SFB 649 Discussion Paper 2006-020

Time Dependent Relative Risk Aversion

Enzo Giacomini*
Michael Handel**
Wolfgang K. Härdle*
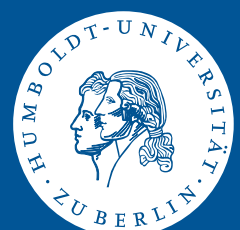

* C.A.S.E. – Center for Applied Statistics and Economics,
Humboldt-Universität zu Berlin, Germany
** Dr. Nagler & Company GmbH, München, Germany

BERLIN

ECONOMIC RISK

SFB 649

Time Dependent Relative Risk Aversion

Enzo Giacomini[1], Michael Handel[2] and Wolfgang K. Härdle[1]

[1] CASE - Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Spandauer Straße 1, 10178 Berlin, Germany

[2] Dr. Nagler & Company GmbH
Maximilanstr. 47, 80538 Munich, Germany

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 and relative risk aversion 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 relative risk aversion is found to be linear. The relation of the same explanatory variables to the principal components of the changes in pricing kernels is found to be log-linear, although this relation is not significant for all of the examined maturities.

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 due to 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. The payoff function of a call option at maturity T is

$$\psi(S_T) = (S_T - K)^+ \overset{\text{def}}{=} \max(S_T - K, 0) \tag{1}$$

where K is the strike price, T is maturity and S_T is the asset's price at maturity.

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

$$C_t = e^{-r(T-t)} \, \mathrm{E}_t^Q[\psi(S_T)] \stackrel{\text{def}}{=} e^{-r(T-t)} \sum_s q_s \psi_s(S_T) \qquad (2)$$

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 section 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.

2 Black and Scholes and 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 with a continuously compounded rate of return r. The stock price follows the stochastic process

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \qquad (3)$$

where μ denotes the drift, σ is the volatility and W_t is a standard Brownian motion. The drift and volatility could be functions of the asset price, time and many other factors. However, for simplicity, they are considered constant in this section. The conditional density of the stock price, which is implied by equation (3), is denoted by $p_t(S_T|S_t)$. In this setting, due to continuous dividend payments, the discounted process with cumulative dividend reinvestments should be a martingale and is denoted by

$$\widetilde{S}_t \stackrel{\text{def}}{=} e^{-(r+\delta)t}S_t \tag{4}$$

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

$$
\begin{aligned}
d\widetilde{S}_t &= d(e^{-rt}S_t) \\
&= -re^{-rt}S_t dt + e^{-rt}dS_t \\
&= -re^{-rt}S_t dt + e^{-rt}[\mu S_t dt + \sigma S_t dW_t] \\
&= (\mu - r)\widetilde{S}_t dt + \sigma \widetilde{S}_t dW_t \\
&= \sigma \widetilde{S}_t d\overline{W}_t
\end{aligned}
\tag{5}
$$

where $\overline{W}_t \stackrel{\text{def}}{=} W_t + \frac{\mu-r}{\sigma}t$ 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 (5) 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 (2). The basic theorem of asset pricing states, that absence of arbitrage implies the existence of a positive linear pricing rule (Cochrane (2001)), and if the market is complete and 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. Under some specific assumptions, it is well known that a representative agent exists. The original representative agent model includes utility functions which are based on consumption (see, for example,

Mas-Colell et al. (1995)). However, introducing labor income or intermediate consumption do not affect the results significantly and hence, without loss of generality, we review the concept of marginal rate of substitution with the help of a simple consumption based asset pricing model. The fundamental desire for more consumption is described by an intertemporal two-periods utility function as

$$U(c_t, c_{s_{t+1}}) = u(c_t) + \beta \operatorname{E}_t[u(c_{s_{t+1}})] \stackrel{\text{def}}{=} u(c_t) + \beta \sum_s u(c_{s_{t+1}}) p_t(s_{t+1}|s_t) \quad (6)$$

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 β 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 ξ is the amount of asset he chooses to buy, then the optimization problem is

$$\max_{\{\xi\}} \{u(c_t) + \operatorname{E}_t[\beta u(c_{s_{t+1}})]\}$$

subject to

$$
\begin{aligned}
c_t &= Y_t - P_t \cdot \xi \\
c_{s_{t+1}} &= Y_{s_{t+1}} + \psi_{s_{t+1}} \cdot \xi
\end{aligned}
$$

The first constraint is the budget contraint 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

$$P_t = \operatorname{E}_t \left[\beta \frac{u'(c_{s_{t+1}})}{u'(c_t)} \psi_{s_{t+1}} \right] \quad (7)$$

We define $\mathrm{MRS}_t \stackrel{\text{def}}{=} \beta \operatorname{E}_t \left[\frac{u'(c_{s_{t+1}})}{u'(c_t)} \right]$ 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 also referred to as a *stochastic discount factor*.

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 before (equation (6)), where period T corresponds to $t+1$ in the previous section. Defining time to maturity as $\tau \overset{\text{def}}{=} T - t$, 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. This price is given by

$$P_t = e^{-r\tau} \int_0^\infty \psi(S_T) \lambda \frac{U'(S_T)}{U'(S_t)} p_t(S_T|S_t) dS_T \tag{8}$$

where $\lambda e^{-r\tau} = \beta$ to correspond to equation (7) and λ being a constant independent of index level, for scaling purposes.

Considering the call option price under the unique risk-neutral probability measure in equation (2) and the existence of a positive linear pricing rule 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) as long as 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 (8) vanishes under Q, and equation (8) can be rewritten as

$$P_t = e^{-r\tau} \int_0^\infty \psi(S_T) q_t(S_T|S_t) dS_T = e^{-r\tau} \, \mathrm{E}_t^Q[\psi(S_T)] \tag{9}$$

where $q_t(S_T|S_t)$ is the State Price Density and the expectation $\mathrm{E}_t^Q[\psi(S_T)]$ 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 (8) and (9) 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

$$M_t(S_T) \overset{\text{def}}{=} \frac{q_t(S_T|S_t)}{p_t(S_T|S_t)} = \lambda \frac{U'(S_T)}{U'(S_t)} \tag{10}$$

and therefore $\mathrm{MRS}_t = e^{-r\tau} \, \mathrm{E}_t[M_t(S_T)]$. Substituting out the $q_t(S_T|S_t)$ in

equation (9) using equation (10) yields the *Lucas asset pricing equation*:

$$
\begin{aligned}
P_t &= e^{-r\tau}\,\mathrm{E}_t^Q[\psi(S_T)] \\
&= e^{-r\tau}\int_0^\infty M_t(S_T)\cdot\psi(S_T)p_t(S_T|S_t)dS_T \\
&= e^{-r\tau}\,\mathrm{E}_t[M_t(S_T)\cdot\psi(S_T)]
\end{aligned}
\tag{11}
$$

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

$$
\rho_t(S_T) \stackrel{\text{def}}{=} -\frac{S_T u''(S_T)}{u'(S_T)}
\tag{12}
$$

According to equation (10) the pricing kernel is related to the marginal utilities as

$$
M_t(S_T) = \lambda\frac{U'(S_T)}{U'(S_t)}
$$
$$
\Rightarrow\quad M_t'(S_T) = \lambda\frac{U''(S_T)}{U'(S_t)}
\tag{13}
$$

Substituting out the first and second derivatives of the utility function in equation (12) using equation (13) yields

$$
\rho_t(S_T) = -\frac{S_T\lambda M_t'(S_T)U'(S_t)}{\lambda M_t(S_T)U'(S_t)} = -\frac{S_T M_t'(S_T)}{M_t(S_T)}
\tag{14}
$$

Using equation (10) we can express the RRA as

$$
\begin{aligned}
\rho_t(S_T) &= -\frac{S_T[q_t(S_T|S_t)/p_t(S_T|S_t)]'}{q_t(S_T|S_t)/p_t(S_T|S_t)} \\
&= -S_T\frac{[q_t'(S_T|S_t)p_t(S_T|S_t) - p_t'(S_T|S_t)q_t(S_T|S_t)]/p_t^2(S_T|S_t)}{q_t(S_T|S_t)/p_t(S_T|S_t)} \\
&= -S_T\frac{q_t'(S_T|S_t)p_t(S_T|S_t) - p_t'(S_T|S_t)q_t(S_T|S_t)}{q_t(S_T|S_t)p_t(S_T|S_t)} \\
&= S_T\left[\frac{p_t'(S_T|S_t)}{p_t(S_T|S_t)} - \frac{q_t'(S_T|S_t)}{q_t(S_T|S_t)}\right]
\end{aligned}
\tag{15}
$$

