Estimating the Pricing Kernel using Butterfly Spreads

J. Koëter (538829)

September 16, 2016

Abstract

I develop a framework to estimate the pricing kernel from butterfly spreads each month. In this thesis, I estimate the pricing kernel on basis of S&P 500 butterflies. In order to estimate the pricing kernel, the return distribution of the S&P 500 is assumed. It turns out that the expected payoff is closest to the realized payoff of the butterflies when I model the returns with a skewed t-distribution with volatility a linear combination of the VIX index. The time-series average of the pricing kernel slopes downward on the return interval [0.95; 1.05], which is similar to the theoretical pricing kernel of a representative agent model with CRRA preferences.

1 Introduction

In this thesis, I estimate the pricing kernel using butterfly spreads and compare it to pricing kernels produced by different representative agent models. A butterfly is a portfolio of three different put or call options, how a butterfly is constructed in detail is explained in Section 2. Different from most papers in the existing literature, I estimate the pricing kernel each month and do not consider a pricing kernel specification. Monthly estimation gives me insights into the general shape of the pricing kernel, but also into differences in the time-series. The pricing kernel is not constant over time and there could be periods that investors pay more for instruments which payout when the stock market performs poorly. If the attitude of the investors changes towards different instruments it is reflected in the level and shape of the pricing kernel. The monthly estimation of the pricing kernel can confirm whether investors have time dependent attitudes towards risk and how it behaves over time.

Butterfly spreads offer a way of identifying the pricing kernel on a narrow grid, as Aït-Sahalia and Lo (2000) show, if the interval on which the payoff of the butterfly is non-zero converges to zero, the instrument converges to an Arrow-Debreu security. Given this property, the preferences of the investors are identified on a narrow grid using butterfly spreads. For example loss aversion can be identified comparing two options. Instrument 1 has a payoff if equity return is more than the reference return (i.e. reference return from Tversky and Kahneman (1992)) but less than 2% plus reference return, i.e. on interval $(r, r + 2\%)$. Instrument 2 hast the same payoff if the return on equity is less than the reference return but not less than the reference return minus 2%, interval $(r - 2\%, r)$. For both instruments the impact of probability weighting on the pricing kernel for these moderate equity returns is small[1] and approximately symmetric. Therefore, any differences in the relative prices of these two instruments is driven by loss aversion and is identified using these instruments.

I derive the pricing kernel from a system of equations based on the general pricing formula: $p = \mathbb{E}(mX)$. In the general asset pricing equation, p is the price of a certain instrument, m is the pricing kernel, X is the payoff of the instrument and the expectation $(\mathbb{E}(\cdot))$ is calculated using a certain probability density function $(f(\cdot))$. Butterfly prices are obtained from data on S&P 500 options from OptionMetrics. To solve the pricing equation for the stochastic discount factor (m), the expected payoff of the butterfly has to be calculated. In order to calculate the expected payoff the distribution of the S&P 500 is assumed. For simplicity, I solve the system for a finite number of discrete states and the probabilities of these states are given by the integral over an interval around the particular state.

In order to estimate the pricing kernel from the butterfly data, the underlying distribution of the S&P 500 is assumed. I start working with the assumption that S&P 500 returns follow a log-normal distribution with time-dependent volatility. As in Boyer and Vorkink (2014), the volatility of the log-normal distribution is equal to the volatility of the returns of the past half year (125 trading days). Different from the Boyer and

[1] Probability weighting influences prices when an instrument has a positive payoff in the tail(s) of the return distribution.

Vorkink (2014) approach is that their aim is to model individual stock returns, whereas I model index returns. It turns out that the log-normal distribution is not able to model the S&P 500 return distribution well, therefore I choose the skewed t-distribution with volatility conditional on implied volatility (VIX) as a benchmark model which does a better job.

The time-series average of the monthly estimated pricing kernel, using the benchmark model for S&P 500 returns, is downward sloping over the return interval [0.95; 1.05]. Only the representative agent model, of the models considered in this thesis, with pure Constant Relative Risk Aversion (CRRA) preferences is able to predict a similar pricing kernel shape. The other models considered in this thesis either predict a flat pricing kernel or predict that the right tail of the pricing kernel is upward sloping. As mentioned, to estimate the pricing kernel the distribution of S&P 500 returns is assumed, it turns out that the skewed t-distribution with volatility a linear combination of VIX is best able to do so.

I compare the empirically estimated pricing kernel estimated to kernels derived from representative agent models. In this thesis, I consider in the thesis three different sets of preferences for the representative agent, namely: Cumulative Prospect Theory from (Tversky and Kahneman, 1992) (CPT), CRRA and risk neutral preferences. Each of the considered preferences have its own implications for the prices of butterfly spreads. Risk neutral pricing kernels are equal to one in each state, whereas the CRRA pricing kernel is a downward-sloping function with a curvature parameter equal to the risk aversion. As Baele et al. (2016) show, models with risk neutral or CRRA preferences are not able to reproduce the expected returns of call and put options empirically observed. Therefore, Baele et al. (2016) propose a new model with CPT preferences and show that this model is able to produce the empirically observed expected returns for both put and call options[2]. The CPT model incorporates probability weighting and loss aversion into the preferences of the investor.

[2]A model with CRRA preferences is only able to fit the expected returns of put or call options, not both

1.1 Literature Overview

Aït-Sahalia and Lo (2000) estimate the implied risk-aversion of a CRRA investor for the total return distribution of the S&P 500 index. It is shown in Aït-Sahalia and Lo (2000) that implied risk-aversion is not equal for each state of the economy. The implied risk-aversion is estimated using the physical return and the risk neutral return distribution of the S&P 500 index. The first is estimated nonparametrically using six-month returns and the latter is estimated semiparametrically using the second derivative with respect to the strike of the Black-Scholes call pricing function. Different from my approach is that Aït-Sahalia and Lo (2000) assume a functional form of the pricing kernel and only estimates the pricing kernel once on basis of the full sample.

Polkovnichenko and Zhao (2013) show that, empirically, pricing kernels are consistent with rank-dependent utility including inverse-S shaped probability weighting (as in Tversky and Kahneman (1992)). The probability weighting function is estimated each month from the physical and risk neutral distribution of the S&P 500 index. Similar to Aït-Sahalia and Lo (2000), the risk neutral distribution is estimated using the second derivatives of call options with respect to the strike. A simulation using the Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) (Nelson (1991)) is carried out to estimate the physical density. The estimation of the probability weighting function, assuming CRRA utility, is done each month. In my specification, no assumptions about the preferences of the investor are assumed to estimate the pricing kernel.

Chaudhuri and Schroder (2015) show that the pricing kernel is strictly monotonic and downward-sloping if and only if expected returns of the 'log-concave' class of option trading strategies are increasing in the strike price. This class includes long calls, puts, butterfly spreads and bullish call spreads, these are use to test if the pricing kernel slopes down monotonically. Chaudhuri and Schroder (2015) find evidence that the pricing kernel as function of S&P 500 returns is not monotonically downward sloping. Using this model-free test of pricing kernel monotonicity, Chaudhuri and Schroder (2015) are only able to say something about the general shape of the pricing kernel over the full sample. Whereas in my specification, I am able to consider the shape of the pricing kernel each month and investigate the dynamics over time.

Boyer and Vorkink (2014) find that stock option portfolios scoring higher on basis of their total skewness measure have, on average, lower returns. The fact that these portfolios with higher skewness have lower returns implies that these options have higher prices. Thus, in the paper evidence is presented that investors have a preference for skewness. In order to calculate the moments of the option portfolios, the distribution of the underlying stocks are assumed to follow a stationary log-normal distribution calibrated on daily data of 6 months prior to the formation date. The distribution of stock returns often tends to be right-skewed, therefore the right-skewed lognormal distribution is supposed to be a good fit. In this thesis, I use butterfly spreads on the S&P 500 to identify the preferences of the investors, in order to do so the distribution of the S&P 500 is assumed. Index returns are often supposed to be left-skewed and for this reason the log-normal would not be a good fit to describe the S&P 500 returns.

Baele et al. (2016) explain the empirical magnitude of the variance premium with a CPT representative investor model, which is the same model I use. Besides this contribution, they also show that the model is able to produce expected option returns similar to returns empirically observed. Empirically, for the same maturity, if a call option's strike increases[3] the expected return of the option declines. The fact that the expected return decreases when the strike increases is the same as saying that the relative price of the option increases. An increase in the relative price of the call option points in the direction that the tail in the good states of nature of the pricing kernel slopes upward, a finding which cannot be explained by CRRA or risk neutral preferences. The opposite holds for put options, empirically the expected return increases for options with higher strikes. Similarly, this increase in expected return points at a relative decrease of put option prices when the strike increases. Given that the relative price increases at the left tail of the return distribution (bad states of nature), it hints at a decreasing pricing kernel in the left tail. A model with a CRRA representative investor also produces a decreasing pricing kernel in the left tail, although the risk aversion of the investor has to be quite large to produce put return in the same order of magnitude as empirically observed. The expected return patterns for put and call options points can only be explained by a pricing kernel downward sloping for bad states of nature and upward sloping for good states of nature, also called U-shaped pricing kernels. For certain parameter values, a U-shaped pricing kernel is implied by CPT preferences, whereas CRRA and risk neutral

[3]Same as options with a higher moneyness level or options which are more out of the money.

preferences are not able to fit U-shaped kernels.

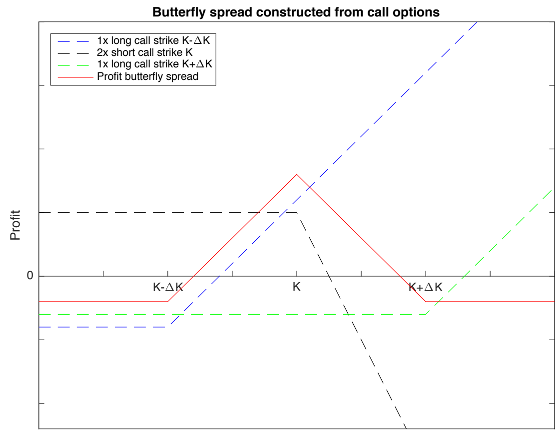
2 Data

In the thesis the pricing kernel solves from a system of equations using prices on butterfly spreads, which are portfolios consisting of three different call (put) options. The data on European options is obtained from OptionMetrics and options are selected with a maturity of three to six weeks. The price of an option is equal to the mid-point of the best closing bid and offer price. All options for which the offer is higher than the bid price are discarded. Similarly, options which do not satisfy the no arbitrage lower bound for European options are excluded from the analysis. These bounds are equal to $S - Ke^{-r\tau}$ for calls and $Ke^{-r\tau} - S$ for puts, where S is the price of the underlying, K is the strike price, r is the risk-free rate and τ are the days to maturity. The risk-free rate used to calculate the no arbitrage bounds is also obtained from OptionMetrics. As the sample includes most options with an maturity of 15 working days, this is chosen as maturity. In the next subsection I will show how a butterfly spread is constructed.

