors. For simplicity, they are considered constant in this section. The conditional density of the stock price, which is implied by equation \(2.1\), is denoted by $p_t(S_T|S_t)$. The riskless bond evolves according to

$$dB_t = rB_t dt \tag{2.2}$$

In this setting, due to continuous dividend payments, the discounted process with cumulative dividend reinvestments should be a martingale and is denoted by

$$\bar{S}_t \overset{\text{def}}{=} e^{-(r+\delta)t}S_t \tag{2.3}$$

Since we are dealing with corrected data and in order to simplify the theoretic explanations, we will consider $\delta = 0$ from now on and omit the dividends from the equations.
Taking the total differential yields

\[ d\tilde{S}_t = d(e^{-rt}S_t) = -re^{-rt}S_t dt + e^{-rt}dS_t = -re^{-rt}S_t dt + e^{-rt}[\mu S_t dt + \sigma S_t dW_t] = (\mu - r)\tilde{S}_t dt + \sigma \tilde{S}_t dW_t \]

\[ = \sigma \tilde{S}_t dW_t \tag{2.4} \]

where \( W_t \overset{\text{def}}{=} W_t + \frac{\mu - r}{\sigma}t \) can be perceived as a Brownian motion on the probability space corresponding to the risk-neutral measure \( Q \). The term \( \frac{\mu - r}{\sigma} \) is called the market price of risk, it measures the excess return per unit of risk borne by the investor and hence it vanishes under \( Q \), justifying the name risk-neutral pricing. Risk-neutral pricing can be understood as the pricing done by a risk-neutral investor, an investor who is indifferent to risk and hence not willing to pay the extra premium. The conditional risk-neutral density of the stock price under \( Q \), implied by equation (2.4) and denoted as \( q_t(S_T|S_t) \), is the state-price density which was described as the continuous counterpart of the Arrow-Debreu prices from equation (1.3). The basic theorem of asset pricing states, that absence of arbitrage implies the existence of a positive linear pricing rule (Cochrane (2001)), which in a continuous time setting can be described by the Radon-Nikodym derivative, or a scaled stochastic discount factor (Fengler (2005)), often written as \( \frac{dQ}{dP} \), where the \( Q \) and \( P \) are defined respectively as the risk-neutral and the subjective probability measures. If the market is indeed arbitrage-free, it can be shown that the risk-neutral measure \( Q \) is unique.

In order to relate the subjective and risk-neutral densities to macroeconomic factors, we first need to review some of the basic concepts and definitions of macroeconomic theory. It is well known, that a representative agent with a utility function \( U \) exists, when the following conditions are met (Mas-Colell et al. (1995)):

- The investors’ preference set is complete and transitive. Completeness essentially means that all possible alternatives can be compared to one another and hence can be evaluated or graded. Transitivity simply means that if \( u(s_1) \geq u(s_2) \) and \( u(s_2) \geq u(s_3) \) then \( u(s_1) \geq u(s_3) \), where \( u(s_j) \) is defined as the utility function of the investor in the state of the world \( s_j \).
- The aggregate wealth is the sum of individual wealth functions which are continuous and homogeneous of degree 1.
• The aggregate demand function is continuous, homogeneous of degree 0 and possesses the Walrasian property, meaning that the endowment is fully used at the end of the time horizon.

The original representative agent model includes utility functions which are based on consumption. However, introducing labor income or intermediate consumption was shown not to affect the results significantly. Hence, without loss of generality, we explain the concept of marginal rate of substitution through a simple consumption based asset pricing model. The fundamental desire for more consumption is described by an intertemporal two-periods utility function as

\[ U(c_t, c_{s,t+1}) = u(c_t) + \beta E_t[u(c_{s,t+1})] = u(c_t) + \beta \sum_s u(c_{s,t+1}) p_t(s_{t+1}|s_t) \] (2.5)

where \( s_t \) denotes the state of the world at time \( t \), \( c_t \) denotes the consumption at time \( t \), \( c_{s,t+1} \) denotes consumption at the unknown state of the world at time \( t + 1 \), \( p_t(s_{t+1}|s_t) \) is the probability of the state of the world at time \( t + 1 \) conditioned on information at time \( t \), \( u(c) \) is the one-period utility of consumption and \( \beta \) is a subjective discount factor. We further assume that an agent can buy or sell as much as he wants from an asset with payoff \( \psi_{s,t+1} \) at price \( P_t \). If \( Y_t \) is the agent’s wealth (endowment) at \( t \) and \( \xi \) is the amount of asset he chooses to buy, then the optimization problem is

\[
\max_{\{\xi\}} \{ u(c_t) + E_t[\beta u(c_{s,t+1})] \}
\]

subject to

\[
\begin{align*}
  c_t &= Y_t - P_t \cdot \xi \\
  c_{s,t+1} &= Y_{s,t+1} + \psi_{s,t+1} \cdot \xi
\end{align*}
\]

The first constraint is the budget constraint at time \( t \), the agent’s endowment at time \( t \) is divided between his consumption and the amount of asset he chooses to buy. The budget constraint at time \( t + 1 \) sustains the Walrasian property, i.e. the agent consumes all of his endowment and asset’s payoff at the last period. The first order condition of this problem yields

\[
P_t = E_t \left[ \beta \frac{u'(c_{s,t+1})}{u'(c_t)} \psi_{s,t+1} \right] \quad (2.6)
\]

We define \( \text{MRS}_t \) as the Marginal Rate of Substitution at \( t \), meaning the rate at which the investor is willing to substitute consumption
at \( t + 1 \) for consumption at \( t \). If consumption at \( t + 1 \) depends on the state of the world (which is the case discussed here), the MRS is referred to as a stochastic discount factor. Cochrane (2001) defines the stochastic discount factor as a random variable that generates today’s prices from future’s payoffs and reflects investors’ preferences for payoffs over different states of the world.

### 2.1 Defining the Pricing Kernel

Famous works like Lucas (1978) or Merton (1973) address the asset pricing models in a more general manner. The utility function depends on the agent’s wealth \( Y_t \) at time \( t \) and the payoff function depends on the underlying asset \( S_t \). According to Merton (1973), in equilibrium, the optimal solution is to invest in the risky stock at every \( t < T \) and then consume the final value of the stock, i.e. \( Y_t = S_t \) for \( \forall t < T \) and \( Y_T = S_T = c_T \). This is a multi-period generalization of the model introduced in the previous section, where period \( T \) corresponds to \( t + 1 \) in the previous section. Defining time to maturity as \( \tau \stackrel{\text{def}}{=} T - t \), the date \( t \) price of an asset with a liquidating payoff of \( \psi(S_T) \) is path independent, as the marginal utilities in the periods prior to maturity cancel out

\[
P_t = e^{-r\tau} \int_0^\infty \psi(S_T) \lambda \frac{U'(S_T)}{U'(S_t)} p_t(S_T|S_t) dS_T
\]

where \( \lambda e^{-r\tau} = \beta \) to correspond to equation (2.6) and \( \lambda \) being a constant independent of index level, for scaling purposes.

Considering the call option price under the risk-neutral probability measure in equation (1.3) and the existence of a positive linear pricing rule (the Radon-Nikodym derivative) in the absence of arbitrage, we argue that the price of any asset can be expressed as a discounted expected payoff, discounted at the risk-free rate, if we calculate the expectation with respect to the risk-neutral density. Since a risk-neutral agent always has the same marginal utility of wealth, the ratio of marginal utilities in equation (2.7) vanishes under \( Q \), and equation (2.7) can be rewritten as

\[
P_t = e^{-r\tau} \int_0^\infty \psi(S_T) q_t(S_T|S_t) dS_T = e^{-r\tau} E_t^Q[\psi(S_T)]
\]

where \( q_t(S_T|S_t) \) is the State Price Density and the expectation \( E_t^Q[\psi(S_T)] \) is taken with respect to the risk-neutral probability measure \( Q \) and not the subjective probability measure, thus reflecting an objective belief about the future states of the world.
Combining equations (2.7) and (2.8) we can define the pricing kernel $M_t(S_T)$, which relates to the the state price density $q_t(S_T|S_t)$, the subjective probability and the utility function as

\[ M_t(S_T) \overset{\text{def}}{=} \frac{q_t(S_T|S_t)}{p_t(S_T|S_t)} = \lambda \frac{U''(S_T)}{U'(S_t)} \] (2.9)

and therefore $\text{MRS}_t = e^{-r\tau} E_t[M_t(S_T)]$. Substituting out the $q_t(S_T|S_t)$ in equation (2.8) using equation (2.9) yields the Lucas asset pricing equation:

\[ P_t = e^{-r\tau} E_t^Q[\psi(S_T)] \]
\[ = e^{-r\tau} \int_0^\infty M_t(S_T) \cdot \psi(S_T)p_t(S_T|S_t)dS_T \]
\[ = e^{-r\tau} E_t[M_t(S_T) \cdot \psi(S_T)] \] (2.10)

2.2 Defining the Investors’ Relative Risk Aversion

The dependence of the pricing kernel on the investor’s utility function has urged researchers to try and estimate distributions based on various utility functions. Arrow (1965) and Pratt (1964) showed a connection between the pricing kernel and the representative agent’s measure of risk aversion. The agent’s risk aversion is a measure of the curvature of the agent’s utility function. The higher the agent’s risk aversion is, the more curved his utility function becomes. If the agent were risk-neutral, the utility function would be linear. In order to keep a fixed scale in measuring the risk aversion, the curvature is multiplied by the level of the asset (the argument of the utility function), i.e. the representative agent’s coefficient of Relative Risk Aversion (RRA) is defined as

\[ \rho_t(S_T) \overset{\text{def}}{=} -\frac{S_T u''(S_T)}{u'(S_T)} \] (2.11)

According to equation (2.9), the pricing kernel is related to the marginal utilities as

\[ M_t(S_T) = \lambda \frac{U''(S_T)}{U'(S_t)} \]
\[ \Rightarrow M_t'(S_T) = \lambda \frac{U''(S_T)}{U'(S_t)} \] (2.12)

Substituting out the first and second derivatives of the utility function in equation (2.11) using equation (2.12) yields

\[ \rho_t(S_T) = -\frac{S_T \lambda M_t'(S_T)U'(S_t)}{\lambda M_t(S_T)U''(S_t)} = -\frac{S_T M_t'(S_T)}{M_t(S_T)} \] (2.13)
Using equation (2.9) we can express the RRA as

\[
\rho_{t}(S_{T}) = -S_{T}\left[\frac{q_{t}(S_{T}|S_{t})}{p_{t}(S_{T}|S_{t})}\right]'
\]

\[
= -S_{T}\left[\frac{q_{t}'(S_{T}|S_{t})p_{t}(S_{T}|S_{t}) - p_{t}'(S_{T}|S_{t})q_{t}(S_{T}|S_{t})}{q_{t}(S_{T}|S_{t})}\right]
\]

\[
= S_{T}\left[\frac{p_{t}'(S_{T}|S_{t})}{p_{t}(S_{T}|S_{t})} - \frac{q_{t}'(S_{T}|S_{t})}{q_{t}(S_{T}|S_{t})}\right]
\]

(2.14)

We now have a method of deriving the investor’s pricing kernel and his risk aversion just by knowing, or being able to estimate, the subjective and the risk-neutral densities. As an example, we consider the popular power utility function, whose properties are defined in the following section.

2.3 Investors with Power Utility

Rubinstein (1976) showed that if the utility function is taken to be a power utility of consumption, i.e.

\[
\begin{array}{ll}
    {u}(c_{t}) = \left\{ \begin{array}{ll}
        \frac{1}{1-\gamma}c_{t}^{1-\gamma} & \text{for } 0 < \gamma \neq 1 \\
        \log(c_{t}) & \text{for } \gamma = 1
    \end{array} \right.
\end{array}
\]

(2.15)

then aggregate consumption is proportional to aggregate wealth, corresponding to the utility of wealth or asset prices discussed above. It can be seen, that as \( \gamma \to 0 \) the utility is reduced to a linear function. The logarithmic utility function when \( \gamma = 1 \) is obtained by applying the L’Hospital rule.

The marginal rate of substitution of an investor with a power utility function is

\[
\text{MRS}_{t} = \beta \mathbb{E}_{t}\left[\frac{u'(cT)}{u'(c_{t})}\right] = \beta \mathbb{E}_{t}\left[\left(\frac{cT}{c_{t}}\right)^{-\gamma}\right]
\]

(2.16)

which means, that it is a function of consumption growth and it is easy to relate it to empirical data. The relative risk aversion of an investor with a power utility can be calculated using equation (2.11), with consumption instead of wealth as an argument, as the utility function is utility of consumption

\[
\rho(c_{T}) = -c_{T}\frac{\gamma(c_{T})^{-\gamma-1}}{(c_{T})^{-\gamma}} = \gamma
\]

(2.17)
This equation shows that the RRA turns out to be a constant, and for the logarithmic utility case, the risk aversion is 1. The next chapter describes the Black & Scholes (1973) model and shows that it corresponds to investors with power utility. The remaining part of this section deals with some problematic properties and implications of the power utility function.

Projecting equation (2.16) onto asset prices, the MRS and thus the pricing kernel $M_t(S_T)$ are nonlinear functions of the return on the underlying asset measured over the life of the option. The two-period version of the Lucas asset pricing equation (2.10) can be written as

$$1 = E_t[M_{t,t+1}R_{t+1}]$$ (2.18)

where $M_{t,t+1} \overset{\text{def}}{=} \frac{1}{1+r_{t,t+1}} M_t(S_{t+1})$ is the one-period pricing kernel including the discount rate $r_{t,t+1}$ for convenience and $R_{t+1} \overset{\text{def}}{=} \psi_{t+1}/P_t$ is the gross return on the asset. Specific important case is when a risk-free asset is discussed. Such asset necessarily has $\psi_{t+1} = P_{t+1}$, and substituting in equation (2.18) leads to the conclusion that the expectation of the pricing kernel is the inverse gross risk-free interest rate

$$E_t[M_{t,t+1}] \overset{\text{def}}{=} (1 + r_{t,t+1})^{-1}$$ (2.19)

Furthermore, we can use a covariance decomposition to write equation (2.18) as

$$1 = E_t[M_{t,t+1}] \cdot E_t[R_{t+1}] + \text{Corr}_{M,R} \cdot \sigma_R \cdot \sigma_M$$ (2.20)

Limiting the correlation between $-1$ and 1 and using the relation in equation (2.19) enable us to bound the asset expected return as follows

$$E_t[R_{t+1}] \in [(1 + r_{t,t+1}) ± (1 + r_{t,t+1})\sigma_R \cdot \sigma_M]$$ (2.21)

where $\sigma_M$ stands for the standard deviation of the pricing kernel. All the variables in equation (2.21) can be estimated empirically, except $\sigma_M$. However, Cochrane (2001) shows that if consumption growth is assumed to be log-normally distributed (an assumption which sustains the B&S model), then $\sigma_M \approx \frac{1}{1+r_{t,t+1}} \cdot \gamma \sigma_{c_{t+1}/c_t}$ and thus all the parameters can be estimated empirically.

If the assumptions were correct, we would expect all assets’ expected returns to be bounded according to equation (2.21). This relation, however, is not supported empirically and was referred to by Mehra & Prescott (1985) as the Equity Premium Puzzle. Actual asset returns are outside the bounded region, usually yielding much higher returns than predicted by equation (2.21).
To solve this puzzle, Weil (1989) tried to expand the return boundaries by adjusting the coefficient of relative risk aversion. Nevertheless, he showed that this automatically leads to an excessive risk free interest rate, because when agents are more risk averse the interest rates are higher. This was referred to as the Risk-free Rate Puzzle.

Possible solutions are suggested in many articles related to this topic. McGrattan & Prescott (2003) discuss the impact of taxes and regulations on the expected dividends. They basically claim that if the tax rate on dividends decreases, the value of expected dividends increases together with asset prices, leading to abnormal returns. Campbell & Cochrane (1999) suggest another logical explanation, which involves a habit persistence. The main idea is that agents have a ”habit level” of consumption, meaning that their utility is not only a function of consumption at $t$, but also a function of their habit level, which depends on past consumption levels and is updated slowly. However, when current consumption is e.g. low relative to the habit level, then the risk aversion will be relatively higher in bad times, meaning that agents hate procyclical assets and demand big risk premia.