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

$$u(c_t) = \begin{cases} \frac{1}{1-\gamma} c_t^{1-\gamma} & \text{for } 0 < \gamma \neq 1 \\ \log(c_t) & \text{for } \gamma = 1 \end{cases} \tag{16}$$

Rubinstein (1976) showed, that for such a utility function, 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

$$\text{MRS}_t = \beta \, \text{E}_t \left[\frac{u'(c_T)}{u'(c_t)} \right] = \beta \, \text{E}_t \left[\left(\frac{c_T}{c_t} \right)^{-\gamma} \right] \tag{17}$$

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 (12), with consumption instead of wealth as an argument, as the utility function is utility of consumption

$$\rho(c_T) = -c_T \frac{-\gamma(c_T)^{-\gamma-1}}{(c_T)^{-\gamma}} = \gamma \tag{18}$$

This equation shows that the RRA turns out to be a constant, and for the logarithmic utility case, the risk aversion is 1.

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 (10). 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.

Under the assumptions of the well-known Black & Scholes (1973) model, the price of a plain vanilla call option with a payoff function as in equation (1) is given by the *Black and Scholes formula*

$$C^{BS}(S_t, t, K, T, \sigma, r, \delta) = e^{-\delta\tau} S_t \Phi(d_1) - e^{-r\tau} K \Phi(d_2) \qquad (19)$$

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

$$d_1 = \frac{\ln(S_t/K) + (r - \delta + 0.5\sigma^2)\tau}{\sigma\sqrt{\tau}} \text{ and } d_2 = d_1 - \sigma\sqrt{\tau} \qquad (20)$$

where we assume $\delta = 0$ for the remaining of this work, as mentioned before. Furthermore, the Black & Scholes (1973) implied volatility is assumed to be constant and the corresponding risk-neutral density is log-normal with mean $(r - 0.5\sigma^2)\tau$ and variance $\sigma^2\tau$.

A famous work by Breeden & Litzenberger (1978) proved the following relation, which also holds when the assumptions of the Black & Scholes (1973) model do not:

$$e^{r\tau} \left. \frac{\partial^2 C(S_t, K, \tau)}{\partial K^2} \right|_{K=S_T} = q_t(S_T) = \text{SPD} \qquad (21)$$

Sustaining the assumptions of the Black & Scholes (1973) model and plugging equation (19) into equation (21) yields

$$q^{BS}(S_T|S_t) = \frac{1}{S_T\sqrt{2\pi\sigma^2\tau}} \cdot e^{-\frac{[\ln(S_T/S_t) - (r - 0.5\sigma^2)\tau]^2}{2\sigma^2\tau}} \qquad (22)$$

meaning that the underlying asset price follows the stochastic process

$$\frac{dS_t}{S_t} = r \cdot dt + \sigma \cdot dW_t \qquad (23)$$

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 μ, plugging the SPD from equation (22) and the

log-normal subjective density into equation (10) yields a closed-form solution for the investor's pricing kernel

$$M_t^{BS}(S_T) = \left(\frac{S_T}{S_t}\right)^{-\frac{\mu-r}{\sigma^2}} \cdot e^{\frac{(\mu-r)(\mu+r-\sigma^2)\tau}{2\sigma^2}} \quad (24)$$

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 (24) is also the ratio of the marginal utility functions (equation (10)), the investor's utility function can be derived by solving the differential equation. If we consider the following constants

$$\begin{aligned} \gamma &= \frac{\mu-r}{\sigma^2} \\ \lambda &= e^{\frac{(\mu-r)(\mu+r-\sigma^2)\tau}{2\sigma^2}} \end{aligned} \quad (25)$$

we can rewrite equation (24) as

$$M_t^{BS}(S_T) = \lambda \left(\frac{S_T}{S_t}\right)^{-\gamma} \quad (26)$$

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

$$u^{BS}(S_t) = \left(1 - \frac{\mu-r}{\sigma^2}\right)^{-1} \cdot S_t^{\left(1-\frac{\mu-r}{\sigma^2}\right)} \quad (27)$$

the subjective discount factor of intertemporal utility is

$$\beta^{BS} = \lambda e^{-r\tau} = e^{\frac{(\mu-r)(\mu+r-\sigma^2)\tau}{2\sigma^2} - r\tau} \quad (28)$$

and the relative risk aversion is constant

$$\rho_t^{BS}(S_T) = \gamma = \frac{\mu-r}{\sigma^2} \quad (29)$$

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 (5), in which the Brownian motion $\overline{W}_t$ 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

$$\overline{W}_t = W_t + \frac{\mu-r}{\sigma}t = W_t + \sigma\gamma t \quad (30)$$

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

$$d\widetilde{S}_t = \sigma\widetilde{S}_t d\overline{W}_t = \sigma\widetilde{S}_t dW_t + \sigma^2\widetilde{S}_t\gamma dt \quad (31)$$

3 A Static Model: Daily 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. Rubinstein (1994) showed, 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 section 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 section is divided into four parts. The first part provides a short description of the database used in this work. The static model for estimating the pricing kernel and relative risk aversion on a daily basis is introduced in the following parts of this section. 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 the next section.

3.1 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 (4)), 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, 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.

3.2   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

$$
\begin{aligned}
\varepsilon_t &= \sigma_t Z_t \\
\sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{32}
$$

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 $\varepsilon_t = \Delta \log(S_t) = \log(S_t) - \log(S_{t-1})$, and this time series together with its daily standard deviation σ_t are the input of the GARCH estimation. The parameters ω, α and β are estimated using the quasi maximum likelihood method, which is an extension of the maximum likelihood measure, when the estimator is not efficient.

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 (32) are used for the simulation, but this time the unknown variables are the time series σ_t and ε_t, while the parameters ω, α and β 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

$$S_t = S_{t-1} e^{\varepsilon_t} \quad \forall t \in \{1, \ldots, T\} \tag{33}$$

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 on the present day. The subjective density is estimated for every trading day included in the database. In figure 1 we plot the simulated subjective densities on four different trading days for four different maturities.

It can be seen in figure 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 Aït Sahalia & Lo (2000) and Brown & Jackwerth (2004).

3.3 State-Price Density Estimation

There is a vast literature on estimating the SPD using nonparametric and semiparametric methods. Aït Sahalia & Lo (2000), for example, suggest a semiparametric approach using 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

$$\widehat{\sigma}(\kappa, \tau) = \frac{\sum_{i=1}^n k_\kappa\left(\frac{\kappa - \kappa_i}{h_\kappa}\right) k_\tau\left(\frac{\tau - \tau_i}{h_\tau}\right) \sigma_i}{\sum_{i=1}^n k_\kappa\left(\frac{\kappa - \kappa_i}{h_\kappa}\right) k_\tau\left(\frac{\tau - \tau_i}{h_\tau}\right)} \tag{34}$$