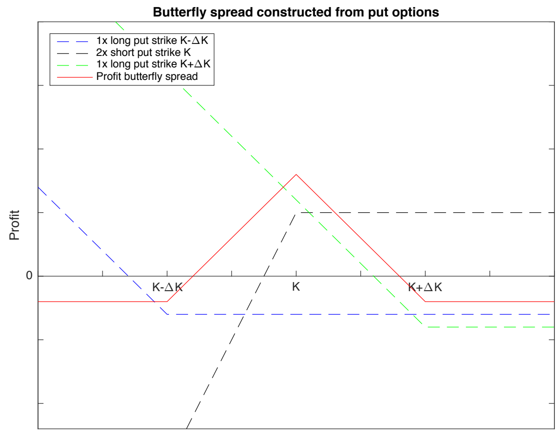
2.1 Butterfly spreads

Butterfly spreads are portfolios of options, to construct a butterfly three different call or put options are needed. In Figure 1 the construction of a butterfly spread with strike K and spread ΔK is explained.

Figure 1: In this figure two portfolios to construct a butterfly spread are presented. The x-axis represents the value of the underlying asset at maturity and the y-axis the profit (including prices of the options) at maturity. The left considers a butterfly constructed from three different call options, the dotted lines in the graph present the profit of the different positions held in these options. Call butterfly spreads (red line) are constructed by a long position in call options with strikes $K - \Delta K$ (blue line) and $K + \Delta K$ (green line) and a short position in two calls with strike K. The put butterfly spread is constructed in the same way, only put instead of call options.



(a) Call butterfly spread.



(b) Put butterfly spread.

As seen in Figure 1(a) a call butterfly spread is constructed by a long position in call options with strikes $K - \Delta K$ (blue line) and $K + \Delta K$ (green line) and a short position in two calls with strike K. The profit of the butterfly spread is given by in the red line in the graph. A put butterfly spread is constructed in exactly the same way as the call butterfly. In the data, the smallest thick size between two consecutive options is 5 index points. The smaller the spread of the butterfly, the better the pricing kernel is identified at the specific point, therefore ΔK is chosen to be equal to 5.

From Figures 1(a) and (b) it is clear that a butterfly, either from put or call options, should have the same price as the payoff is identical. If the prices are not identical an arbitrage opportunity exists. In the data prices of put or call butterflies often differ, therefore if two identical butterflies exist on the same trading day, the price of the butterfly is replaced with the average of the two.

3 Methodology

In this section, I explain the methodology into more details to estimate the pricing kernel using butterfly spreads. From the data of OptionMetrics on S&P500 index options butterflies are constructed, it could be the case that the amount of butterflies is not sufficient to calculate the pricing kernel on a certain grid. If so, inter- and extrapolation techniques on expected returns of the butterflies are used to construct butterflies with different strikes. It is known from the first fundamental theory of asset pricing that when a market does not allow arbitrage and is complete, a unique positive pricing kernel exists. In order for a market to be complete, the number of linearly independent financial assets has to be equal to the amount of possible states of nature. As the returns on the S&P500 are assumed to follow an assumed continuous distribution, infinitely many possible returns (states of nature) exist. To have a complete the market, the same amount of butterflies has to be available. To avoid the problem of an incomplete market, the payoff space (and corresponding probabilities) is discretized. In case of discrete states, a finite number of butterflies is sufficient to determine the pricing kernel. If the amount of discrete states is larger than the amount of observed butterflies, missing butterflies are constructed using inter- and extrapolation of the expected returns of observed butterflies. The interpolation is based on the linear combination of the expected returns of the observed butterflies. The weights are chosen in such a way that the weighted sum of the strikes is equal to the desired strike. For the extrapolation, the expected return of the newly constructed butterfly is equal to the expected return of the closest observed butterfly.

The next section consists of an example on how the pricing kernel is estimated from the observed instruments. Furthermore, the example shows how the inter- and extrapolation works. Note, in the example payoffs of 5 are chosen to bridge between the example and the real problem.

3.1 Example

Consider an economy with an asset paying $S \in \{90, 95, 100, 105, 110\}$, with probabilities $q \in \{0.1, 0.2, 0.4, 0.2, 0.1\}$. Two financial instruments trade on the market which have the following prices (p) and payoffs (X):

$$\begin{bmatrix} p_2 \\ p_4 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix}, \qquad \begin{bmatrix} X_2 \\ X_4 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \end{bmatrix}$$

Given these payoffs and probabilities, the expected returns on the instruments are given by:

$$r_2 = \frac{\mathbb{E}(X_2)}{p_2} = \frac{0.2 \cdot 5}{\frac{4}{5}} = \frac{5}{4}, \qquad r_4 = \frac{5}{3}.$$

To determine the pricing kernel from these instruments, prices are needed for the same instruments paying 5 in the remainder of the states. Let me start with interpolation to get the price of the instrument paying 5 when $S = 100$. The strike is in the middle of the other two instruments, so it is assumed that the return is a linear combination (in this case with equal weights) of the two. Therefore, the interpolated instrument is constructed in the following way:

$$r_3' = \frac{1}{2}r_1 + \frac{1}{2}r_2 = 1.43, \qquad \Rightarrow \qquad p_3' = \frac{\mathbb{E}(X_3)}{r_3'} = \frac{0.4 \cdot 5}{1.43} = 1.40.$$

Note, the $'$ indicates that the price and return are constructed. In this way an instrument between two others is constructed using interpolation. Using extrapolation instruments paying at the tails of the state-distribution (with strikes 90 and 110) are constructed. It is assumed that the expected return on the extrapolated instruments is equal to the expected return of the closest instrument[4]. Using this assumption, the prices of the instruments can be constructed in the following way:

[4] In a later stage different extrapolation techniques will be investigated. Might be insightful given that in Prospect Theory the tails are interesting.

$$
\begin{aligned}
r_2 = r_1' = \frac{\mathbb{E}(X_1)}{p_1'} \qquad &\Rightarrow \qquad p_1' = \frac{0.1 \cdot 5}{\frac{5}{4}} = \frac{2}{5} \\
r_4 = r_5' = \frac{\mathbb{E}(X_5)}{p_5'} \qquad &\Rightarrow \qquad p_5' = \frac{0.1 \cdot 5}{\frac{5}{3}} = \frac{3}{10}.
\end{aligned}
$$

Given the fact that I work under the assumption of no-arbitrage and a complete market, implying a positive unique pricing kernel, the pricing kernel can be derived from the following system of equations:

$$
\begin{bmatrix} \frac{2}{5} \\ \frac{4}{5} \\ \frac{7}{5} \\ \frac{3}{5} \\ \frac{3}{10} \end{bmatrix} = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 & m_5 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad \Rightarrow \quad \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} \\ \frac{4}{5} \\ \frac{7}{10} \\ \frac{3}{5} \\ \frac{3}{5} \end{bmatrix}.
$$

Where the square matrix is the probability weighted payoff-space.

3.2 Analysis

The example is the general idea of how to derive the pricing kernel from the actual data on butterflies. Butterflies in the analysis have a maturity of 15 working days, spread (ΔK) of 5 index points. For each trading day with butterflies available, the pricing kernel can be derived on a specified domain. In principle it is possible to determine the pricing kernel on a domain as large as wanted by extrapolation observed butterflies. For now, the estimation domain of the pricing kernel is equal to smallest($\underline{b}$) to largest($\bar{b}$) moneyness level of the butterflies in the sample on a certain trading day. In order to estimate the pricing kernel, it is needed to discretize the continuous payoff-space. Therefore, the states on which the pricing kernel is considered are equal to the lower till upper bound of the domain in steps of $0.01S_0$. Basically, the payoff-space is discretized using steps of 1% in return of the underlying on the considered trading day. To have a complete market the number of states (N) has to be equal to the amount of financial instruments, in this case butterflies with ΔK of 5 index points. As it is not possible for butterflies to have payoffs outside the payoff-space, N butterflies are considered with strikes equally distributed between $\underline{b}+5$ and $\bar{b}-5$. The amount of observed butterflies is usually not sufficient to estimate the pricing kernel in a complete market setting. Therefore, the remaining butterflies spreads are constructed using interpolation.

Using these constructed butterflies the pricing kernel can be derived, only the matrix with the probability weighted payoffs are needed for each butterfly. The derivation of this matrix is best illustrated using an example: Consider an economy with N states $S \in \{S_1, \cdots, S_N\}$, with S continuously distributed with probability density function $f(S)$ on domain $(0, \infty)$. Suppose that $dS = S_i - S_{i-1} = S_j - S_{j-1} \quad \forall i, j$. Let the payoff of a butterfly, $P(S, K)$, with strike K be a continuous function of S. For this economy, an entry of the probability weighted payoff space is given by:

$$s_{i,j} = \int_{S_j - \frac{1}{2}dS}^{S_j + \frac{1}{2}dS} P\left(x, K_i\right) f(x) dx. \tag{1}$$

Using equation (1), the total matrix can be calculated to derive the pricing kernel.

3.3 Distributional assumption

In order to estimate the pricing kernel, a return distribution for the underlying asset, in my case the S&P 500, is assumed. In this section, I explore the implications of the S&P 500 return distribution and explain which distribution fits the data best. I present estimators for three distributions to model the S&P 500 returns. As a starting point, the normal and log-normal distribution are considered and in a later stage a skewed t-distribution. Note, all distributions considered in this thesis have time-varying volatility and it is explained how it is modeled in this section. The distributions are estimated on S&P 500 returns from 1990 to 2015, for the log-normal and normal I use daily returns. Whereas for the skewed t-distribution, returns over 15 working days are used.