Jackwerth (2000) argues that due to the risk aversion of the investor with a power utility function, the pricing kernel is a monotonically decreasing function of aggregate wealth. He estimates $q$ and $p$ using data on the S&P500 index returns, as it is common to assume that this index represents the aggregate wealth held by investors, and computes the pricing kernel according to equation (2.9). However, he finds out that the pricing kernel is not a monotonically decreasing function as expected. Plotted against the return on the S&P500, the pricing kernel according to Jackwerth (2000) is locally increasing, implying an increasing marginal utility and a convex utility function. It is referred to as the Pricing Kernel Puzzle. The shape of the pricing kernel does not correspond to the basic assumption of asset pricing theory. Although Jackwerth (2000) tends to rule out methodological errors, he never proves that the ratio of two estimators equals the estimate of the ratio. He assumes that if $q$ and $p$ are estimated correctly, then their ratio should yield a good estimator for the pricing kernel. This assumption still needs to be proved, but dealing with it is beyond the scope of this work. Rubinstein (1994) shows, that any two of the following imply the third:

- The preferences of the representative agent;
- The subjective probability of the representative agent;
- The risk neutral SPD.
Therefore, we conclude that for a good estimation of the pricing kernel, a good estimation of the subjective probability $p$ and the risk-neutral probability $q$ is required. The next chapter discusses the properties of the Black & Scholes (1973) model, shows its equivalence to the power utility investor and arrives at the same conclusions regarding pricing kernel estimation.
3. THE PRICING OF OPTIONS

3.1 Black and Scholes and Implied Volatility

One of the most important milestones in modern financial theory is the well-known Black & Scholes (1973) model. The key assumptions of the model are:

- The price of the underlying instrument is a geometric Brownian motion with constant drift and volatility;
- Short selling of the underlying stock is allowed;
- There are no riskless arbitrage opportunities;
- Trading in the stock is continuous;
- There are no transaction costs or taxes;
- All securities are perfectly divisible;
- The risk-free interest rate is constant across time and maturities.

Under the assumptions of the model, the price of a plain vanilla call option with a payoff function as in equation (1.1) is given by the Black and Scholes formula

\[ C^{BS}(S_t, t, K, T, \sigma, \delta) = e^{-\delta \tau} S_t \Phi(d_1) - e^{-r \tau} K \Phi(d_2) \]  

where \( \delta \) is the continuous dividend rate, \( r \) is a constant riskless interest rate, \( \tau \) is time to maturity, \( \Phi(u) \) is the cumulative standard normal distribution function and

\[ d_1 = \frac{\ln(S_t/K) + (r - \delta + 0.5\sigma^2)\tau}{\sigma \sqrt{\tau}} \quad \text{and} \quad d_2 = d_1 - \sigma \sqrt{\tau} \]
3. The Pricing of Options

A put option on the same underlying asset with equal strike and time to maturity has a payoff according to equation (1.2) and hence could be priced using the Put-Call Parity (for a complete proof we refer to Theorem 2.3 in Franke et al. (2004))

\[ C(S_t, K, \tau, \sigma, r, \delta) + e^{-r\tau} K = P(S_t, K, \tau, \sigma, r, \delta) + e^{-\delta\tau} S_t \]  
(3.3)

where \( P(S_t, K, \tau, \sigma, r, \delta) \) is the price of the put option. As stated in the previous chapter, we assume \( \delta = 0 \) for the remaining of this work.

In hedging and risk management, the derivatives of the Black and Scholes formula with respect to the other parameters, also known as the greeks, play an important role. The first derivative with respect to the stock price, called Delta, determines the number of underlying assets to be held in a hedge portfolio. The second derivative with respect to stock price, called Gamma, measures the convexity of the price function. For a full description of the greeks and their properties we refer to Chapter 6 in Franke et al. (2004). Particularly, the second derivative of the option price with respect to the strike was shown to have a unique property. As shown by Breeden & Litzenberger (1978), this derivative, when normalized in order to sum up to 1, is the SPD. A simple example would demonstrate this concept. Consider a portfolio which consists of two short calls with strike \( K \), one long call with strike \( K + \varepsilon \) and one long call with strike \( K - \varepsilon \). This portfolio is called a butterfly spread and its payoff is depicted in figure 3.1. This spread pays nothing outside the interval \([K - \varepsilon, K + \varepsilon]\). If we take \(1/\varepsilon^2\) shares of this portfolio and let \( \varepsilon \) tend to zero, the butterfly spread payoff function converges to a Dirac delta function with a mass at \( K \), becoming an Arrow-Debreu security paying 1$ if \( S_T = K \) and 0 otherwise. Therefore, as in equation (2.8), the price \( P_{butterfly} \) of the butterfly spread when \( \varepsilon \to 0 \) should be equal to

\[
\lim_{\varepsilon \to 0} P_{butterfly} = \lim_{\varepsilon \to 0} e^{-\tau r} \int_{K-\varepsilon}^{K+\varepsilon} \psi_{butterfly}(S_T) q_t(S_T) dS_T = e^{-\tau r} q_t(K) 
\]  
(3.4)

where \( q_t(S_T) \) is the risk neutral density (the SPD). In addition, if \( C(S_t, K, \tau) \) is the price of a call option, then the price of the butterfly spread is expressed by

\[
P_{butterfly} = \frac{1}{\varepsilon^2} [C(S_t, K - \varepsilon, \tau) + C(S_t, K + \varepsilon, \tau) - 2C(S_t, K, \tau)] 
\]  
(3.5)
3. The Pricing of Options 25

Fig. 3.1: Butterfly spread payoff function consisting of two short calls with strike \( K \), a long call with strike \( K - \varepsilon \) and a long call with strike \( K + \varepsilon \) (\( K = 100, \varepsilon = 10 \)). When \( \varepsilon \to 0 \) it describes the Arrow-Debreu security.

Taking the limit of the price when \( \varepsilon \to 0 \) and comparing the two equations yield

\[
e^{rt} \lim_{\varepsilon \to 0} P_{\text{butterfly}} = e^{rt} \left. \frac{\partial^2 C(S_t, K, \tau)}{\partial K^2} \right|_{K = S_T} = q_t(S_T) = \text{SPD} \quad (3.6)
\]

All the parameters in the Black and Scholes formula (equation (3.1)) are known, except for the volatility. Hence, the volatility needs to be estimated empirically. However, a more common concept is to derive the volatility which is implied by the observed option price on the market, and therefore known as the implied volatility (IV). Given observed market prices for call options \( \widetilde{C}_t \), the implied volatility is the \( \widetilde{\sigma} \) which satisfies

\[
C^{BS}(S_t, t, K, \tau, \widetilde{\sigma}, r) = \widetilde{C}_t \quad (3.7)
\]

There are numerous algorithms for estimating the IV numerically based on options data. The IV is in fact a mapping of strike prices, time and expiry.
3. The Pricing of Options

3. The Pricing of Options

3. The Pricing of Options

3. The Pricing of Options

The implied volatility surface (IVS). It is more convenient to refer to time to maturity and strikes in relative terms, i.e. using time to maturity $\tau$ and forward moneyness measure

$$\kappa \equiv \frac{K}{F_t} = \frac{K}{e^{r\tau}S_t}$$  \hspace{1cm} (3.8)

The moneyness scale reduces the number of parameters in the formula and makes the surface independent of large moves in the price of the underlying asset. The reason for using forward prices rather than underlying asset prices in the denominator is that traders usually use the index futures market and not the cash market for hedging because cash markets lag future prices by a few minutes due to delays in reporting transactions of the constituent stocks in the index (Jackwerth (2000), Fengler (2005)).

Huynh et al. (2002) offer a thorough review of various algorithms for extracting the SPD from the IVS, as well as many practical examples for implementing them in $XploRe$. Some of the methods will be reviewed in the next chapter.

3.2 The Pricing Kernel under Black and Scholes Assumptions

Under the assumptions of the Black & Scholes (1973) model, the IV is a straight line (as it is assumed to be constant) and the corresponding risk-neutral density is log-normal with mean $(r - 0.5 \sigma^2)\tau$ and variance $\sigma^2\tau$. Plugging equation (3.1) into equation (3.6) yields

$$q^{BS}(S_T|S_t) = \frac{1}{S_T \sqrt{2\pi \sigma^2 \tau}} \cdot e^{-\frac{\ln(S_T/S_t) - (r - 0.5 \sigma^2)\tau}{2\sigma^2 \tau}}$$  \hspace{1cm} (3.9)

meaning that the underlying asset price follows the stochastic process

$$\frac{dS_t}{S_t} = r \cdot dt + \sigma \cdot dW_t$$  \hspace{1cm} (3.10)

i.e., the stock price in a Black & Scholes (1973) world follows a geometric Brownian motion under both probability measures, only with different drifts. Since the subjective probability under the Black & Scholes (1973) is also log-normal but with drift $\mu$, plugging the SPD from equation (3.9) and the log-normal subjective density into equation (2.9) yields a closed-form solution for the investor’s pricing kernel

$$M^{BS}_t(S_T) = \left( \frac{S_T}{S_t} \right)^{\frac{\mu - r}{\sigma^2 \tau}} \cdot e^{-\frac{(\mu - r)(\mu + r - \sigma^2)\tau}{2\sigma^2 \tau}}$$  \hspace{1cm} (3.11)
3. The Pricing of Options

The only non constant term in this expression is \(\frac{S_T}{S_t}\), which corresponds to consumption growth in a pure exchange economy. Since the pricing kernel in equation (3.11) is also the ratio of the marginal utility functions (equation (2.9)), the investor’s utility function can be derived by solving the differential equation. If we consider the following constants

\[
\gamma = \frac{\mu - r}{\sigma^2} \\
\lambda = e^{\frac{(\mu-r)(\mu+r-\sigma^2)}{2\sigma^2}t}
\]

we can rewrite equation (3.11) as

\[
M_t^{BS}(S_T) = \lambda \left(\frac{S_T}{S_t}\right)^{-\gamma}
\]

which corresponds to a power utility function. The B&S utility function is therefore

\[
u^{BS}(S_t) = \left(1 - \frac{\mu - r}{\sigma^2}\right)^{-1} \cdot S_t^{(1 - \frac{\mu - r}{\sigma^2})}
\]

the subjective discount factor of intertemporal utility is

\[
\beta^{BS} = \lambda e^{-r\tau} = e^{\frac{(\mu-r)(\mu+r-\sigma^2)}{2\sigma^2} - r\tau}
\]

and the relative risk aversion is constant

\[
\rho_t^{BS}(S_T) = \gamma = \frac{\mu - r}{\sigma^2}
\]

The above equations prove that a constant RRA utility function sustains the Black & Scholes (1973) model, as was shown by Rubinstein (1976), Breeden & Litzenberger (1978) and many others.

Referring again to the stochastic process in equation (2.4), in which the Brownian motion \(W_t\) is defined on the probability space corresponding to the risk-neutral measure, the Brownian motion under the assumptions of the Black & Scholes (1973) model with a constant RRA can be expressed as

\[
\tilde{W}_t = W_t + \frac{\mu - r}{\sigma} t = W_t + \sigma \gamma t
\]

whereas the stochastic process of the corrected stock price can be expressed as a direct function of the investor’s relative risk aversion

\[
d\tilde{S}_t = \sigma \tilde{S}_t d\tilde{W}_t = \sigma \tilde{S}_t dW_t + \sigma^2 \tilde{S}_t \gamma dt
\]
4. METHODS FOR PRICING KERNEL ESTIMATION

It is well known that the assumptions of the Black & Scholes (1973) model do not hold in practice. Transaction costs, taxes, restrictions on short-selling and non-continuous trading violate the model’s assumptions. Moreover, the stochastic process does not necessarily follow a Brownian motion and the implied volatility is not constant and experiences a smile. Consequently, the SPD does not have a closed form solution and has to be estimated numerically. Fengler (2005) states some stylized facts of the implied volatility surface, such as a more pronounced smile for short maturities, a global minimum of the smile function near at-the-money call options, a leverage effect (returns of the underlying asset and returns of implied volatility are negatively correlated), mean reversion and correlated implied volatility shocks.

Generalizing the models described in the previous chapter to fit real data and bearing the pricing kernel definition in mind, a good estimation of the pricing kernel can be achieved by estimating the subjective density ($p$) and the state-price density ($q$) empirically.

4.1 Subjective Probability Estimation

The first component to be estimated is the conditional density of the data generating process. Various parametric and nonparametric methods for estimating the subjective density are discussed in the literature. In the following, some of them are reviewed, together with their advantages and disadvantages.

Parametric Estimation

The most common parametric approach is to assume that the distribution belongs to a certain type, and the only parameters to be estimated are those characterizing the distribution. A prevailing distribution in such case is a log-normal distribution, the parameters of which are estimated from the data.

An example for a slightly different parametric estimation of $p$ is introduced
by Rosenberg & Engle (2002), yielding a dynamic time-varying pricing kernel (which they name \textit{empirical pricing kernel}) projected onto asset return states. They assume certain specifications for the projected pricing kernel function and estimate a subjective probability based on a stochastic volatility model. They assess their estimate by fitting their model to options data. As a subjective probability they show that the following asymmetric GARCH model leads to the best fit to the data

\[
\ln \left( \frac{S_t}{S_{t-1}} \right) - r^f = \mu + \varepsilon_t, \quad \text{where } \varepsilon_t \sim f(0, \sigma_{t|t-1}^2) \tag{4.1}
\]

\[
\sigma_{t|t-1}^2 = \omega_1 + \omega_2 I + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1|t-2}^2 + \delta \max[0, -\varepsilon_{t-1}]^2 \tag{4.2}
\]

The conditional return variance \( \sigma_{t|t-1}^2 \) is a function of two constants (allowing a shift in long run volatility), the lagged square innovation \( \varepsilon_{t-1}^2 \), the lagged conditional variance and lagged returns. The empirical innovation density \( f \) is modeled by factoring the conditional variance into a standardized innovation and a conditional standard deviation \( \frac{\varepsilon_t}{\sigma_{t|t-1}} \cdot \sigma_{t|t-1} \). The GARCH estimation process, based on significance tests for the parameters in equations (4.1) and (4.2), leads to the conclusion, that the best fit to the data is obtained by fitting a TGARCH(1,1) model. Estimating and testing the model parameters, \( \beta \) and \( \delta \) are significantly different from zero. In addition, the standardized innovations and the squared standardized innovations processes are not autocorrelated.

The main advantages of this method are, it allows dynamics in the pricing kernel estimation and it is independent of risk-neutral distribution characteristics. Rosenberg & Engle (2002) estimate \( p \) and make assumptions regarding the investors’ preferences before deriving the risk-neutral distribution and comparing it to options data, without violating non-arbitrage constraints when estimating \( q \). The problem of a parametric estimator is, as usual, its dependence on the assumed distribution.

\textbf{Nonparametric Estimation}

Aït Sahalia & Lo (2000) suggest a nonparametric kernel estimator for the density of the stock returns, i.e. they define \( u_\tau = \ln(S_T/S_t) \) and estimate the density function of the continuously compounded returns as

\[
\hat{g}(u_\tau) = \frac{1}{N \cdot h_u} \sum_{i=1}^{N} K_u \left( \frac{u_\tau - u_{t_i,\tau}}{h_u} \right) \tag{4.3}
\]
where \( N \) is the number of observations and \( h_u \) is an appropriate bandwidth, thus yielding a straightforward estimator for the subjective probability \( p \)

\[
\hat{p}_t(S_T) = \frac{\tilde{g}(u_t) \cdot \ln(S_T/S_t)}{S_T}
\]  

(4.4)

Brown & Jackwerth (2004) use historical time series of the underlying index to estimate the subjective distribution. They calculate 31-day, nonoverlapping returns from an arbitrarily chosen period of 4-year sample using a Gaussian kernel function with bandwidth \( h = 1.8\sigma/\sqrt{n} \), where \( n \) is the number of observations and \( \sigma \) is the standard deviation of the sample returns. Since the kernel density is continuous and the data is discretized, the density is then discretized onto the state space \( S_j \) by numerically integrating around these states.