where $\kappa \stackrel{\text{def}}{=} \frac{K}{e^{r\tau} S_t}$ is future moneyness, τ is time to maturity and σ_i is the implied volatility. The kernel functions k_κ and k_τ together with the appropriate bandwidths h_κ and h_τ are chosen such that the asymptotic properties
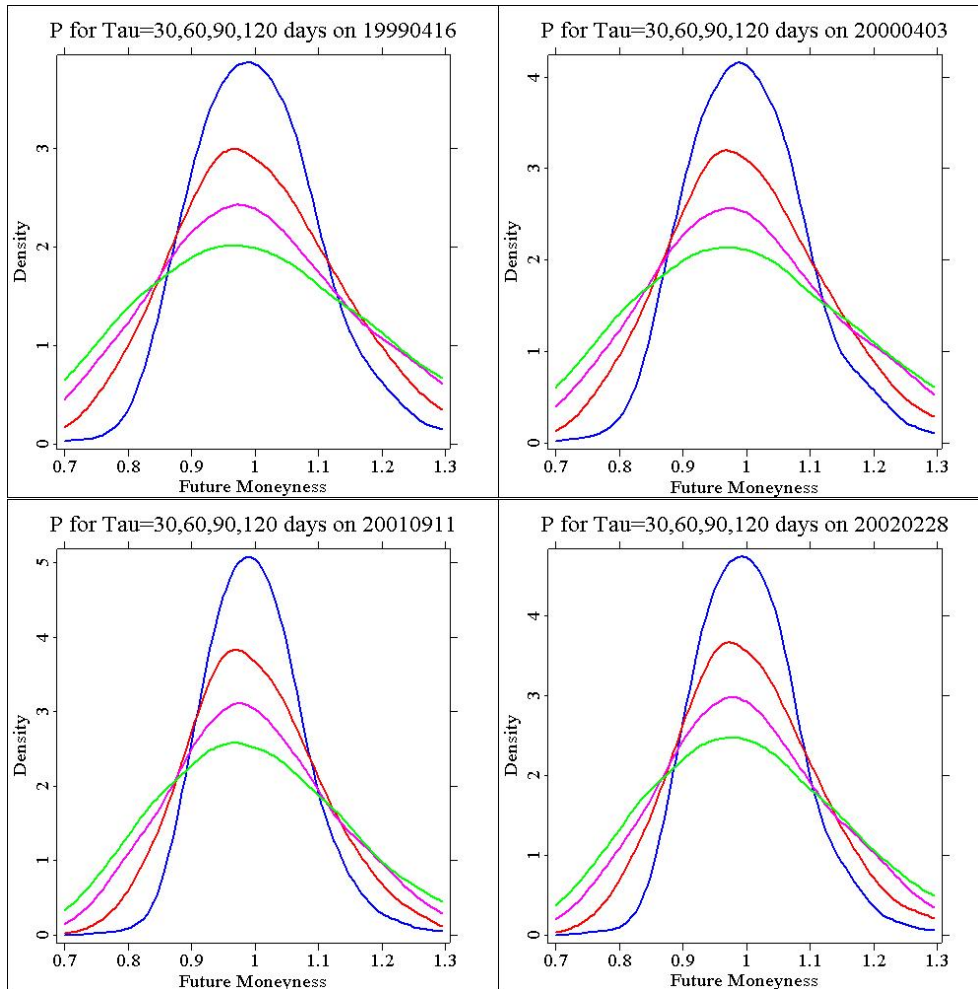
Figure 1: Subjective density for different maturities (30,60,90,120 days) on different trading days.

EPKdailyprocess.xpl

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 (21).

A major advantage of such a method 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

14

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.

The state-price density in this work 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 Aït Sahalia & Lo (2000), 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 σ is the unknown). Then a local polynomial regression is used to smooth the implied volatility 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.

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

$$\widehat{\sigma} = \phi(\kappa, \tau) + \sigma^{BS}(\kappa, \tau) * \varepsilon \tag{35}$$

where $\phi(\kappa, \tau)$ is an unknown function, which is three times continuously differentiable, and ε is a Gaussian white noise. Then a Taylor expansion for

15

the function $\phi(\kappa, \tau)$ in the neighborhood of (κ_0, τ_0) is

$$
\begin{aligned}
\phi(\kappa, \tau) \approx \phi(\kappa_0, \tau_0) \quad &+ \quad \left.\frac{\partial \phi}{\partial \kappa}\right|_{\kappa_0, \tau_0} (\kappa - \kappa_0) + \frac{1}{2} \left.\frac{\partial^2 \phi}{\partial \kappa^2}\right|_{\kappa_0, \tau_0} (\kappa - \kappa_0)^2 \\
&+ \quad \left.\frac{\partial \phi}{\partial \tau}\right|_{\kappa_0, \tau_0} (\tau - \tau_0) + \frac{1}{2} \left.\frac{\partial^2 \phi}{\partial \tau^2}\right|_{\kappa_0, \tau_0} (\tau - \tau_0)^2 \\
&+ \quad \frac{1}{2} \left.\frac{\partial^2 \phi}{\partial \kappa \partial \tau}\right|_{\kappa_0, \tau_0} (\kappa - \kappa_0)(\tau - \tau_0) \qquad (36)
\end{aligned}
$$

Minimizing the expression

$$
\begin{aligned}
\sum_{j=1}^{n} \Big\{ \sigma^{BS}(\kappa_j, \tau_j) &- [\beta_0 + \beta_1(\kappa_j - \kappa_0) + \beta_2(\kappa_j - \kappa_0)^2 + \beta_3(\tau_j - \tau_0) \\
&+ \beta_4(\tau_j - \tau_0)^2 + \beta_5(\kappa_j - \kappa_0)(\tau_j - \tau_0)] \Big\}^2 K_h(\kappa - \kappa_0)(\tau - \tau_0) \qquad (37)
\end{aligned}
$$

yields the estimated implied volatility surface and its first two derivatives at the same time, as $\left.\widehat{\frac{\partial \phi}{\partial \kappa}}\right|_{\kappa_0, \tau_0} = \widehat{\beta_1}$ and $\left.\widehat{\frac{\partial^2 \phi}{\partial \kappa^2}}\right|_{\kappa_0, \tau_0} = 2\widehat{\beta_2}$. 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 $\frac{\partial^2 C}{\partial K^2}$ (used for the SPD according to Breeden & Litzenberger (1978)) as a function of $\frac{\partial \sigma}{\partial \kappa}$ and $\frac{\partial^2 \sigma}{\partial \kappa^2}$ (which are obtained from the implied volatility surface estimation) is given, for example, by Huynh et al. (2002).

The estimated risk neutral densities for the same dates and the same maturities as in figure 1 are depicted in figure 2. The SPD is estimated on a future moneyness scale, thus reducing the number of parameters that need to be estimated.

One of the trading days plotted in figure 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.
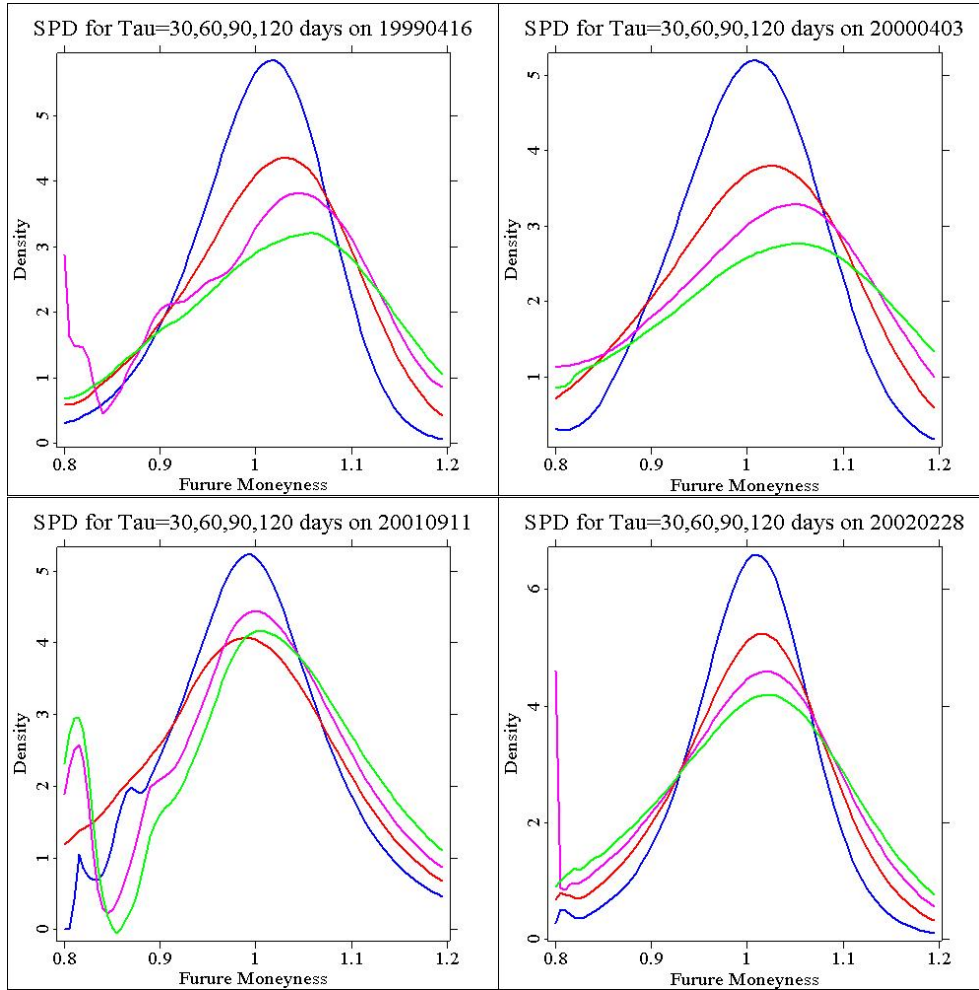
Figure 2: State-Price density for different maturities (30,60,90,120 days) on different trading days.