3.3.1 Normal and Log-normal distribution

In this section, I present the estimators of the parameters of the normal and log-normal distribution. The returns are modeled with time-varying volatility and, as in Boyer and Vorkink (2014), the volatility is estimated with the standard deviation of the returns in the past half year (125 working days). Furthermore, the average of the normal distribution is estimated with the sample average over the total sample. In sum:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t^s, \qquad \hat{\sigma_t}^2 = \frac{1}{124}\sum_{i=t-125}^{t-1} \left(r_i^s - \bar{r}^s\right)^2. \tag{2}$$

Assuming a normal distribution for returns, the estimators depend on simple returns (indicated by the 's' in r_t^s).

Similarly, the estimators for the log-normal distribution are given by:

$$\hat{\mu}_t = \frac{1}{T}\sum_{i=1}^{T} r_i^l, \qquad \hat{\sigma_t}^2 = \frac{1}{124}\sum_{i=t-125}^{t-1} \left(r_i^l - \bar{r}^l\right)^2. \tag{3}$$

Assuming a log-normal distribution for returns, the estimators use log returns returns (indicated by the 'l' in r_t^l).

3.3.2 Skewed t-distribution

I introduce the Skewed t-distribution and report later the Maximum Likelihood Estimates of the parameters of the distribution. The background information about the skewed t-distribution is taken from Baele et al. (2016), Lambert and Laurent (2001) and Bauwens and Laurent (2002). Different from these papers, I consider the distribution with time-varying volatility. Skewed t-distributions support left-skew and heavy tailed distributions, which are often ought to be needed in order to be a good fit for index returns. As highlighted before, the standard deviation of returns is not constant over time and the changes of this variable has a significant impact on prices of butterflies. I assume that the returns of the S&P500 follow a skewed t-distribution and the standard deviation is a linear function of the implied volatility index (VIX index). The VIX index is equal to the yearly implied volatility of S&P 500 options, for interpretation the yearly volatility (250 days) is scaled to 15 day volatility.

The assumed return distribution uses a standardized skewed t-distribution. The return on the S&P 500 follows the process:

$$r_t^s = \mu + \epsilon_t \tag{4}$$

$$\sigma_t = \alpha + \beta VIX_t \tag{5}$$

$$\epsilon_t = \sigma_t \varsigma_t, \tag{6}$$

where r_t^s are simple 15-day returns and the standard deviation is a linear function of the VIX at time t. Moreover, the random variable ς_t is $SKST(0, 1, \xi, v)$; it follows a standardized skewed t-distribution with parameters $v > 2$ (degrees of freedom) and $\xi > 0$ (skewness parameter). The density is given by:

$$f(\varsigma_t|\xi, v) = \begin{cases} \frac{2}{\xi + \frac{1}{\xi}} sg\left[\xi(s\varsigma_t + m)|v\right] & \text{if} \quad \varsigma_t < -\frac{m}{s} \\ \frac{2}{\xi + \frac{1}{\xi}} sg\left[\frac{1}{\xi}(s\varsigma_t + m)|v\right] & \quad\quad \varsigma_t \geq -\frac{m}{s}, \end{cases} \tag{7}$$

where $g(\cdot|v)$ is a symmetric (zero mean and unit variance) Student t-distribution with v degrees of freedom, defined by:

$$g(x|v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi(v-2)}\Gamma\left(\frac{v}{2}\right)} \left[1 + \frac{x^2}{v-2}\right]^{-(v+1)/2}, \tag{8}$$

where $\Gamma(\cdot)$ is Euler's gamma function. The constants $m = m(\xi, v)$ and $s = \sqrt{s^2(\xi, v)}$ are the mean and standard deviation of the non-standardized skewed t-distribution, $SKST(m, s^2, \xi, v)$, and given by:

$$m(\xi, v) = \frac{\Gamma\left(\frac{v-1}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\Gamma\left(\frac{v}{2}\right)} \left(\xi - \frac{1}{\xi}\right), \tag{9}$$

$$s^2(\xi, v) = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2. \tag{10}$$

As mentioned before, the skewed t-distribution supports some important stylized facts of index returns. Namely, left-skew and fat tailed distributions. The ξ is a skewness related parameters and when $\xi < 1$, the distribution is left skewed, for $\xi > 1$ the distribution is right skewed and for $\xi = 1$ the distribution is not skewed. Fat tails are related to the parameter of the degrees of freedom v, the lower the parameter the fatter the tails and when $v \to \infty$ the distribution converges to a skewed normal distribution.

I estimate the parameters of equations (4)-(10) using Maximum Likelihood on 15 working day simple returns of the S&P 500 from 1990 to 2015. The estimates are represented in Table 1.

Table 1: Maximum Likelihood Estimates of the parameters of the skewed t-distribution with volatility a linear combination of the VIX. The parameters correspond to the equations of (4)-(6). These parameters are estimated on basis on 15-day simple returns of the S&P 500 from 1990-2015.

μ	α	β	ξ	v
0.0043^{***}	-0.0054	0.8712^{***}	0.7384^{***}	16.9797^{*}

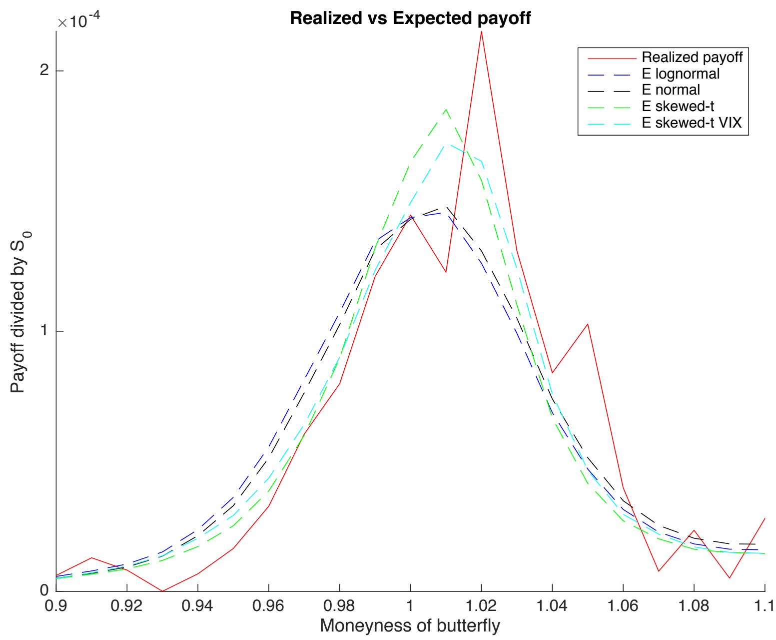
The μ of Table 1 is approximately equal to the sample average of the return sample. As the estimate of β is significant at a 1% level, it supports the hypothesis that volatility varies over time and the VIX is a good predictor of future volatility. Given the estimates of α and β, it can be concluded that the overall level of the VIX is higher than the realized volatility of the S&P 500 index, which Baele et al. (2016) discuss as well. The estimate of ξ is significantly smaller than 1 and it thus indicates that the return distribution is left-skewed. Furthermore, some evidence of fat tails is found as the estimate of the degrees of freedom v is not too large. The model using the estimates of Table 1 will be referred to as the benchmark model throughout the thesis.

3.3.3 Implications of distributional differences

In Figure 2 the expected and realized payoff of butterflies are presented. The expected value of the butterfly spreads is calculated with respect to the distributions of the S&P 500 return discussed in the previous section. All expected and realized payoffs are normalized with the price of the S&P 500 at the time of calculation. Basically the graphs in Figure 2 represent the time-series average of the expected payoff, under different return distributions, per S&P 500 index point is calculated.

The blue line is the time-series average of the expected payoff of the butterflies with S&P 500 return log normally distributed. As visible in Figure 2, the expected and realized payoff is largest around a moneyness level of 1.00 (at the money). This makes sense as the maturity of the butterflies for which the payoffs are calculated is 3 weeks (or 15 working days). In 3 weeks the S&P500 is unlikely to move a lot and, therefore, butterflies at the money are most likely to have positive payoff at maturity. To justify if the log-normal distribution is a good fit for calculating the expected payoff of a butterfly, the realized payoff of the same butterflies is presented in the same figure (red line). Although the fit seems reasonable, there is still some room for improvement. Interestingly, for a horizon of 15 trading days the distributions of the log-normal and normal are approximately the same, therefore the estimated pricing kernels under these assumptions are approximately equal.

Figure 2: In the figure the time-series averages, over the total sample, of expected and realized payoff standardized with S_0 are presented. Time-series averages for four different distributions are presented in the figure. The average of the expected payoff, over the total sample, is equal to the time-series average of the expected value of the butterflies at a certain moneyness level. Realized and expected payoffs are normalized with the current stock level in order to measure the payoff per index point. The blue line corresponds to log-normal distributed returns, the black line to normal distributed returns and the light-blue and green correspond to a skewed t-distribution specifications. The light-blue is the benchmark model for S&P 500 returns with the volatility as a linear function of VIX. Whereas the green corresponds with a similar model, only the parameters of the same model as in equations (4)-(6) is estimated with constant volatility over the total sample. When calculating the expected payoff of a certain butterfly, the volatility is replaced by the volatility of the returns of the past half year.



Similar as in the (log-)normal case, the expected payoff highest when the butterflies are at the money when returns are modeled with the benchmark model. It seems that the skewed t-distribution with conditional volatility fits the realized payoffs best. This is in line with expectation as this is the only model which accounts for a left-skew and kurtosis in the return distribution of the S&P 500. The green line in Figure 2 is equal to a skewed t-distribution estimated with unconditional volatility, but replacing the volatility parameter with the standard deviation of the past half year. From Figure 2 it visible that the resulting distributions are quite close, therefore the results are robust to both underlying distributions.

4 Results

In this section, I show results of the general shape of the pricing kernel and investigate the impact of differences in the return distribution assumption. As explained in section 3, I want to inter- and extrapolate on basis of the expected return of the butterfly spread. Inter- and extrapolation on basis of expected returns is not possible for prices smaller or equal to zero (arbitrage prices). In order to exclude the arbitrage in the butterfly spreads, all butterflies with prices smaller or equal to zero should be discarded. By excluding butterflies which have a negative or price of zero, a significant pricing error is made. Given the amount of arbitrage prices observed in the sample, it is unlikely that investors were able to trade these butterflies, a more likely explanation would be that the option database contains substantial noise in the prices.

