

Empirical Pricing Kernels and Investor Preferences

K. Detlefsen¹, W. K. Härdle², R. A. Moro³,

¹CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany; e-mail: detlefsen@wiwi.huberlin.de; phone: +49(0)30 2093-5807

²CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany; e-mail: haerdle@wiwi.huberlin.de; phone: +49(0)30 2093-5630

³German Institute for Economic Research, Königin-Luise-Straße 5, 14195 Berlin, Germany; e-mail: rmoro@diw.de; phone: +49(0)30 8978-9262 and CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin

Abstract

This paper analyzes empirical market utility functions and pricing kernels derived from the DAX and DAX option data for three market regimes. A consistent parametric framework of stochastic volatility is used. All empirical market utility functions show a region of risk proclivity that is reproduced by adopting the hypothesis of heterogeneous individual investors whose utility functions have a switching point between bullish and bearish attitudes. The inverse problem of finding the distribution of individual switching points is formulated in the space of stock returns by discretization as a quadratic optimization problem. The resulting distributions vary over time and correspond to different market regimes.

JEL classification: G12, G13, C50

Keywords: Utility function, pricing kernel, behavioral finance, risk aversion, risk proclivity, Heston model

1 Introduction

Numerous attempts have been undertaken to describe basic principles on which the behaviour of individuals are based. Expected utility theory was originally proposed by J. Bernoulli in 1738. In his work J. Bernoulli used such terms as risk aversion and risk premium and proposed a concave (logarithmic) utility function, see Bernoulli (1956). The utilitarianism theory that emerged in the 18th century considered utility maximization as a principle for the organisation of society. Later the expected utility idea was applied to game theory and formalized by von Neumann and Morgenstern (1944). A utility function relates some observable variable, in most cases consumption, and an unobservable utility level that this consumption delivers. It was suggested that individuals' preferences are based on this unobservable utility: such bundles of goods are preferred that are associated with higher utility levels. It was claimed that three types of utility functions – concave, convex and linear – correspond to three types of individuals – risk averse, risk neutral and risk seeking. A typical economic agent was considered to be risk averse and this was quantified by coefficients of relative or absolute risk aversion. Another important step in the development of utility theory was the prospect theory of Kahneman and Tversky (1979). By behavioural experiments they found that people act risk averse above a certain reference point and risk seeking below it. This implies a concave form of the utility function above the reference point and a convex form below it.

Besides these individual utility functions, market utility functions have recently been analyzed in empirical studies by Jackwerth (2000), Rosenberg and Engle (2002) and others. Across different markets, the authors observed a common pattern in market utility functions: There is a reference point near the initial wealth and in a region around this reference point the market utility functions are convex. But for big losses or gains they show a concave form – risk aversion. Such utility functions disagree with the classical utility functions of von Neumann and Morgenstern (1944) and also with the findings of Kahneman and Tversky (1979). They are however in concordance with the utility function form proposed by Friedman and Savage (1948).

In this paper, we analyze how these market utility functions can be explained by aggregating individual investors' attitudes. To this end, we first determine empirical pricing kernels from DAX data. Our estimation procedure is based on historical and risk neutral densities and these distributions are derived with stochastic volatility models that are widely used in industry. From these pricing kernels we construct the corresponding market utility functions. Then we describe our method of aggregating individual utility functions to a market utility function. This leads to an inverse problem for

the density function that describes how many investors have the utility function of each type. We solve this problem by discrete approximation. In this way, we derive utility functions and their distribution among investors that allow to recover the market utility function. Hence, we explain how (and what) individual utility functions can be used to form the behaviour of the whole market.

The paper is organized as follows: In section 2, we describe the theoretical connection between utility functions and pricing kernels. In section 3, we present a consistent stochastic volatility framework for the estimation of both the historical and the risk neutral density. Moreover, we discuss the empirical pricing kernel implied by the DAX in 2000, 2002 and 2004. In section 4, we explain the utility aggregation method that relates the market utility function and the utility functions of individual investors. This aggregation mechanism leads to an inverse problem that is analyzed and solved in this section. In section 5, we conclude and discuss related approaches.

2 Pricing kernels and utility functions

In this section, we derive the fundamental relationship between utility functions and pricing kernels. It describes how a representative utility function can be derived from historical and risk-neutral distributions of assets. In the following sections, we estimate the empirical pricing kernel and observe in this way the market utility function.

First, we derive the price of a security in an equilibrium model: we consider an investor with a utility function U who has as initial endowment one share of stock. He can invest into the stock and a bond up to a final time when he can consume. His problem is to choose a strategy that maximizes the expected utility of his initial and terminal wealth. In continuous time, this leads to a well known optimization problem introduced by Merton (1973) for stock prices modelled by diffusions. In discrete time, it is a basic optimization problem, see Cochrane (2001).

From this result, we can derive the asset pricing equation

$$P_0 = E^P [\psi(S_T)M_T]$$

for a security on the stock (S_t) with payoff function ψ at maturity T . Here, P_0 denotes the price of the security at time 0 and E^P is the expectation with respect to the real/historical measure P . The stochastic discount factor M_T is given by

$$M_T = \beta U'(S_T)/U'(S_0) \tag{1}$$

where β is a fixed discount factor. This stochastic discount factor is actually the projection of the general stochastic discount factor on the traded asset (S_t). The stochastic discount factor can depend on more variables in general. But as discussed in Cochrane (2001) this projection has the same interpretation for pricing as the general stochastic discount factor.

Besides this equilibrium based approach, Black and Scholes (1973) derived the price of a security relative to the underlying by constructing a perfect hedge. The resulting continuous delta hedging strategy is equivalent to pricing under a risk neutral measure Q under which the discounted price process of the underlying becomes a martingale. Hence, the price of a security is given by an expected value with respect to a risk neutral measure Q :

$$P_0 = E^Q [\exp(-rT)\psi(S_T)]$$

If p denotes the historical density of S_T (i.e. $P(S_T \leq s) = \int_{-\infty}^s p(x) dx$) and q the risk neutral density of S_T (i.e. $Q(S_T \leq s) = \int_{-\infty}^s q(x) dx$) then we get

$$\begin{aligned} P_0 &= \exp(-rT) \int \psi(x)q(x)dx \\ &= \exp(-rT) \int \psi(x) \frac{q(x)}{p(x)} p(x)dx \\ &= E^P \left[\exp(-rT)\psi(S_T) \frac{q(S_T)}{p(S_T)} \right] \end{aligned} \tag{2}$$

Combining equations (1) and (2) we see

$$\beta \frac{U'(s)}{U'(S_0)} = \exp(-rT) \frac{q(s)}{p(s)}.$$

Defining the pricing kernel by $K = q/p$ we conclude that the form of the market utility function can be derived from the empirical pricing kernel by integration:

$$\begin{aligned} U(s) &= U(S_0) + \int_{S_0}^s U'(S_0) \frac{\exp(-rT)}{\beta} \frac{q(x)}{p(x)} dx \\ &= U(S_0) + \int_{S_0}^s U'(S_0) \frac{\exp(-rT)}{\beta} K(x) dx \end{aligned}$$

because S_0 is known.

As an example, we consider the model of Black and Scholes (1973) where the stock follows a geometric Brownian motion

$$dS_t/S_t = \mu dt + \sigma dW_t \quad (3)$$

Here the historical density p of S_t is log-normal, i.e.

$$p(x) = \frac{1}{x} \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} \exp\left\{-\frac{1}{2}\left(\frac{\log x - \tilde{\mu}}{\tilde{\sigma}}\right)^2\right\}, \quad x > 0$$

where $\tilde{\mu} = (\mu - \sigma^2/2)t + \log S_0$ and $\tilde{\sigma} = \sigma\sqrt{t}$. Under the risk neutral measure Q the drift μ is replaced by the riskless interest rate r , see e.g. Harrison and Pliska (1981). Thus, also the risk neutral density q is log-normal. In this way, we can derive the pricing kernel

$$K(x) = \left(\frac{x}{S_0}\right)^{-\frac{\mu-r}{\sigma^2}} \exp\{(\mu - r)(\mu + r - \sigma^2)T/(2\sigma^2)\}.$$

This pricing kernel has the form of a derivative of a power utility

$$K(x) = \lambda \left(\frac{x}{S_0}\right)^{-\gamma}$$

where the constants are given by $\lambda = e^{\frac{(\mu-r)(\mu+r-\sigma^2)T}{2\sigma^2}}$ and $\gamma = \frac{\mu-r}{\sigma^2}$. This gives a utility function corresponding to the underlying (3)

$$U(S_T) = \left(1 - \frac{\mu - r}{\sigma^2}\right)^{-1} S_T^{(1 - \frac{\mu-r}{\sigma^2})}$$

where we ignored additive and multiplicative constants. In this power utility function the risk aversion is not given by the market price of risk $(\mu - r)/\sigma$. Instead investors take the volatility more into account. The expected return $\mu - r$ that is adjusted by the riskfree return is related to the variance. This results in a higher relative risk aversion than the market price of risk.