EPKdailyprocess.xpl

3.4   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 (10), 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 p and q are estimated on the moneyness grid, the estimated pricing kernel is

17

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 (14).

The estimated pricing kernels depicted in figure 3 for different trading days and different maturities bear similar characteristics to those reported by Aït 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 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.

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 first observe historical returns without considering crash possibilities, and only afterwards 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.
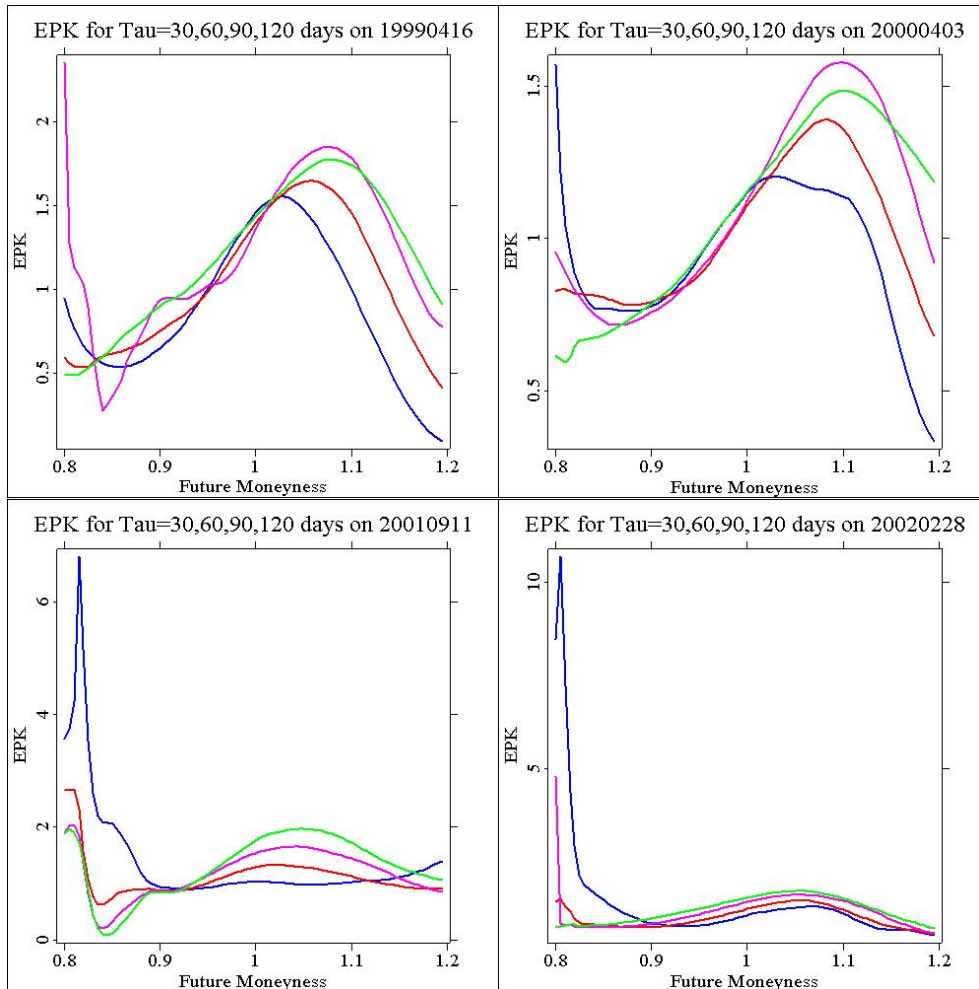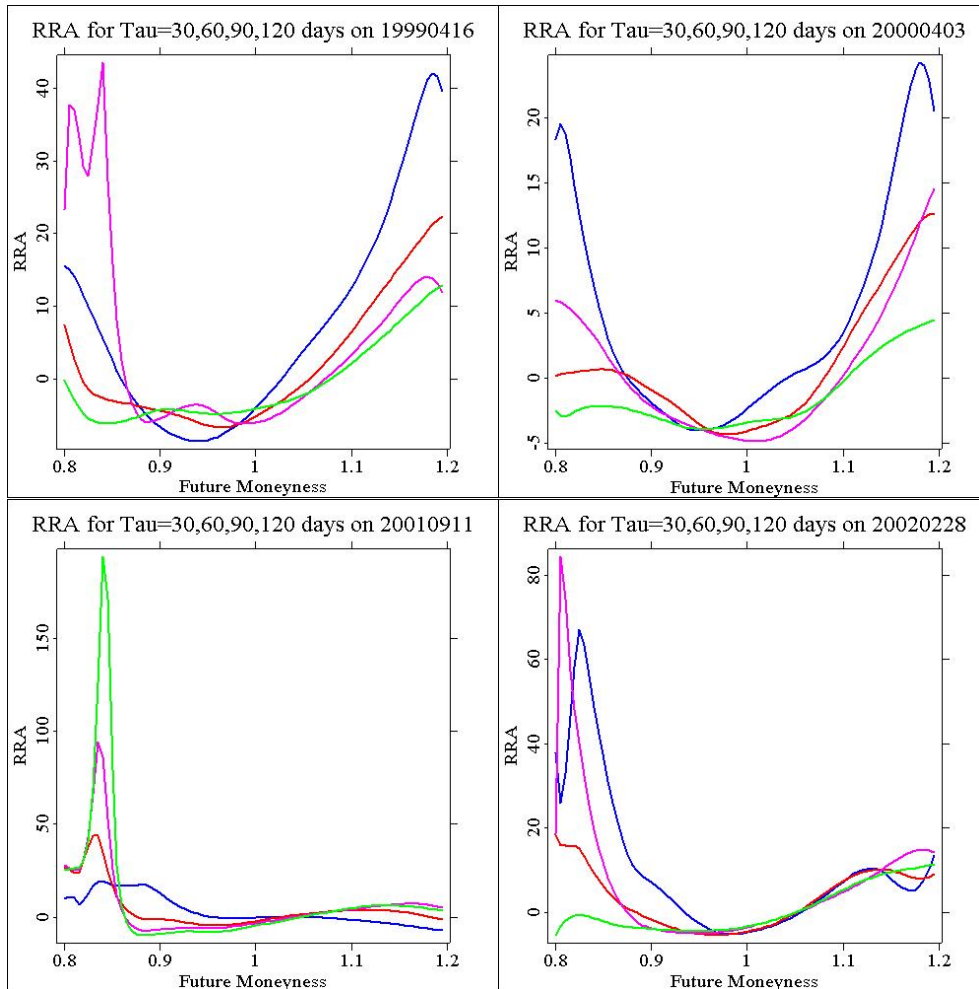
18

Figure 3: Estimated Pricing Kernel for different maturities (30,60,90,120 days) on different trading days.

EPKdailyprocess.xpl

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 following section, a dynamic analysis of the pricing kernel and relative risk aversion is conducted along the three-year time frame.

Figure 4: Estimated relative risk aversion for different maturities
(30,60,90,120 days) on different trading days.

4 A Dynamic Model: Time-Series Analysis

Since the process described above 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 section we will analyze these
time-series and show their moments. A principal component analysis will
be conducted on the stationary series and the principal components will be
tested as response variables in a GLS regression.