4.1 Pricing error

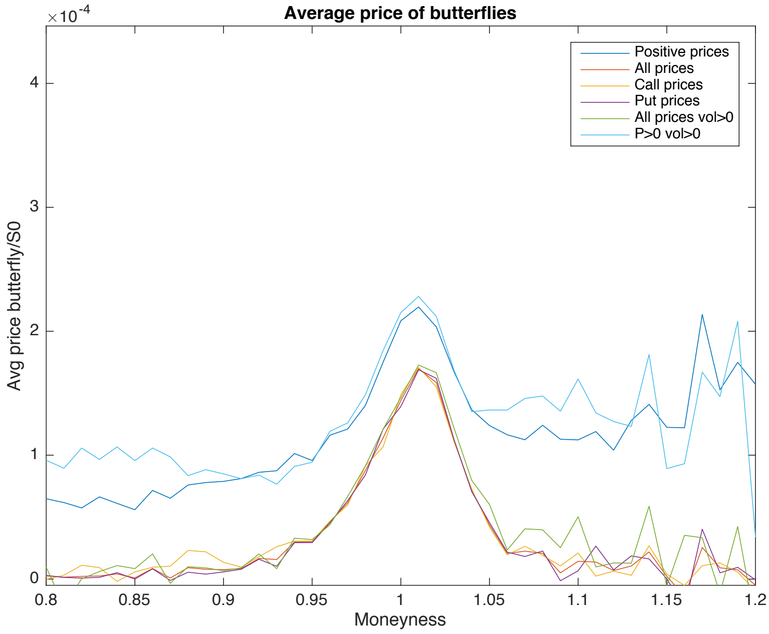
In this section some insights are given into the pricing error when excluding arbitrage prices. Lets assume that the arbitrage prices are a result of noise in the option database. If the assumption is correct, I observe the price of the butterfly in the following way:

$$\hat{p_{i,t}} = p_{i,t} + \epsilon_{i,t}. \tag{11}$$

The observed price ($\hat{p_{i,t}}$) is equal to the actual price of the butterfly ($p_{i,t}$) plus noise ($\epsilon_{i,t}$), for each moneyness level i at time t. If I work under the assumption that the time-series average of the noise equals zero, the expected value of the observed price is equal to the actual price. The observed price can only become negative for a substantial negative value of noise. If the prices with substantial negative noise are excluded from the sample, automatically the time-series average of the noise is positive. When the time-series average of the noise is positive, the time-series average of the observed price is greater than the average of the actual price. The pricing error when excluding arbitrage prices is illustrated with Figure 3.

In Figure 3 the two blue lines are the average butterfly prices when discarding

Figure 3: Time-series averages of the price per index point of S&P 500. Butterfly prices are averaged over time, for a certain moneyness level. The blue lines (light and dark) represent average butterfly prices excluding arbitrage prices. The average prices of all the butterflies is represented with the red line, the other lines are averages of just call or put butterflies or butterflies based on options with positive volume.



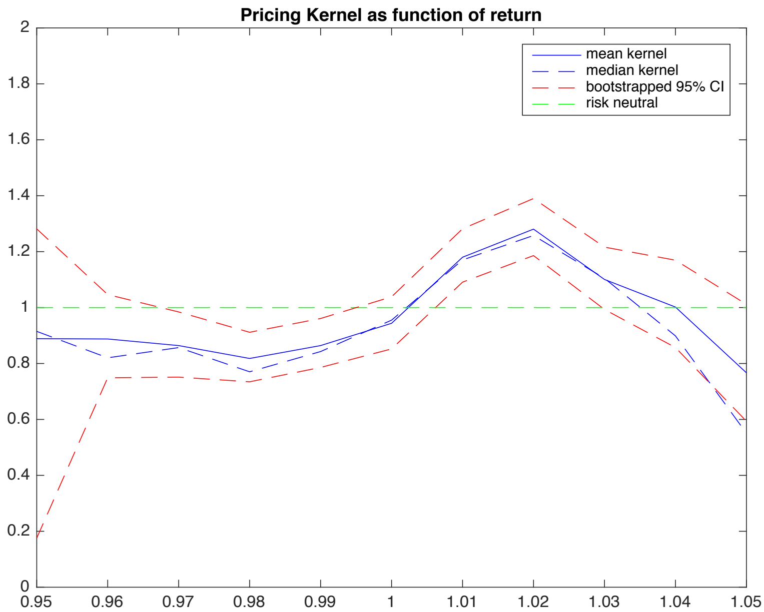
arbitrage prices. It is visible that for each moneyness level there is a positive pricing error and the error becomes larger towards the tails of the distribution, when comparing to the average price of the total sample (green line). Therefore, discarding the arbitrage prices leads to a significant increase in average prices of the butterflies. If prices are observed with positive noise in the data, also the pricing kernel is estimated with a positive bias. Especially in the tails, as the moneyness level decreases (or increases) the price of the butterfly remains approximately equal, but the expected payoff of the butterfly falls leading to extreme positive values of the pricing kernel. Due to this pricing error it is not possible to inter- and extrapolate on basis of expected returns, because this approach only makes sense when prices are positive. Therefore, I adjust the methodology in such a way that I only interpolate butterflies on basis of prices and not returns. Extrapolating prices of butterflies towards the tails does not make sense and it biases the estimation of the pricing kernel in the tails heavily.

The light-blue line is the time-series average of excluding arbitrage prices constructed with options which have a positive volume on the formation date. Only using options which have a positive volume on a certain trading day does not solve the positive pricing bias as it is fairly close to the line including all options (dark-blue). The remainder of the lines are the butterflies based on call options, put options, put and call options and on options with positive volume on the formation date. As these lines are all fairly close to each other it is concluded that the difference in input options is not reflected on the average prices of butterfly spreads for the full distribution of moneyness. In Figure 3 it is also seen that the prices of the butterflies primarily positive in the moneyness region from 0.95 to 1.05 and therefore is the main region of interest to look at the pricing kernel.

4.2 Pricing Kernel under log-normal

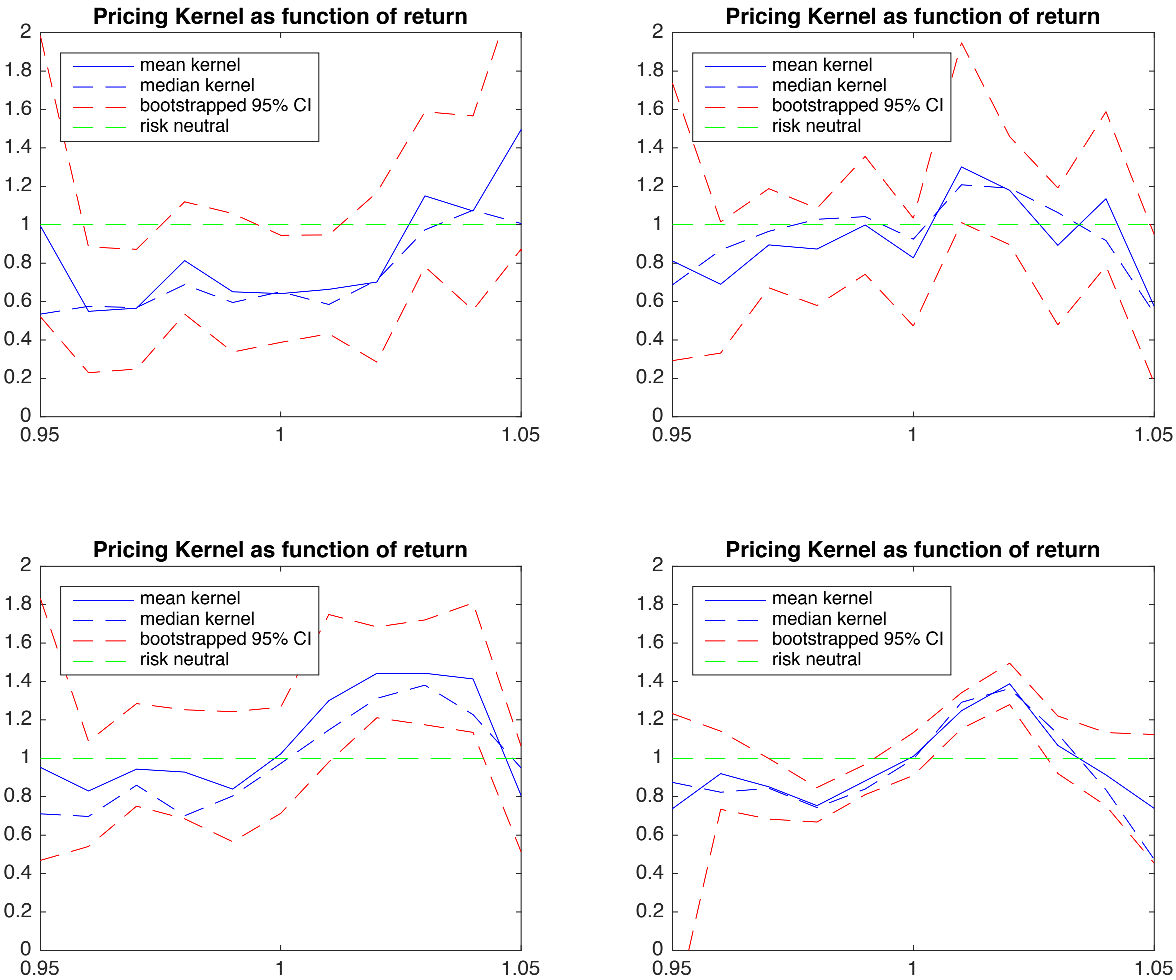
In this subsection, I estimate the pricing kernel using the log-normal distribution to model the S&P 500 returns. Due to the significant positive pricing error when discarding arbitrage prices, the pricing kernel will be estimated with the same bias. For this reason the analysis is performed on basis of all butterfly prices. As explained before, when negative and prices of zero are present in the data it is not possible to interpolate the expected returns of butterflies. Therefore the results make use of interpolation with respect to prices of the butterflies. As explained in Section 2, if on a certain trading day two butterflies exist which should have the same price, the prices are replaced by the mean of the two. In Figure 4 the time-series average, over the total sample, of the pricing kernel is plotted as a function of return. The solid (dotted) blue line represents the (median) average over time of the pricing kernel. As these lines are fairly close to each other, it is concluded that the distribution of stochastic discount factors over time is symmetric. Risk neutral preferences are represented in the figure by the green dotted line. The observed pricing kernel appears to be lower than 1 for outside the money butterflies and higher than 1 for in the money butterflies and towards the tail it declines again. This could either be explained by the fact that the expected payoff for butterflies outside the money is too high, which can be seen in Figure 2 as the expected is higher than the realized payoff. If this is the case, then it is simply a distributional problem, i.e. the log-normal distribution does not reflect the S&P 500 return distribution well. Alternatively, the result is obtained from the preferences of the investor, implying that it is willing to pay less for instruments paying in bad states of nature. The last explanation contradicts most of the theoretical asset pricing models and, therefore, the explanation of the distributional problem seems more likely.

