The Term Structure and Time-Series Variation of the Pricing Kernel

Joost Driessen, Joren Koëter and Ole Wilms*

Preliminary and incomplete.

December, 2018

Abstract

We estimate the pricing kernel from portfolios of options on the S&P 500 index for different horizons and over time. This allows us to compare short- and long-term pricing kernels and check its time-series variation. We document a U-shaped short-term pricing kernel for which the U-shape is more pronounced in good times, a finding that is at odds with leading asset pricing models. However, we find the U-shape to be a unique feature of the short-term pricing kernel and once we look at longer horizons, the pricing kernel is monotonically decreasing with only little time-series variation. We show that leading asset pricing models capture the level, shape and time-series variation of the long-term kernel well.

1 Introduction

The pricing kernel represents the preferences of investors over risky outcomes, therefore estimating the pricing kernel is of interest to academics and practitioners. Starting with the seminal work of Jackwerth (2000) and Aït-Sahalia and Lo (2000), many studies have used option pricing information to estimate pricing kernels for the U.S. equity market[1]. The main finding in this literature is that the pricing kernel is not strictly downward sloping when it is projected on the equity index return. Rosenberg and Engle (2002) documents that the pricing kernel has an upward sloping region, similarly Jackwerth (2000) reports a U-shape which are both in contrast to a monotonically decreasing

*All authors are at the Department of Finance at Tilburg University.

[1]See Cuesdeanu and Jackwerth (2018a) for an overview of this literature.

kernel as predicted by classical finance theory. This result is often referred to as the "pricing kernel puzzle". We contribute to this literature by documenting two new and surprising stylized facts of the pricing kernel. First, by exploiting multiple option maturities, we analyze the term structure of the pricing kernel. Second, we present evidence with respect to the time-series variation of the pricing kernel. In order to compare short- and long-term pricing kernels, we define a forward kernel which does not contain any short-term information. In the following, we refer to one-month pricing kernel as the short-term pricing kernel and to the expected forward pricing kernel with horizon longer than six months as the long-term pricing kernel. The results we discuss have important implications on the evaluation of risks by investors, both over time and different horizons. Furthermore, we asses whether leading asset pricing models can explain the new empirical evidence.

The first new stylized fact concerns the term structure of the pricing kernel. Existing work focuses on the pricing kernel for a horizon of one month, while we use options with maturities beyond one month to estimate the pricing kernel for longer horizons. In line with existing empirical work, for a one-month horizon we find a strong U-shaped pricing kernel. The U-shape implies that investors are willing to pay more than the expected value for securities paying in bad states and in good states of the economy– a feature that is hard to reconcile with leading consumption-based asset pricing models[2]. We show that standard calibration of the asset pricing models, for standard calibrations, produce a monotonically decreasing pricing kernel when projecting the pricing kernel onto the market return. In contrast to the results on the short-term kernels and new in the literature, we find that for longer horizons, the empirical pricing kernel is monotonically downward sloping. Hence, the pricing kernel puzzle vanishes when looking at longer maturities and we show that the theoretical asset pricing models match both, the shape and level of the long-term pricing kernels well. We discuss two potential reasons for the differences we document between short-term and long-term pricing kernels. First, the finding is in line with a representative agent which prices short-term risk different from long-term risk. A phenomenon which is also reported in Dew-Becker et al. (2017), who find that hedging short-term realized volatility is more costly than hedging shocks to expected long-term

[2]In particular, we consider the habit formation model of Campbell and Cochrane (1999), the rare disaster model of Wachter (2013), the long-run risks model of Bansal and Yaron (2004), and the time-varying recovery model of Gabaix (2012).

volatility. Second, the results can be explained by maturity segmentation in the option market. It could be that speculative investors with non-standard preferences mostly trade short-term options whereas investors with more standard preferences trade long-term options. We leave it to future research to examine such segmentation in option markets.