The advantage of the two methods mentioned above is that they do not depend on an assumed distribution. There is a vast literature criticizing parametric estimators for producing a poor fit to financial data. The problem with the mentioned nonparametric approaches is that, both A¨ıt Sahalia & Lo (2000) and Brown & Jackwerth (2004) estimate average \( p \) and \( q \) over the sample period, which is never shorter than one year. Therefore, their estimate for the pricing kernel could also be interpreted as a measure for the average pricing kernel over the sample period.

### 4.2 SPD Estimation

The second component to be estimated is the state-price density. There is a vast literature on estimating the SPD using nonparametric and semiparametric methods. In the following we will again review some of them and state their advantages and disadvantages.

#### Semiparametric Approach

A¨ıt Sahalia & Lo (2000) suggest a semiparametric approach to the nonparametric kernel regression discussed in Härdle (1990). They propose a call pricing function according to Black & Scholes (1973), but with a nonparametric function for the volatility. The volatility is estimated using a two dimensional kernel estimator

\[
\hat{\sigma}(\kappa, \tau) = \frac{\sum_{i=1}^{n} k_x(\delta_n \kappa_i / h_x) k_y(\delta_n \tau_i / h_y) \sigma_i}{\sum_{i=1}^{n} k_x(\delta_n \kappa_i / h_x) k_y(\delta_n \tau_i / h_y)}
\]  

(4.5)
where \( \sigma_t \) is the implied volatility. The kernel functions \( k_\kappa \) and \( k_\tau \) together with the appropriate bandwidths \( h_\kappa \) and \( h_\tau \) are chosen such that the asymptotic properties of the second derivative of the call price are optimized. The kernel function measures the drop of likelihood, that the true density function goes through a certain point, when it does not coincide with a certain observation. The price of the call is then calculated using the Black & Scholes (1973) formula but with the estimated volatility, and the SPD is estimated using equation (3.6).

Jackwerth (2000) proposes a discretisation of the future index value to coincide with the strike prices and estimates the implied volatility using a scaled trade off between balancing smoothness and fit to data. The second derivative is approximated numerically (using the discretisation) and the remaining part of the process is identical to the one described above.

A major advantage of these methods comparing to nonparametric ones is that, only the volatility needs to be estimated using a nonparametric regression. The other variables are parametric, thus reducing the size of the problem significantly. Other important qualities of kernel estimators are a well developed and tractable statistical inference and the fact that kernel estimators take advantage of past data, as well as future data, when estimating the current distribution.

The problem of kernel based SPDs is that they could, for certain dates, yield a poor fit to the cross-section of option prices, although for other dates the fit could be quite good.

Dynamic Semiparametric Factor Model for IV

Fengler (2005) introduces a dynamic semiparametric factor model (DSFM) with time-varying coefficients for estimating the IVS, from which the SPD can be derived. This methodology could be perceived as a combination of principal component analysis, nonparametric curve estimation and backfitting additive models. The estimator of the log of the implied volatility is regressed on forward moneyness \((\kappa_{i,j})\) and time to maturity \((\tau_{i,j})\), where \(i\) denotes the day index and \(j\) denotes the intraday observation index. Logs are taken since the data is less skewed and outliers are scaled down. The DSFM is defined as

\[
\sum_{l=0}^{L} \beta_{i,l} m_l(\kappa_{i,j}, \tau_{i,j})
\]

where \( \beta_{i,l} \) are the factor loadings and \( \beta_{i,0} = 1 \). These \( \beta_{i,l} \) form a multivariate time-series model and can be approximated together with the basic smooth
functions $m_l(\kappa_{i,j}, \tau_{i,j})$ by minimizing the following least squares criterion

$$
\min_{\hat{m}_l, \hat{\beta}_{i,l}} \sum_{i=1}^{I} \sum_{j=1}^{J_l} \int \left[ \ln\{\hat{\sigma}_{i,j}(\kappa, \tau)\} - \sum_{l=0}^{L} \hat{\beta}_{i,l} \hat{m}_l(u) \right]^2 K_h\{u-(\kappa_{i,j}, \tau_{i,j})^\top\} du \quad (4.7)
$$

where the integrand is $u = (u_\kappa, u_\tau)^\top$, the bandwidth is $h = (h_\kappa, h_\tau)^\top$ and the two-dimensional kernel function $K_h$ is the product of two one-dimensional kernel functions. It is an iterative algorithm, which begins with initial values for $\hat{\beta}_{i,l}$, e.g. corresponding to piecewise constants on time intervals (equal 1 for a certain day and 0 otherwise) and runs over all functions $\hat{m}_l : \mathbb{R}^2 \to \mathbb{R}$, then fixes the function $\hat{m}_l$ and runs across all $\hat{\beta}_{i,l} \in \mathbb{R}$, fixes a $\hat{\beta}_{i,l}$, runs again across all functions and so on until minor changes occur.

This method results in smaller bias than the naive trader models, but it might, in some cases, violate the non-arbitrage constraints.

**Discretized Nonparametric Approach**

Härdle & Hlávka (2005) propose a method of estimating SPD satisfying the non-arbitrage constraints. They suggest a discretisation of $p$ distinct strike prices such that the distances between two adjacent discretized prices are equal to 1. Then a nonlinear regression model is fitted

$$
C(K) = X_\Delta(K) \beta + \epsilon \quad (4.8)
$$

$$
\beta_i(\xi) = \frac{e^{\xi_i}}{\sum_{j=1}^{p} e^{\xi_j}}, \forall i \in [0, p-1] \quad (4.9)
$$

The model is based on a design matrix $X_\Delta(K)$ with elements corresponding to distances between strike prices. $K$ denotes the vector of strike prices and the $\beta_i(\xi)$ can be perceived as point estimates of the SPD and are estimated using the maximum likelihood method. Imposing the equality

$$
\epsilon_p \left[ \sum_{j=1}^{p-1} e^{\xi_j} \right]^{-1} = 1 - \left[ \sum_{j=1}^{p-1} \beta_j(\xi) \right]^{-1} \quad (4.10)
$$

guarantees the distribution property for $\beta_j(\xi)$, as they sum up to one when $\xi_p \to -\infty$.

The main contribution of this model is that it provides an arbitrage free estimation of the SPD. In addition, the covariance structure proposed in this work and allowing for heteroscedasticity yields a good fit to observed option prices.
Implied Binomial Trees

Another common method is the implied binomial trees (IBT), originally introduced by Rubinstein (1994) and discussed by Derman & Kani (1994) and others. The tree is equally spaced and the risk neutral transition probability and Arrow-Debreu prices are recovered from option data. For a detailed description and practical implementation we refer to Härdle & Zheng (2002).

The advantage of this method is that it yields a consistent estimator of the distribution at each date. The disadvantages of the IBT comparing to kernel estimators are, trees require a prior distribution for the SPD (while kernel regression does not) and they are estimated for cross sections separately, meaning they yield indeed a consistent estimator with all option prices at each date, but consistency across time is not necessarily achieved.

The current chapter reviewed some common methods for estimating the subjective and state-price densities, in order to derive the pricing kernel and relative risk aversion functions. In the next chapter the chosen estimating methods as well as the statistical model are described in detail.
5. A STATIC MODEL: DAILY ESTIMATION

As stated at the end of Chapter 2, Rubinstein (1994) has shown, that an estimated subjective probability together with a good estimation of the SPD enable an assessment of the representative agent’s preferences. Hence, the model presented in this chapter aims at estimating the pricing kernel using the ratio between the subjective density and the SPD, and it disregards the issue of whether a ratio of two estimates is a good approximation for the estimated ratio itself.

This chapter is divided into two parts. The first part provides a detailed description of the database used in this work. In the second section, the static model for estimating the pricing kernel and relative risk aversion on a daily basis is introduced. When the densities and preferences are known for every day, the dynamics of the time-series can be examined. The results of this examination are reported in Chapter 6.

5.1 The Database

The database used for this work consists of intraday DAX and options data which has undergone a thorough preparation scheme. The data was obtained from the MD*Base, maintained at the Center for Applied Statistics and Economics (CASE) at the Humboldt-University of Berlin. The first trading day in the database is January 4th 1999 and the last one is April 30th 2002, i.e. more than three years of intraday data and 2,921,181 observations. The options data contains tick statistics on the DAX index options and is provided by the German-Swiss Futures Exchange EUREX. Each single contract is documented and contains the future value of the DAX (corresponding to the maturity and corrected for dividends according to equation (2.3)), the strike, the interest rate (linearly interpolated to approximate a “riskless” interest rate for the specific option’s time to maturity), the maturity of the contract, the closing price, the type of the option, calculated future moneyness, calculated Black and Scholes implied volatility as in equation (3.7), and the exact
time of the trade (in hundredths of seconds after midnight), the number of contacts and the date.

In order to exclude outliers at the boundaries, only observations with a maturity of more than one day, implied volatility of less than 0.7 and future moneyness between 0.74 and 1.22 are considered, remaining with 2,719,640 observations on 843 trading days. For every single trading day starting April 1999, the static model described in the following section is run and the results are collected. The daily estimation begins three months after the first trading day in the database because part of the estimation process is conducted on historical data, and the history "window" is chosen to be three months, as explained in the next section.

The daily estimation process described in this chapter is implemented in \textit{XploRe} using the quantlet \texttt{EPKdailyprocess.xpl}.

5.2 Subjective Density Estimation

The subjective density is estimated using a simulated GARCH model, the parameters of which are estimated based on historical data. This method was shown by [Jackwerth (2000)] and others to resemble the actual subjective density.

The first step is to extract the data from the three months preceding the date of the daily assessment. That is the reason for starting the daily process in April instead of January 1999. The intraday options data from the preceding three months are replaced by daily averages of the stock index and the interest rate, averaged over the specific day. When we have a three months history of daily asset prices, we can fit a GARCH (1,1) model to the data. A strong GARCH (1,1) model is described by

\begin{align*}
\epsilon_t &= \sigma_t Z_t \\
\sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2
\end{align*}

(5.1)

where $Z_t$ is an independent identically distributed innovation with a standard normal distribution. The logarithmic returns of the daily asset prices are calculated according to $\epsilon_t = \Delta \log(S_t) = \log(S_t) - \log(S_{t-1})$, and this time series together with its daily standard deviation $\sigma_t$ are the input of the GARCH estimation. The parameters $\omega$, $\alpha$ and $\beta$ are estimated using the quasi maximum likelihood method, as implemented in the \textit{XploRe} function \texttt{garchest}. 
The quasi maximum likelihood method is an extension of the maximum likelihood measure, when the estimator is not efficient. The likelihood function for a GARCH (1,1) is defined as

\[
l_b(\theta) = \sum_{t=2}^{n} l_t(\theta) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^{n} \log(\sigma_t^2) - \frac{1}{2} \sum_{t=2}^{n} \frac{\varepsilon_t + 1}{\sigma_t^2} \tag{5.2}
\]

where \( \theta \) is the parameter vector \( \theta = (\omega, \alpha, \beta)^\top \). The first order condition imposes

\[
\sum_{t=2}^{n} \frac{\partial l_t}{\partial \theta} = 0 \tag{5.3}
\]

where

\[
\frac{\partial l_t}{\partial \theta} = \frac{1}{2\sigma_t^2} \cdot \frac{\partial \sigma_t^2}{\partial \theta} \left( \frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right)
\]

\[
\frac{\partial \sigma_t^2}{\partial \theta} = (1, \varepsilon_{t-1}^2, \sigma_{t-1}^2)^\top + \frac{\partial \sigma_{t-1}^2}{\partial \theta} \tag{5.4}
\]

After the parameters of the GARCH process have been estimated, a simulation of a new GARCH process is conducted, starting on the date of the daily assessment. Equations (5.1) are used for the simulation, but this time the unknown variables are the time series \( \sigma_t \) and \( \varepsilon_t \), while the parameters \( \omega, \alpha \) and \( \beta \) are the ones estimated from the historical data. The simulation creates a \( T \) days long time series, and is run \( N \) times. The simulated DAX is calculated as

\[
S_t = S_{t-1} e^{\varepsilon_t} \quad \forall t \in \{1, \ldots, T\} \tag{5.5}
\]

where \( S_0 \) is the present level of the index on the day of the daily assessment.

Our aim is to estimate the subjective density in some fixed time points, which correspond to specific maturities used for the SPD estimation discussed next. Therefore, after the simulation has been completed, the simulated data on the dates, which correspond to the desired maturities, is extracted, and the daily subjective density is estimated using a kernel regression on the desired moneyness grid, which corresponds to the asset’s gross return. The transformation from the simulated \( S_t \) to the moneyness grid is achieved using \( e^{-rT} \frac{S_T}{S_0} \) for each desired horizon \( T \), where \( r \) is the daily average risk-free rate at the present day. The subjective density is estimated for every trading day included in the database. In figure 5.1 we plot the simulated subjective densities on four different trading days for four different maturities.
Fig. 5.1: Subjective density for different maturities (30, 60, 90, 120 days) on different trading days.
It can be seen in figure 5.1, that the distribution resembles a log normal distribution, which is more spread the longer the maturity is. A well known feature of financial data is that equity index return volatility is stochastic, mean-reverting and responds asymmetrically to positive and negative returns, due to the leverage effect. Therefore, this GARCH (1,1) model estimation, which experiences a slight positive skewness, is an adequate measure for the index returns, and it resembles the nonparametric subjective densities, which were estimated by Ait Sahalia & Lo (2000) and Brown & Jackwerth (2004).

5.3 State-Price Density Estimation

The state-price density is estimated using a local polynomial regression as proposed by Rookley (1997) and described thoroughly in Huynh et al. (2002). The choice of Nadaraya-Watson type smoothers, used by Ait Sahalia & Lo (2000) and described in the previous chapter, is inferior to local polynomial kernel smoothing. More accurately, the Nadaraya-Watson estimator is actually a local polynomial kernel smoother of degree 0. If we use higher order polynomial smoothing methods, we can obtain better estimates of the functions. Local polynomial kernel smoothing also provides a convenient and effective way to estimate the partial derivatives of a function of interest, which is exactly what we look for when estimating SPDs.

The first step is to calculate the implied volatility for each given maturity and moneyness in the daily data (based on the B&S formula when prices are given and \( \sigma \) is the unknown). Then a local polynomial regression is used to smooth the IV points and to create the implied volatility surface from which the SPD can be derived. The basic idea of local polynomial regression is based on a locally weighted least squares regression, where the weights are determined by the choice of a kernel function, the distance of an observation from a certain estimated point defining the surface/line at this coordinate and the chosen bandwidth vector. The use of the moneyness measure and time to maturity reduces the regression to two dimensions and enables freedom in estimating the surface in fictional points that do not exist in the database.