Figure 4: In the figure the time-series average (solid) and median (dotted) Pricing Kernel are plotted in blue. The green line represents the Pricing Kernel of the risk neutral investor. A bootstrapped 95% confidence interval of the time-series average are represented with the dotted red lines. The bootstrap is based on a 1,000 samples.



In Figure 4 the pricing kernel is calculated using data on the total sample. Interestingly, it is possible to calculate the pricing kernel each month and to see deviations on a monthly basis. As there is a trade-off between noise and sample splits, I choose to split the sample in quarters. The pricing kernels based on approximately 5 years of data are represented in Figure 5.

Figure 5: In these figures the pricing kernel is calculated based on five years of data. The upper left figure pricing kernel is calculated using data from approximately 1996 till 2000, upper right 2001 till 2005, bottom left 2006 till 2010 and bottom right 2011 till 2015. In these figures the time-series averages (solid) and medians (dotted) Pricing Kernel are plotted in blue. The green line represents the Pricing Kernel of the risk neutral investor. A bootstrapped 95% confidence interval of the time-series average are represented with the dotted red lines. The bootstrap is based on a 1,000 samples.

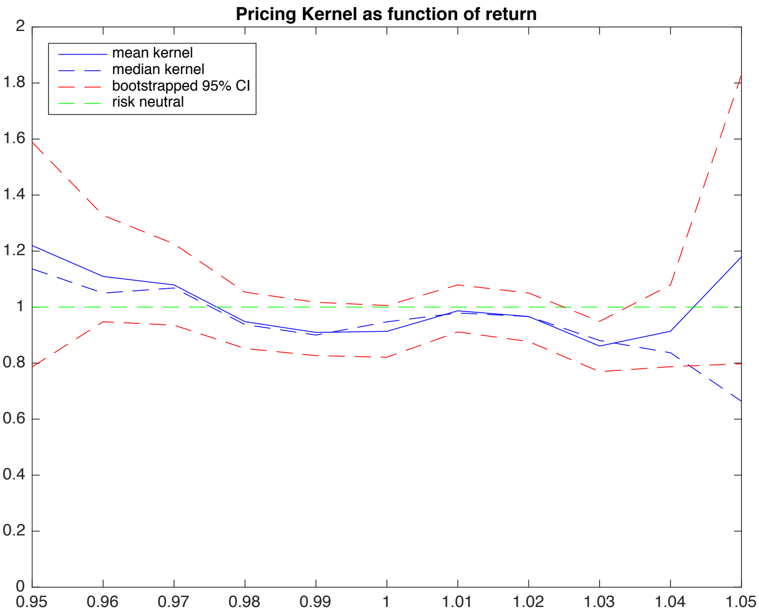


Especially the shape of the pricing kernel based on the oldest data differs from the other three. In all of the others the hump-shape around 1.02 is visible and therefore it seems that the results are robust to these sample periods. Furthermore, the first three kernels are more noisy than the latter which is caused by the fact that more options became available around 2012, since the amount of exercise dates each month increased to four. The first graph, which contains the oldest data, shows a U-shaped pricing kernel which could either imply that the representative investor distorts probabilities in the tails and this effect vanishes over time. Or it could mean that the little data that is available of this period contains significant noise, the second explanation seems more likely given the amount of data in the sample.

4.3 Pricing Kernel under skewed t-distribution

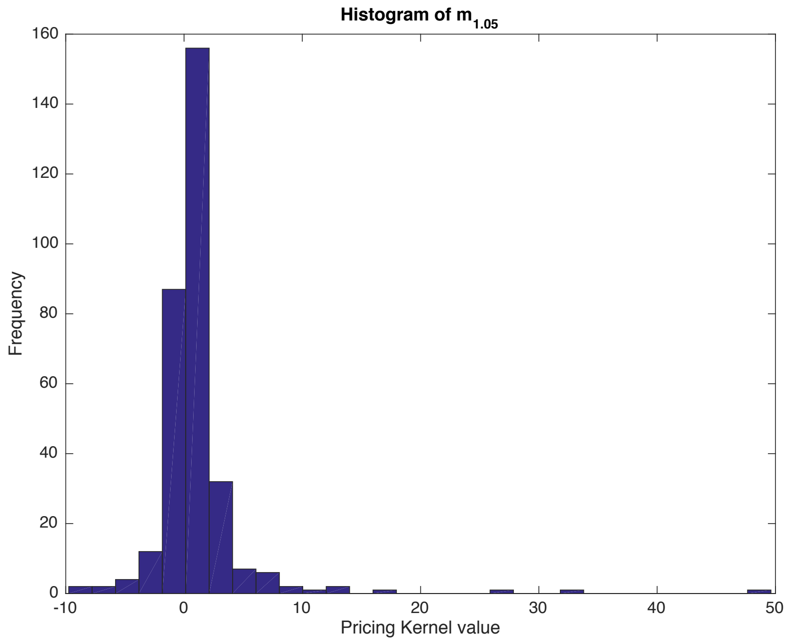
In this section the same results are presented as in the previous one, only here I assume that the S&P500 returns are skewed t-distributed with volatility conditional on VIX. To derive the pricing kernel, the estimates of section 3.3.2 are used. The average pricing kernel derived from the full sample is given in Figure 6.

Figure 6: In the figure the time-series average (solid) and median (dotted) Pricing Kernel are plotted in blue. The green line represents the Pricing Kernel of the risk neutral investor. A bootstrapped 95% confidence interval of the time-series average are represented with the dotted red lines. The bootstrap is based on a 1,000 samples.



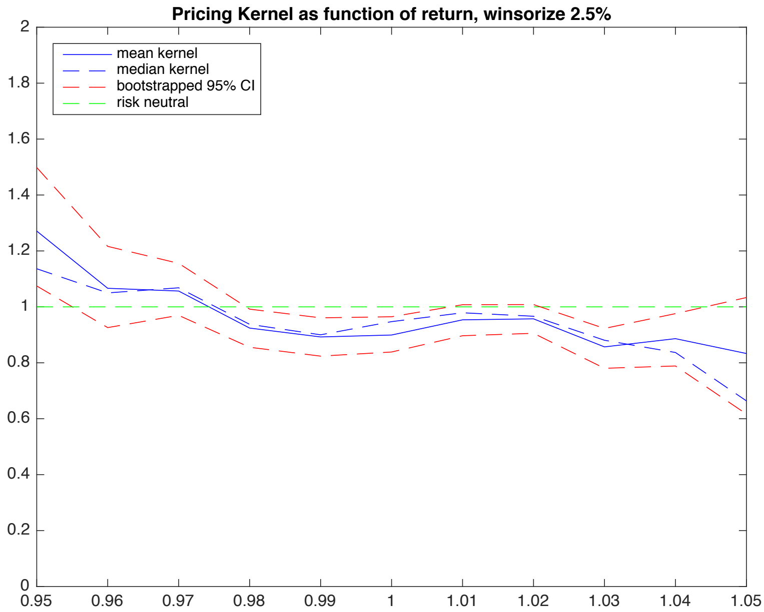
In Figure 6 the pricing kernel is larger than 1 around 0.98, which is both predicted by CRRA preferences, in case of sufficient risk aversion, and by CPT preferences due to loss aversion and/or probability weighting. Interestingly, in the right tail of Figure 6 the average sharply increases pointing at probability weighting, as the median continues the decline it is possibly driven by positive outliers. The deviation of the average from the median is most extreme around the return of 1.05. Note, in the left tail a small deviation of the average from the median is observed as well. It makes sense that in the tails the largest outliers exist, if the price of the butterfly is observed with noise the pricing kernel is amplified a lot as the expected payoffs of the butterflies are relatively small in the tails. From the fact that the average is larger than the median, it is concluded that the positive outliers are larger than the negative outliers. Consequently, it appears that the noise in the observed price is, on average, positive for moneyness levels of butterflies in the tails of the distribution. To investigate the degree to which the noise in the observed price affects the pricing kernel distribution, the distribution of the pricing kernel at a return level of 1.05 is presented. Figure 7 represents the distribution of the pricing kernel at 1.05.

Figure 7: The distribution of the pricing kernel values is represented in this figure, i.e. the values of the pricing kernel each month. As visible there are some severe positive outliers which influence the average significantly.



From Figure 7 it can be concluded that severe outliers are present in the distribution of pricing kernel values each month. Pricing kernel values are usually distributed around one, given the small expected values of the butterflies around moneyness of 1.05, small deviations due to noise in the observed price can lead to extreme pricing kernel values. Nevertheless, pricing kernel values of 50 are unreasonably high. It seems that the distribution of pricing kernel values is not symmetrically affected by the noise in the observed prices of the butterflies. For this reason it makes sense to winsorize the distribution above and below a certain quantile. After winsorizing the pricing kernel value distribution for each return level at 2.5%, the analogue to Figure 6 looks like Figure 8.