The second new stylized fact concerns the time-series variation of the pricing kernel. We find that the U-shape of the short-term pricing kernel is more pronounced in good times compared to bad times, where we define good and bad times according to the level of the CFNAI[3]. The stronger U-shaped pricing kernel in good times indicates that investors are more sensitive towards large negative and positive returns in good times. We show that the time-series variation is most pronounced for the short-term pricing kernel, while the time-series variation of the long-term kernel is significantly smaller. The time-series variation of the pricing kernel in the asset pricing models is, in contrast to the data on the short-term pricing kernel, quantitatively small[4]. Therefore, the fact that we document large time-series variation for the short-term pricing kernel is form a theoretical prespective puzzling. Given that the time-series variation of the long-term pricing kernel is much smaller, leading asset pricing models are able to match the data. In terms of shape and in terms of the magnitude of the time-series variation theoretical asset pricing models match the data of the long-term pricing kernel well. Therefore, the main result of this paper is that the pricing kernel puzzle disappears in the long run and leading asset pricing models are able to match the data on long-term pricing kernels, in terms of shape, level and time-series variation. However, for the short horizon, the pricing kernel is not only puzzling with regards to its u-shape, but also with regards to its time-series variation.

Behavioral models with probability weighting from Kahneman and Tversky (1979) can potentially explain the violation of monotonicity and the time-series variation of the short-term pricing kernel. Due to probability weighting, an agent overweights small probabilities of extreme good or bad events and Baele et al. (2016) show it yields a U-shaped pricing kernel. Furthermore, the finding of the U-shape being more pronounced in good times than bad times is explained by probability weighting. In good times volatility is lower, therefore returns smaller in magnitude have lower probabilities

[3]Chicago Fed National Activity Index.

[4]Whether the curvature of the pricing kernel increases or decreases in good times is dependent on the calibration, however for reasonable calibrations the time-series variation is quantitatively small.

and for this reason overweighted due to probability weighting. Ultimately resulting in a more-pronounced U-shape in good times. The implications of probability weighting with respect to the horizon are unknown and therefore an interesting topic for future research.

Our empirical procedure is conceptually similar, among others[5], to Aït-Sahalia and Lo (2000). Aït-Sahalia and Lo (2000) exploit the result of Breeden and Litzenberger (1978) to estimate the risk-neutral distribution and model the underlying distribution non-parametrically. Our method uses a discrete-state analogue of the result of Breeden and Litzenberger (1978), which is a portofolio of options on the S&P 500. Using options on the S&P 500 is the standard in the literature, as it is a reasonable and commonly-used proxy for the market return and the options are among the most liquid traded options. We estimate the pricing kernel using portfolios of options called butterfly spreads. The prices of butterfly spreads capture the dynamics of the risk-neutral distribution. Roughly speaking, the pricing kernel follows from the quotient of the price of the butterfly spread and the expected payoff. In order to calculate the expected return, we need to model the returns of the underlying S&P 500 index. We assume a skewed t-distribution with time-varying volatility. An advantage of our method is the proxy we use for the risk-neutral distribution, as it is an investable portfolio it facilitates economic interpretation. Given our methodology, we estimate the pricing kernel as a function of the market return. We do this for each day in our sample period, and for different horizons, ranging from one to twelve months.

1.1 Literature

Jackwerth (2000), Aït-Sahalia and Lo (2000) and Rosenberg and Engle (2002) document that the empirical pricing kernel is not monotonically declining and exhibits a S-shape or U-shape depending on the sample period. Chabi-Yo (2012) and Song and Xiu (2016) confirm the earlier results over longer sample periods. All of the aforementioned studies use the result of Breeden and Litzenberger (1978) to obtain an estimate of the risk-neutral distribution and model the real-world return distribution using parametric

[5]Jackwerth (2000), Rosenberg and Engle (2002), Polkovnichenko and Zhao (2013) and Song and Xiu (2016) all use conceptually the same method as ours to estimate the pricing kernel.

or nonparametric ways.

Chabi-Yo et al. (2007) and Bakshi et al. (2010) try to reconcile the findings of the non-monotonic pricing kernel with either a regime-switching model or a model with disagreeing investors. Christoffersen et al. (2013) show that a reduced form pricing kernel, which allows for a U-shape due to the negative price of variance risk, captures the options data well. Polkovnichenko and Zhao (2013) estimate probability weighting functions from option data and find that the empirical probability weighting functions are such that investors over-weight small probabilities as in prospect theory by Kahneman and Tversky (1979). Baele et al. (2016) show that a model with probability weighting can reconcile a U-shaped pricing kernel and explains the returns on put/call options and the variance premium well.

We also contribute to the literature on term structures of several risk premiums. van Binsbergen et al. (2012) look at the term structure of equity and Dew-Becker et al. (2017) describe the term structure of variance price risk. See van Binsbergen and Koijen (2017) for a recent overview and evidence of other asset classes. We are the first to document stylized facts with respect to the term structure of the pricing kernel.