A utility function corresponding to the Black-Scholes model is shown in the upper panel of figure 1 as a function of returns. In order to make different market situations comparable we consider utility functions as functions of (half year) returns $R = S_{0.5}/S_0$. We chose the time horizon of half a year ahead for our analysis. Shorter time horizons are interesting economically and moreover the historical density converges to the Dirac measure so that results become trivial (in the end). Longer time horizons are economically

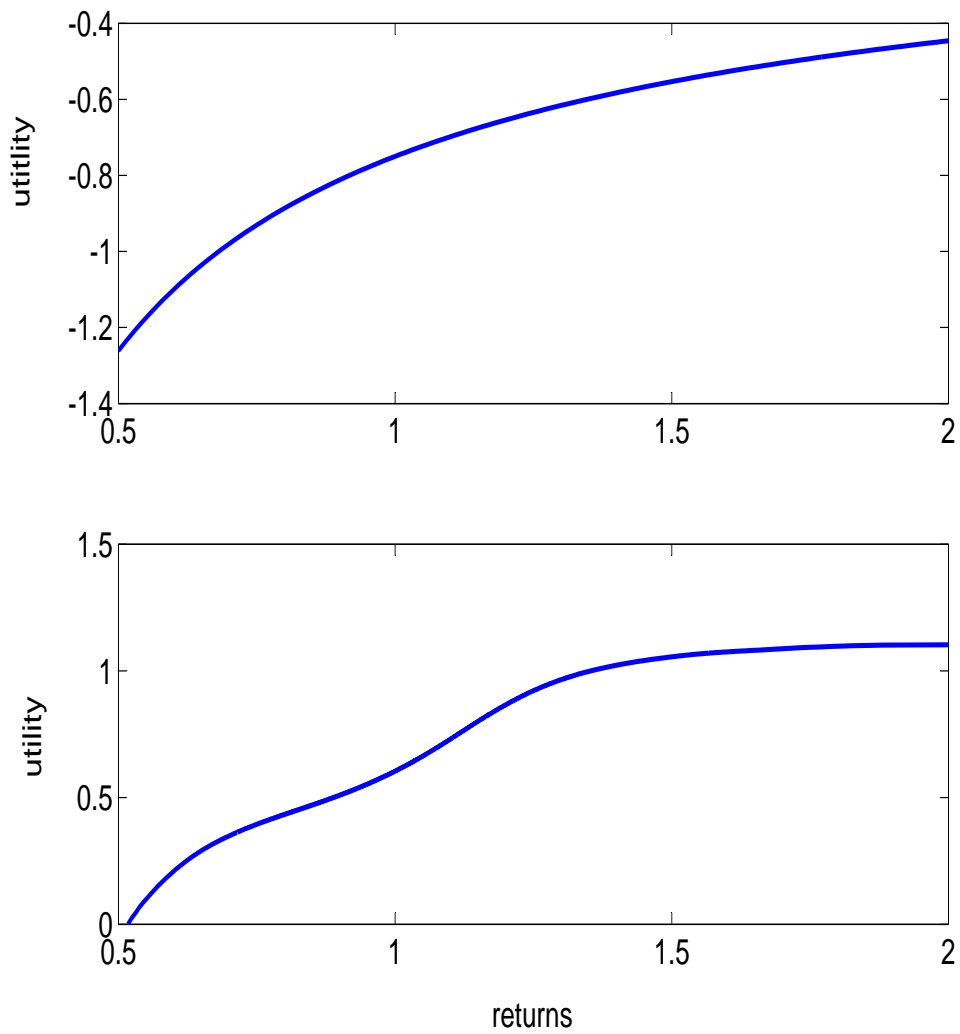


Figure 1: up: Utility function in the Black Scholes model for $T = 0.5$ years ahead and drift $\mu = 0.1$, volatility $\sigma = 0.2$ and interest rate $r = 0.03$. down: Market utility function on 06/30/2000 for $T = 0.5$ years ahead.

more interesting but it is hardly possible to estimate the historical density for a long time ahead. It neither seems realistic to assume that investors have clear ideas where the DAX will be in e.g. 10 years. For these reasons we use half a year as future horizon. Utility functions \tilde{U} of returns are defined by:

$$\tilde{U}(R) := U(RS_0), \quad R > 0$$

where S_0 denotes the value of the DAX on the day of estimation. Because of $U' = cK$ for a constant c we have $\tilde{U}'(R) = cK(RS_0)S_0$ and we see that also utility functions of returns are given as integrals of the pricing kernel. The change to returns allows us to compare different market regimes independently of the initial wealth. In the following we denote the utility functions of returns by the original notation U . Hence, we suppress in the notation the dependence of the utility function U on the day of estimation t .

The utility function corresponding to the model of Black and Scholes (1973) is a power utility, monotonically increasing and concave. But such classical utility functions are not observed on the market. Parametric and nonparametric models that replicate the option prices all lead to utility functions with a hump around the initial wealth level. This is described in detail later but is shown already in figure 1. The upper panel presents the utility function corresponding to Black-Scholes model with a volatility of 20% and an expected return of 10%. The function is concave and implies a constant relative risk aversion. The utility function estimated on the bullish market in summer 2000 is presented in the lower panel. Here, the hump around the money is clearly visible. The function is no more concave but has a region where investors are risk seeking. This risk proclivity around the money is reflected in a negative relative risk aversion.

3 Estimation

In this section, we start by reviewing some recent approaches for estimating the pricing kernel. Then we describe our method that is based on estimates of the risk neutral and the historical density. The risk neutral density is derived from option prices that are given by an implied volatility surface and the historical density is estimated from the independent data set of historical returns. Finally, we present the empirical pricing kernels and the inferred utility and relative risk aversion functions.

3.1 Estimation approaches for the pricing kernel

There exist several ways and methods to estimate the pricing kernel. Some of these methods assume parametric models while others use nonparametric techniques. Moreover, some methods estimate first the risk neutral and subjective density to infer the pricing kernel. Other approaches estimate directly the pricing kernel.

Ait-Sahalia and Lo (1998) derive a nonparametric estimator of the risk neutral density based on option prices. In Ait-Sahalia and Lo (2000), they consider the empirical pricing kernel and the corresponding risk aversion using this estimator. Moreover, they derive asymptotic properties of the estimator that allow e.g. the construction of confidence bands. The estimation procedure consists of two steps: First, the option price function is determined by nonparametric kernel regression and then the risk neutral density is computed by the formula of Breeden and Litzenberger (1978). Advantages of this approach are the known asymptotic properties of the estimator and the few assumptions necessary.

Jackwerth (2000) analyses risk aversion by computing the risk neutral density from option prices and the subjective density from historical data of the underlying. For the risk neutral distribution, he applies a variation of the estimation procedure described in Jackwerth and Rubinstein (1996): A smooth volatility function derived from observed option prices gives the risk neutral density by differentiating it twice. The subjective density is approximated by a kernel density computed from historical data. In this method bandwidths have to be chosen as in the method of Ait-Sahalia and Lo (1998).

Rosenberg and Engle (2002) use a different approach and estimate the subjective density and directly (the projection of) the pricing kernel. This gives the same information as the estimation of the two densities because the risk neutral density is the product of the pricing kernel and the subjective density. For the pricing kernel, they consider two parametric specifications as power functions and as exponentials of polynomials. The evolution of the underlying is modelled by GARCH processes. As the parametric pricing kernels lead to different results according to the parametric form used this parametric approach appears a bit problematic.

Chernov (2003) also estimates the pricing kernel without computing the risk neutral and subjective density explicitly. Instead of assuming directly a parametric form of the kernel he starts with a (multi dimensional) modified model of Heston (1993) and derives an analytic expression for the pricing kernel by the Girsanov theorem, see Chernov (2000) for details. The ker-

nel is estimated by a simulated method of moments technique from equity, fixed income and commodities data and by reprojection. An advantage of this approach is that the pricing kernel is estimated without assuming an equity index to approximate the whole market portfolio. But the estimation procedure is rather complex and model dependent.

In a recent paper, Barone-Adesi et al. (2004) price options in a GARCH framework allowing the volatility to differ between historical and risk neutral distribution. This approach leads to acceptable calibration errors between the observed option prices and the model prices. They estimate the historical density as a GARCH process and consider the pricing kernel only on one day. This kernel is decreasing which coincides with standard economic theory. But the general approach of changing explicitly the volatility between the historical and risk neutral distribution is not supported by the standard economic theory.