Let us first review the concept of local polynomial estimation. The input data at this stage is a trivariate data, a given grid of moneyness (\( \kappa \)), time to maturity (\( \tau \)) and the implied volatility (\( \sigma^{BS}(\kappa, \tau) \)). We now consider the following process for the implied volatility surface

\[
\tilde{\sigma} = \phi(\kappa, \tau) + \sigma^{BS}(\kappa, \tau) \ast \epsilon \tag{5.6}
\]
where $\phi(\kappa, \tau)$ is an unknown function, which is three times continuously differentiable, and $\varepsilon$ is a Gaussian white noise. Then a Taylor expansion for the function $\phi(\kappa, \tau)$ in the neighborhood of $(\kappa_0, \tau_0)$ is

$$
\phi(\kappa, \tau) \approx \phi(\kappa_0, \tau_0) + \frac{\partial \phi}{\partial \kappa} \bigg|_{\kappa_0, \tau_0} (\kappa - \kappa_0) + \frac{1}{2} \frac{\partial^2 \phi}{\partial \kappa^2} \bigg|_{\kappa_0, \tau_0} (\kappa - \kappa_0)^2 
$$

$$
+ \frac{\partial \phi}{\partial \tau} \bigg|_{\kappa_0, \tau_0} (\tau - \tau_0) + \frac{1}{2} \frac{\partial^2 \phi}{\partial \tau^2} \bigg|_{\kappa_0, \tau_0} (\tau - \tau_0)^2 
$$

$$
+ \frac{1}{2} \frac{\partial^2 \phi}{\partial \kappa \partial \tau} \bigg|_{\kappa_0, \tau_0} (\kappa - \kappa_0)(\tau - \tau_0) 
$$

(5.7)

Minimizing the expression

$$
\sum_{j=1}^{n} \left\{ \sigma_{BS}(\kappa_j, \tau_j) - \left[ \beta_0 + \beta_1(\kappa_j - \kappa_0) + \beta_2(\kappa_j - \kappa_0)^2 + \beta_3(\tau_j - \tau_0) 
\right.
\right.
\left. + \beta_4(\tau_j - \tau_0)^2 + \beta_5(\kappa_j - \kappa_0)(\tau_j - \tau_0) \right\}^2 K_h(\kappa - \kappa_0)(\tau - \tau_0) 
$$

(5.8)

yields the estimated IVS and its first two derivatives at the same time, as

$$
\frac{\partial \phi}{\partial \kappa} \bigg|_{\kappa_0, \tau_0} = \hat{\beta}_1 \quad \text{and} \quad \frac{\partial^2 \phi}{\partial \kappa^2} \bigg|_{\kappa_0, \tau_0} = 2\hat{\beta}_2.
$$

This is a very useful feature, as the second derivative is used to calculate the SPD for a certain fixed maturity. A detailed derivation of $\frac{\partial^2 \phi}{\partial \kappa^2}$ (used for the SPD according to Breeden & Litzenberger (1978)) as a function of $\frac{\partial \phi}{\partial \kappa}$ and $\frac{\partial^2 \phi}{\partial \kappa^2}$ (which are obtained from the IVS estimation) is given by Huynh et al. (2002). A full implementation of this process is to be found in the quantlet EPKsystemfiles.xpl. This quantlet includes a procedure, which is based on the function spdb1 and performs the whole process, including the IVS estimation and the derivation. This procedure’s output is the estimated SPD for the given maturity on the desired moneyness grid. The chosen maturities are similar to those used to estimate the subjective probability.

The estimated risk neutral densities for the same dates and the same maturities as in figure 5.1 are depicted in figure 5.2. As stated above, the SPD is estimated on a future moneyness scale, thus reducing the number of parameters that need to be estimated.
Fig. 5.2: State-Price density for different maturities \((30, 60, 90, 120\) days) on different trading days.
One of the trading days plotted in figure 5.2 is September 11th 2001. It is interesting to see that the options data on this trading day reflects some increased investors’ beliefs, that the market will go down in the long run. Similar behavior is found in the trading days following that particular day as well as in other days of crisis. The highly volatile SPD for negative returns, which could be explained, for example, by the leverage effect or the correlation effect, could reflect a dynamic demand for insurance against a market crash. This phenomenon is more apparent in days of crisis and was reported by Jackwerth (2000) as well.

5.4 Deriving the Pricing Kernel and Risk Aversion

At this stage, we have the estimated subjective and state-price densities for the same maturities and spread over the same grid. The next step is to calculate the daily estimates for the pricing kernel and risk aversion.

The pricing kernel is calculated using equation (2.9), where the estimated subjective density and the estimated SPD replace $p(S_T|S_t)$ and $q(S_T|S_t)$ in the equation respectively. Since the grid is a moneyness grid, and the estimated $p$ and $q$ are estimated on the moneyness grid, what we get is actually $M_t(\kappa_T)$. The coefficient of relative risk aversion is then computed by numerically estimating the derivative of the estimated pricing kernel with respect to the moneyness and then according to equation (2.13).

The estimated pricing kernels depicted in figure 5.3 for different trading days and different maturities bear similar characteristics to those reported by Aıt Sahalia & Lo (2000), Jackwerth (2000), Rosenberg & Engle (2002) and others, who conducted a similar process on the S&P500 index. The pricing kernel is not a monotonically decreasing function, as suggested in classic macroeconomic theory. It is more volatile and steeply upward sloping for large negative return states, and moderately downward sloping for large positive return states. Moreover, the pricing kernel contains a region of increasing marginal utility at the money (around $\kappa = 1$), implying a negative risk aversion. This feature can clearly be seen in figure 5.4, which depicts the coefficient of relative risk aversion and shows clearly, that the minimal risk aversion is obtained around the ATM region and the relative risk aversion is negative. The negative risk aversion around the ATM region implies the possible existence of risk seeking investors, whose utility functions are locally convex.
Fig. 5.3: Estimated Pricing Kernel for different maturities (30, 60, 90, 120 days) on different trading days.
Fig. 5.4: Estimated relative risk aversion for different maturities (30, 60, 90, 120 days) on different trading days.

EPKdailyprocess.xpl
Jackwerth (2000) named this phenomenon the *pricing kernel puzzle* and suggested some possible explanations to it. One possible explanation is that, a broad index (DAX in this work, S&P500 in his work) might not be a good proxy for the market portfolio and as such, the results are significantly different than those implied in the standard macroeconomic theory. In addition to the poor fit of the index, the assumptions for the existence of a representative agent might not hold, meaning that markets are not complete or the utility function is not strictly state-independent or time-separable.

Another possibility is that, historically realized returns are not reliable indicators for subjective probabilities, or that the subjective distribution is not well approximated by the actual one. This deviation stems from the fact that investors observe historical returns without considering crash possibilities, and then incorporate crash possibilities, which make their subjective distribution look quite different than the one estimated here. The historical estimation or the log-normal distribution assumptions ignore the well-known volatility clustering of financial data.

Looking from another interesting point of view, investors might make mistakes in deriving their own subjective distributions from the actual objective one, thus leading to mispricing of options. Jackwerth (2000) claims, that mispricing of options in the market is the most plausible explanation to the negative risk aversion and increasing marginal utility function.

This work does not aim, however, at finding a solution to the pricing kernel puzzle. The implicit assumption in this work is that, some frictions in the market lead to the contradicting of standard macroeconomic theory, resulting in a region of increasing marginal utility. In the next chapter, a dynamic analysis of the pricing kernel and relative risk aversion is conducted along the three-year time frame.
6. A DYNAMIC MODEL: TIME-SERIES ANALYSIS

The previous chapter provided a detailed description of the daily process. Since the process is conducted on a daily basis and in most of the trading days, the GARCH and local polynomial estimations produce a good fit to the data, three-year long time-series data of pricing kernel and relative risk aversion are obtained. In this chapter we will analyze these time-series and show their moments. A principal component analysis will be conducted on the stationary series and the PCs will be tested as a response variable in a GLS regression. The collection of the data is implemented in XploRe using the quantlet EPKmain.xpl, whereas the time-series analysis is done with EPKtimeseries.xpl.

6.1 Moments of the Pricing Kernel and Relative Risk Aversion

In order to explore the characteristics of the pricing kernel and the relative risk aversion, their moments at any trading day have to be computed. Each of the estimates (pricing kernel and relative risk aversion) is a function of moneyness and time to maturity. Let \( \hat{f}_t(\kappa, \tau) \) be the estimate at time point \( t \) as a function of the moneyness and time to maturity. We further choose \( \tau = (\tau_1, \tau_2, \tau_3, \tau_4) \top \) to be a vector of four predetermined maturities, and as in the previous section we concentrate on \( \tau = (30, 60, 90, 120) \top \) days. \( \hat{f}_t \) can be the daily estimated pricing kernel or the daily estimated relative risk aversion. For each of them we focus on five functions:

- The estimator’s value when the moneyness equals 1:
  \[
  \hat{f}_t^{\text{ATM}} = \hat{f}_t(\kappa = 1, \tau)
  \]
  (6.1)

- The mean of the estimate:
  \[
  \mu_t(\hat{f}_t) = \int_{-\infty}^{\infty} \kappa \hat{f}_t(\kappa, \tau) d\kappa
  \]
  (6.2)
6. A Dynamic Model: Time-Series Analysis

- The standard deviation of the estimate:
  \[ \sigma_t(\hat{f}_t) = \sqrt{\int_{-\infty}^{\infty} \{\kappa - \mu_t(\hat{f}_t)\}^2 \hat{f}_t(\kappa, \tau) d\kappa} \]  
  (6.3)

- The skewness of the estimate:
  \[ Skew_t(\hat{f}_t) = \frac{1}{\sigma^3_t(\hat{f}_t)} \int_{-\infty}^{\infty} \{\kappa - \mu_t(\hat{f}_t)\}^3 \hat{f}_t(\kappa, \tau) d\kappa \]  
  (6.4)

- The kurtosis of the estimate:
  \[ Kurt_t(\hat{f}_t) = \frac{1}{\sigma^4_t(\hat{f}_t)} \int_{-\infty}^{\infty} \{\kappa - \mu_t(\hat{f}_t)\}^4 \hat{f}_t(\kappa, \tau) d\kappa \]  
  (6.5)

Since we are interested in stationary data, it is reasonable to calculate the first differences of the time-series as well as the logarithmic differences. For a time-series \( \{X_t\}_{t=1}^n \), the first differences are defined as \( \Delta X_t = X_t - X_{t-1} \) for \( t \in \{2, \ldots, n\} \) and the differences of the logarithms of the time-series are \( \Delta \log(X_t) = \log(X_t) - \log(X_{t-1}) \) for \( t \in \{2, \ldots, n\} \). The figures in the following pages depict the time-series of the different moments of the pricing kernel and the relative risk aversion, each estimated for four different maturities on 589 trading days between April 1999 and April 2002. The trading days, on which the GARCH model does not fit the data, or the local polynomial estimation experiences some negative volatilities, were dropped. The time-series of the differences and the differences of the logarithms are shown in Appendix A. Descriptive statistics of the all time-series are shown in tables 6.1, 6.2 and 6.3, which follow the figures.

The plots in the next pages show, that the pricing kernel at the money (figure 6.1) behaves similarly across different maturities and bears similar characteristics to its general mean (figure 6.2). This result implies, that characterizing the pricing kernel using the four first moments of its distribution is adequate. Contrary to the pricing kernel, the relative risk aversion at the money (figure 6.6) looks quite different than its general mean (figure 6.7). The ATM RRA is mostly negative, as detected already in the daily estimated RRA. The RRA mean, however, is mostly positive. Another feature of the RRA is that, it becomes less volatile the longer the maturity is, implying the existence of more nervous investors for assets with short maturities. The main conclusion we can draw from the RRA plots is that, the four first moments of the distribution do not necessarily represent all the features of the RRA correctly, and the collection of the extra details regarding the ATM behavior is justified, as it will be shown by the principal component analysis.
6. A Dynamic Model: Time-Series Analysis

Fig. 6.1: ATM Pricing Kernel for different maturities (30, 60, 90, 120 days).

Fig. 6.2: Mean of Pricing Kernel for different maturities (30, 60, 90, 120 days).
Fig. 6.3: Standard Deviation of Pricing Kernel for different maturities (30, 60, 90, 120 days).

Fig. 6.4: Skewness of Pricing Kernel for different maturities (30, 60, 90, 120 days).
6. A Dynamic Model: Time-Series Analysis

Fig. 6.5: Kurtosis of Pricing Kernel for different maturities (30, 60, 90, 120 days).

Fig. 6.6: ATM Relative Risk Aversion for different maturities (30, 60, 90, 120 days).
Fig. 6.7: Mean of Relative Risk Aversion for different maturities (30, 60, 90, 120 days).

Fig. 6.8: Standard deviation of Relative Risk Aversion for different maturities (30, 60, 90, 120 days).
6. A Dynamic Model: Time-Series Analysis

**Fig. 6.9:** Skewness of Relative Risk Aversion for different maturities (30, 60, 90, 120 days).

**Fig. 6.10:** Kurtosis of Relative Risk Aversion for different maturities (30, 60, 90, 120 days).
### Tab. 6.1: Descriptive statistics: Moments of pricing kernel and relative risk aversion

<table>
<thead>
<tr>
<th>Time-Series</th>
<th>Pricing Kernel Mean</th>
<th>STD</th>
<th>Risk Aversion Mean</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATM</td>
<td></td>
<td>ATM</td>
<td></td>
</tr>
<tr>
<td>τ = 30 days</td>
<td>1.24</td>
<td>0.31</td>
<td>-3.38</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>µ_t</td>
<td>1.15</td>
<td>0.40</td>
<td>5.96</td>
</tr>
<tr>
<td></td>
<td>σ_t</td>
<td>1.00</td>
<td>0.74</td>
<td>17.63</td>
</tr>
<tr>
<td></td>
<td>Skewit</td>
<td>2.03</td>
<td>1.97</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>Kurtit</td>
<td>11.61</td>
<td>13.15</td>
<td>6.88</td>
</tr>
<tr>
<td>τ = 60 days</td>
<td>1.34</td>
<td>0.32</td>
<td>-4.20</td>
<td>1.73</td>
</tr>
<tr>
<td></td>
<td>µ_t</td>
<td>1.13</td>
<td>0.21</td>
<td>3.40</td>
</tr>
<tr>
<td></td>
<td>σ_t</td>
<td>0.67</td>
<td>0.54</td>
<td>16.45</td>
</tr>
<tr>
<td></td>
<td>Skewit</td>
<td>1.30</td>
<td>2.17</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>Kurtit</td>
<td>9.98</td>
<td>14.18</td>
<td>10.08</td>
</tr>
<tr>
<td>τ = 90 days</td>
<td>1.43</td>
<td>0.35</td>
<td>-3.71</td>
<td>2.55</td>
</tr>
<tr>
<td></td>
<td>µ_t</td>
<td>1.26</td>
<td>0.25</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td>σ_t</td>
<td>0.65</td>
<td>0.57</td>
<td>16.25</td>
</tr>
<tr>
<td></td>
<td>Skewit</td>
<td>0.95</td>
<td>2.08</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Kurtit</td>
<td>8.68</td>
<td>13.04</td>
<td>11.58</td>
</tr>
<tr>
<td>τ = 120 days</td>
<td>1.46</td>
<td>0.50</td>
<td>-2.54</td>
<td>5.06</td>
</tr>
<tr>
<td></td>
<td>µ_t</td>
<td>1.34</td>
<td>0.30</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>σ_t</td>
<td>0.65</td>
<td>0.58</td>
<td>15.37</td>
</tr>
<tr>
<td></td>
<td>Skewit</td>
<td>0.82</td>
<td>2.15</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>Kurtit</td>
<td>8.76</td>
<td>12.85</td>
<td>12.36</td>
</tr>
</tbody>
</table>

EPKtimeseries.xpl
\[ \begin{array}{cccccc}
| \text{Time-Series} | \text{Pricing Kernel Mean} & \text{STD} & \text{Risk Aversion Mean} & \text{STD} |
|-------------------|----------------|--------|----------------|--------|
| \tau = 30 \text{ days} | \Delta \text{ (ATM)} & 1.05 \times 10^{-3} & 0.16 & -1.58 \times 10^{-3} & 2.60 |
| & \Delta \mu_t & 1.00 \times 10^{-4} & 0.33 & -1.23 \times 10^{-4} & 7.18 |
| & \Delta \sigma_t & -2.42 \times 10^{-4} & 0.83 & -1.16 \times 10^{-2} & 20.13 |
| & \Delta \text{Skew}_t & -1.17 \times 10^{-3} & 2.46 & 9.52 \times 10^{-4} & 2.65 |
| & \Delta \text{Kurt}_t & 3.17 \times 10^{-3} & 16.99 & 1.22 \times 10^{-4} & 16.91 |
| \tau = 60 \text{ days} | \Delta \text{ (ATM)} & 5.61 \times 10^{-4} & 0.21 & 3.76 \times 10^{-4} & 1.77 |
| & \Delta \mu_t & 2.87 \times 10^{-4} & 0.24 & 8.36 \times 10^{-4} & 8.52 |
| & \Delta \sigma_t & -2.50 \times 10^{-5} & 0.74 & -2.71 \times 10^{-3} & 26.39 |
| & \Delta \text{Skew}_t & -3.34 \times 10^{-4} & 2.97 & 7.76 \times 10^{-4} & 3.50 |
| & \Delta \text{Kurt}_t & 3.36 \times 10^{-3} & 19.70 & -8.30 \times 10^{-4} & 21.07 |
| \tau = 90 \text{ days} | \Delta \text{ (ATM)} & 3.19 \times 10^{-4} & 0.28 & 9.72 \times 10^{-4} & 3.21 |
| & \Delta \mu_t & 4.77 \times 10^{-4} & 0.28 & 3.22 \times 10^{-3} & 7.90 |
| & \Delta \sigma_t & 3.66 \times 10^{-4} & 0.74 & 2.48 \times 10^{-3} & 28.73 |
| & \Delta \text{Skew}_t & -1.59 \times 10^{-4} & 2.64 & 7.52 \times 10^{-4} & 3.88 |
| & \Delta \text{Kurt}_t & 4.38 \times 10^{-3} & 17.09 & -1.41 \times 10^{-3} & 24.08 |
| \tau = 120 \text{ days} | \Delta \text{ (ATM)} & 4.47 \times 10^{-4} & 0.28 & 5.81 \times 10^{-3} & 6.99 |
| & \Delta \mu_t & 9.98 \times 10^{-4} & 0.32 & -1.72 \times 10^{-2} & 7.98 |
| & \Delta \sigma_t & -3.53 \times 10^{-5} & 0.76 & 5.45 \times 10^{-2} & 28.47 |
| & \Delta \text{Skew}_t & -3.53 \times 10^{-4} & 2.91 & -1.23 \times 10^{-2} & 4.02 |
| & \Delta \text{Kurt}_t & -1.29 \times 10^{-1} & 17.49 & 1.45 \times 10^{-1} & 23.03 |
\end{array} \]