Linn et al. (2017) argue that the violation of monotonicity documented by earlier work in this literature is spurious. The pricing kernel follows from the real-world and risk-neutral distribution, and, as the risk-neutral distribution is derived from options, it incorporates all information upon time t. In some of these earlier mentioned studies, the real-world probabililty density is estimated non-parametrically on a sample of past data and therefore unable to capture all information available on time t. Linn et al. (2017) claim that if both the real-world and risk-neutral distribution contain all information on time t, the violation of monotonicity disappears. However, Cuesdeanu and Jackwerth (2018b) show that even when the pricing kernel is estimated from forward-looking data only it is U-shaped. The difference in results compared to Linn et al. (2017) are, according to Cuesdeanu and Jackwerth (2018b), a result of lack of scaling of the subjective densities and their particular choice of two-step optimization.

In line with Cuesdeanu and Jackwerth (2018b), we find a U-shape even when we condition on forward-looking infomation, the VIX in our case. The level of VIX is forward-looking as it represents the integrated risk-neutral volatility over the following month. We run several specifications of our volatility model, e.g. adding lagged volatility or squared VIX

term, however our results are unaffected.

2 Methodology

We estimate the pricing kernel using butterfly spreads, which is a portfolio of options. We use European options on the S&P500 as it is regarded as a reasonable proxy for the market portfolio and the option contracts are among the most traded options. First, we explain how to construct butterfly spreads from options and later on we discuss the details of our estimation procedure. A Butterfly spread can be constructed using three different call (or put) options with equal distance in strike. To form a butterfly spread from call options with strike K and spread ΔK, one has to take a long position in call options with strikes $K-\Delta K$ and $K+\Delta K$ and a short position in two calls with strike K. Similarly, it is possible to construct a butterfly spread from put options, by taking the same positions as for the call butterfly. The payoff of a butterfly is only positive on the interval $[K-\Delta K, K+\Delta K]$, therefore if the spread of the butterfly is small it identifies the pricing kernel well over this interval. It is possible to form an identical butterfly from call or put options. By no-arbitrage the butterflies should have the same price as the payoff is identical. As we will explain later on, we use out of the money options to construct the butterflies as these contracts are the most liquid.

2.1 Estimation procedure

The method we use to estimate the pricing kernel exploits the discrete-state analogue of the result by Breeden and Litzenberger (1978). In the paper they prove that the price of an Arrow-Debreu security equals:

$$P(K,T) = \lim_{\Delta K \to 0} \frac{c(K+\Delta K,T) + c(K-\Delta K,T) - 2c(K,T)}{(\Delta K)^2} = \frac{\partial^2 c(K,T)}{\partial K^2}, \qquad (1)$$

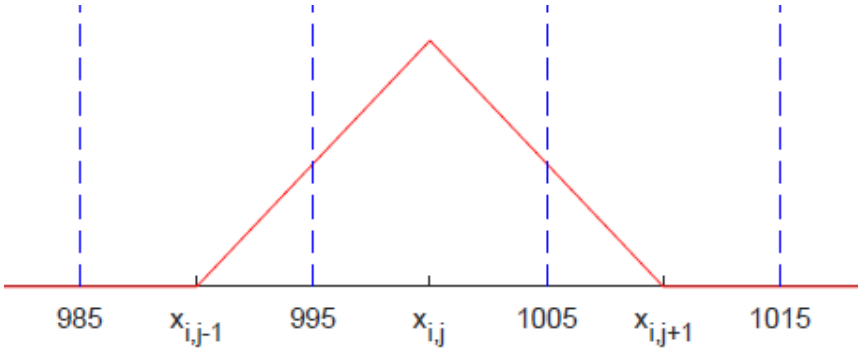
where $P(K,T)$ represents the price of an Arrow-Debreu security paying one if the stock price at maturity equals the strike ($S_T = K$). $c(K,T)$ is the price of a call option, for that reason the numerator of equation (1) corresponds to the price of a butterfly spread

with strike K, spread ΔK and maturity T. Aït-Sahalia and Lo (2000) utilize the result to estimate the risk-neutral probability density semiparametrically. They assume the price of a call option to equal the Black-Scholes equation with volatility modeled nonparametrically, by differentiating twice they obtain the risk-neutral distribution. Together with a non-parametric model of the S&P 500 return distribution, it allows them to tease out the pricing kernel. Conceptually our methodology is similar, but we use butterfly spreads to capture the risk-neutral distribution and model the S&P 500 parameterically. The parametric model we assume for the return distribution of the S&P 500 allows us to calculate expected payoffs of butterfly spreads. From the price of the butterfly spread and the expected payoff, we are able to estimate the pricing kernel. The advantage of our methodology is that the proxy for the risk-neutral distribtution is an investable portofolio which facilitates economic intuition. I.e. high levels of the pricing kernel in the states over which a butterlfy spread pays, correspond to low expected returns of the butterfly spread.