We estimate the pricing kernel in this paper by estimating the risk neutral and the subjective density and then deriving the pricing kernel. This approach does not impose a strict structure on the kernel. Moreover, we use accepted parametric models because nonparametric techniques for the estimation of second derivatives depend a lot on the bandwidth selection although they yield the same pricing kernel behaviour over a wide range of bandwidths. For the risk neutral density we use a stochastic volatility model that is popular both in academia and in industry. The historical density is more difficult to estimate because the drift is not fixed. Hence, the estimation depends more on the model and the length of the historical time series. In order to get robust results we consider different (discrete) models and different lengths. In particular, we use a GARCH model that is the discrete version of the continuous model for the risk neutral density. In the following, we describe these models, their estimation and the empirical results.

3.2 Estimation of the risk neutral density

Stochastic volatility models are popular in industry because they replicate the observed smile in the implied volatility surfaces (IVS) rather well and moreover imply rather realistic dynamics of the surfaces. Nonparametric approaches like the local volatility model of Dupire (1994) allow a perfect fit to observed price surfaces but their dynamics are in general contrary to the market. As Bergomi (2005) points out the dynamics are more important for modern products than a perfect fit. Hence, stochastic volatility models are popular.

We consider the model of Heston (1993) for the risk neutral density be-

cause it can be interpreted as the limit of GARCH models. The Heston model has been refined further in order to improve the fit, e.g. by jumps in the stock price or by a time varying mean variance level. We use the original Heston model in order to maintain a direct connection to GARCH processes. Although it is possible to estimate the historical density also with the Heston model e.g. by Kalman filter methods we prefer more direct approaches in order to reduce the dependence of the results on the model and the estimation technique.

The stochastic volatility model of Heston (1993) is given by the two stochastic differential equations:

$$\frac{dS_t}{S_t} = rdt + \sqrt{V_t}dW_t^1$$

where the variance process is modelled by a square-root process:

$$dV_t = \xi(\eta - V_t)dt + \theta\sqrt{V_t}dW_t^2$$

and W^1 and W^2 are Wiener processes with correlation ρ and r is the risk free interest rate. The first equation models the stock returns by normal innovations with stochastic variance. The second equation models the stochastic variance process as a square-root diffusion.

The parameters of the model all have economic interpretations: η is called the long variance because the process always returns to this level. If the variance V_t is e.g. below the long variance then $\eta - V_t$ is positive and the drift drives the variance in the direction of the long variance. ξ controls the speed at which the variance is driven to the long variance. In calibrations, this parameter changes a lot and makes also the other parameters instable. To avoid this problem, the reversion speed is kept fixed in general. We follow this approach and choose $\xi = 2$ as Bergomi (2005) does. The volatility of variance θ controls mainly the kurtosis of the distribution of the variance. Moreover, there are the initial variance V_0 of the variance process and the correlation ρ between the Brownian motions. This correlation models the leverage effect: When the stock goes down then the variance goes up and vice versa. The parameters also control different aspects of the implied volatility surface. The short (long) variance determines the level of implied volatility for short (long) maturities. The correlation creates the skew effect and the volatility of variance controls the smile.

The variance process remains positive if the volatility of variance θ is small enough with respect to the product of the mean reversion speed ξ and

the long variance level η (i.e. $2\xi\eta > \theta^2$). As this constraint leads often to significantly worse fits to implied volatility surfaces it is in general not taken into account and we follow this approach.

The popularity of this model can probably be attributed to the semiclosed form of the prices of plain vanilla options. Carr and Madan (1999) showed that the price $C(K, T)$ of a European call option with strike K and maturity T is given by

$$C(K, T) = \frac{\exp\{-\alpha \ln(K)\}}{\pi} \int_0^{+\infty} \exp\{-\mathbf{i}v \ln(K)\} \psi_T(v) dv$$

for a (suitable) damping factor $\alpha > 0$. The function ψ_T is given by

$$\psi_T(v) = \frac{\exp(-rT) \phi_T\{v - (\alpha + 1)\mathbf{i}\}}{\alpha^2 + \alpha - v^2 + \mathbf{i}(2\alpha + 1)v}$$

where ϕ_T is the characteristic function of $\log(S_T)$. This characteristic function is given by

$$\begin{aligned} \phi_T(z) &= \exp\left\{\frac{-(z^2 + \mathbf{i}z)V_0}{\gamma(z) \coth \frac{\gamma(z)T}{2} + \xi - \mathbf{i}\rho\theta z}\right\} \\ &\times \frac{\exp\left\{\frac{\xi\eta T(\xi - \mathbf{i}\rho\theta z)}{\theta^2} + \mathbf{i}zTr + \mathbf{i}z \log(S_0)\right\}}{\left(\cosh \frac{\gamma(z)T}{2} + \frac{\xi - \mathbf{i}\rho\theta z}{\gamma(z)} \sinh \frac{\gamma(z)T}{2}\right)^{\frac{2\xi\eta}{\theta^2}}} \end{aligned} \quad (4)$$

where $\gamma(z) \stackrel{\text{def}}{=} \sqrt{\theta^2(z^2 + \mathbf{i}z) + (\xi - \mathbf{i}\rho\theta z)^2}$, see e.g. Cizek et al. (2005).

For the calibration we minimize the absolute error of implied volatilities based on the root mean square error:

$$\text{ASE}_t \stackrel{\text{def}}{=} \sqrt{\sum_{i=1}^n n^{-1} \{IV_i^{\text{mod}}(t) - IV_i^{\text{mar}}(t)\}^2}$$

where *mod* refers to a model quantity, *mar* to a quantity observed on the market and $IV(t)$ to an implied volatility on day t . The index i runs over all n observations of the surface on day t .

It is essential for the error functional ASE_t which observed prices are used for the calibration. As we investigate the pricing kernel for half a year to maturity we use only the prices of options that expire in less than 1.5 years. In order to exclude liquidity problems occurring at expiry we consider for the

calibration only options with more than 1 month time to maturity. In the moneyness direction we restrict ourselves to strikes 50% above or below the spot for liquidity reasons.

The risk neutral density is derived by estimation of the model parameters by a least squares approach. This amounts to the minimization of the error functional ASE_t . Cont and Tankov (2004) provided evidence that such error functionals may have local minima. In order to circumvent this problem we apply a stochastic optimization routine that does not get trapped in a local minimum. To this end, we use the method of differential evolution developed by Storn and Price (1997).

Having estimated the model parameters we know the distribution of $X_T = \log S_T$ in form of the characteristic function ϕ_T , see (4). Then the corresponding density f of X_T can be recovered by Fourier inversion:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} \phi_T(t) dt,$$

see e.g. Billingsley (1995). This integral can be computed numerically.

Finally, the risk neutral density q of $S_T = \exp(X_T)$ is given as a transformed density:

$$q(x) = \frac{1}{x} f\{\log(x)\}.$$

This density q is risk neutral because it is derived from option prices and options are priced under the risk neutral measure. This measure is applied because banks replicate the payoff of options so that no arbitrage conditions determine the option price, see e.g. Rubinstein (1994). An estimated risk neutral density is presented in figure 2. It is estimated from the implied volatility shown in figure 3 for the day 24/03/2000. The distribution is right skewed and its mean is fixed by the martingale property. This implies that the density is low for high profits and high for high losses. Moreover, the distribution is not symmetrical around the neutral point where there are neither profits nor losses. For this and all the following estimations we approximate the risk free interest rates by the EURIBOR. On each trading day we use the yields corresponding to the maturities of the implied volatility surface. As the DAX is a performance index it is adjusted to dividend payments. Thus, we do not have to consider dividend payments explicitly.

3.3 Estimation of the historical density

While the risk neutral density is derived from option prices observed on the day of estimation we derive the subjective density from the historical time

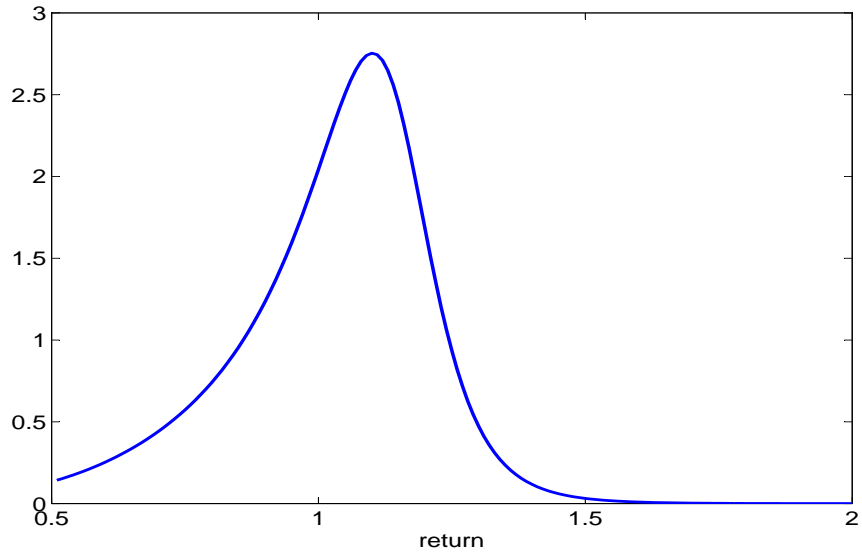


Figure 2: Risk neutral density on 24/03/2000 half a year ahead.