Tab. 6.2: Descriptive statistics: Differences of the moments of pricing kernel and relative risk aversion
### Tab. 6.3: Descriptive statistics: Differences of the logarithms of the moments of pricing kernel and relative risk aversion. The logarithms are not always defined, as relative risk aversion could be zero or negative, and crossing the zero line results in an infinite logarithm.
6.2 Stationarity Tests

After depicting and describing the characteristics of the different time-series, and before we concentrate on specific time-series for further analysis, it is essential to determine which of the time-series are stationary. The test chosen to check for stationarity is the KPSS test, originally suggested by Kwiatkowski et al. (1992). The null hypothesis of this statistic test is that the time-series is stationary. The test starts by forming the following regression:

\[ X_t = c + \nu t + \phi \sum_{i=1}^{t} \xi_i + \eta_t \]  

(6.6)

where \( \nu \) is a linear time trend, \( \eta_t \) is stationary and \( \xi_i \) is white noise. The null hypothesis is therefore \( H_0 : \phi = 0 \) and since white noise does not affect the process, it is stationary due to \( \eta_t \). The residuals \( \eta_t \) are obtained by running the regression under the null hypothesis, and thus the partial sum of residuals \( V_t = \sum_{i=1}^{t} \eta_i \) grows linearly in time. The KPSS statistic is:

\[ \text{KPSS}_t = \frac{\sum_{t=1}^{n} V_t^2}{n^2 \hat{\omega}_T^2} \]  

(6.7)

where

\[ \hat{\omega}_T^2 = \hat{\sigma}_\eta^2 + 2 \sum_{\tau=1}^{T} \left( 1 - \frac{\tau}{T-1} \right) \hat{\gamma}_\tau \]  

(6.8)

is a spectral density at a frequency of zero, \( \hat{\sigma}_\eta^2 \) is the estimate of the variance of \( \eta_t \) and \( \hat{\gamma}_\tau = \frac{1}{n} \sum_{t=\tau+1}^{T} \hat{\eta}_t \hat{\eta}_{t-\tau} \) is the covariance estimator. The test is sensitive to the choice of \( T \), as well as to the existence of a time linear trend. If \( T \) is too small and autocorrelation is evident, the test is biased, whereas for large \( T \), the test loses power. Therefore, the test has to be conducted with various reference points \( T \) and with or without a trend, and the results can be quite different. As an example, the KPSS tests for the mean of pricing kernel, its differences and the differences of its logarithms are shown in table 6.4.
TAB. 6.4: KPSS test for stationarity of the mean pricing kernels, their differences and differences of their logarithms for different maturities. The values in red indicate higher values than the critical values, for which the null hypothesis is rejected at a confidence level of 90%. Conclusions, whether and under which conditions a time-series is stationary, are summarized in the right-most column.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>No Linear Trend</th>
<th>With Linear Trend</th>
<th>Stationary</th>
</tr>
</thead>
<tbody>
<tr>
<td>T=0</td>
<td>T=7</td>
<td>T=21</td>
<td>T=0</td>
</tr>
<tr>
<td>$\tau = 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_t(PK)$</td>
<td>0.883</td>
<td>0.135</td>
<td>0.078</td>
</tr>
<tr>
<td>$\Delta \mu_t(PK)$</td>
<td>0.003</td>
<td>0.022</td>
<td>0.036</td>
</tr>
<tr>
<td>$\Delta \log \mu_t(PK)$</td>
<td>0.003</td>
<td>0.022</td>
<td>0.037</td>
</tr>
<tr>
<td>$\tau = 60$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_t(PK)$</td>
<td>0.281</td>
<td>0.107</td>
<td>0.067</td>
</tr>
<tr>
<td>$\Delta \mu_t(PK)$</td>
<td>0.002</td>
<td>0.014</td>
<td>0.037</td>
</tr>
<tr>
<td>$\Delta \log \mu_t(PK)$</td>
<td>0.002</td>
<td>0.014</td>
<td>0.038</td>
</tr>
<tr>
<td>$\tau = 90$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_t(PK)$</td>
<td>1.351</td>
<td>0.442</td>
<td>0.250</td>
</tr>
<tr>
<td>$\Delta \mu_t(PK)$</td>
<td>0.002</td>
<td>0.013</td>
<td>0.037</td>
</tr>
<tr>
<td>$\Delta \log \mu_t(PK)$</td>
<td>0.002</td>
<td>0.014</td>
<td>0.039</td>
</tr>
<tr>
<td>$\tau = 120$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_t(PK)$</td>
<td>2.454</td>
<td>0.632</td>
<td>0.344</td>
</tr>
<tr>
<td>$\Delta \mu_t(PK)$</td>
<td>0.002</td>
<td>0.018</td>
<td>0.050</td>
</tr>
<tr>
<td>$\Delta \log \mu_t(PK)$</td>
<td>0.002</td>
<td>0.014</td>
<td>0.051</td>
</tr>
</tbody>
</table>

Altogether we have five functions, four maturities and three time-series to test (original series, differences of the moments and the differences of the logarithms of the moments), that means a total of 60 time-series for each of the estimates (PK and RRA). A thorough test of stationarity was conducted for each of those time-series, and the main results of this test can be shown in table 6.5. In this table, the conditions for stationarity can be seen, i.e. under which conditions the null hypothesis of the KPSS test is not rejected.
Tab. 6.5: Conditions for stationarity of the moments of pricing kernel and relative risk aversion. The time-series of the differences of the moments are always stationary, the time-series of the differences of the logarithm of the moments are stationary, where the logarithm is defined and the test is applicable.

After conducting stationarity tests for the various functions, we found that the moments themselves are in most of the cases not stationary, and the logarithmic differences of the moments are not always defined, due to the existence of negative values. Therefore, we will from now on concentrate only on the absolute differences of the moments, which have always been found to be stationary.
6.3 Principal Component Analysis

In the previous section, the dynamics of the pricing kernel and relative risk aversion were examined using a stationarity test. This section will focus on a principal component analysis (PCA) of the time-series in order to try and explain the variation of the time-series using a small number of influential factors. As stated before, the only time-series to be considered are the differences of the moments, found to be stationary. The PCA process starts with the definition of the following data matrix for pricing kernel differences

\[
X = \begin{pmatrix}
\Delta PK_{2ATM} & \Delta \mu_2 & \Delta \sigma_2 & \Delta Skew_2 & \Delta Kurt_2 \\
\Delta PK_{3ATM} & \Delta \mu_3 & \Delta \sigma_3 & \Delta Skew_3 & \Delta Kurt_3 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\Delta PK_{nATM} & \Delta \mu_n & \Delta \sigma_n & \Delta Skew_n & \Delta Kurt_n
\end{pmatrix}
\]  

(6.9)

for each of the desired maturities. A similar matrix is defined for the differences of the relative risk aversion. PCA can be conducted either on the covariance matrix of the variables or on their correlation matrix. If the variation is of the same scale, the covariance matrix can be used for the PCA, but if the data is not scale-invariant, a standardized PCA must be applied, i.e. conducting the PCA on the correlation matrix. Table 6.6 contains, as an example, the covariance matrix of the pricing kernel differences for maturity of 60 days. Looking at this table it is obvious, that a PCA based on the covariance matrix will not be effective due to the different scales of the variables.

<table>
<thead>
<tr>
<th>(\Delta PK_{iATM}^{\tau})</th>
<th>(\Delta \mu_i)</th>
<th>(\Delta \sigma_i)</th>
<th>(\Delta Skew_i)</th>
<th>(\Delta Kurt_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta PK_{2ATM})</td>
<td>0.042</td>
<td>0.011</td>
<td>0.016</td>
<td>0.006</td>
</tr>
<tr>
<td>(\Delta \mu_i)</td>
<td>0.011</td>
<td>0.060</td>
<td>0.158</td>
<td>0.464</td>
</tr>
<tr>
<td>(\Delta \sigma_i)</td>
<td>0.016</td>
<td>0.158</td>
<td>0.541</td>
<td>1.684</td>
</tr>
<tr>
<td>(\Delta Skew_i)</td>
<td>0.006</td>
<td>0.464</td>
<td>1.684</td>
<td>8.866</td>
</tr>
<tr>
<td>(\Delta Kurt_i)</td>
<td>-0.036</td>
<td>2.460</td>
<td>10.190</td>
<td>55.301</td>
</tr>
</tbody>
</table>

Tab. 6.6: Pricing kernel differences covariance matrix (\(\tau=60\) days)

\[
R = GLG^T
\]

(6.10)

The next step is, therefore, to calculate the empirical correlation matrix \(R\) and decompose it using the spectral decomposition
where $\mathcal{R} \overset{\text{def}}{=} \text{diag}(l_1, l_2, \ldots, l_5)$ is a diagonal matrix containing the eigenvalues of $\mathcal{R}$ in a descending order and $\mathcal{G} \overset{\text{def}}{=} (g_1, g_2, \ldots, g_5)$ is a matrix containing the corresponding eigenvectors. By multiplying the original data with the eigenvectors matrix, we get the standardized principal components (PCs)

$$\mathcal{Y} = \mathcal{X} \mathcal{G}$$

(6.11)

The principal components can explain the variability of the data. The proportion of variance explained by a certain PC is the ratio of the corresponding eigenvalue to the sum of all eigenvalues, whereas the proportion of variance explained by the first few PCs is the sum of the proportions of variance explained by each of them. Tables 6.7 and 6.8 show the explained variance by the PCs of the differences of pricing kernel and relative risk aversion, respectively. Maturity of 120 days was dropped, since the kurtosis is not defined on some trading days.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau = 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.81</td>
<td>56.20%</td>
<td>56.20%</td>
</tr>
<tr>
<td>2</td>
<td>1.11</td>
<td>22.27%</td>
<td>78.47%</td>
</tr>
<tr>
<td>3</td>
<td>0.95</td>
<td>18.95%</td>
<td>97.42%</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>1.82%</td>
<td>99.24%</td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>0.76%</td>
<td>100.00%</td>
</tr>
<tr>
<td>$\tau = 60$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3.24</td>
<td>64.76%</td>
<td>64.76%</td>
</tr>
<tr>
<td>2</td>
<td>1.06</td>
<td>21.27%</td>
<td>86.03%</td>
</tr>
<tr>
<td>3</td>
<td>0.56</td>
<td>11.26%</td>
<td>97.29%</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>1.92%</td>
<td>99.21%</td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>0.79%</td>
<td>100.00%</td>
</tr>
<tr>
<td>$\tau = 90$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3.03</td>
<td>60.57%</td>
<td>60.57%</td>
</tr>
<tr>
<td>2</td>
<td>1.23</td>
<td>24.67%</td>
<td>85.24%</td>
</tr>
<tr>
<td>3</td>
<td>0.57</td>
<td>11.41%</td>
<td>96.65%</td>
</tr>
<tr>
<td>4</td>
<td>0.12</td>
<td>2.31%</td>
<td>98.96%</td>
</tr>
<tr>
<td>5</td>
<td>0.05</td>
<td>1.04%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Tab. 6.7: Principal components analysis of pricing kernel differences
### Principal Components Analysis of Relative Risk Aversion Differences

<table>
<thead>
<tr>
<th>Principal Component</th>
<th>Eigenvalue</th>
<th>Explained Variance</th>
<th>Cum. Expl. Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau = 30 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.87</td>
<td>37.48%</td>
<td>37.48%</td>
</tr>
<tr>
<td>2</td>
<td>1.63</td>
<td>32.55%</td>
<td>70.03%</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>19.95%</td>
<td>89.98%</td>
</tr>
<tr>
<td>4</td>
<td>0.36</td>
<td>7.23%</td>
<td>97.21%</td>
</tr>
<tr>
<td>5</td>
<td>0.14</td>
<td>2.79%</td>
<td>100.00%</td>
</tr>
<tr>
<td>( \tau = 60 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.88</td>
<td>37.58%</td>
<td>37.58%</td>
</tr>
<tr>
<td>2</td>
<td>1.77</td>
<td>35.45%</td>
<td>73.03%</td>
</tr>
<tr>
<td>3</td>
<td>0.98</td>
<td>19.53%</td>
<td>92.56%</td>
</tr>
<tr>
<td>4</td>
<td>0.23</td>
<td>4.58%</td>
<td>97.14%</td>
</tr>
<tr>
<td>5</td>
<td>0.14</td>
<td>2.86%</td>
<td>100.00%</td>
</tr>
<tr>
<td>( \tau = 90 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.17</td>
<td>43.46%</td>
<td>43.46%</td>
</tr>
<tr>
<td>2</td>
<td>1.37</td>
<td>27.33%</td>
<td>70.79%</td>
</tr>
<tr>
<td>3</td>
<td>0.97</td>
<td>19.37%</td>
<td>90.16%</td>
</tr>
<tr>
<td>4</td>
<td>0.27</td>
<td>5.46%</td>
<td>95.62%</td>
</tr>
<tr>
<td>5</td>
<td>0.22</td>
<td>4.38%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Tab. 6.8: Principal components analysis of relative risk aversion differences

By observing tables 6.7 and 6.8, it can be seen, that in order to explain at least 85% of the variance of both the pricing kernel and the relative risk aversion, three principal components are needed. The \( j^{th} \) eigenvector expresses the weights used in the linear combination of the original data in the \( j^{th} \) PC. Since we are considering only three PCs, the first three eigenvectors are of interest. More specifically, we can construct the first PCs for each of the examined time-series. The following demonstrates the weights of the moments in the PCs of the differences of the pricing kernel with a maturity of 60 days:

\[
\begin{align*}
y_{1,t}(\tau = 60) &= 0.06 \Delta \text{PK}^{\text{ATM}}_t + 0.92 \Delta \mu_t + 0.38 \Delta \sigma_t + 0.05 \Delta \text{Skew}_t - 0.03 \Delta \text{Kurt}_t \\
y_{2,t}(\tau = 60) &= 0.47 \Delta \text{PK}^{\text{ATM}}_t + 0.24 \Delta \mu_t - 0.58 \Delta \sigma_t - 0.54 \Delta \text{Skew}_t + 0.29 \Delta \text{Kurt}_t \\
y_{3,t}(\tau = 60) &= 0.52 \Delta \text{PK}^{\text{ATM}}_t + 0.06 \Delta \mu_t - 0.35 \Delta \sigma_t + 0.75 \Delta \text{Skew}_t - 0.21 \Delta \text{Kurt}_t
\end{align*}
\]

It can clearly be seen, that the dominant factors in the first PC are the changes in expectation and standard deviation, whereas the dominant factors in the second PC are the changes in skewness and standard deviation. The
dominant factors in the third PC are the changes in skewness and the ATM pricing kernel. As for the moments of the relative risk aversion, the first PC is dominated solely by the changes in standard deviation, the second PC is mainly dominated by the change in ATM RRA and the third PC is mainly dominated by the change in the RRA kurtosis. The dynamics of the first three PCs of the pricing kernel and relative risk aversion differences are presented in figure 6.11.