The pricing kernel follows from the classical asset pricing equation: $p = \mathbb{E}(mX)$. Under the assumption of market completeness, the pricing kernel is unique. In order to solve for the discrete state pricing kernel, we write the price of a butterfly spread as a linear combination of expected payoff in each state times the corresponding discount factor. To do so, we discretize the outcome space and assume that stochastic discount factor is constant within each state. As we will see, the states correspond to return levels of the S&P 500. Put differently, we estimate the pricing kernel as a function of the return on the S&P 500.

We provide a brief example of how we discretize the state-space. Consider an index with current price $S_0 = 1000$ and a butterfly i with strike $K_i = 1000$ and spread $\Delta K = 10$. Assume for now that the state space is discretized with distance between states equal to $1\% \cdot S_0 = 10$, which we graphically represented in 1. The state where the return over horizon T is 0% corresponds to the interval $[995; 1005]$.

Figure 1: In the graph the example is graphically illustrated, a butterfly which has a payoff on the range $[990; 1010]$ is represented. Three states are important for the valuation of the instrument, namely $x_{i,j-1}$, $x_{i,j}$ and $x_{i,j+1}$.



Three non-zero expected payoffs over the states are important for the valuation of the butterfly spread. Expected payoff $x_{i,j-1}$ calculated over the range $[985; 995]$, expected payoff $x_{i,j}$ calculated over range $[995; 1005]$ and expected payoff $x_{i,j+1}$ calculated over range $[1005; 1015]$. The expected payoff over the middle state is given by:

$$x_{i,j} = \int_{995}^{1005} X_i(S_T) \cdot f(S_T) dS_T.$$

Where $X_i(S_T)$ is the payoff function of a butterfly with strike $K_i = 1000$. To calculate expected payoffs $x_{i,j-1}$ and $x_{i,j+1}$ the intervals are adjusted accordingly. Ultimately, the butterfly spread is valued as if there are only three states possible where the instrument has a positive payoff.

On a certain trading day we observe several butterfly spreads with positive payoffs on different regions of the outcome space. For now, similar to the example given above, the butterflies all have three discrete states in which the payoff is positive. For each butterfly spread, we write the price as a linear combination of the expected payoff in that state times the stochastic discount factor of the particular state. In matrix notation, it is represented as follows, where b_j, b_{j+1} represent the bounds of the state:

$$\begin{pmatrix} p_1 \\ \vdots \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & 0 \\ 0 & x_{2,2} & x_{2,3} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & x_{n-1,n-2} & x_{n-1,n-1} & 0 \\ 0 & \cdots & x_{n,n-2} & x_{n,n-1} & x_{n,n} \end{pmatrix} \cdot \begin{pmatrix} m_1 \\ \vdots \\ \vdots \\ m_n \end{pmatrix}$$
$$b_1 \quad b_2 \quad \cdots \quad \cdots \quad b_n \quad b_{n+1}$$

The observed price of butterfly i is given by p_i. Furthermore, the stochastic discount factor corresponding to state j is given by m_j. In general the expected payoff of butterfly i in state j, with interval $[b_j, b_{j+1}]$ is given by:

$$x_{i,j} = \int_{b_j}^{b_{j+1}} X_i(S_T) \cdot f(S_T) dS_T,$$

where $X_i(S_T)$ is the payoff function of butterfly i. When the number of states (j) equals the number of butterflies (i), the expected payoff matrix is square, and invertible, making pricing kernel (m) unique.

To calculate the expected payoff matrix, we have to model the return distribution of the S&P 500. We assume the return distribution to follow a Skewed-t distribution with the volatility a function of the volatility index (VIX). The details of the model are discussed in section 2.3.