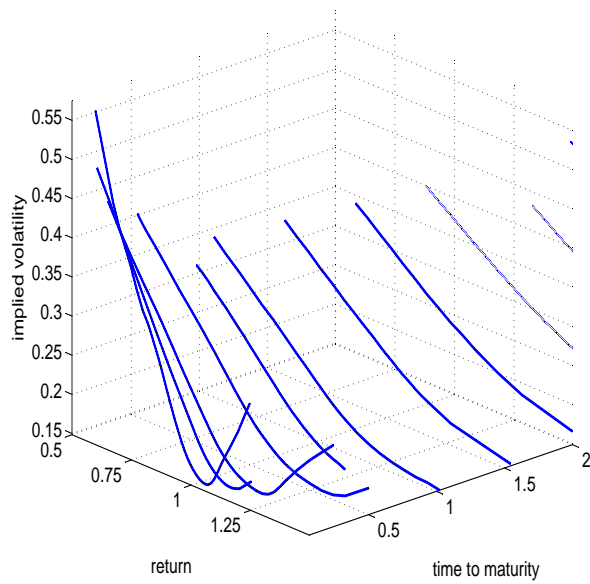


Figure 3: Implied volatility surface on 24/03/00.

model	time period
GARCH in mean	2.0y
discrete Heston	2.0y
observed returns	1.0y

Table 1: Models and the time periods used for their estimation.

series of the index. Hence, the two data sets are independent in the sense that the option prices reflect the future movements and the historical time series the past.

The estimation of the historical density seems more difficult than the estimation of the risk neutral density because the drift is not fixed and it depends in general on the length of the time series. Because of these difficulties we use different models and time horizons for the historical density: First, we estimate a GARCH in mean model for the returns. Returns are generally assumed to be stationary and we confirmed this at least in the time intervals we consider. The mean component in the GARCH model is important to reflect different market regimes. We estimate the GARCH model from the time series of the returns of the last two year because GARCH models require quite long time series for the estimation in order to make the standard error reasonably small. We do not choose longer time period for the estimation because we want to consider special market regimes. Besides this popular model choice we apply a GARCH model that converges in the limit to the Heston model that we used for the risk neutral density. As this model is also hard to estimate we use again the returns of the last 2 years for this model. Moreover, we consider directly the observed returns of the last year. The models and their time period for the estimation are presented in table 1. All these models give by simulation and smoothing the historical density for half a year ahead.

The GARCH estimations are based on the daily log-returns

$$R_i = \log(S_{t_i}) - \log(S_{t_{i-1}})$$

where (S_t) denotes the price process of the underlying and t_i , $i = 1, 2, \dots$ denote the settlement times of the trading days. Returns of financial assets have been analyzed in numerous studies, see e.g. Cont (2001). A model that has often been successfully applied to financial returns and their stylized facts

is the GARCH(1,1) model. This model with a mean is given by

$$\begin{aligned} R_i &= \mu + \sigma_i Z_i \\ \sigma_i^2 &= \omega + \alpha R_{i-1}^2 + \beta \sigma_{i-1}^2 \end{aligned}$$

where (Z_i) are independent identically distributed innovations with a standard normal distribution, see e.g. Franke et al. (2004). On day t_j the model parameters μ, ω, α and β are estimated by quasi maximum likelihood from the observations of the last two years, i.e. R_{j-504}, \dots, R_j assuming 252 trading days per year.

After the model parameters have been estimated on day t_j from historical data the process of logarithmic returns (R_i) is simulated half a year ahead, i.e. until time $t_j + 0.5$. In such a simulation μ, ω, α and β are given and the time series (σ_i) and (R_i) are unknown. The values of the DAX corresponding to the simulated returns are then given by inverting the definition of the log returns:

$$S_{t_i} = S_{t_{i-1}} \exp(R_i)$$

where we start with the observed DAX value on day t_j . Repeating the simulation N times we obtain N samples of the distribution of $S_{t_j+0.5}$. We use $N = 2000$ simulations because tests have shown that the results become robust around this number of simulations.

From these samples we estimate the probability density function of $S_{t_j+0.5}$ (given $(S_{t_j-126}, \dots, S_{t_j})$) by kernel density estimation. We apply the Gaussian kernel and choose the bandwidth by Silverman's rule of thumb, see e.g. Silverman (1986). This rule provides a trade-off between oversmoothing – resulting in a high bias – and undersmoothing – leading to big variations of the density. We have moreover checked the robustness of the estimate relative to this bandwidth choice. The estimation results of a historical density are presented in figure 4 for the day 24/03/2000. This density that represents a bullish market is has most of its weight in the profit region and its tail for the losses is relatively light.

As we use the Heston model for the estimation of the risk neutral density we consider in addition to the described GARCH model a GARCH model that is a discrete version of the Heston model. Heston and Nandi (2000) show that the discrete version of the square-root process is given by

$$V_i = \omega + \beta V_{i-1} + \alpha(Z_{i-1} - \gamma\sqrt{V_{i-1}})$$

and the returns are modelled by

$$R_i = \mu - \frac{1}{2}V_i + \sqrt{V_i}Z_i$$

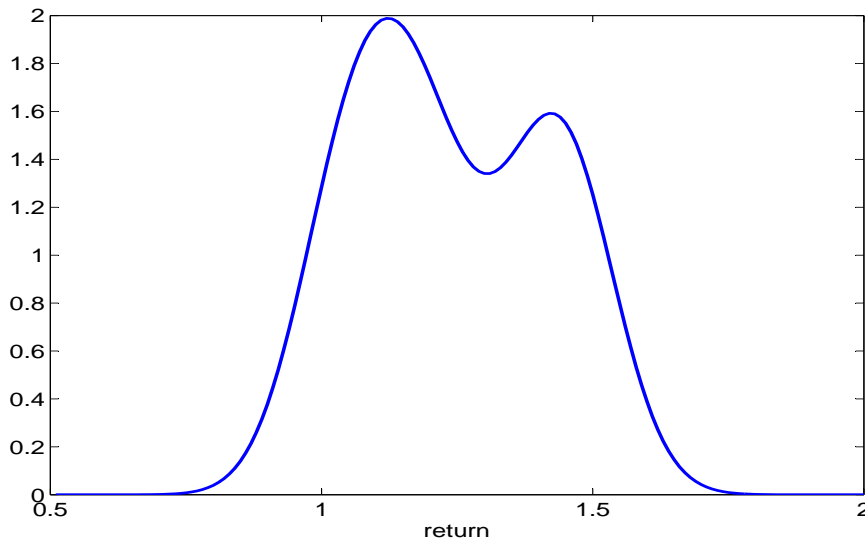


Figure 4: Historical density on 24/03/2000 half a year ahead.

where (Z_i) are independent identically distributed innovations with a standard normal distribution. Having estimated this model by maximum likelihood on day t_j we simulate it half a year ahead and then smooth the samples of $S_{t_j+0.5}$ in the same way as in the other GARCH model.

In addition to these parametric models, we consider directly the observed returns over half a year

$$\tilde{R}_i = S_{t_i}/S_{t_i-126}.$$

In this way, we interpret these half year returns as samples from the distribution of the returns for half a year ahead. Smoothing these historical samples of returns gives an estimate of the density of returns and in this way also an estimate of the historical density of $S_{t_j+0.5}$.

3.4 Empirical pricing kernels

In contrast to many other studies that concentrate on the S&P500 index we analyze the German economy by focusing on the DAX, the German stock index. This broad index serves as an approximation to the German economy. We use two data sets: A daily time series of the DAX for the estimation of the subjective density and prices of European options on the DAX for the estimation of the risk neutral density.