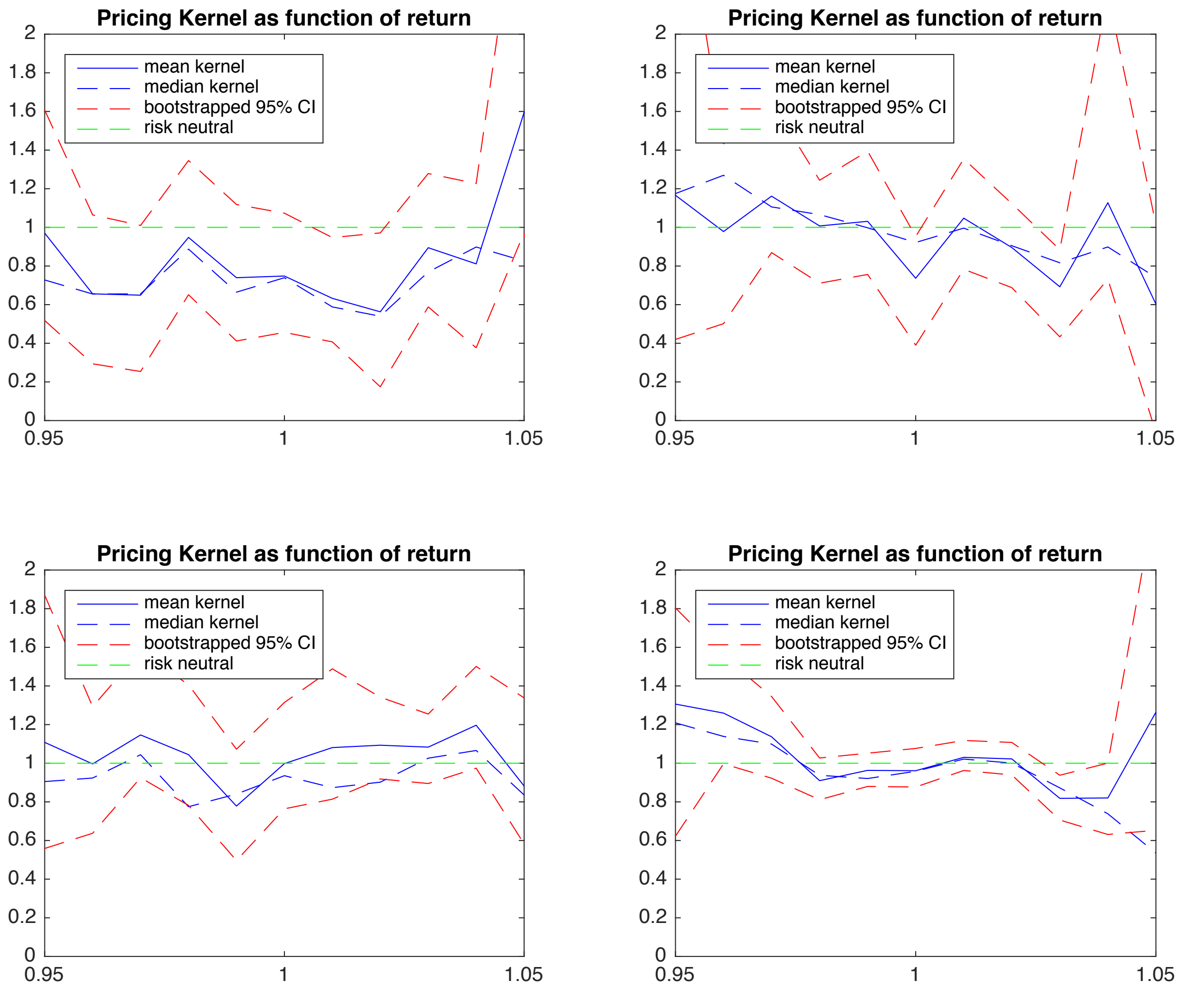
Figure 8: In the figure the time-series average (solid) and median (dotted) Pricing Kernel are plotted in blue. The green line represents the Pricing Kernel of the risk neutral investor. A bootstrapped 95% confidence interval of the time-series average are represented with the dotted red lines. The bootstrap is based on a 1,000 samples. Different from the previous figures, the distribution of the pricing kernel is winsorized at 2.5% for each return level.



In Figure 8 the median of the pricing kernel is in the bootstrapped confidence interval for each return value, indicating that winsorizing the distribution of the pricing kernel makes it more symmetric. The distribution of the pricing kernel is winsorized at the 2.5% upper and lower quantile for each return level. As the distance between the median and the mean of the pricing kernel has become smaller at 1.05 return, I conclude that positive outliers where driving the mean upwards. Over the interval $[0.95; 1.05]$ the general shape of the pricing kernel is downward sloping, which is theoretically implied by a representative agent model with CRRA utility. Interestingly, on average no evidence of an upward sloping pricing kernel is found in Figure 8 for the best states of nature. It could be due to the fact that investors price butterfly spreads with a downwards sloping pricing kernel. Alternatively, the noise in the observed prices could be interfering with the analysis, i.e. small price deviations lead to extreme pricing kernel values for low probability events. In the data, prices of butterflies with high moneyness levels seem to suffer more from noise as the prices are more often smaller or equal to zero. To what extend the noise is interfering with the results is an interesting research avenue for the future. In future research, I could use butterflies with $\Delta K = 10$ instead of 5. The use of these butterflies is interesting because the noise will have a smaller impact on the pricing kernel given that the expected value of these butterflies is larger.

As before, the sample is now split into quarters to check the robustness of the result.

Figure 9: In these figures the pricing kernel is calculated based on five years of data. The upper left figure pricing kernel is calculated using data from approximately 1996 till 2000, upper right 2001 till 2005, bottom left 2006 till 2010 and bottom right 2011 till 2015. In these figures the time-series averages (solid) and medians (dotted) Pricing Kernel are plotted in blue. The green line represents the Pricing Kernel of the risk neutral investor. A bootstrapped 95% confidence interval of the time-series average are represented with the dotted red lines. The bootstrap is based on a 1,000 samples.

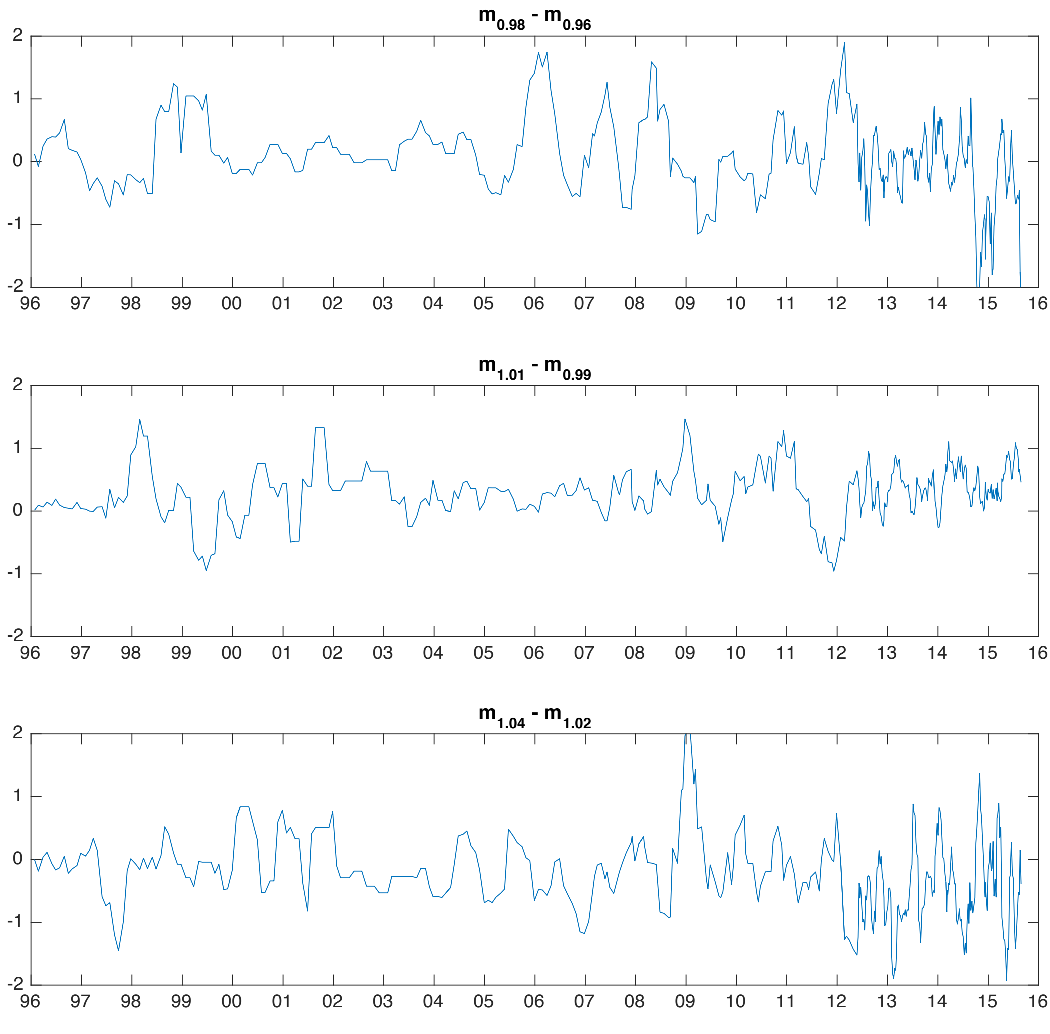


The pricing kernel based on the oldest data (1996-2000), upper left figure, appears to differ the most with respect to the others. Also when the S&P 500 return distribution is modeled with a log-normal distribution this is the case, as seen in Figure 5. For the other three sample periods, the general shape of the pricing kernel is (monotonically) downward sloping. Except for the pricing kernel values in the most recent sample period (2011-2015), bottom right, here the average pricing kernel is not downward sloping around 1.05. As explained above, the increase in average pricing kernel and deviation from the median clearly hints at positive outliers driving this result.