4.1 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 first four moments at any trading day have to be computed, i.e. the mean (μ_t), standard deviation (σ_t), skewness ($Skew_t$) and kurtosis ($Kurt_t$) of the functions across the moneyness grid. In addition, the daily values of the estimated functions at the money (ATM) are calculated and analyzed. Including this additional moment could prove essential as it was shown before that the functions behave quite differently at the money than in other regions. Each of the estimates (pricing kernel and relative risk aversion) is a function of moneyness and time to maturity, which was chosen to be a vector of four predetermined maturities, and as in the previous section we concentrate on $\tau = (30, 60, 90, 120)^\top$ days.

The figures in the following pages depict the time-series of the ATM values and mean values 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. Time-series of the daily standard deviation, skewness and kurtosis, as well as the differences time-series, were collected but not included in this paper.

The plots in the next pages show, that the pricing kernel at the money (figure 5) behaves similarly across different maturities and bears similar characteristics to its general mean (figure 6). 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 7) looks quite different than its general mean (figure 8). The ATM relative risk aversion is mostly negative, as detected already in the daily estimated relative risk aversion. The mean relative risk aversion, however, is mostly positive. Another feature of the relative risk aversion 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 relative risk aversion plots is that the four first moments of the distribution do not necessarily represent all the features of the relative risk aversion correctly, and the collection of the extra details regarding the ATM behavior is justified, as it will be shown by the principal component analysis.
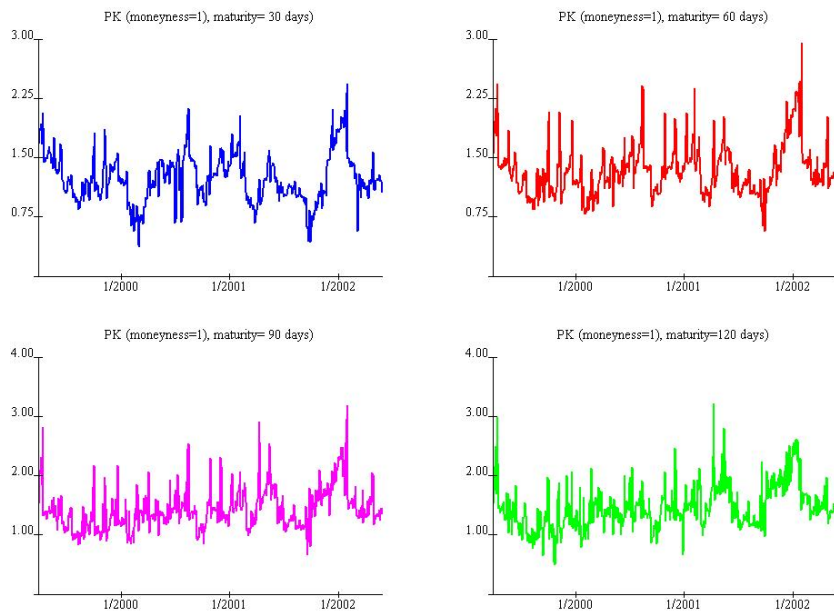
Figure 5: ATM Pricing Kernel for different maturities (30,60,90,120 days).
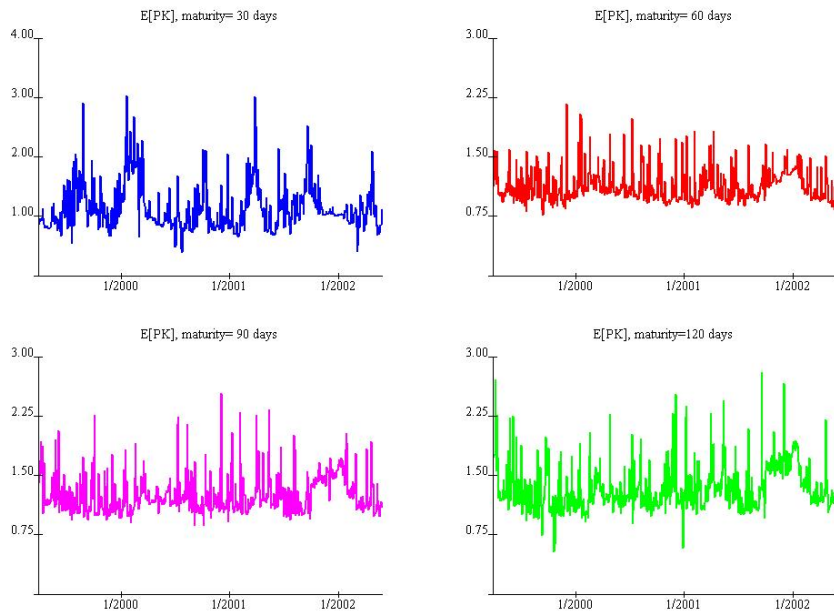
Q EPKtimeseries.xpl

Figure 6: Mean of Pricing Kernel for different maturities (30,60,90,120 days).
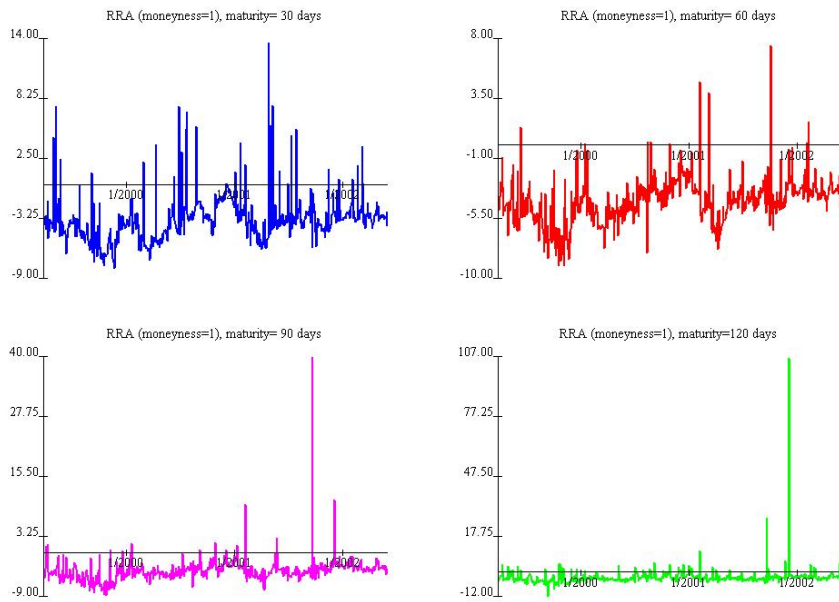
Q EPKtimeseries.xpl

Figure 7: ATM Relative Risk Aversion for different maturities (30,60,90,120 days).
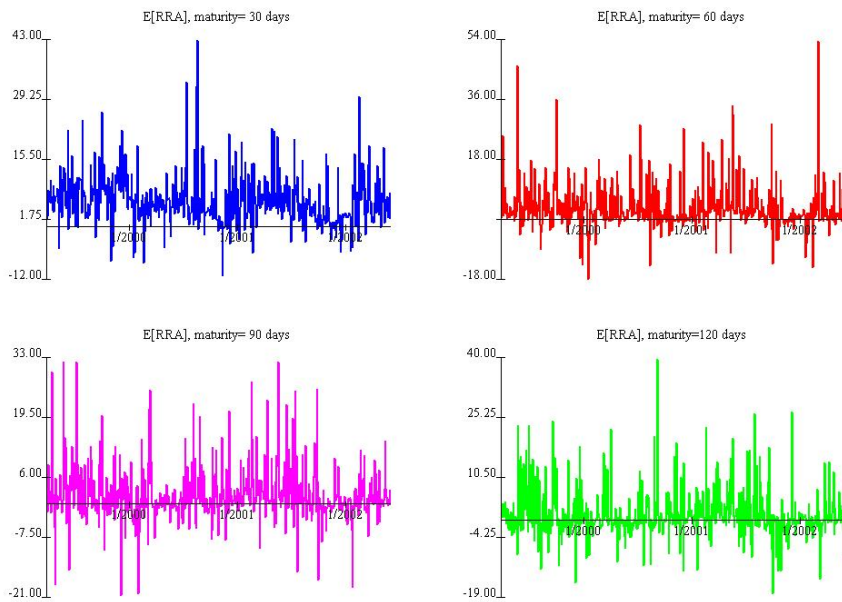
EPKtimeseries.xpl

Figure 8: Mean of Relative Risk Aversion for different maturities (30,60,90,120 days).