2.2 Forward pricing kernel

We are interested in pricing kernels with different horizons, and we want to observe its differences. Our methodology allows us to estimate a one-period pricing kernel ($m_{t,t+1}$) and a two-period pricing kernel ($m_{t,t+2}$) from the following equations, where X_{t+1} and X_{t+2} are the payoffs at the end of the first and second period, respectively.

$$P_{t,t+1} = \mathbb{E}_t\big(m_{t,t+1} X_{t+1}\big), \qquad P_{t,t+2} = \mathbb{E}_t\big(m_{t,t+2} X_{t+2}\big)$$

We estimate the one- and two-period pricing kernel as a function of the one- and two-period return. Given that the distributions of the one- and two-period return differ, it is difficult to observe the differences. In order to do this in an accesible way, we define the forward pricing kernel. Estimating the forward pricing kernel allows us to compare a one-month pricing kernel dependent on one-month returns with a forward pricing kernel dependent on one-month forward return. As the one-month return distribution and the one-month forward distribution are arguably similar, it allows us to compare the sensitivity of the investor towards one-month risks and one-month forward risks.

We estimate n-period pricing kernels as a function of the n-period return. There-

fore, we can write the two-period pricing kernel as in equation (2). From the two-period pricing kernel we can derive the forward pricing kernel:

$$m_{t,t+2}(r_{t,t+1} \cdot r_{t+1,t+2}) = m_{t,t+1}(r_{t,t+1}) \cdot m_{t+1,t+2}(r_{t+1,t+2}). \tag{2}$$

We estimate the two-period pricing kernel $m_{t,t+2}(\cdot)$ and the one-period pricing kernel $m_{t,t+1}(\cdot)$ empirically of equation (2). The one-period forward pricing kernel is represented by $m_{t+1,t+2}(\cdot)$. As seen in equation (2) the forward pricing kernel is dependent on both the first and second-period return. To obtain the forward pricing kernel as a function of the forward return only, we calculate the expectation of the forward pricing kernel with respect to first period return, conditional on the return in the second period.

$$\mathbb{E}_t\big(m_{t+1,t+2}(r_{t+1,t+2})|r_{t+1,t+2} = r\big) = \int_0^\infty \frac{m_{t,t+2}(r_{t,t+1} \cdot r)}{m_{t,t+1}(r_{t,t+1})} f_t(r_{t,t+1}) dr_{t,t+1}, \tag{3}$$

where $r_{t,t+k} = \frac{P_{t+k}}{P_t}$ and $f_t(r_{t,t+1})$ denotes the one-period return distribution conditional on time t and second period return $r_{t+1,t+2}$.

2.3 Return distribution of the S&P 500

In order to estimate the pricing kernel, we model the return distribution of the S&P 500. The distribution is estimated on S&P 500 returns from 1990 to 2016 using Maximum-Likelihood. We use the the skewed t-distribution proposed by Bauwens and Laurent (2002) to model S&P 500 returns. Instead of the constant volatility used in the original paper, we introduce time-varying volatility to adequately capture volatility dynamics of the stock market:

$$r_{t,t+n} = \mu + \sigma_t \epsilon_{t,t+n}, \tag{4}$$

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

where $r_{t,t+n}$ are n-month returns and the standard deviation is a linear function of the VIX at time t, this will be the benchmark specification[6]. The random variable $\epsilon_{t,t+n} \sim$

[6]We modeled volatility in several ways, e.g. adding a squared VIX term or lagged realized volatility.

$SKST(0,1,\xi,v)$ follows a standardized skewed t-distribution with parameters $v > 2$ (degrees of freedom) and $\xi > 0$ (skewness parameter). The density is given by:

$$f(\varsigma_t|\xi,v) = \begin{cases} \frac{2}{\xi+\frac{1}{\xi}} \cdot s \cdot g\left[\xi(s\varsigma_t + m)|v\right] & \text{if} \quad \varsigma_t < -\frac{m}{s} \\ \frac{2}{\xi+\frac{1}{\xi}} \cdot s \cdot g\left[\frac{1}{\xi}(s\varsigma_t + m)|v\right] & \quad\quad \varsigma_t \geq -\frac{m}{s}, \end{cases} \tag{6}$$

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

$$g(x|v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi(v-2)}\Gamma\left(\frac{v}{2}\right)}\left[1+\frac{x^2}{v-2}\right]^{-(v+1)/2}, \tag{7}$$