4.4 Slopes over time

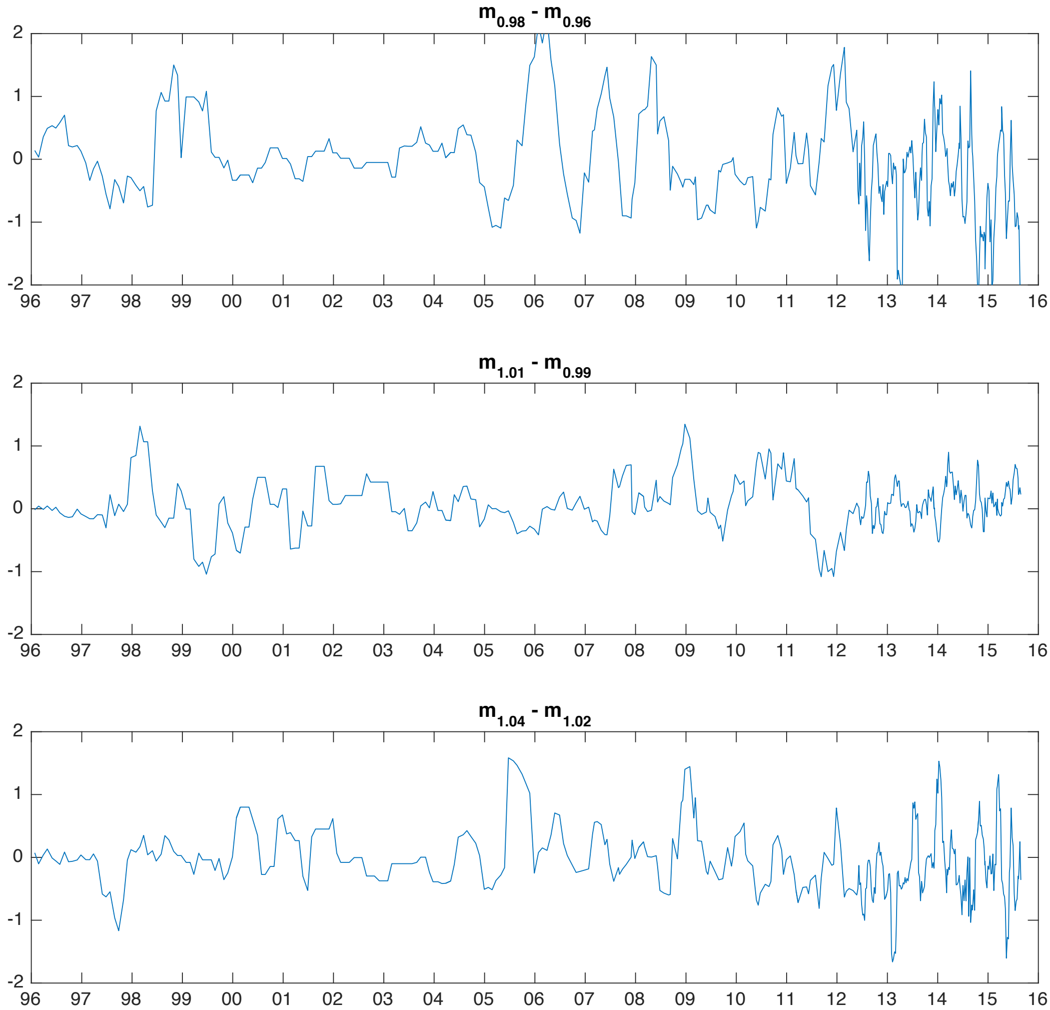
In this section I will show plots of the slopes of the pricing kernel over time. Basically, I calculate the difference between points of the pricing kernel, i.e. 0.99 and 1.01 and observe the sign and magnitude of the slope. This is done at different moneyness levels to observe potential differences. The analysis carried-out is basically a finite difference of the form $m(r_2) - m(r_1)$, where r_1, r_2 are returns of which the pricing kernel is a function $m(\cdot)$. Now I present the differences of the pricing kernel at different moneyness levels over time. Note, the lines represent a six-month moving average to account for some noise.

Figure 10: In this figure the six-month rolling average three finite differences of pricing kernel values are presented. The pricing kernel is estimated on basis of S&P 500 returns modeled with the log-normal distribution. The first graph shows the result for $m(0.98) - m(0.96)$, the second $m(1.01) - m(0.99)$ and the third $m(1.04) - m(1.02)$.



These represent the differences of the pricing kernel points computed using the log normal distribution as underlying. As of the first of June 2012 there are 4 exercise dates each month instead of 1 (that's why it seems more noisy). Below the same graphs in case of the skewed t-distribution.

Figure 11: In this figure the six-month rolling average three finite differences of pricing kernel values are presented. The pricing kernel is estimated on basis of S&P 500 returns modeled with the benchmark model. The first graph shows the result for $m(0.98) - m(0.96)$, the second $m(1.01) - m(0.99)$ and the third $m(1.04) - m(1.02)$.



In sum, little can be concluded from these graphs except for the fact that the pricing kernel estimated each month includes significant noise. It is interesting to investigate the slopes and changes of the slopes over time into more details, to see whether risk attitudes change over time and especially look at crises periods. Polkovnichenko and Zhao (2013) look at estimates of the slope of the pricing kernel as well and find evidence in favor of U-shaped pricing kernels. The estimation of the slope of the pricing kernel is done in a different way. Polkovnichenko and Zhao (2013) estimate the slope with the ratio of the area under the physical pricing kernel to the area under the risk-neutral pricing kernel

5 Theoretical Pricing Kernel

In this section I present a static representative agent model, which is used to derive the theoretically predicted pricing kernel of Figure 12. The model is the same as the model in Baele et al. (2016) and contains Cumulative Prospect Theory (CPT) from Tversky and Kahneman (1992). In CPT small probabilities of very good or bad states are distorted with an inverse S-shaped weighting function. Effectively, the representative investor overweights these small probabilities of a very good/bad state to occur. The representative investor suffers from loss aversion, i.e. the utility loss is larger than the utility gain for the same financial gain or loss. Where the loss aversion influences the utility is dependent on a reference return, the representative investor is loss averse with respect to states which have a lower return than the reference return. I choose the reference return to be equal to the risk-free rate as in Baele et al. (2016).

5.1 Preferences of the investor

The preferences of the representative agent are given by the total utility function, $\Psi\left(W_T, X_T\right)$, which has two separable components. The components of the utility function are defined as standard CRRA utility over final wealth W_T and standard CPT value function over gains and losses X_T. In total, the utility function is defined as:

$$\Psi\left(W_T, X_T\right) = U(W_T) + bV(X_T), \tag{12}$$

where b is a scaling term which governs the relative importance of the CPT part. The utility framework of equation (12) is equal to standard CRRA for $b = 0$. A CRRA utility function, $U(\cdot)$, is given by:

$$U(W_T) = \begin{cases} \frac{W_T^{1-\gamma}}{1-\gamma} & \text{if} \quad \gamma \neq 1 \\ \log W_T & \quad\quad\; \gamma = 1 \end{cases}, \tag{13}$$

where γ is the risk aversion coefficient. The CPT part of the utility function defined of gains and losses is given by:

$$V(X_T) = \begin{cases} X_T & \text{if} \quad X_T \geq 0 \\ \lambda X_T & \quad\quad\; X_T < 0, \end{cases} \tag{14}$$

where λ represents the degree of loss aversion. Note, $X_T := W_T - W_{ref}$ is defined relative to a reference level W_{ref} associated with the reference point of the CPT framework.

The preferences from equations (12)-(14) are the same as in Barberis et al. (2001), except their framework is extended with the addition of probability weighting. Equation (14) abstracts from the S-shaped value function as in the pure CPT value function and therefore, $\Psi(W_T, X_T)$, is concave. Because of concavity of the total utility function standard maximization techniques can be used to analyze its properties.

5.2 Equilibrium Pricing Kernel

I consider a simple market in this setup with the following assets: (i) a risk-free asset with constant gross return R_f; (ii) equity that pays x_i^E in state i at time T, which occurs with probability p_i; and (iii) M derivatives on equity, where $m \in [1, M]$ pays $x_i^{D,m}$ in state i at time T. The derivatives are in zero net supply, and the supply of equity is normalized to one. In order to derive the equilibrium pricing kernel a one-period optimal portfolio choice problem is considered. The representative investor chooses its positions in the risky asset, α_E, each derivative m, $\alpha_{D,m}$ and the risk-free asset such that she maximizes:

$$\max_{\alpha_E, \alpha_{D,m}} \mathbb{E}[\Psi(W_T, X_T)], \tag{15}$$

with respect to the budget constraint:

$$W_{T,i} = \left[(1 - \alpha_E - \sum_m \alpha_{D,m}) R_f + \alpha_E \frac{x_i^E}{s_0^E} + \sum_m \alpha_{D,m} \frac{x_i^{D,m}}{s_0^{D,m}} \right] W_0. \tag{16}$$

In equation (16) W_0 is the wealth at time zero, s_0^E and $s_0^{D,m}$ denote time-zero prices of equity and the derivatives, respectively.

As stated before, the risk-free asset and the derivatives are in zero net supply, consequently the market clearing conditions yield $\alpha_E = 1$ and $\alpha_{D,m} = 0$ for all m. Plugging these equilibrium conditions into the budget constraint of (16) and the wealth constraint into the first order constraint of the objective function, (15), yields:

$$0 = \sum_i p_i \left(\frac{x_i^T}{s_0^E} - R_f \right) \left[\left(W_0 \frac{x_i^T}{s_0^E} \right)^{-\gamma} + b \frac{\pi_i}{p_i} (1 + (\lambda - 1) \mathbb{1}_{X_i<0}) \right]. \tag{17}$$

The part between brackets of this equation can be interpreted as the theoretical pricing kernel. In equation (17), $\mathbb{1}_{X_i<0}$ is an indicator function which equals 1 for losses and π_i are the distorted probabilities as in Tversky and Kahneman (1992). Furthermore, I assume, just as in Barberis et al. (2001), that the parameter b in equation (17) is a linear combination of $W_0^{-\gamma}$. In case b is a linear combination of $W_0^{-\gamma}$ the pricing kernel will not depend on wealth. As the first order condition of equation (17) is derived with respect to the excess return, the pricing kernel in the equation only says something about the shape and not about the level. The level of the pricing kernel can be determined if one considers the exogenous risk-free rate. As in many other representative agent models, $1/R_f = \mathbb{E}(m)$ has to hold and can in this case identify the level of the pricing kernel. For now it is beyond the scope of this thesis, but it is interesting to investigate in the future. Furthermore, the pure CRRA case is embedded in this model, when $b = 0$ the investor has CRRA preferences. The probability distortion in (17), happens in the $\frac{\pi_i}{p_i}$ part of the CPT value function. This distortion is given by:

$$\frac{\pi_i}{p_i} = \frac{w(p_1 + \cdots + p_{i+1}) - w(p_1 + \cdots + p_i)}{p_{i+1}}, \tag{18}$$

where $w(\cdot)$ is the probability weighting function of TK given by:

$$w(p) = \frac{p^{\delta}}{(p^{\delta} + (1-p)^{\delta})^{\frac{1}{\delta}}}. \tag{19}$$

When $b = 0$, the representative investor has pure CRRA preferences and the pricing equation of (17) is equal to:

$$0 = \sum_i p_i \left(\frac{x_i^T}{s_0^E} - R_f \right) \left[\left(\frac{x_i^T}{s_0^E} \right)^{-\gamma} \right]. \tag{20}$$

In equation (20) the pricing kernel from a representative investor with CRRA preferences is given by the part between brackets. The theoretical pricing kernel of equation (20) embeds the pricing kernel for a risk neutral representative investor for $\gamma = 0$.