![Figure 6.11](EPKtimeseries.xpl)

The correlation of the principal components with the original moments can be calculated using the following relation:

\[ R_{X,Y} = \mathcal{G}\sqrt{L} \]  

(6.12)
where $L$ is standardized, i.e., the correlation between the $i^{th}$ moment and the $j^{th}$ PC is

$$r_{X_i,Y_j} = g_{ij} \sqrt{\frac{l_j}{s_{X_i}X_i}} \quad (6.13)$$

where $g_{ij}$ is the $i^{th}$ element of the $j^{th}$ eigenvector, $l_j$ is the corresponding eigenvalue and $s_{X_i}X_i$ is the standard deviation of $X_i$. Note that

$$\sum_{j=1}^{5} r_{X_i,Y_j}^2 = \sum_{j=1}^{5} \frac{g_{ij}^2 l_j}{s_{X_i}X_i} = \frac{s_{X_i}X_i}{s_{X_i}X_i} = 1 \quad (6.14)$$

meaning essentially, that for every two PCs $Y_{j_1}$ and $Y_{j_2}$, $(j_1, j_2) \in \{1, \ldots, 5\}$ the following has to hold

$$r_{X_i,Y_{j_1}}^2 + r_{X_i,Y_{j_2}}^2 \leq 1 \quad (6.15)$$

so that the points are always inside a unit circle. The closer the variables are to the circumference of the circle, the higher the correlation between them and the PCs, which are represented by the axes, is. Figure 6.12 depicts the correlations of the first three PCs with the original moments for a maturity of 60 days. The moments highly correlated with the PCs are, not surprisingly, the ones which were reported to be dominant when constructing the PCs. Descriptive statistics of the PC time-series and their correlations with the moments are given in tables 6.9 and 6.10 for the PK and RRA respectively. The means of the PCs are very close to zero, as they are linear combinations of the differences of the moments, which are themselves very close to zero.

We conclude therefore, that the variation of the pricing kernel and relative risk aversion differences can be explained by three factors. The first factor of pricing kernel differences can be perceived as a central mass movement factor, consisting of the changes in expectation and standard deviation. The second factor can be perceived as a change of tendency factor, consisting of changes in skewness and standard deviation. The third factor is mainly dominated by the change in the ATM pricing kernel, meaning the shifts in investors’ beliefs where the market is the most sensitive.

The principal components of the relative risk aversion are a little different. The first factor can be perceived as a dispersion change factor, dominated by the change in standard deviation. The second factor is dominated by the change in relative risk aversion of the investors at the money. The third factor is again a dispersion change factor, dominated by the change in kurtosis. The mean RRA seems to play no role in examining the variability of the relative risk aversion.
Fig. 6.12: Correlation of the three principal components of the differences of pricing kernel (left panel) and relative risk aversion (right panel) with the original moments ($\tau = 60$ days)
### Tab. 6.9: Descriptive Statistics: The PCs of the PK differences

<table>
<thead>
<tr>
<th>Principal Component</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>$\Delta P K^A T_M$</th>
<th>Correlation with $\Delta \mu_t$, $\Delta \sigma_t$, $\Delta Skew_t$, $\Delta Kurt_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau = 30$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_{1,t}$</td>
<td>$-2.46 \times 10^{-4}$</td>
<td>0.76</td>
<td>-0.02</td>
<td>0.42, 0.62, 0.02, -0.02</td>
</tr>
<tr>
<td>$y_{2,t}$</td>
<td>$-4.39 \times 10^{-4}$</td>
<td>4.15</td>
<td>0.21</td>
<td>0.25, -0.16, 0.29, 0.08</td>
</tr>
<tr>
<td>$y_{3,t}$</td>
<td>$3.36 \times 10^{-5}$</td>
<td>3.51</td>
<td>0.23</td>
<td>0.15, -0.09, -0.33, -0.04</td>
</tr>
<tr>
<td>$\tau = 60$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_{1,t}$</td>
<td>$4.34 \times 10^{-4}$</td>
<td>0.44</td>
<td>0.06</td>
<td>0.74, 0.30, 0.04, -0.03</td>
</tr>
<tr>
<td>$y_{2,t}$</td>
<td>$8.53 \times 10^{-4}$</td>
<td>4.06</td>
<td>0.22</td>
<td>0.11, -0.27, -0.25, 0.13</td>
</tr>
<tr>
<td>$y_{3,t}$</td>
<td>$-1.24 \times 10^{-3}$</td>
<td>2.27</td>
<td>0.17</td>
<td>0.02, -0.12, 0.25, -0.07</td>
</tr>
<tr>
<td>$\tau = 90$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_{1,t}$</td>
<td>$2.80 \times 10^{-4}$</td>
<td>0.55</td>
<td>0.09</td>
<td>-0.61, 0.46, 0.11, -0.05</td>
</tr>
<tr>
<td>$y_{2,t}$</td>
<td>$9.20 \times 10^{-4}$</td>
<td>2.04</td>
<td>0.23</td>
<td>-0.19, -0.21, -0.32, 0.11</td>
</tr>
<tr>
<td>$y_{3,t}$</td>
<td>$6.90 \times 10^{-4}$</td>
<td>0.72</td>
<td>0.18</td>
<td>-0.04, -0.15, 0.24, -0.03</td>
</tr>
</tbody>
</table>

### Tab. 6.10: Descriptive Statistics: The PCs of the RRA differences

<table>
<thead>
<tr>
<th>Principal Component</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>$\Delta R R A^A T_M$</th>
<th>Correlation with $\Delta \mu_t$, $\Delta \sigma_t$, $\Delta Skew_t$, $\Delta Kurt_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau = 30$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_{1,t}$</td>
<td>$1.15 \times 10^{-2}$</td>
<td>14.75</td>
<td>0.03</td>
<td>0.04, 0.61, 0.00, 0.01</td>
</tr>
<tr>
<td>$y_{2,t}$</td>
<td>$5.46 \times 10^{-4}$</td>
<td>9.36</td>
<td>0.33</td>
<td>-0.22, -0.02, -0.32, 0.26</td>
</tr>
<tr>
<td>$y_{3,t}$</td>
<td>$-1.32 \times 10^{-3}$</td>
<td>11.75</td>
<td>0.29</td>
<td>0.14, -0.01, -0.05, -0.31</td>
</tr>
<tr>
<td>$\tau = 60$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_{1,t}$</td>
<td>$-2.57 \times 10^{-3}$</td>
<td>26.90</td>
<td>0.10</td>
<td>0.04, 0.60, -0.02, 0.03</td>
</tr>
<tr>
<td>$y_{2,t}$</td>
<td>$1.60 \times 10^{-3}$</td>
<td>13.75</td>
<td>0.36</td>
<td>0.20, -0.06, -0.24, -0.35</td>
</tr>
<tr>
<td>$y_{3,t}$</td>
<td>$4.95 \times 10^{-4}$</td>
<td>12.87</td>
<td>0.03</td>
<td>0.32, -0.04, -0.10, 0.29</td>
</tr>
<tr>
<td>$\tau = 90$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_{1,t}$</td>
<td>$1.72 \times 10^{-3}$</td>
<td>28.60</td>
<td>-0.08</td>
<td>0.15, 0.63, 0.05, 0.04</td>
</tr>
<tr>
<td>$y_{2,t}$</td>
<td>$3.71 \times 10^{-3}$</td>
<td>9.22</td>
<td>0.18</td>
<td>0.36, -0.05, -0.27, 0.20</td>
</tr>
<tr>
<td>$y_{3,t}$</td>
<td>$8.51 \times 10^{-4}$</td>
<td>19.11</td>
<td>-0.19</td>
<td>0.25, -0.07, 0.01, -0.29</td>
</tr>
</tbody>
</table>
Tables 6.9 implies an inconsistent behavior of the different moments across maturities. The first PCs of the PK differences (the first rows for each of the maturities in table 6.9) are positively correlated with the changes in mean and standard deviation (the dominating moments) for short term maturities, but negatively correlated with the mean differences of 90 days maturity pricing kernels. The second PCs of pricing kernel differences (the second rows for each of the maturities in table 6.9) are negatively correlated with the change of standard deviation for all maturities, but their correlations with the change of skewness are not consistent across maturities, implying a bad fit. Since the first PC of the pricing kernel differences could explain approximately 60% of the variability, whereas the second factor can explain only 20%, the inconsistent behavior could be justified by the poor contribution of the second PC to the variability.

The correlations of the first and second PCs of the relative risk aversion differences with their dominant factors (table 6.10) are found to be consistent across maturities. The first PC is positively correlated with its most dominant moment, the changes in the RRA standard deviation. This correlation means essentially, that the less homoscedastic the RRA is, i.e. the larger the changes in standard deviation are, the larger the first PC of RA differences become. The second PC of RRA differences is positively correlated with its most dominant moment, the behavior at the money. The more volatile the ATM RRA is, the higher the second PC is. Both PCs of the RRA differences contribute more than 30% of the variability and imply a good fit of the PCs to the data.
6.4 A Time-Series Model for the Principal Components

The former section focused on explaining the variability of the differences of pricing kernel and relative risk aversion using a small numbers of factors. We found, that three principal components can explain most of the variability of the data. In this section, we will try to fit a time-series model to the PCs, which were estimated in the previous section.

If the principal components were indeed orthogonal to each other, a univariate analysis of each of the PCs would be sufficient for estimating an adequate model. However, the third PC for both PK and RRA differences is found to be highly correlated with the second PC. Therefore, although the third PC helps in explaining the variability of the data, it is dropped from the univariate analysis performed in this section, leaving only the first two PCs to be analyzed.

The first step is to check the autocorrelation and the partial autocorrelation functions of the time-dependent PCs. This is illustrated in figures 6.13 and 6.14 for the PK and RRA differences respectively. The lower panels of the figures depict the ACF and PACF of the third PC, and it is seen to have slightly different characteristics than the first two. The third PC is depicted here only to give an impression of its behavior, it will not be analyzed as mentioned above. In addition, since the PCs have similar autocorrelation and partial autocorrelation functions for different maturities, we concentrate in this section on a maturity of 60 days and report, that the findings are similar when other maturities are considered. For the first two PCs, the ACF drops abruptly after the first order autocorrelation whereas the PACF decays gradually. These characteristics imply a MA(1) behavior (Chapter 11 in [Franke et al., 2004]) and we therefore concentrate on fitting a model with a moving average component to the PCs.

Since the ACF and PACF plots are qualitative measures, we need to check more accurately, which model fits the PCs best. If the fitted model is adequate, the residuals should be approximately white noise. Therefore, we should first check the residuals mean and autocorrelation. The key instruments are the timeplot, the ACF and the PACF of the residuals, as well as the Ljung-Box statistic, checking for autocorrelation, the Akaike Information Criterion (AIC) and the Schwarz Information Criterion (SIC) checking for model adequacy.
Fig. 6.13: Autocorrelation function (left panel) and partial autocorrelation function (right panel) of the principal components of pricing kernel differences ($\tau = 60$ days)
Fig. 6.14: Autocorrelation function (left panel) and partial autocorrelation function (right panel) of the principal components of relative risk aversion differences ($\tau = 60$ days)
The Ljung-Box statistic $Q_{LB}$ checks for autocorrelation of the residuals. The null hypothesis is that, a set of $M$ residuals are not autocorrelated, i.e. $H_0: \rho_1 = \rho_2 = \cdots = \rho_M = 0$ and the statistic can be calculated as:

$$Q_{LB} = T(T + 2) \sum_{j=1}^{M} \frac{r_j^2}{T-j}$$

(6.16)

where $T$ is the length of the period and $r_j$ is the sample autocorrelation of order $j$. The statistic is $\chi^2(M)$ distributed under $H_0$ and should therefore be as small as possible for the model to be adequate. The information criteria AIC and SIC are monotonic functions of the standard deviation of errors of the proposed ARMA($p,q$) model, and therefore, the smaller the criteria’s values are, the better model is.

$$AIC = \ln(\hat{\sigma}^2_{\varepsilon}) + \frac{2(p + q)}{T}$$ $$SIC = \ln(\hat{\sigma}^2_{\varepsilon}) + \frac{p + q}{T} \ln(T)$$

(6.17)

As stated before, we focus on the first two PCs of the PK and RRA differences for a maturity of 60 days. The characteristics of the two first PCs for other maturities were found to resemble those reported next. Table 6.11 shows the AIC and SIC values as means of comparison between the various suggested time-series models for the first two PCs. Many researchers tend to prefer the more parsimonious SIC in case of contradiction between the two criteria, and the selected models are therefore determined according to the SIC. According to Table 6.11, we can conclude, that the first PCs of PK and RRA differences follow an ARMA process (ARMA(1,1) and ARMA(1,2) respectively), whereas the second PCs follow a MA(1) process. All PCs have an autocorrelated error term as estimated at the beginning of this section. Using the conditional non linear least squares estimation method, we obtain

$$y_{1,t}^{\Delta PK} = 0.141y_{1,t-1}^{\Delta PK} - 0.923\varepsilon_{t-1} + \varepsilon_t$$ $$y_{2,t}^{\Delta PK} = -0.952\varepsilon_{t-1} + \varepsilon_t$$ $$y_{1,t}^{\Delta RRA} = -0.309y_{1,t-1}^{\Delta RRA} - 0.555\varepsilon_{t-2} - 0.418\varepsilon_{t-1} + \varepsilon_t$$ $$y_{2,t}^{\Delta RRA} = -0.978\varepsilon_{t-1} + \varepsilon_t$$

(6.18)
### Tab. 6.11: Comparison between different time-series models for the PCs of PK and RRA differences based on Akaike and Schwarz information criteria. The minimal values are bolded, best model is selected based on SIC in case of contradiction.

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC 1.PC</th>
<th>AIC 2.PC</th>
<th>SIC 1.PC</th>
<th>SIC 2.PC</th>
<th>AIC 1.PC</th>
<th>AIC 2.PC</th>
<th>SIC 1.PC</th>
<th>SIC 2.PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.090</td>
<td>2.148</td>
<td>6.044</td>
<td>4.653</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.082</td>
<td>2.155</td>
<td>6.051</td>
<td><strong>4.660</strong></td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>-2.103</strong></td>
<td>2.150</td>
<td>6.034</td>
<td><strong>4.649</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>-2.088</strong></td>
<td>2.165</td>
<td>6.049</td>
<td>4.664</td>
</tr>
<tr>
<td>AR(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.854</td>
<td>2.471</td>
<td>6.392</td>
<td>5.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.846</td>
<td>2.478</td>
<td>6.400</td>
<td>5.016</td>
</tr>
<tr>
<td>ARMA(2,1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.101</td>
<td><strong>2.145</strong></td>
<td>6.027</td>
<td>4.654</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.079</td>
<td>2.167</td>
<td>6.050</td>
<td>4.677</td>
</tr>
<tr>
<td>ARMA(1,2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.101</td>
<td>2.145</td>
<td><strong>6.026</strong></td>
<td>4.654</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.078</td>
<td>2.168</td>
<td><strong>6.049</strong></td>
<td>4.676</td>
</tr>
<tr>
<td>ARMA(2,2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.098</td>
<td>2.148</td>
<td>6.029</td>
<td>4.657</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.068</td>
<td>2.178</td>
<td>6.059</td>
<td>4.687</td>
</tr>
<tr>
<td><strong>Best Fit</strong></td>
<td>ARMA(1,1)</td>
<td>MA(1)</td>
<td>ARMA(1,2)</td>
<td>MA(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. Dynamic Model: Time-Series Analysis

6.5 GLS Regression Model for the Principal Components

Conditional least squares estimation does not always yield accurate estimates when moving average models are involved, as the least squares estimates become non-linear. Therefore, an alternative model is introduced in this section. This model tests the existence of a relation between the principal component and easily observed data, such as changes in the DAX level and implied volatility at the money.