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

$$m(\xi,v) = \frac{\Gamma\left(\frac{v-1}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\Gamma\left(\frac{v}{2}\right)}\left(\xi - \frac{1}{\xi}\right), \tag{8}$$

$$s^2(\xi,v) = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2. \tag{9}$$

The volatility of returns changes over time and have a significant impact on prices of options. We assume that the standard deviation of the S&P 500 return distribution is a linear function of the model-free risk-neutral volatility introduced by Britten-Jones and Neuberger (2000). The CBOE introduced the VIX index where it uses the result of Britten-Jones and Neuberger (2000) to estimate risk-neutral volatility from monthly options. In the estimation of our model for the S&P 500, we scale the yearly VIX (252 trading days) to the corresponding maturity.

A skewed t-distribution supports important stylized facts of index returns. Namely, left-skew and fat-tailed distributions. The ξ is a skewness related parameter and facilitates economic interpretation. The probability mass right from the mode divided by the probability mass left from the mode equals $\frac{1}{\xi^2}$. Therefore, when $\xi < 1$, the distribution is left skewed and when $\xi > 1$ the distribution is right skewed. Fat tails are related to the parameter of the degrees of freedom v corresponding to the t-distribution, the lower the parameter the fatter the tails and when $v \to \infty$ the distribution converges

to a skewed normal distribution.

The parameters of equations (4)-(5) we estimate using Maximum Likelihood on n-month simple overlapping returns of the S&P 500 from 1990 to 2016. Our methodology allows us to estimate the pricing kernel for different maturities if we adjust the maturity of the options and horizon of return distribution accordingly. As we are interested in the term structure of the pricing kernel, we estimate the pricing kernel for different horizons n, where n ranges from one up to twelve months. The estimates for the one month S&P 500 return distribution 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 first line with estimates correspond to the equations of (4)-(5). The second line adds a squared VIX term (γ) to the volatility equation. These parameters are estimated using monthly overlapping simple returns of the S&P 500 from 1990-2016.

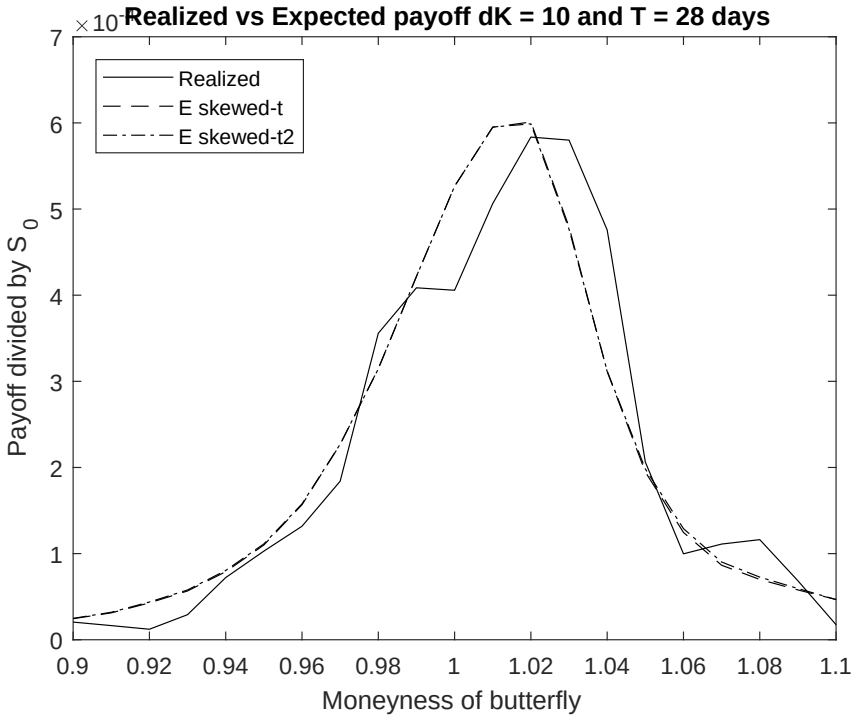
μ	α	β	γ	ξ	v
0.0058	-0.0066	0.8488	-	0.7022	14.0910
0.0058	-0.0148	1.1629	-2.6792	0.7028	14.6734

The parameter μ captures the sample average of the returns. Given the estimates of α and β of the first row, we conclude that the overall level of the VIX is larger than the realized volatility of the S&P 500 index. This phenomenom is known as the variance premium and documented in the literature by Bollerslev et al. (2009). The estimate of $\xi < 1$ indicates that the return distribution is left-skewed and this left-skew is economically sizable. Furthermore, there is evidence of fat tails given the estimate of the degrees of freedom v.