EPKtimeseries.xpl

After 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).

Conducting stationarity tests for the various functions has shown, that the moments of the time-series 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. Contrary to that, the absolute differences of all moments and across all maturities were found to be stationary. Therefore, we concentrate from now on only on the absolute differences of the moments.

4.2   Principal Component Analysis

In the following, we 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

$$
\mathcal{X} = \begin{pmatrix} \Delta PK_2^{ATM} & \Delta\mu_2 & \Delta\sigma_2 & \Delta Skew_2 & \Delta Kurt_2 \\ \Delta PK_3^{ATM} & \Delta\mu_3 & \Delta\sigma_3 & \Delta Skew_3 & \Delta Kurt_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Delta PK_n^{ATM} & \Delta\mu_n & \Delta\sigma_n & \Delta Skew_n & \Delta Kurt_n \end{pmatrix} \tag{38}
$$

for each maturity 30, 60 and 90 days, where the differences are defined e.g. as $\Delta\mu_t \overset{\text{def}}{=} \mu_t - \mu_{t-1}$ and similarly for the other columns of the matrix $\mathcal{X}$. 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 were of the same scale, the covariance matrix could be used for the PCA. However, the data is not scale-invariant, hence a standardized PCA must be applied, i.e. conducting the PCA on the correlation matrix.

The principal components can explain the variability of the data. The proportion of variance explained by a certain principal component is the ratio of the corresponding eigenvalue of the correlation matrix to the sum of all eigenvalues, whereas the proportion of variance explained by the first few

principal components is the sum of the proportions of variance explained by each of them.

The principal component analysis shows, that three principal components could explain about 85% of the total variability. Nevertheless, the second and third principal components were found to be correlated, and in order to perform a univariate analysis on the principal components, they have to be orthogonal to each other. Therefore, only the first two principal components of the pricing kernel and relative risk aversion differences are considered from now on. The first two principal components explain approximately 80% of the variability of the pricing kernel differences (the first factor explains 60% and the second explains 20%), and approximately 70% of the variability of the relative risk aversion differences (divided equally among the two factors).

The jth eigenvector expresses the weights used in the linear combination of the original data in the jth principal component. Since we are considering only two principal components, the first two eigenvectors are of interest. More specifically, we can construct the first principal components for each of the examined time-series. The following demonstrates the weights of the moments in the principal components of the differences of the pricing kernel with a maturity of 60 days

$$
\begin{aligned}
y_{1,t}(\tau = 60) &= 0.06\Delta PK_t^{AT} + 0.92\Delta\mu_t + 0.38\Delta\sigma_t + 0.05\Delta Skew_t \\
&\quad -0.03\Delta Kurt_t \\
y_{2,t}(\tau = 60) &= 0.47\Delta PK_t^{ATM} + 0.24\Delta\mu_t - 0.58\Delta\sigma_t - 0.54\Delta Skew_t \\
&\quad +0.29\Delta Kurt_t
\end{aligned}
$$

It can clearly be seen, that the dominant factors in the first principal component are the changes in mean and standard deviation, whereas the dominant factors in the second principal component are the changes in skewness and standard deviation. The equations do not change much when other maturities are considered. As for the moments of the relative risk aversion, the first principal component is dominated solely by the changes in standard deviation and the second principal component is mainly dominated by the change in relative risk aversion at the money.

We conclude therefore, that the variation of the pricing kernel and relative risk aversion differences can be explained by two factors. The first factor of pricing kernel differences explains 60% of the variability and can be perceived as a central mass movement factor, consisting of the changes in expectation

and standard deviation. The second factor explains additional 20% of the variability and can be perceived as a change of tendency factor, consisting of changes in skewness and standard deviation. The principal components of the relative risk aversion are a little different. The first one explains approximately 35% of the variability and can be perceived as a dispersion change factor, dominated by the change in standard deviation. The contribution of the second principal component to the total variability is 35% as well and it is dominated by the change in relative risk aversion of the investors at the money. The mean of relative risk aversion differences seems to play no role in examining the variability of the relative risk aversion.

The correlation between the ith moment and the jth principal component is calculated as

$$r_{X_i,Y_j} = g_{ij}\sqrt{\frac{l_j}{s_{X_i X_i}}} \tag{39}$$

where g_{ij} is the ith element of the jth eigenvector, l_j is the corresponding eigenvalue and $s_{X_i X_i}$ is the standard deviation of the ith moment X_i.

Descriptive statistics of the principal components time-series and their correlations with the moments are given in tables 1 and 2 for the pricing kernel and relative risk aversion respectively. The means of the principal components are very close to zero, as they are linear combinations of the differences of the moments, which are themselves approximately zero mean.

Principal Component	Mean $\times 10^4$	Standard Deviation	Correlation with				
			ΔPK_t^{ATM}	$\Delta\mu_t$	$\Delta\sigma_t$	$\Delta Skew_t$	$\Delta Kurt_t$
$\tau = 30$							
$y_{1,t}$	-2.46	0.76	-0.02	0.42	0.62	0.02	-0.02
$y_{2,t}$	-4.39	4.15	0.21	0.25	-0.16	0.29	0.08
$\tau = 60$							
$y_{1,t}$	4.34	0.44	0.06	0.74	0.30	0.04	-0.03
$y_{2,t}$	8.53	4.06	0.22	0.11	-0.27	-0.25	0.13
$\tau = 90$							
$y_{1,t}$	2.80	0.55	0.09	-0.61	0.46	0.11	-0.05
$y_{2,t}$	9.20	2.04	0.23	-0.19	-0.21	-0.32	0.11

Table 1: Descriptive statistics, principal components of the pricing kernel differences.

The moments highly correlated with the principal components are, not surprisingly, the ones which were reported to be dominant when constructing the

Principal	Mean	Standard	Correlation with				
Component	$\times 10^3$	Deviation	ΔRRA_t^{ATM}	$\Delta \mu_t$	$\Delta \sigma_t$	$\Delta Skew_t$	$\Delta Kurt_t$
$\tau = 30$							
$y_{1,t}$	11.5	14.75	0.03	0.04	0.61	0.00	0.01
$y_{2,t}$	0.55	9.36	0.33	-0.22	-0.02	-0.32	0.26
$\tau = 60$							
$y_{1,t}$	-2.57	26.90	0.10	0.04	0.60	-0.02	0.03
$y_{2,t}$	1.60	13.75	0.36	0.20	-0.06	-0.24	-0.35
$\tau = 90$							
$y_{1,t}$	1.72	28.60	-0.08	0.15	0.63	0.05	0.04
$y_{2,t}$	3.71	9.22	0.18	0.36	-0.05	-0.27	0.20

Table 2: Descriptive statistics, principal components of the relative risk aversion differences.

principal components. Nevertheless, table 1 implies an inconsistent behavior of the different moments across maturities. The first principal components of the pricing kernel differences (the first rows for each of the maturities in table 1) 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 principal components of pricing kernel differences (the second rows for each of the maturities in table 1) 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 principal component 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 principal component to the total variability.