The simplest relation between an explanatory variable and a response variable can be described and examined using a simple linear regression model

\[ y = X\beta + \epsilon \]  \hspace{1cm} \text{(6.19)}

where \( y \) is a \( n \times 1 \) response vector, \( X \) is a \( n \times p \) explanatory matrix, \( \beta \) is a \( p \times 1 \) vector of parameters to estimate and \( \epsilon \) is a \( n \times 1 \) vector of errors. If the errors were normally distributed and uncorrelated, i.e. \( \epsilon \sim N_n(0,\sigma^2 I_n) \) then the regression would result in the familiar ordinary least squares (OLS) estimator

\[ \hat{\beta}_{OLS} = (X'X)^{-1}X'y \]  \hspace{1cm} \text{(6.20)}

with a covariance matrix

\[ \text{Cov}(\hat{\beta}_{OLS}) = \sigma^2(X'X)^{-1} \]  \hspace{1cm} \text{(6.21)}

Introducing autocorrelated errors as in the moving average model described in the previous section, the relation between the explanatory variable and the response variable can be modeled using the generalized least squares (GLS) estimator. In the previous section, we found evidence of autocorrelated errors of order 1, meaning that the error process could be modeled using the following AR(1) process

\[ \epsilon_t = \rho \epsilon_{t-1} + u_t \]  \hspace{1cm} \text{(6.22)}

for all \( t \in \{1,\ldots,n\} \) with \( u_t \sim N_n(0,\sigma^2 I_n) \) as i.i.d. white noise and \( |\rho| < 1 \) for stability. We could choose autoregressive processes of higher order, but since most principal components of PK and RRA differences were found to have an autocorrelated error term of order 1, we concentrate here on AR(1) processes.

Iterating equation (6.22) from time 0 onwards yields

\[ \epsilon_t = \lim_{n \to \infty} (\rho^{n+1} \epsilon_{t-n-1} + \sum_{s=0}^{n-1} \rho^s u_{t-s}) = \sum_{s=0}^{\infty} \rho^s u_{t-s} \]  \hspace{1cm} \text{(6.23)}
and hence

\[
\begin{align*}
E[\epsilon_t] &= 0 \\
\text{Var}(\epsilon_t) &= \sigma_u^2 \sum_{s=0}^{\infty} (\rho^2)^s = \frac{\sigma_u^2}{1 - \rho^2} \\
\text{Cov}(\epsilon_t, \epsilon_{t+\tau}) &= \sum_{s=0}^{\infty} \rho^{2s+\tau} \sigma_u^2 = \rho^r \frac{\sigma_u^2}{1 - \rho^2}
\end{align*}
\]

(6.24)

so the covariance matrix of the error term is

\[
\text{Cov}(\epsilon) = \sigma_u^2 \Omega = \frac{\sigma_u^2}{1 - \rho^2} \begin{pmatrix}
1 & \rho & \rho^2 & \ldots & \rho^{n-1} \\
\rho & 1 & \rho & \ldots & \rho^{n-2} \\
\rho^2 & \rho & 1 & \ldots & \rho^{n-3} \\
& \vdots & \vdots & \ddots & \vdots \\
\rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \ldots & 1
\end{pmatrix}
\]

(6.25)

However, in a real application like the model discussed in this work, the error-covariance matrix is not known and must be estimated from the data along with the regression coefficients \( \hat{\beta} \). If the generating process is stationary, which is the case in the model discussed here, a commonly used algorithm for estimating these errors is normally referred to as the Prais & Winsten (1954) procedure. This algorithm begins with running a standard OLS regression and examining the residuals. The errors vector of the OLS regression is obtained simply by plugging \( \hat{\beta} \) in equation (6.19). Considering the residuals’ first order autocorrelations from the preliminary OLS regression can suggest a reasonable form for the error-generating process. These first order autocorrelations can be estimated as

\[
\hat{\rho} = \frac{\sum_{t=2}^{n} \epsilon_t \epsilon_{t-1}}{\sum_{t=1}^{n} \epsilon^2_t}
\]

(6.26)

Replacing the \( \rho \)’s in equation (6.25) with the \( \hat{\rho} \)’s from equation (6.26) results in the estimated matrix \( \hat{\Omega} \). The best linear unbiased estimator in that case would be the estimated generalized least squares estimator

\[
\hat{\beta}_{\text{GLS}} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y
\]

(6.27)

The Prais & Winsten (1954) algorithm may seem as a simple model, but it involves a computationally challenging estimation of \( \hat{\Omega} \). Therefore, an alternative algorithm is suggested and discussed by Sen & Srivastava (1990).
We define the following matrix as

\[
\hat{\Psi} = \begin{pmatrix}
\sqrt{1 - \hat{\rho}^2} & 0 & 0 & \ldots & 0 & 0 & 0 \\
-\hat{\rho} & 1 & 0 & \ldots & 0 & 0 & 0 \\
0 & -\hat{\rho} & 1 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & -\hat{\rho} & 1 & 0 \\
0 & 0 & 0 & \ldots & 0 & -\hat{\rho} & 1 \\
\end{pmatrix}
\]

(6.28)

and hence

\[
\frac{1}{1 - \hat{\rho}^2} \hat{\Psi}'\hat{\Psi} = \begin{pmatrix}
\frac{1}{1 - \hat{\rho}^2} & -\hat{\rho} & 0 & 0 & \ldots & 0 & 0 & 0 \\
-\hat{\rho} & \frac{1 + \hat{\rho}^2}{1 - \hat{\rho}^2} & -\hat{\rho} & 0 & \ldots & 0 & 0 & 0 \\
0 & \frac{1}{1 - \hat{\rho}^2} & \frac{1 + \hat{\rho}^2}{1 - \hat{\rho}^2} & -\hat{\rho} & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \ldots & \frac{1 + \hat{\rho}^2}{1 - \hat{\rho}^2} & -\hat{\rho} & 0 \\
0 & 0 & 0 & 0 & \ldots & \frac{1}{1 - \hat{\rho}^2} & -\hat{\rho} & 0 \\
0 & 0 & 0 & 0 & \ldots & 0 & \frac{1}{1 - \hat{\rho}^2} & \frac{1 - \hat{\rho}^2}{1 - \hat{\rho}^2} \\
\end{pmatrix}
\]

Multiplying with \(\hat{\Omega}\) (which is defined in equation (6.25)) yields

\[
\frac{1}{1 - \hat{\rho}^2} \hat{\Psi}'\hat{\Psi} \hat{\Omega} = I_n
\]

and hence the matrix \(\hat{\Psi}\) has the following property

\[
\hat{\Psi}'\hat{\Psi} = (1 - \hat{\rho}^2)\hat{\Omega}^{-1}
\]

(6.29)

Since least squares estimation is not affected by scalar multiplication, we multiply the regression model by \(\sqrt{1 - \hat{\rho}^2}\). Expressing \(\hat{\Omega}^{-1}\) in equation (6.27) using equation (6.29) leads to the the following GLS estimator

\[
\hat{\beta}_{GLS} = (X'\hat{\Psi}'\hat{\Psi}X)^{-1}X'\hat{\Psi}'\hat{\Psi}y = ((\hat{\Psi}X)'(\hat{\Psi}X))^{-1}((\hat{\Psi}X)'(\hat{\Psi}y))
\]

(6.30)

which is actually an OLS estimator of the original variables multiplied by a scalar. The transformed model can be described as

\[
y_t - \hat{\rho}y_{t-1} = \sum_{j=0}^{p}(x_{tj} - \hat{\rho}x_{t-1,j})\beta_j + u_t
\]

(6.31)
for \( t \in \{2, \ldots, n\} \) whereas for \( t = 1 \) it is simply

\[
\sqrt{1 - \hat{\rho}^2} y_1 = \sqrt{1 - \hat{\rho}^2} \sum_{j=0}^p \beta_j x_{1j} + \sqrt{1 - \hat{\rho}^2} \epsilon_1
\]  

(6.32)

As stated in the beginning of the current section, the changes in the DAX level \( (S_t) \) and the changes of ATM implied volatility \( (IV_{t\,ATM}^M) \) were chosen to be tested as explanatory variables \( (X) \), whereas the first two principal components of the PK and RRA differences for a different maturities were the dependent variables for the different models \( (y) \). Since the dependency on the explanatory variable does not have to be linear, different functions of the explanatory variables were tested. For each of the explanatory variables the differences, the squared differences, the logarithmic differences and the squared logarithmic differences were tested. The examined models consisted of all possible combinations between the functions stated above, as well as checking for interactions in each of the proposed models. Since no interaction was ever found to be significant, they were dropped from the model. The criterion for choosing the best model was a maximal value of the F-statistic.

The GLS regression was implemented in EPKregression.xpl. In the following, the best fitted models for each of the PCs are presented (based on equation (6.31)), where \( \Delta IV_{t\,ATM}^M = IV_{t\,ATM}^M - IV_{t-1\,ATM}^M \) and \( \Delta S_t = S_t - S_{t-1} \). For this analysis, we consider a confidence level of 95%, i.e. any regression or regression coefficient yielding a Pvalue > 5% is regarded as non significant. The Pvalues for the regression’s coefficients appear in brackets.

The regression for the first PC of the PK differences for a maturity of 30 days is:

\[
y_{\Delta PK, t} + 0.426 y_{\Delta PK, t-1} = 0.00 - 1.80 \left[ \log \frac{S_t}{S_{t-1}} + 0.426 \log \frac{S_{t-1}}{S_{t-2}} \right]_{(0.994)} + 0.289 + 1.76 \left[ \log \frac{IV_{t\,ATM}^M}{IV_{t-1\,ATM}^M} + 0.426 \log \frac{IV_{t-1\,ATM}^M}{IV_{t-2\,ATM}^M} \right]_{(0.000)} + \epsilon_t
\]

with \( F = 18.958 \) (Pvalue = 0.000).
The regression for the first PC of the PK differences for a maturity of 60 days is:

\[ y_{1,t}^{\Delta PK} + 0.468y_{1,t-1}^{\Delta PK} = 0.00 + 2.71 \left( \log \frac{S_t}{S_{t-1}} + 0.468 \log \frac{S_{t-1}}{S_{t-2}} \right) + 0.00 + 2.71 \left( \log \frac{IV_{ATM,t}}{IV_{ATM,t-1}} + 0.468 \log \frac{IV_{ATM,t-1}}{IV_{ATM,t-2}} \right) + u_t \]

with \( F = 10.784 \) (\( P \text{value} = 0.000 \)).

The regression for the first PC of the PK differences for a maturity of 90 days was not found to be significant and hence not stated.

The first PC of the PK differences, which was described in Section 6.3 as a central mass movement factor, dominated by the changes in expected pricing kernel and the pricing kernel’s standard deviation, is found to depend significantly on the logarithmic differences of ATM implied volatility. This regression is only significant for short term maturities, and the impact of the explanatory variables is positive and log-linear. The impact of the DAX log return is not significant for a short term maturity, meaning the first PC of the PK differences is mainly influenced by the logarithmic changes in the implied volatility at the money. Therefore, we can deduce the following: The larger the changes in ATM implied volatility are and the higher the DAX log returns are (only for maturities of 60 days), the more volatile the pricing kernel becomes, with bigger daily changes in its expectation and standard deviation.

The regression for the second PC of the PK differences for a maturity of 30 days is:

\[ y_{2,t}^{\Delta PK} + 0.473y_{2,t-1}^{\Delta PK} = 0.00 + 30.21 \left( \log \frac{S_t}{S_{t-1}} + 0.473 \log \frac{S_{t-1}}{S_{t-2}} \right) + u_t \]

with \( F = 20.718 \) (\( P \text{value} = 0.000 \)).

The regressions for the second PCs of the PK differences for the maturities of 60 and 90 days were not found to be significant and hence not stated.
We could not find a significant relationship between the second PC of PK differences and the explanatory variables (other than for very short maturities), a result that supports the second PC’s smaller contribution to the variance of PK differences. The pricing kernel differences have one dominant factor which explains approximately 60% of their variance and depends mainly on the logarithmic changes of the ATM implied volatility. The regression coefficients are positive, as are the correlations of the first PC with $\Delta \mu_t(PK)$ and $\Delta \sigma_t(PK)$ for the respective maturities.

The results regarding the PCs of RRA differences are quite different. These PCs are related to the absolute changes in the DAX level and in ATM implied volatility. The dependence is not log-linear, but strictly linear.

The regression for the first PC of the RRA differences for a maturity of 30 days is:

$$y_{t,t}^{\Delta RRA} + 0.544y_{t-1,t}^{\Delta RRA} = -0.03 + 0.03 [\Delta S_t + 0.544\Delta S_{t-1}] +$$

$$+ 145.34 [\Delta IV_{t,ATM} + 0.544\Delta IV_{t-1,ATM}] + u_t$$

with $F = 11.562$ (Pvalue = 0.000).

The regression for the first PC of the RRA differences for a maturity of 60 days is:

$$y_{t,t}^{\Delta RRA} + 0.457y_{t-1,t}^{\Delta RRA} = -0.02 + 0.03 [\Delta S_t + 0.457\Delta S_{t-1}] +$$

$$+ 286.43 [\Delta IV_{t,ATM} + 0.457\Delta IV_{t-1,ATM}] + u_t$$

with $F = 18.048$ (Pvalue = 0.000).

The regression for the first PC of the RRA differences for a maturity of 90 days is:

$$y_{t,t}^{\Delta RRA} + 0.510y_{t-1,t}^{\Delta RRA} = -0.01 + 0.02 [\Delta S_t + 0.510\Delta S_{t-1}] +$$

$$+ 224.27 [\Delta IV_{t,ATM} + 0.510\Delta IV_{t-1,ATM}] + u_t$$

with $F = 10.666$ (Pvalue = 0.000).
According to table 6.10, the correlations of the first PCs of RRA differences with their dominant moments are positive. The first PC of RRA differences is a dispersion factor, dominated by the change in the relative risk aversion standard deviation. According to the regression, large changes in the DAX level and the ATM implied volatility yield a larger PC, which is associated with a larger change in risk aversion standard deviation. This result implies the existence of more uncertain investors with a more heteroscedastic risk aversion, when the DAX level and ATM implied volatility are more time-varying. This relation could be explained by the dispersion of information sets among investors. Veldkamp (2005) examines the impact of information markets on assets prices. She basically claims, that information markets, not assets markets, are the source of frenzies and herds in assets prices. However, the price fluctuations on the market affect these information sets and determine the information prices, which are incorporated in the investors’ subjective beliefs. More volatile markets lead necessarily to a higher risk and to less information, which increases the demand for information in a competitive market. Hence, more volatile markets cause more information to be provided at a lower price. When less information is involved, individual agents are willing to pay for information, and the information sets of the individual agents become more dispersed. More dispersed information sets could increase heteroscedasticity of the aggregate relative risk aversion as a function of assets’ returns.

The results regarding the second PC of the RRA differences are slightly different. Regressing the second PC of the RRA differences for a maturity of 30 days on the same explanatory variables was not found to be significant and hence not stated here.

The regression for the second PC of the RRA differences for a maturity of 60 days is:

\[
y_{2,t}^{\Delta RRA} + 0.460y_{2,t-1}^{\Delta RRA} = 0.01 - 0.01 \left[ \Delta S_t + 0.460 \Delta S_{t-1} \right] - 92.15 \left[ \Delta IV_t^{ATM} + 0.460 \Delta IV_{t-1}^{ATM} \right] + u_t
\]

with \( F = 7.217 \) (\( P \)value = 0.001).
The regression for the second PC of the RRA differences for a maturity of 90 days is:

\[ y_{2,t}^{\Delta RRA} + 0.497y_{2,t-1}^{\Delta RRA} = 0.00 + 0.01[\Delta S_t + 0.497\Delta S_{t-1}] + 0.00 + 0.01[\Delta S_t + 0.497\Delta S_{t-1}] + u_t \]

\[ + 35.72[\Delta IV_t^{ATM} + 0.497\Delta IV_{t-1}^{ATM}] + u_t \]

with \( F = 4.026 \) (Pvalue = 0.018).

The second PC's of RRA differences are positively correlated to the change of relative risk aversion at the money (table 6.10). Nevertheless, the linear regression is not significant for a very short term maturity of 30 days. For long term maturities the coefficients of the regression are positive, whereas for medium term maturities, they are negative. That could be interpreted as follows: When the changes in DAX level and ATM implied volatility are larger, the relative risk aversion at the money is more volatile for long term maturities, but is less volatile for the medium term maturities.