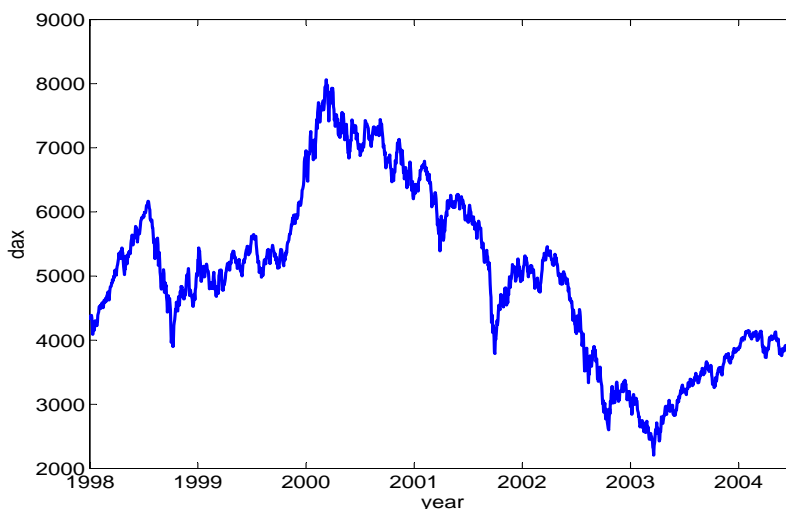


Figure 5: DAX, 1998 - 2004.

	1.0y	2.0y
03/2000	1.63	1.57
07/2002	0.66	0.54
06/2004	1.11	0.98

Table 2: Market regimes in 2000, 2002 and 2004 described by the return $S_0/S_{0-\Delta}$ for periods $\Delta = 1.0y, 2.0y$.

In figure 5, we present the DAX in the years 1998 to 2004. This figure shows that the index reached its peak in 2000 when all the internet firms were making huge profits. But in the same year this bubble burst and the index fell afterwards for a long time. The historical density is estimated from the returns of this time series. We analyze the market utility functions in March 2000, July 2002 and June 2004 in order to consider different market regimes. We interpret 2000 as a bullish, 2002 as a bearish and 2004 as a sideways market. These interpretations are based on table 2 that describes the changes of the DAX over the preceding 1 or 2 years. (In June 2004 the market went up by 11% in the last 10 months.)

A utility function derived from the market data is a market utility function. It is estimated as an aggregate for all investors as if the representative investor existed. A representative investor is however just a convenient con-

struction because the existence of the market itself implies that the asset is bought and sold, i.e. at least two counterparties are required for each transaction.

In section 2 we identified the market utility function (up to linear transformations) as

$$U(R) = \int_{R_0}^R K(x)dx$$

where K is the pricing kernel for returns. It is defined by

$$K(x) = q(x)/p(x)$$

in terms of the historical and risk neutral densities p and q of returns. Any utility function (both cardinal and ordinal) can be defined up to a linear transformation, therefore we have identified the utility functions sufficiently. In section 3.3 we proposed different models for estimating the historical density. In figure 6 we show the pricing kernels resulting from the different estimation approaches for the historical density. The figure shows that all three kernels are quite similar: They have the same form, the same characteristic features like e.g. the hump and differ in absolute terms only a little. This demonstrates the economic equivalence of the three estimation methods on this day and this equivalence holds also for the other days. In the following we work with historical densities that are estimated by the observed returns.

Besides the pricing kernel and the utility function we consider also the risk attitudes in the markets. Such risk attitudes are often described in terms of relative risk aversion that is defined by

$$RRA(R) = -R \frac{U''(R)}{U'(R)}.$$

Because of $U' = cK = cq/p$ for a constant c the relative risk aversion is also given by

$$RRA(R) = -R \frac{q'(R)p(R) - q(R)p'(R)}{p^2(R)} \bigg/ \frac{q(R)}{p(R)} = R \left(\frac{p'(R)}{p(R)} - \frac{q'(R)}{q(R)} \right).$$

Hence, we can estimate the relative risk aversion from the estimated historical and risk neutral densities.

In figure 7 we present the empirical pricing kernels in March 2000, July 2002 and June 2004. The dates represent a bullish, a bearish and a sideways markets, see table 2. All pricing kernels have a proclaimed hump located

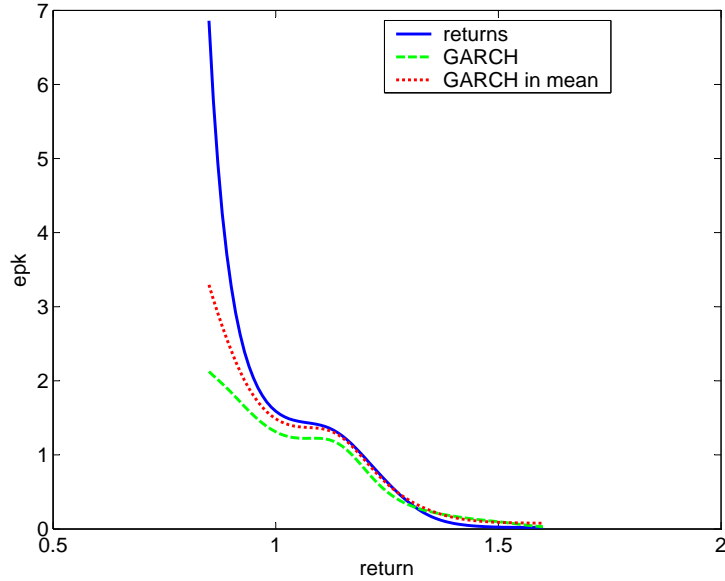


Figure 6: Empirical pricing kernel on 24/03/2000.

at small profits. Hence, the market utility functions do not correspond to standard specification of utility functions. We present the pricing kernels only in regions around the initial DAX (corresponding to a return of 1) value because the kernels explode outside these regions. This explosive behaviour reflects the typical pricing kernel form for losses. The explosion of the kernel for large profits is due to numerical problems in the estimation of the very low densities in this region. But we can see that in the sideways market the kernel is concentrated on a small region while the bullish and bearish markets have wider pricing kernels. The hump of the sideways market is also narrower than in the other two regimes. The bullish and bearish regimes have kernels of similar width but the bearish kernel is shifted to the loss region and the bullish kernel is located mainly in the profit area. Moreover, the figures show that the kernel is steeper in the sideways markets than in the other markets. But this steepness cannot be interpreted clearly because pricing kernels are only defined up to a multiplicative constant.

The pricing kernels are the link between the relative risk aversion and the utility functions that are presented in figure 8. These utility functions are only defined up to linear transformations, see section 2. All the utility functions are increasing but only the utility function of the bullish market is concave. This concavity can be seen from the monotonicity of the kernel, see figure 7. Actually, this non convexity can be attributed to the quite special

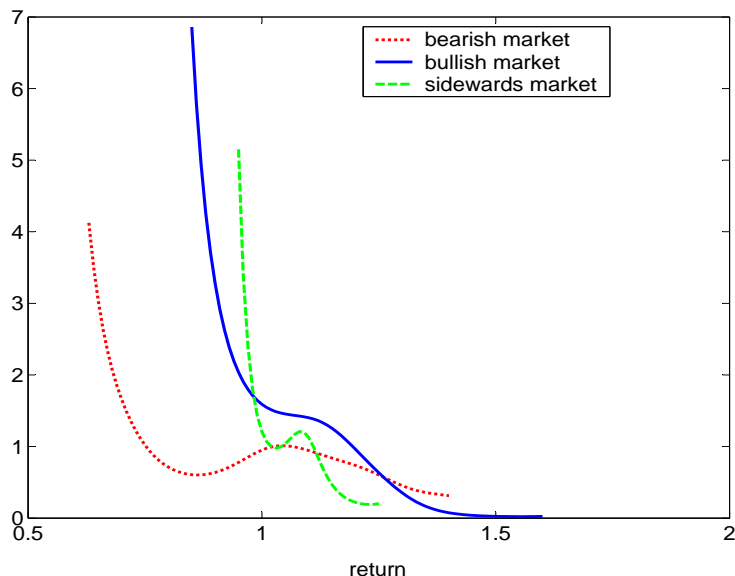


Figure 7: Empirical pricing kernel on 24/03/2000, 30/07/2002 and 30/06/2004.

form of the historical density which has two modes on this date, see figure 4. Hence, we presume that also this utility function has in general a region of convexity. The other two utility functions are convex in a region of small profits where the bullish utility is almost convex. The derivatives of the utility functions cannot be compared directly because utility functions are identified only up to multiplicative constants. But we can compare the ratio of the derivatives in the loss and profit regions for the three dates because the constants cancel in these ratios. We see that the derivatives in the loss region are highest in the bullish and lowest in the bearish market and vice versa in the profit region. Economically these observations can be interpreted in such a way that in the bullish market a loss (of 1 unit) reduces the utility stronger than in the bearish market. On the other hand, a gain (of 1 unit) increases the utility less than in the bearish market. The sideways market shows a behaviour between these extreme markets. Hence, investors fear in a good market situation losses more than in a bad situation and they appreciate profits in a good situation less than in a bad situation.