The correlations of the first and second principal components of the relative risk aversion differences with their dominant factors (table 2) are found to be consistent across maturities. The first principal component is positively correlated with its most dominant moment, the changes in the relative risk aversion standard deviation. This correlation means essentially, that the less homoscedastic the relative risk aversion is, i.e. the larger the changes in standard deviation are, the larger the first principal component of the relative risk aversion differences becomes. The second principal component of the relative risk aversion differences is positively correlated with its most dominant moment, the behavior at the money. The more volatile the relative risk
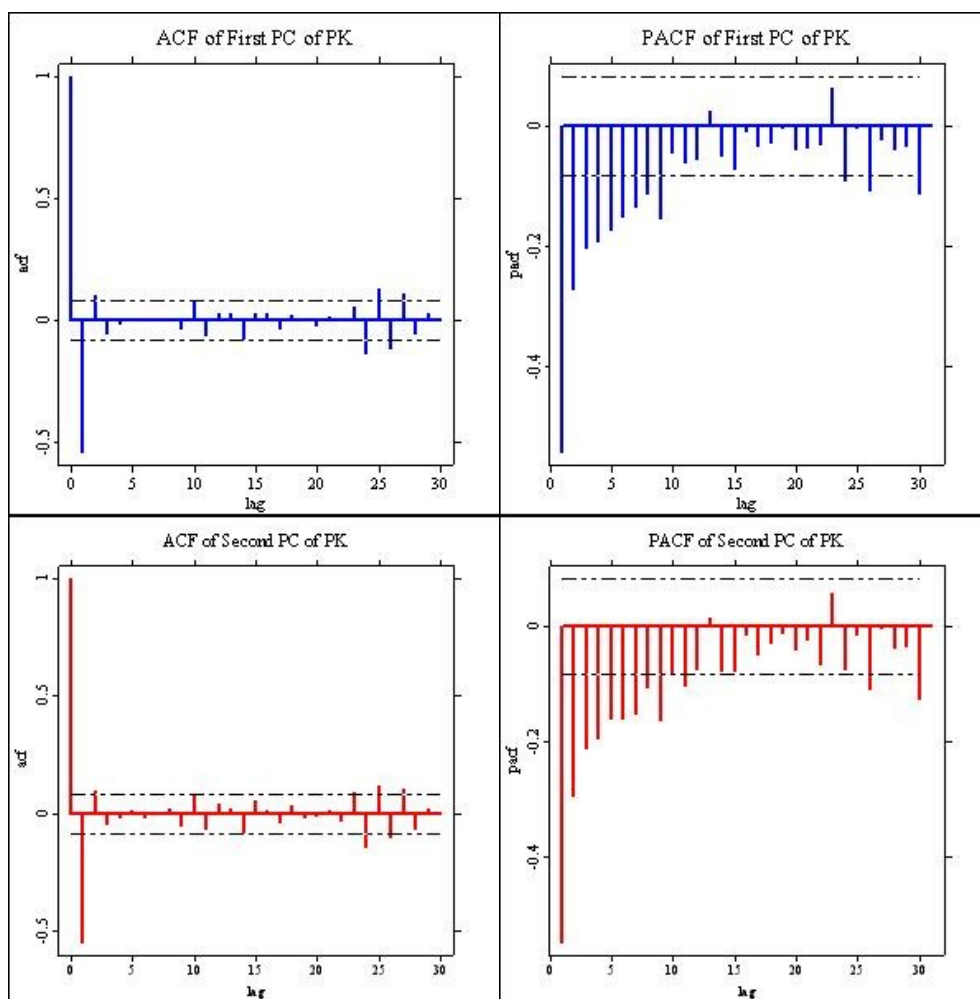
Figure 9: Autocorrelation function (left panel) and partial autocorrelation function (right panel) of the principal components of pricing kernel differences ($\tau = 60$ days). The autocorrelation functions of the principal components of relative risk aversion differences behave similarly exhibiting a MA(1) process.

🔍 EPKtimeseries.xpl

aversion at the money is, the higher the second principal component is. Both principal components of the relative risk aversion differences contribute more than 30% of the variability and imply a good fit of the principal components to the data.

After constructing principal components, which explain the variability of the time-series, it is essential to check the autocorrelation and the partial auto-correlation functions of the time-dependent principal components. This is illustrated in figure 9 for the pricing kernel differences. The same functions for the principal components of the relative risk aversion differences have similar characteristics and hence not reported here. Since the principal components have similar autocorrelation and partial autocorrelation functions for all different maturities, a maturity of 60 days was arbitrarily chosen to be presented. It can be seen, that the autocorrelation function drops abruptly after the first order autocorrelation whereas the partial autocorrelation function 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 principal components. A calculation of the Akaike and Schwarz information criteria confirms, that the best-fitted models for the first principal components are ARMA (1,1), whereas the second principal components follow a MA(1) process. As expected, all principal components have an autocorrelated error term.

4.3 GLS Regression Model for the Principal Components

The last test conducted in this work is to detect a possible relation between the principal components and easily observed data, such as changes in the DAX level and in implied volatility at the money. It is well known, that the simplest relation between an explanatory variable and a response variable can be described and examined using a simple linear regression model

$$y = X\beta + \epsilon \tag{40}$$

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

$$\widehat{\beta}_{OLS} = (X'X)^{-1}X'y \tag{41}$$

with a covariance matrix

$$\mathrm{Cov}(\widehat{\beta}_{OLS}) = \sigma^2(X'X)^{-1} \tag{42}$$

Introducing autocorrelated errors as described above, 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

$$\epsilon_t = \rho\epsilon_{t-1} + u_t \tag{43}$$

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

Iterating equation (43) from time 0 onwards yields

$$\epsilon_t = \lim_{n\to\infty} (\rho^{n+1}\epsilon_{t-n-1} + \sum_{s=0}^{n} \rho^s u_{t-s}) = \sum_{s=0}^{\infty} \rho^s u_{t-s} \tag{44}$$

and hence $E[\epsilon_t] = 0$ and the covariance matrix of the error term is

$$\text{Cov}(\epsilon) = \sigma_u^2 \Omega = \frac{\sigma_u^2}{1 - \rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{n-1} \\ \rho & 1 & \rho & \ldots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \ldots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \ldots & 1 \end{pmatrix} \tag{45}$$

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 $\widehat{\beta}$. 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 $\widehat{\beta}$ in equation (40). 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

$$\widehat{\rho} = \frac{\sum_{t=2}^{n} \epsilon_t \epsilon_{t-1}}{\sum_{t=1}^{n} \epsilon_t^2} \tag{46}$$

Replacing the ρ's in equation (45) with the $\widehat{\rho}$'s from equation (46) results in the estimated matrix $\widehat{\Omega}$. The best linear unbiased estimator in that case would be the estimated generalized least squares estimator

$$\widehat{\beta}_{GLS} = (X'\widehat{\Omega}^{-1}X)^{-1}X'\widehat{\Omega}^{-1}y \tag{47}$$

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

$$
\widehat{\Psi} = \begin{pmatrix}
\sqrt{1-\widehat{\rho}^2} & 0 & 0 & \ldots & 0 & 0 & 0 \\
-\widehat{\rho} & 1 & 0 & \ldots & 0 & 0 & 0 \\
0 & -\widehat{\rho} & 1 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & -\widehat{\rho} & 1 & 0 \\
0 & 0 & 0 & \ldots & 0 & -\widehat{\rho} & 1
\end{pmatrix}
\tag{48}
$$

It can be shown, that this matrix, multiplied by its transpose and the matrix $\widehat{\Omega}$ (which is defined by equation (45)) is proportional to the unit matrix

$$
\frac{1}{1-\widehat{\rho}^2} \widehat{\Psi}' \widehat{\Psi} \widehat{\Omega} = I_n
$$

and hence the matrix $\widehat{\Psi}$ has the following property

$$
\widehat{\Psi}' \widehat{\Psi} = (1-\widehat{\rho}^2)\widehat{\Omega}^{-1}
\tag{49}
$$

Since least squares estimation is not affected by scalar multiplication, we multiply the regression model by $\sqrt{1-\widehat{\rho}^2}$. Expressing $\widehat{\Omega}^{-1}$ in equation (47) using equation (49) leads to the the following GLS estimator

$$
\widehat{\beta}_{GLS} = (X'\widehat{\Psi}'\widehat{\Psi}X)^{-1}X'\widehat{\Psi}'\widehat{\Psi}y = [(\widehat{\Psi}X)'(\widehat{\Psi}X)]^{-1}[(\widehat{\Psi}X)]'(\widehat{\Psi}y)
\tag{50}
$$

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

$$
y_t - \widehat{\rho}y_{t-1} = \sum_{j=0}^{p}(x_{tj} - \widehat{\rho}x_{t-1,j})\widehat{\beta}_j + u_t
\tag{51}
$$

for $t \in \{2,\ldots,n\}$, u_t being a Gaussian noise. For $t = 1$ it is simply

$$
\sqrt{1-\widehat{\rho}^2}y_1 = \sqrt{1-\widehat{\rho}^2}\sum_{j=0}^{p}\widehat{\beta}_j x_{1j} + \sqrt{1-\widehat{\rho}^2}\epsilon_1
\tag{52}
$$

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}) were chosen to be tested as explanatory variables (X), whereas the first two principal

components of the pricing kernel and relative risk aversion differences for 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.