5.2.1 Implications of preferences

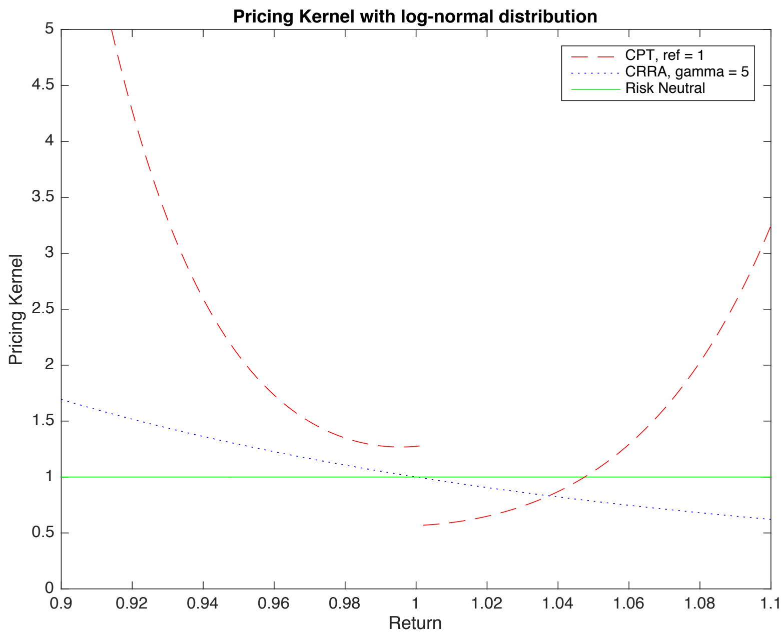
I consider three different preferences of the representative investor. The preferences have different implications for the shape of the pricing kernel and consequently on prices of financial instruments. A risk-neutral investor maximizes expected returns and the states of nature do not influence the decision making. Consequently, the pricing kernel appears flat and always equal to one, as all states of nature are valued in the same way.

A model with a pure CRRA investor predicts a downward sloping pricing kernel. Marginal utility is larger in bad compared to good states of nature, therefore the demand for instruments paying in these states is higher and consequently prices as well. Baele et al. (2016) show that the expected returns of options predicted by the model with a CRRA investor are not in line with what is empirically observed.

The model with CPT representative investor is able to predict expected option returns aligned with empirically observed option returns. This model implicates a pricing kernel which has three regions which are specific for the CPT investor. In the CPT framework probabilities are distorted, especially at the tails of the distribution. These small probabilities are overweighted and due to this overweighting instruments paying in the tails of the distribution are overpriced. Consequently, the pricing kernel in the left tail is (strongly) downward sloping, whereas in the right tail it slopes (strongly) upward. Especially the second implication is rather unique, classical asset pricing models do not predict high prices for instruments paying in the best states of nature when marginal utility is low. Furthermore, the CPT framework consists of loss aversion, therefore the model predict a downward jump of the pricing kernel around the reference point[5]. The implications of the loss aversion are observed in the data if a (strong) negative slope is present around the reference return.

The dynamics of the pricing kernels of the representative agent with different preferences is best explained based on Figure 12.

Figure 12: In the figure the pricing kernels for pure CPT, CRRA and Risk Neutral preferences are represented. Note the results are based using a risky asset log-normally distributed with μ and σ estimated from S&P500 data from 1990-2015 for a horizon of three weeks. In case of CRRA, the pricing kernel is derived with $\gamma = 5$. The parameters of the CPT framework are chosen as in Tversky and Kahneman (1992)



First, the risk neutral preferences are represented by the green line in figure 12 and is simply always equal to one since its marginal utility is equal in every state of nature. Therefore, in an economy with a risk neutral representative agent the prices of the butterfly spreads in different states of the world (e.g. bad, mediocre and good) would be, relative to its probabilities, equal. Second, the agent with CRRA preferences has a relatively high pricing kernel in bad states of the world, when marginal utility is high and a

[5]It is not clear where this reference point is, in this thesis I assume the reference return is equal to the risk-free rate.

low pricing kernel in good states. This leads to relatively high prices for butterfly spreads paying in bad states of the world, whereas prices of butterfly spreads paying in good states are lower. Third, in case of pure CPT preferences two main implications of CPT can be identified in figure 12. Note, the pricing kernel is represented only for the CPT part of the pricing kernel of equation (17). Around the reference point (1 in this case) a jump in the pricing kernel is observed as a result of loss aversion. Instruments that pay in bad states of the world are therefore relatively expensive. Furthermore, in Figure 12 a relatively high pricing kernel is obtained at the tails of the distribution as a result of probability weighting. The effect of probability weighting works at the tails of the distribution, since in CPT small probabilities are overweighted. In the worst states of nature, the pricing kernel is relatively high when marginal utility is high and the effect is amplified because the small probabilities are overweighted. Similarly, the pricing kernel at the other tail of the distribution, the probabilities for the best states of nature are overweighted and therefore instruments paying in this state are relatively expensive.

Given the fact that the empirically estimate of the winsorized average pricing kernel slopes downward over the interval [0.95; 1.05] when S&P 500 returns are modeled with the benchmark model (Figure 8), I conclude that a representative investor with pure CRRA preferences is best able to match such a pricing kernel. The conclusion is drawn upon the fact that a model with CPT preferences predicts an upward sloping pricing kernel in the region with the best states of nature. Note, implicitly it assumed that the parameters are in the order of magnitude of Tversky and Kahneman (1992) in order to produce such an upward sloping pricing kernel. This preliminary conclusion can change if the results are not robust to estimating the pricing kernel with butterflies having a wider spread, e.g. $\Delta K = 10$. Furthermore, it will be interesting in future research to estimate the parameters of the CPT representative investor on basis of the obtained pricing kernel. As the pure CRRA case is embedded in the CPT framework, it will be possible to reach a conclusion on which model is better able to explain the empirical pricing kernel.

6 Conclusion

In this thesis, I develop a framework in order to estimate the pricing kernel using data on butterfly spreads. For S&P 500 butterfly spreads with a fixed maturity of 15 days, I estimate the pricing kernel each month. The average pricing kernel, estimated using a skewed t-distribution with volatility a linear function of the VIX to model S&P 500 returns, has a negative slope on the return interval of [0.95; 1.05]. This empirically observed shape is in line with a theoretically predicted pricing kernel of a representative investor model with pure CRRA preferences.

The S&P 500 returns are modeled in this thesis using three different distributions, namely: log-normal, normal and skewed t-distribution. As the volatility of the S&P 500 returns is important in the valuation of derivatives, the three distributions have time-varying volatility. It turns out that the skewed t-distribution with time-varying volatility is best able to match the time-series average of expected with the realized payoff of the butterflies.

An important issue with the observed prices of the butterflies is noise. Many butterfly prices in the data violate the no arbitrage rule, i.e. prices that are smaller or equal to zero. If the butterflies with arbitrage prices are discarded, a significant positive pricing error is made. This pricing error is observed as the average butterfly price, excluding the butterflies with arbitrage prices, is significantly larger than the average prices including the arbitrage butterflies over the total distribution of moneyness. Furthermore, the noise is clearly visible when the slopes of the pricing kernel are analyzed with finite differences. I do not observe a clear pattern in the slopes of the pricing kernel over time, due to the noisy sample.

6.1 Future Research

In future research, I want to estimate the parameters in the CPT representative agent model on basis of the monthly sample of the pricing kernel. The pure CRRA model is a special case of the CPT model, therefore I would be able to present more convincing evidence of which model is best able to explain the empirical pricing kernel.

Furthermore, I would like to check how robust the findings are. For instance, by estimating the pricing kernel on basis of butterflies with a spread (ΔK) equal to ten instead of five. Another advantage of using butterflies with a wider spread is that the expected payoff is higher, for the same pricing error the estimation error in the pricing kernel estimate will be less severe.

As it is interesting to see the evolution of the slope of the pricing kernel over time, the six month rolling average of the finite difference estimate seems very noisy. Polkovnichenko and Zhao (2013) estimate slopes with the area under the physical pricing kernel divided by the area under the risk neutral kernel. If the estimate of the slope is more stable, it will be interesting to see the change in attitudes towards risk over time.

7 Bibliography

Aït-Sahalia, Y. and Lo, A.W. (2000). Nonparametric risk management and implied risk aversion. *Journal of econometrics*, 94(1):9–51.

Baele, L., Driessen, J., Ebert, S., Londono, J. M. and Spalt, O. G. (2016). Cumulative prospect theory, option prices, and the variance premium. *Working paper*.

Barberis, N., Huang, M. and Santos, T. (2001). Prospect theory and asset prices. *Journal of Financial Economics*, 116(1):1–53.

Bauwens, L. and Laurent, S. (2002). A new class of multivariate skew densities, with applications to garch models. *Working paper*.

Boyer, B. H. and Vorkink, K. (2014). Stock options as lotteries. *Journal of Finance*, 69(4):1485–1527.

Chaudhuri, R. and Schroder, M. (2015). Monotonicity of the stochastic discount factor and expected option returns. *Review of Financial Studies*, 28(5):1462–1505.

Lambert, P. and Laurent, S. (2001). Modelling financial time series using garch-type models and a skewed student density. *Working paper*.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370.

Polkovnichenko, V. and Zhao, F. (2013). Probability weighting functions implied in options prices. *Journal of Financial Economics*, 107(3):580–609.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.