Finally, we consider the relative risk aversions in the three market regimes. These risk aversions are presented in figure 9, they do not depend on any constants but are completely identified. We see that the risk aversion is smallest in all markets for a small profit that roughly corresponds to the

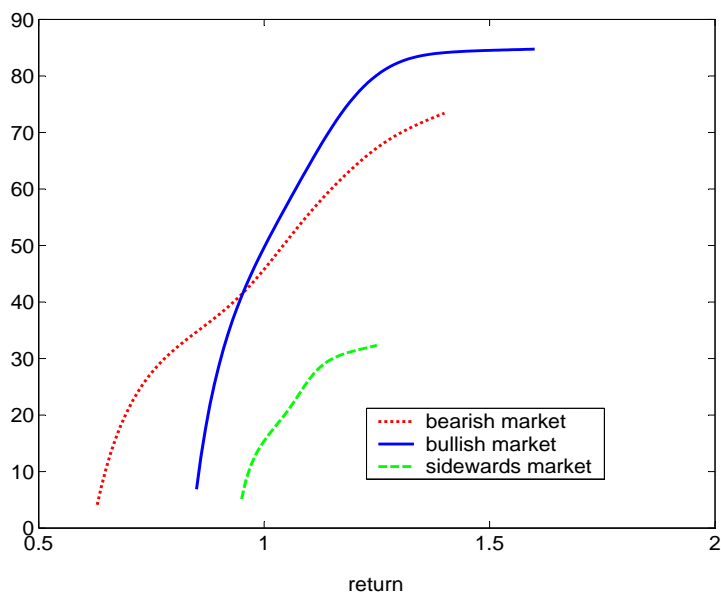


Figure 8: Market utility functions on 24/03/2000, 30/07/2002 and 30/06/2004.

initial value plus a riskless interest on it. In the sideways regime the market is risk seeking in a small region around this minimal risk aversion. But then the risk aversion increases quite fast. Hence, the representative agent in this market is willing to take small risks but is sensitive to large losses or profits. In the bullish and bearish regimes the representative agent is less sensitive to large losses or profits than in the sideways market. In the bearish situation the representative agent is willing to take more risks than in the bullish regime. In the bearish regime the investors are risk seeking in a wider region than in the sideways regime. In this sense they are more risk seeking in the bearish market. In the bullish market – on the other hand – the investors are never risk seeking so that they are less risk seeking than in the sideways market.

The estimated utility functions most closely follow the specification proposed by Friedman & Savage (1948). The utility function proposed by Kahneman & Tversky (1979) consists of one concave and one convex segment and is less suitable for describing the observed behaviour, see figure 10. Both utility functions were proposed to account for two opposite types of behaviour with respect to risk attitudes: buying insurance and gambling. Any utility function that is strictly concave fails to describe both risk attitudes. Most notable examples are the quadratic utility function with the linear pricing

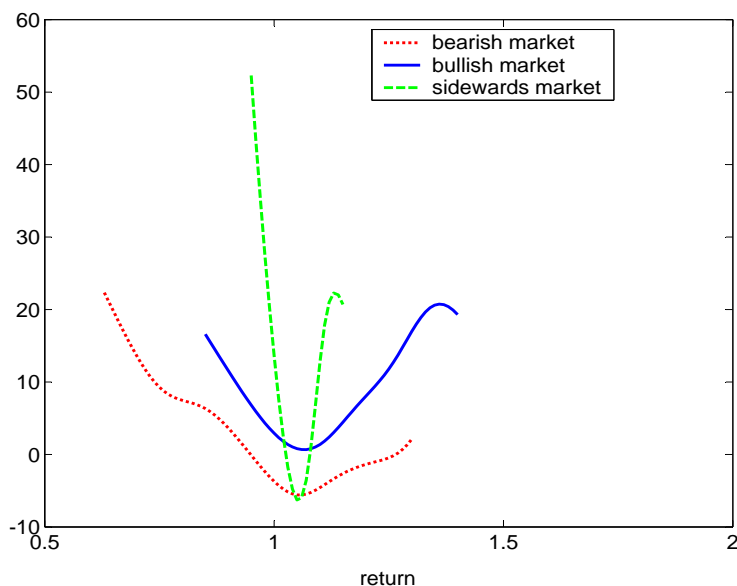


Figure 9: Relative risk aversions on 24/03/2000, 30/07/2002 and 30/06/2004.

kernel as in the CAPM model and the CRRA utility function. These functions are presented in figure 10. Comparing this theoretical figure with the empirical results in figure 7 we see clearly the shortcoming of the standard specifications of utility functions to capture the characteristic hump of the pricing kernels.

4 Individual investors and their utility functions

In this section, we introduce a type of utility function that has two regions of different risk aversion. Then we describe how individual investors can be aggregated to a representative agent that has the market utility function. Finally, we solve the resulting estimation problem by discretization and estimate the distribution of individual investors.

4.1 Individual Utility Function

We learn from figures 10 and 7 that the market utility differs significantly from the standard specification of utility functions. Moreover, we can observe from the estimated utility functions 8 that the loss part and the profit part

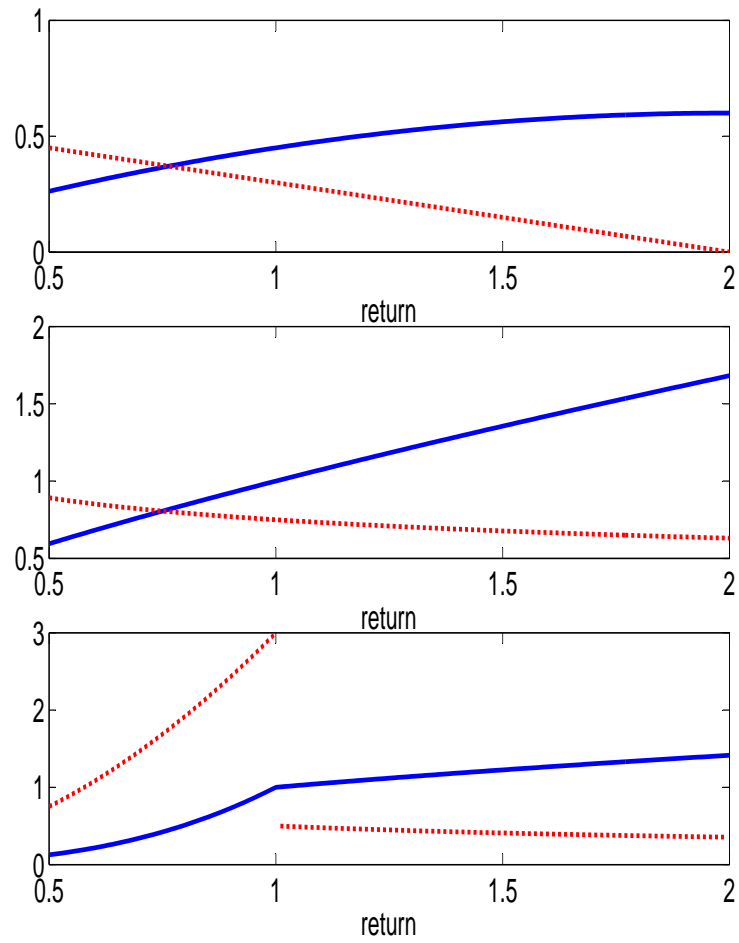


Figure 10: Common utility functions (solid) and their pricing kernels (dotted) (upper: quadratic, middle: power, lower panel: Kahneman and Tversky utility function).

of the utility functions can be quite well approximated with shifted CRRA functions, $k = 1, 2$:

$$U^{(k)}(R) = a_k \frac{(R - c_k)^{1-\gamma_k}}{1 - \gamma_k} + b_k,$$

where the shift parameter is c_k . These power utility functions become infinitely negative for $R = c_k$ and can be extended by $U^{(k)}(R) = -\infty$ for $R \leq c_k$, i.e. investors will avoid by all means the situation when $R \leq c_k$. The standard CRRA utility function has $c_k = 0$.

We try to reconstruct the market utility of the representative investor by individual utility functions and hence assume that there are many investors on the market. Investor i will be attributed with a utility function that consists of two CRRA functions:

$$U_i(R) = \begin{cases} \max \{U(R, \theta_1, c_1); U(R, \theta_2, c_{2,i})\}, & \text{if } R > c_1 \\ -\infty, & \text{if } R \leq c_1 \end{cases}$$

where $U(R, \theta, c) = a \frac{(R-c)^{1-\gamma}}{1-\gamma} + b$, $\theta = (a, b, \gamma)^\top$, $c_{2,i} > c_1$. If $a_1 = a_2 = 1$, $b_1 = b_2 = 0$ and $c_1 = c_2 = 0$, we get the standard CRRA utility function.