Table 3 describes the best fitted models for each of the principal components (based on equation (51)). 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 regressions' coefficients appear in brackets.

The first principal component of the pricing kernel differences, which was described before as a central mass movement factor, dominated by the changes in the mean 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 principal component of the pricing kernel 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 mean and standard deviation.

We can not find a significant relationship between the second principal component of the pricing kernel differences and the explanatory variables (other than for very short maturities), a result that supports the second principal component's smaller contribution to the variability of pricing kernel 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 principal component with $\Delta\mu_t(PK)$ and $\Delta\sigma_t(PK)$ for the respective maturities.

Pricing Kernel Differences

PC	Maturity	$\widehat{\rho}$	$\widehat{\beta}_1$	$x_{t,1}$	$\widehat{\beta}_2$	$x_{t,2}$	F
1	30	-0.426	-1.80	$\log \frac{S_t}{S_{t-1}}$	1.76	$\log \frac{IV_t^{ATM}}{IV_{t-1}^{ATM}}$	18.958
			(0.289)		(0.000)		(0.000)
1	60	-0.468	2.71	$\log \frac{S_t}{S_{t-1}}$	0.98	$\log \frac{IV_t^{ATM}}{IV_{t-1}^{ATM}}$	10.784
			(0.005)		(0.001)		(0.000)
1	90		N o t		S i g n i f i c a n t		
2	30	-0.473	30.21	$\log \frac{S_t}{S_{t-1}}$	12.77	$\log \frac{IV_t^{ATM}}{IV_{t-1}^{ATM}}$	20.718
			(0.001)		(0.000)		(0.000)
2	60		N o t		S i g n i f i c a n t		
2	90		N o t		S i g n i f i c a n t		

Relative Risk Aversion Differences

PC	Maturity	$\widehat{\rho}$	$\widehat{\beta}_1$	$x_{t,1}$	$\widehat{\beta}_2$	$x_{t,2}$	F
1	30	-0.544	0.03	ΔS_t	145.34	ΔIV_t^{ATM}	11.562
			(0.000)		(0.000)		(0.000)
	60	-0.457	0.03	ΔS_t	286.43	ΔIV_t^{ATM}	18.048
			(0.001)		(0.000)		(0.000)
	90	-0.510	0.02	ΔS_t	224.27	ΔIV_t^{ATM}	10.666
			(0.028)		(0.000)		(0.000)
2	30		N o t		S i g n i f i c a n t		
	60	-0.460	-0.01	ΔS_t	-92.15	ΔIV_t^{ATM}	7.217
			(0.042)		(0.000)		(0.000)
	90	-0.497	0.01	ΔS_t	35.72	ΔIV_t^{ATM}	4.026
			(0.020)		(0.011)		(0.018)

Table 3: The estimated parameters of the suggested regression model in equation (51): $y_t - \widehat{\rho} y_{t-1} = \widehat{\beta}_0 + \widehat{\beta}_1 (x_{t,1} - \widehat{\rho} x_{t-1,1}) + \widehat{\beta}_2 (x_{t,2} - \widehat{\rho} x_{t-1,2}) + u_t$ where $\Delta IV_t^{ATM} \stackrel{\text{def}}{=} IV_t^{ATM} - IV_{t-1}^{ATM}$, $\Delta S_t \stackrel{\text{def}}{=} S_t - S_{t-1}$ and $\widehat{\beta}_0 = 0$ because the constant is never significant due to the zero mean property of the principal components.

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

According to table 2 in the previous section, the correlations of the first principal components of the relative risk aversion differences with their dominant moments are positive. The first principal component 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 principal component, 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 principal component of the relative risk aversion differences are slightly different. The second principal components are positively correlated to the change of relative risk aversion at the money. 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

principal component of pricing kernel differences for short and medium term maturities, and to both principal components of relative risk aversion differences. 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.

5 Final Statements

This work focused on estimating the subjective density and the state-price density of the stochastic process associated with the DAX. Based on the work of Rubinstein (1994), a good estimation of those two measures is sufficient for deriving the investors' preferences. However, 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 section 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 daily estimated pricing kernel and relative risk aversion were found to have similar characteristics to those reported by Jackwerth (2000) and Aït Sahalia & Lo (2000). The pricing kernel was shown not to be a strictly decreasing function as suggested by classical macroeconomic theory, and the relative risk aversion 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 variability of the stationary daily changes in pricing kernel and relative risk aversion was found to be well explained by two factors. Since the factors experienced some evident autocorrelation, the principal components were tested as the response variable in a GLS regression model, which regressed each of the principal components on the daily changes in the DAX and in ATM implied volatility.

We found that large changes in ATM implied volatility lead to a more volatile and time-varying pricing kernel. The absence of a significant fitted regression model for the second principal component of the pricing kernel differences was in accordance with its smaller contribution to the explained variability.

35

In addition, 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.

References

Aït Sahalia, Y. & Lo, A. W. (2000), 'Nonparametric risk management and implied risk aversion', *Journal of Econometrics* **94**, 9–51.

Arrow, K. J. (1964), 'The role of securities in the optimal allocation of risk bearing', *Review of Economic Studies* **31**, 91–96.

Arrow, K. J. (1965), 'Aspects of the theory of risk-bearing', *Yrjö Hahnsson Foundation, Helsinki* .

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabililties', *Journal of Political Economy* **81**, 637–654.

Breeden, D. & Litzenberger, R. (1978), 'Prices of state-contingent claims implicit in option prices', *Journal of Business* **51**, 621–651.

Brown, D. P. & Jackwerth, J. C. (2004), 'The pricing kernel puzzle: Reconciling index option data and economic theory', *Working Paper, University of Konstanz / University of Wisconsin* .

Cochrane, J. H. (2001), 'Asset pricing', *Princeton University Press* .

Debreu, G. (1959), 'The theory of value', *Wiley, New York* .

Franke, J., Härdle, W. & Hafner, C. (2004), 'Statistics of financial markets', *Springer Verlag, Heidelberg* .

Härdle, W. (1990), 'Applied nonparametric regression', *Cambridge University Press, Cambridge* .

Huynh, K., Kervella, P. & Zheng, J. (2002), 'Estimating state-price densities with nonparametric regression', *in: W. Härdle, T. Kleinow and G. Stahl (eds.), Applied Quantitative Finance, Springer, Ch. 8* .

Jackwerth, J. C. (2000), 'Recovering risk aversion from option prices and realized returns', *Review of Financial Studies* **13**, 433–451.

Kwiatkowski, D., Phillips, P., Schmidt, P. & Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic series have a unit root', *Journal of Econometrics* **54**, 159–178.

Lucas, R. E. (1978), 'Asset prices in an exchange economy', *Econometrica* **46**, 1429–1446.

Mas-Colell, A., Whinston, M. & Green, J. (1995), 'Microeconomic theory', *Oxford University Press* .

Merton, R. (1973), 'Rational theory of option pricing', *Journal of Economics and Management Science* **4**, 141–183.

Prais, S. J. & Winsten, C. B. (1954), 'Trend estimators and serial correlation', *Cowles Commission Discussion Paper 383, Chicago* **383**.

Pratt, J. (1964), 'Risk aversion in the small and in the large', *Econometrica* **32**.

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

Rosenberg, J. V. & Engle, R. F. (2002), 'Empirical pricing kernels', *Journal of Financial Economics* **64**, 341–372.

Rubinstein, M. (1976), 'The valuation of uncertain income streams and the pricing of options', *Bell Journal of Economics* **7**, 407–425.

Rubinstein, M. (1994), 'Implied binomial trees', *Journal of Finance* **49**, 771–818.

Sen, A. & Srivastava, M. (1990), 'Regression analysis: Theory, methods and applications', *Springer Verlag, New-York* .

Veldkamp, L. (2005), 'Media frenzies in markets for financial information', *American Economic Review, forthcoming* .

SFB 649 Discussion Paper Series 2006

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.