From this section we can conclude, that the principal components model fits the relative risk aversion differences better than it fits the pricing kernel differences. We were able to fit an autocorrelated regression model to the first PC of PK differences for short and medium term maturities, and to both PCs of RRA differences. All of these models do not have a significant constant term \( \beta_0 \), which is in accordance with the mean-reversion property of the moments differences that define the PCs. The autocorrelation is indeed found to be quite large (approximately -0.5) for all of the above models, implying the existence of an autocorrelated error as detected already in the previous section.
7. FINAL STATEMENTS

Risk managers can often extract useful information about market expectations and investors’ behavior from option prices and stock indices. This work focused on estimating the subjective density and the state-price density of the stochastic process associated with the DAX. According to [Rubinstein (1994)], a good estimation of those two measures is sufficient for deriving the investors’ preferences. This work did not include a direct approximation of the utility function based on empirical data, but rather an estimation of the pricing kernel and the relative risk aversion as functions of the return states. The utility function could be approximated numerically by solving the differential equations discussed in Chapter 2, after the pricing kernel and relative risk aversion function have been estimated. Nevertheless, this work aimed at examining the dynamics of these two measures, characterizing the investors’ behavior, rather than deriving their implied utility function.

The use of a simulated GARCH (1,1) model to estimate the investors’ subjective density, and the concept of local polynomial estimation to derive the state-price density from the implied volatilities, resulted in a daily estimation of the pricing kernel and the relative risk aversion across return states. The PK and RRA were estimated with respect to four different maturities, in order to check the consistency of the results. The daily estimated PK and RRA were found to have similar characteristics to those reported by [Jackwerth (2000) and Aıt Sahalia & Lo (2000)]. The PK was shown not to be a strictly decreasing function as suggested by classical macroeconomic theory, and the RRA experienced some negative values at the money. These findings were apparent throughout the three year long database, implying existence of risk seeking investors with a locally convex utility function, possibly due to some frictions in the representative agent’s model.

The daily PK and RRA were characterized by their first four moments (with respect to the return states) as well as their values at the money. The absolute changes in those moments were found to be stationary and served as the starting point of a principal component analysis. The variability of the changes in PK and RRA was found to be well explained by two factors.
For the PK differences, one factor dominated the analysis and could explain some 60% of the variance. This factor was perceived as a dispersion change factor, relating to the daily changes in the mean and standard deviation of the PK. The RRA differences were found to be well explained by two factors, each contributing around 35% to the explained variance. The first factor was associated with the changes in the standard deviation of the RRA, whereas the second factor was dominated by the change in RRA at the money. This result substantiated the importance of collecting the ATM values.

The principal components have undergone a time-series analysis and were found to have a significant first order autocorrelation. ARMA models were fitted to the first PCs of PK and RRA differences, whereas the second PCs were best described by MA models. The evident first order autocorrelation was estimated and incorporated in a GLS regression model, which regressed each of the principal components on the daily changes in the DAX and in ATM implied volatility.

The first PC of the PK differences, which was perceived as a central mass movement factor, was found to depend mainly on the logarithmic differences of ATM implied volatility. However, this dependency was only found to be significant for short term maturities. For a maturity of 60 days, the DAX log return had a significant impact, but this result was not consistent across other maturities. Since the PC was found to be positively correlated to the changes in mean and standard deviation of the PK, we concluded that large changes in ATM implied volatility lead to a more volatile and time-varying PK. The absence of a significant fitted regression model for the second PC of the PK differences was in accordance with its less significant contribution to the explained variability.

The first PC of RRA differences, dominated by the daily changes in the RRA standard deviation, was found to depend significantly on the absolute changes in the DAX level and the ATM implied volatility. We found evidence for the existence of more uncertain investors with a more heteroscedastic risk aversion, when the daily changes in the DAX and the ATM implied volatility were larger. This result was explained by possibly more dispersed information sets among investors. The second PC of RRA differences, dominated by the changes in RRA at the money, was also found to depend significantly on the changes in the DAX level and the ATM implied volatility. Large changes in the DAX level and the ATM implied volatility result in a more volatile RRA at the money for long term maturities, but a less volatile RRA at the money for short term maturities.
A. SOME MORE GRAPHICAL ILLUSTRATIONS

In this appendix the time-series formed by the differences and the differences of the logarithms of the different functions mentioned in Chapter 6 are presented for both estimates, the pricing kernel and the relative risk aversion, each estimated at four different maturities.

![Graphs of ATM Pricing Kernel for different maturities](EPKtimeseries.xpl)

*Fig. A.1: Differences of ATM Pricing Kernel for different maturities (30, 60, 90, 120 days)*
A. Some More Graphical Illustrations

Fig. A.2: Differences of mean Pricing Kernel for different maturities (30, 60, 90, 120 days).

Fig. A.3: Differences of Standard Deviation of Pricing Kernel for different maturities (30, 60, 90, 120 days).
Fig. A.4: Differences of Skewness of Pricing Kernel for different maturities (30, 60, 90, 120 days).

Fig. A.5: Differences of Kurtosis of Pricing Kernel for different maturities (30, 60, 90, 120 days).
Fig. A.6: Differences of ATM Relative Risk Aversion for different maturities (30, 60, 90, 120 days).

Fig. A.7: Differences of mean Relative Risk Aversion for different maturities (30, 60, 90, 120 days).
Fig. A.8: Differences of Standard Deviation of Relative Risk Aversion for different maturities (30, 60, 90, 120 days).

Fig. A.9: Differences of Skewness of Relative Risk Aversion for different maturities (30, 60, 90, 120 days).
Fig. A.10: Differences of Kurtosis of Relative Risk Aversion for different maturities (30, 60, 90, 120 days).

Fig. A.11: Log differences ATM Pricing Kernel for different maturities (30, 60, 90, 120 days).
Fig. A.12: Log differences of Mean Pricing Kernel for different maturities (30, 60, 90, 120 days).

Fig. A.13: Log differences of Standard Deviation of Pricing Kernel for different maturities (30, 60, 90, 120 days).
Fig. A.14: Log differences of Skewness of Pricing Kernel for different maturities (30,60,90,120 days).

Fig. A.15: Log differences of Kurtosis of Pricing Kernel for different maturities (30,60,90,120 days).
Fig. A.16: Log differences of Mean Risk Aversion for different maturities (30, 60, 90, 120 days).

Fig. A.17: Log differences of Standard Deviation of Risk Aversion for different maturities (30, 60, 90, 120 days).
A. Some More Graphical Illustrations

Fig. A.18: Log differences of Skewness of Risk Aversion for different maturities (30, 60, 90, 120 days).

Fig. A.19: Log differences of Kurtosis of Risk Aversion for different maturities (30, 60, 90, 120 days).
B. XPLORE® QUANTLETS DOCUMENTATION

The following is a detailed description of the XploRe quantlets which were developed in order to run the model described in this work. Should these quantlets be executed again, the following can be used as a guide for the right order of execution as well as code lines which have to be altered. The quantlets list which follows is sorted according to the proper execution sequence. In all quantlets, the first lines which have to be altered are the lines defining the path for reading and writing data.

**EPKDataconstruct.xpl**

This quantlet is used to read the data and construct the database, and hence it has to be the first file executed. The reading of the data was implemented separately in order to avoid reading the data again each time the program is to be executed. The path for reading and writing the dat files has to be altered according to the specific user’s computer and properties. The quantlet transforms the time to maturity so that it is expressed in days, creates an array of all trading days for looping purposes and deletes outliers from the data as mentioned in the description of the data in Chapter 5. Running time of this quantlet is approximately four minutes for the three-year long data.

**EPKsystemfiles.xpl**

This quantlet contains altered versions of procedures which are already implemented in XploRe. Altered versions of the procedures spdbl.xpl, spd1p.xpl, genarch.xpl, acfplot.xpl and pacfplot.xpl are included in this file. The original procedures were manipulated in order to avoid crashing in case of errors in the data. The current versions of these procedures, which bear the same names plus a prefix EPK, skip the error messages in case of inadequate data, and simply inserts an artificial line in the database, signalling an inadequate observation to be removed from the data at a later stage. The graphical output of the procedures acfplot.xpl and pacfplot.xpl was altered to include more plots in one picture. This file has to be executed either before or right after the data is read.
This quantlet includes the daily estimation process and has to be executed before the main program is executed. This quantlet is a collection of the following procedures:

1. **EPKPEst**:
   This procedure calculates and presents the daily subjective density. The input parameters of this procedure are:
   - *OpData*: the 11 column database corresponding to the structure read by the previous quantlet;
   - *Today*: the current date;
   - *Tau*: a vector of four different maturities to check;
   - *MGrid*: a chosen moneyness grid, on which the density is to be estimated;
   - *graphshow*: a binary variable for showing or suppressing graphs.

   It is possible to alter the desired number of runs for the simulated GARCH (*MaxRep*) and the desired length of simulated period in days (*SimPeriod*). The daily stock price is calculated for the three months preceding the daily estimation date (*Today*) and the parameters of the GARCH process are estimated. Then a GARCH process is simulated based on the estimated parameters, and the subjective density is fitted for the predetermined maturities (*Tau*) on the desired moneyness grid (*MGrid*). If the indicator *graphshow* is set to 1, a graph is plotted. The output of this procedure is the estimated daily subjective density for each of the maturities on the desired moneyness grid.

2. **EPKQEst**:
   This procedure calculates and presents the daily state-price density. The input parameters of this procedure are similar to those of the subjective density estimation mentioned above. The first step is to restructure the data so that it would fit the altered *XploRe* procedures *EPKspdbl.xpl* and *EPKspdlp.xpl* which compute the local polynomials. When the restructuring is done properly, *XploRe* provides the estimated SPD on the desired moneyness grid for the desired maturities, which is the output of the procedure. As in the subjective density estimation, if the indicator *graphshow* is set to 1, a graph is plotted.

3. **EPKcalcdaily**:
   This procedure calculates the daily PK and RRA for the different maturities, presents them if desired and collects the five values mentioned in
Chapter 6 (around ATM value, mean across moneyness, standard deviation, skewness and kurtosis). The input parameters of this procedure are again similar to the parameters of the subjective density estimation mentioned above. This procedure calls the procedures EPKPEst and EPKQEst mentioned above, gets their output and calculates the pricing kernel and relative risk aversion functions on the moneyness grid for the desired maturities, as described in Chapter 5. In case the GARCH estimation does not fit the daily data, the subjective density is set to 0, and this line is later on erased from the data. Outliers of PK and RRA (at specific moneyness values) are also deleted in order to avoid biased moment calculations. The output is the five desired values of the PK and RRA for each of the desired maturities on the desired moneyness grid.

**EPKmain.xpl**

This quantlet includes the main program as well as a procedure written for the looping purposes and called EPKdynamics.xpl. This procedure is responsible for calculating the dynamics of the pricing kernel and relative risk aversion. It executes the daily PK and RRA calculation for each day of the database and collects the results, saving them in a dat file. Similarly to other quantlets, it is necessary to specify the proper path for writing and reading the dat files. The input parameters of this procedure are:

- **OrData**: the 11 column database corresponding to the structure read by the previous quantlets;
- **Tau**: a vector of four different maturities to check;
- **MG**: a chosen moneyness grid, on which the density is to be estimated;
- **StartDate**: The first day of calculation;
- **EndDate**: the last day of calculation;
- **graphshow**: a binary variable for showing or suppressing graphs.

This procedure runs a loop on all desired trading days and its final output is the daily PK and RRA for each of the maturities on the desired moneyness grid, for each trading day between **StartDate** and **EndDate**. The data is written into a dat file at the end of each iteration of the loop (i.e. before the program moves to the next trading day), in order to avoid the loss of data in case the program crashes.
The second part of this quantlet is the main program which executes the whole code. The desired vector of maturities and the moneyness grid are defined at this point (and remain the same throughout the whole process), as well as the first and last day of trading. After these definitions the program calls **EPKdynamics.xpl** which starts the computation process. A daily process could last between 20 and 50 minutes, depending on the daily number of contracts which appear in the database. Such a long daily process results in an extremely long running time for the whole three years (about two weeks!). If more than one trading day is desired (i.e. \( \text{StartDate} \neq \text{EndDate} \)), the **graphshow** indicator should be set to 0, as the daily graphs are overwritten on the next day.

After calculating the PK and RRA moments for the whole database, two dat files exist in the specific folder specified in advance. These dat files contain the complete information about each of the desired moments of the PK and RRA for each of the maturities, for each of the trading days. At this point the quantlets which have been described so far are not being used anymore. The time-series saved in these files are being used from now on as an input for the next quantlets.

**EPKtimeseries.xpl**

This quantlet reads the dat files, which contain the daily data for the three year long period, builds the time-series, their differences and their logarithmic differences, plots them, checks for stationarity and runs the principal component analysis. The plotting commands and the KPSS tests are commented out in order not to overload the screen with plots and numbers, and should be included if new dat files are considered. The PCA, as explained in Chapter 6, is conducted only on the PK and RRA differences, which were found to be stationary. The results of the PCA are collected for later use in the time-series and regression analysis. The ACF and PACF plots of the PCs (done with **EPKacfplot.xpl** and **EPKpacfplot.xpl** respectively) are also commented out for the same reason specified above, and can be activated if desired. A proper path for reading and writing the data is essential for a successful run of this quantlet.

**EPKtimeplot.xpl**, **EPKtimeplotPCA.xpl**

These quantlets are used to plot the different time-series of the different moments of PK and RRA and the time-series of their PCs, respectively. Both are based on **timeplot.xpl** but have different colors, titles and axes names.
This quantlet conducts the PCA and plots the correlation circles between the different PCs. It is invoked by \EPKtimeseries.xpl for each of the desired time-series. The input parameters of this procedure are:

- \textit{momentsmatrix}: the desired moments to be analyzed;
- \textit{name}: graph title.

The output is the eigenvalues and eigenvectors of the input matrix as well as the correlations between the variables and the PCs with a graphical representation.

This last quantlet fits a time-series model according to the AIC and SIC criteria and fits a GLS regression model, when the explanatory variables are the DAX level and the implied volatilities at the money, as explained in Chapter 6. This quantlet contains two procedures:

1. \textbf{EPKcalcdailyfigures}: This procedure is run only once, in order to calculate daily averages of the intraday DAX data. The outcome is saved in a dat file which is read in all other instances of this quantlet. The proper specification of the path is therefore important in this quantlet as well.

2. \textbf{EPKPsiReg}: This procedure runs a linear regression on the transformed model according to \cite{Sen & Srivastava 1990} to deal with an autocorrelated error of order 1, as explained in Chapter 6. The input parameters of this model are the desired explanatory variables and the response variable. The first step is a standard OLS regression using the quantlet \linreg.xpl. The OLS regression is followed by the calculation of the residuals and the estimation of the first order autocorrelation according to equation (6.26). When $\hat{\rho}$ is obtained, the $\Psi$ matrix is calculated according to equation (6.28). Then a standard OLS regression is run on the transformed model, and its output parameters are the output of the whole procedure. The output of this procedure is therefore a GLS regression of the original variables, considering the autocorrelated error term.

The main program which follows these two procedures expects one data file for each of the desired maturities (currently these are 30, 60 and 90
days), and in each data file the date and the values of the three PCs constitute the columns, whereas the rows correspond to the different trading days. Each of the input dat files has already been written by the quantlet `EPKtimeseries.xpl` and at this point contains the date and the three PCs of PK/RRA stationary differences at the corresponding date. After the dat files have been read, a time-series model is fitted using the quantlet `arimacls.xpl`. The models which produced a poorer fit to the data based on the information criteria were omitted from the quantlet, which contains only the best-fitted models, though they are currently commented out. The last part of the quantlet is the specification of the explanatory variables for the GLS regression, which are combinations of log-returns and absolute differences of the DAX level and ATM implied volatilities, as described in Chapter 6.
BIBLIOGRAPHY


Pratt, J. (1964), ‘Risk aversion in the small and in the large’, *Econometrica*.

Rookley, C. (1997), ‘Fully exploiting the information content of intra day option quotes: Applications in option pricing and risk management’, *University of Arizona*. 