The parameters θ_1 and θ_2 and c_1 are the same for all investors who differ only with the shift parameter c_2 . θ_1 and c_1 are estimated from the lower part of the utility market function, where all investors probably agree that the market is “bad”. θ_2 is estimated from the upper part of the utility function where all investors agree that the state of the world is “good”. The distribution of c_2 uniquely defines the distribution of switching points and is computed in section 4.3. In this way a bear part $U_{bear}(R) = U(R, \theta_1, c_1)$ and a bull part $U_{bull}(R) = U(R, \theta_1, c_2)$ can be estimated by least squares.

The individual utility function can then be denoted conveniently as:

$$U_i(R) = \begin{cases} \max \{U_{bear}(R); U_{bull}(R, c_i)\}, & \text{if } R > c_1; \\ -\infty, & \text{if } R \leq c_1. \end{cases} \quad (5)$$

Switching between U_{bear} and U_{bull} happens at the *switching point* z , whereas $U_{bear}(z) = U_{bull}(z, c_i)$. The switching point is uniquely determined by $c_i \equiv c_{2,i}$. The notations *bear* and *bull* have been chosen because U_{bear} is activated when returns are low and U_{bull} when returns are high.

Each investor is characterised by a switching point z . The smoothness of the market utility function is the result of the aggregation of different attitudes. U_{bear} characterizes more cautious attitudes when returns are low and U_{bull} describes the attitudes when the market is booming. Both U_{bear}

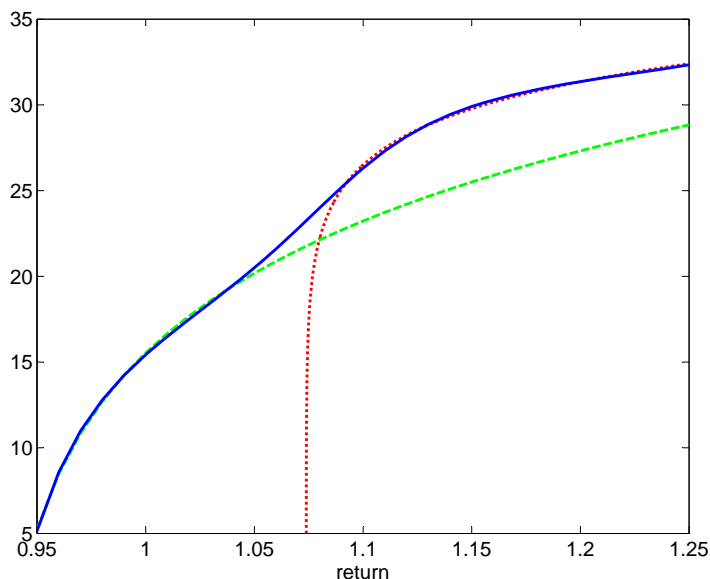


Figure 11: Market utility function (solid) with bearish (dashed) and bullish (dotted) part of an individual utility function 5 estimated in the sideways market of 30/06/2004.

and U_{bull} are concave. However, due to switching the total utility function can be locally convex.

These utility functions are illustrated in figure 11 that shows the results for the sideways market. We observe/estimate the market utility function that does not correspond to standard utility approaches because of the convex region. We propose to reconstruct this phenomenon by individual utility functions that consist of a bearish part and a bullish part. While the bearish part is fixed for all investors the bullish part starts at the switching point that characterizes an individual investor. By aggregating investors with different switching points we reconstruct the market utility function. We describe the aggregation in section 4.2 and estimate the distribution of switching points in section 4.3. In this way we explain the special form of the observed market utility functions.

4.2 Market Aggregation Mechanism

We consider the problem of aggregating individual utility functions to a representative market utility function. A simple approach to this problem is to identify the market utility function with an average of the individual utility functions. To this end one needs to specify the *observable* states of the world

in the future by returns R and then find a weighted average of the utility functions for each state. If the importance of the investors is the same, then the weights are equal:

$$U(R) = \frac{1}{N} \sum_{i=1}^N U_i(R),$$

where N is the number of investors. The problem that arises in this case is that utility functions of different investors can not be summed up since they are incomparable.

Therefore, we propose an alternative aggregation technique. First we specify the *subjective* states of the world given by utility levels u and then aggregate the outlooks concerning the returns in the future R for each perceived state. For a *subjective* state described with the utility level U , such that

$$u = U_1(R_1) = U_2(R_2) = \dots = U_N(R_N)$$

the aggregate estimate of the resulting returns is

$$R_A(u) = \frac{1}{N} \sum_{i=1}^N U_i^{-1}(u) \tag{6}$$

if all investors have the same market power. The market utility function U_M resulting from this aggregation is given by the inverse R_A^{-1} .

In contrast to the naive approach described at the beginning of this section, this aggregation mechanism is consistent under transformations: if all individual utility functions are changed by the same transformation then the resulting market utility is also given by the transformation of the original aggregated utility. We consider the individual utility functions U_i and the resulting aggregate U_M . In addition, we consider the transformed individual utility functions $U_i^\phi(x) = \phi\{U_i(x)\}$ and the corresponding aggregate U_M^ϕ where ϕ is a transformation. Then the aggregation is consistent in the sense that $U_M^\phi = \phi(U_M)$. This property can be seen from

$$\begin{aligned} (U_M^\phi)^{-1}(u) &= \frac{1}{N} \sum_{i=1}^N (U_i^\phi)^{-1}(u) \\ &= \frac{1}{N} \sum_{i=1}^N U_i^{-1}\{\phi^{-1}(u)\} \\ &= U_M^{-1}\{\phi^{-1}(u)\} \end{aligned}$$

The naive aggregation is not consistent in the above sense as the following example shows: We consider the two individual utility functions $U_1(x) = \sqrt{x}$

and $U_2(x) = \sqrt{x}/2$ under the logarithmic transformation $\phi = \log$. Then the naively aggregated utility is given by $U_M(x) = 3\sqrt{x}/4$. Hence, the transformed aggregated utility is $\phi\{U_M(x)\} = \log(3/4) + \log(x)/2$. But the aggregate of the transformed individual utility functions is

$$\begin{aligned} U_M^\phi(x) &= \frac{1}{2} \{ \log(\sqrt{x}) + \log(\sqrt{x}/2) \} \\ &= \frac{1}{2} \log\left(\frac{1}{2}\right) + \log(x)/2. \end{aligned}$$

This implies that $U_M^\phi \neq \phi(U_M)$ in general.

This described aggregation approach can be generalized in two ways: If the individual investors have different market power then we use the corresponding weights w_i in the aggregation (6) instead of the uniform weights. As the number of market participants is in general big and unknown it is better to use a continuous density f instead of the discrete distributions given by the weights w_i . These generalizations lead to the following aggregation

$$R_A(u) = \int U^{-1}(\cdot, z)(u) f(z) dz$$

where $U(\cdot, z)$ is the utility function of investor z . We assume in the following that the investors have utility function of the form described in section 4.1. In the next section we estimate the distribution of the investors who are parametrized by z .

4.3 The Estimation of the Distribution of Switching Points

Using the described aggregation procedure, we consider now the problem of replicating the market utility by aggregating individual utility functions. To this end, we choose the parametric utility functions $U(\cdot, z)$ described in 4.1 and try to recover with them the market utility U_M . We do not consider directly the utility functions but minimize instead the distance between the inverse functions:

$$\min_f \left\| \int U^{-1}(\cdot, z) f(z) dz - U_M^{-1} \right\|_{L^2(\tilde{P})} \quad (7)$$

where \tilde{P} is image measure of the historical measure P on the returns under the transformation U_M . As the historical measure has the density p the

transformation theorem for densities implies that \tilde{P} has the density

$$\tilde{p}(u) = p\{U_M^{-1}(u)\}/U'_M\{U_M^{-1}(u)\}.$$

With this density the functional to be minimized in problem (7) can be stated as

$$\begin{aligned} & \int \left(\int U^{-1}(u, z) f(z) dz - U_M^{-1}(u) \right)^2 \tilde{p}(u) du \\ &= \int \left(\int U^{-1}(u, z) f(z) dz - U_M^{-1}(u) \right)^2 p\{U_M^{-1}(u)\}/U'_M\{U_M^{-1}(u)\} du \\ &= \int \left(\int U^{-1}(u, z) f(z) dz - U_M^{-1}(u) \right)^2 p\{U_M^{-1}(u)\}(U_M^{-1})'(u) du \end{aligned}$$

because the derivative of the inverse is given by $(g^{-1})'(y) = 1/g'\{g^{-1}(y)\}$. Moreover, we can apply integration by substitution to simplify this expression further

$$\begin{aligned} & \int \left(\int U^{-1}(u, z) f(z) dz - U_M^{-1}(u) \right)^2 p\{U_M^{-1}(u)\}(U_M^{-1})'(u) du \\ &= \int \left(\int U^{-1}\{U_M(x), z\} f(z) dz - x \right)^2 p(x) dx. \end{aligned}$$

For replicating the market utility by minimizing (7) we observe first that we have samples of the historical distribution with density p . Hence, we can replace the outer integral by the empirical expectation and the minimization problem can be restated as

$$\min_f \frac{1}{n} \sum_{i=1}^n \left(\int g\{U_M(x_i), z\} f(z) dz - x_i \right)^2$$

where x_1, \dots, x_n are the samples from the historical distribution and $g = U^{-1}$.

Replacing the density f by a histogram $f(z) = \sum_{j=1}^J \theta_j I_{B_j}(z)$ with bins B_j the problem is transformed into

$$\min_{\theta_j} \frac{1}{n} \sum_{i=1}^n \left\{ \sum_{j=1}^J \tilde{g}(i, j) \theta_j - x_i \right\}^2$$

where $\tilde{g}(i, j) = \int_{B_j} g\{U_M(x_i), z\} dz$.

Hence, the distribution of switching points can be estimated by solving the quadratic optimization problem

$$\min_{\theta_j} \frac{1}{n} \sum_{i=1}^n \left\{ \sum_{j=1}^J \tilde{g}(i, j) \theta_j - x_i \right\}^2, \\ \text{s.t.} \quad \theta_j \geq 0, \\ \sum_{j=1}^J \theta_j = 1.$$

Such quadratic optimization problems are well known and their solutions can be obtained using standard techniques, see e.g. Mehrotra (1992) or Wright (1998).

We present in figure 12 the estimated distribution of switching points of the sideways market. For this picture we have replaced the density f by a histogram with bins of width 0.01. We checked the robustness of our results by considering also smaller and bigger bandwidths. Smaller bandwidths led to rather ragged densities and bigger bandwidths masked too much structure. The density presented in figure 12 shows that the distribution is concentrated in a narrow region. Moreover, it is left-skewed. As the switching points lie in the area of 5% to 10% returns the investors do not seem to use the riskless return as reference. Rather they use a higher returns that could correspond to the expected return of the DAX. The long right tail of the distribution means that there are some relatively pessimistic investors that switch to the bullish regime only for high returns. The switching points are all located to the right from the riskless return, hence all investors want in a sideways situation more than only positive returns in order to show bullish attitudes. The corresponding distributions of the bullish and bearish markets of 24/03/2000 and 30/07/2002 show similar characteristics only the region where the densities are concentrated are bigger.

5 Conclusion

We have analyzed in this paper empirical pricing kernels in three market regimes using data on the German stock index and options on this index. In a bullish, a bearish and a neutral market regime we estimate the pricing kernel and derive the corresponding utility functions and relative risk aversions.

In the sideways market of June 2004, the market investor is risk seeking in a small region around the riskless return but risk aversion increases fast for

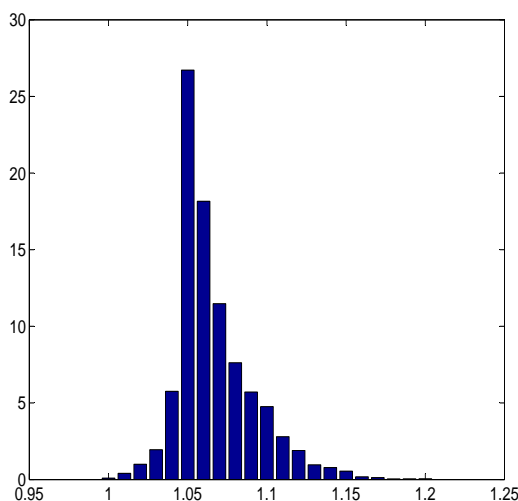


Figure 12: The density of the distribution of the reference points in the sideways market of 24/03/2000.

high absolute returns. In the bullish market of March 2000, the investor is on the other hand never risk seeking while he becomes more risk seeking in the bearish market of July 2002. Before the stock market crash in 1987 European options did not show the smile and the Black-Scholes model captured the data quite well. Hence, utility functions could be estimated at that times by power utility functions with a constant positive risk aversion. Our analysis shows that this simple structure does not hold anymore and discusses different structures corresponding to different market regimes.

The empirical pricing kernels of all market regimes demonstrate that the corresponding utility functions do not correspond to standard specifications of utility functions including Kahneman and Tversky (1979). The observed utility functions are closest to the general utility functions of Friedman and Savage (1948). We propose a parametric specification of these functions, estimate it and explain the observed market utility function by aggregating individual utility functions. In this way, we can estimate a distribution of individual investors.

The proposed aggregation mechanism is based on homogeneous investors in the sense that they differ only with switching points. Future research can reveal how nonlinear aggregation procedures could be applied to heterogeneous investors.

6 Acknowledgements

The research work of R. A. Moro was supported by the German Academic Exchange Service (DAAD). K. Detlefsen was supported by Bankhaus Sal. Oppenheim. This research was supported by Deutsche Forschungsgemeinschaft through the SFB 649 “Economic Risk”.

References

- Ait-Sahalia, Y. and A. Lo, 1998: Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, **53**(2).
- Ait-Sahalia, Y. and A. Lo, 2000: Nonparametric risk-management and implied risk aversion. *Journal of Econometrics*, **94**(9).
- Barone-Adesi, G., R. Engle, and L. Mancini, 2004: Garch options in incomplete markets. working paper, University of Lugano.
- Bergomi, L., 2005: Smile dynamics 2. *Risk*, **18**(10).
- Bernoulli, D., 1956: Exposition of a new theory on the measurement of risk. *Econometrica*, **22**, 23–36.
- Billingsley, P., 1995: *Probability and Measure*. Wiley-Interscience.
- Black, F. and M. Scholes, 1973: The pricing of options and corporate liabilities. *Journal of Political Economy*, **81**, 637–659.
- Breedon, D. and R. Litzenberger, 1978: Prices of state-contingent claims implicit in option prices. *Journal of business*, **51**, 621–651.
- Carr, P. and D. Madan, 1999: Option valuation using the fast fourier transform. *Journal of Computational Finance*, **2**, 61–73.
- Chernov, M., 2000: Essays in financial econometrics. Phd thesis, Pennsylvania State University.
- Chernov, M., 2003: Empirical reverse engineering of the pricing kernel. *Journal of Econometrics*, **116**, 329–364.
- Cizek, P., W. Härdle, and R. Weron, 2005: *Statistical Tools in Finance and Insurance*. Springer, Berlin.
- Cochrane, J., 2001: *Asset Pricing*. Princeton University Press.

- Cont, R., 2001: Empirical properties of asset returns: stylized facts and statistical issues. 223-349.
- Cont, R. and P. Tankov, 2004: Nonparametric calibration of jump-diffusion option pricing models. *Journal of Computational Finance*, **7**(3), 1–49.
- Dupire, B., 1994: Pricing with a smile. *Risk*, **7**, 327–343.
- Franke, J., W. Härdle, and C. Hafner, 2004: *Statistics of Financial Markets*. Springer Verlag, Berlin.
- Friedman, M. and L. P. Savage, 1948: The utility analysis of choices involving risk. *Journal of Political Economy*, **56**, 279–304.
- Harrison, M. and S. Pliska, 1981: Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, **11**, 215–260.
- Heston, S., 1993: A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, **6**(2), 327–343.
- Heston, S. and S. Nandi, 2000: A closed form GARCH option pricing model. *Review of Financial Studies*, **13**, 585–625.
- Jackwerth, J., 2000: Recovering risk aversion from option prices and realized returns. *Review of Financial Studies*, **13**(2), 433–451.
- Jackwerth, J. and M. Rubinstein, 1996: Recovering probability distributions from option prices. *Journal of Finance*, **51**(5), 1611–1631.
- Kahneman, D. and A. Tversky, 1979: Prospect theory: An analysis of decision under risk. *Econometrica*, **47**, 263–291.
- Mehrotra, S., 1992: On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, **2**(4), 575–601.
- Merton, R. C., 1973: An intertemporal capital asset pricing model. *Econometrica*, **41**(5), 867–887.
- Rosenberg, J. and R. Engle, 2002: Empirical pricing kernels. *Journal of Financial Economics*, **64**(7), 341–372.
- Rubinstein, M., 1994: Implied binomial trees. *Journal of Finance*, **69**, 771–818.

- Silverman, B., 1986: *Density Estimation*. Chapman and Hall, London.
- Storn, R. and K. Price, 1997: Differential evolution - a simple and efficient heuristic for global optimization over continuous space. *Journal of Global Optimization*, **11**, 341–359.
- von Neumann, J. and O. Morgenstern, 1944: *The Theory of Games and Economic Behavior*. Princeton University Press.
- Wright, S., 1998: Primal-dual interior-point methods. *Mathematics of Computation*, **67**(222), 867–870.