We model the return distribution in order to calculate expected payoffs of butterfly spreads. In order to check the validity of the underlying model, we compare expected with realized payoffs of butterflies in our sample from 1996-2016. We construct butterflies using raw option data from OptionMetrics with $\Delta K = 10$. A butterfly is constructed with three options with consecutive difference of 10 index points in strike. The moneyness of the butterfly is calculated dividing the strike of the butterfly by the current index value (S_0). For the butterflies we observe in the sample we can calculate the expected payoff using our return distribution of the S&P 500. The expected payoff of a butterfly with fixed spread is decreasing in S_0, we account for this by normalizing realized payoff and

expected payoff with S_0. In order to asses the overall ability of our return distribution to capture the expected payoff of S&P 500 butterfly spreads, we calculate for each moneyness level the time-series average of realized payoff divided by S_0 and the time-series average of the expected payoff divided by S_0. The results are given in Figure 2.

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 two 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 solid black line represents the time-series average of the realized payoff. The dashed black line corresponds to the skewed t-distribution specification with volatillity a linear function of VIX. Whereas the dash-dot line adds the squared VIX term.



Two things are important from Figure 2. First, the expected payoff matches the realized payoff quite well. Second, the expected payoff of the butterflies is more or less equal when we add a squared term to the volatility equation. The only difference is that the squared term in the volatility equation makes expected volatility lower for extreme values of VIX.

As discussed in section 2.2, we extract the one-month forward pricing kernel from the ratio of the two-month pricing kernel and the one-month pricing kernel. We present evidence of the expected forward pricing kernel, by calculating the expectation with respect to the return in the first month, conditional on the return in the second month. The simplest setup for doing this, would be to assume independence of the returns. However, this assumption is quite restrictive and given that variance of the stock

market is persistent it does not hold up in the data. This persistence yields the following intuition, if return in the second period is large in magnitude, it is more likely that the return is drawn from a distribution with large variance, implying that the variance of the first period return is also likely to be larger. For the first n-month return, we assume the model of (4)-(5). To model the dependence in variance of the first n-months returns and the one-month return n-periods forward, we assume the following:

$$r_{t+n,t+n+1} = \mu + \sigma_{t+n}\epsilon_{t+n,t+n+1}, \tag{10}$$

$$\sigma_{t+n} = \alpha + \beta VIX_{t+n}, \tag{11}$$

$$VIX_{t+n} = \mu_\sigma + \lambda \cdot (VIX_t - \mu_\sigma) + \rho_1\epsilon_{t,t+n} + \rho_2\epsilon^2_{t,t+n} + \sigma_u u_{t+n}. \tag{12}$$

In order to calculate the conditional distribution, we assume that $\epsilon_{t,t+n} \sim \text{SKST}(0, 1, \xi_n, v_n)$, $\epsilon_{t+n,t+n+1} \sim \text{SKST}(0, 1, \xi_1, v_1)$, $u_{t+n} \sim N(0, 1)$ and all errors are independently distributed. The intuition of the conditioning works as follows, if $r_{t+n,t+n+1}$ is large in magnitude, it is likely drawn from a distribution with large volatility. In the model volatility is driven by VIX_{t+n}, in order for the VIX to be large, $\epsilon_{t,t+n}$ has to be large implying that the volatility of $r_{t,t+n}$ is large.

3 Data

We use European options on the S&P 500 available on OptionMetrics to construct the volatility surface from 1996-2016. One of the reasons to make use of the surface is to address noise in the data and make our methodology less vulnerable to outliers. Furthermore, the volatility surface allows us to obtain a daily estimate of the n-month pricing kernel. We estimate the pricing kernel from a system of equations using prices of butterfly spreads, which are portfolios of three different call (put) options. In the volatility surface of OptionMetrics options with a delta of 20% to 80% for call and -20% to -80% for put options are represented. One of our goals is to explore the behavior of the pricing kernel in the tails of the distribution, therefore we use the same kernel smoothing algorithm to obtain estimates up to a delta (-)1 basis point. For the construction of the butterfly

spreads we make use of out of the money options as these are the most liquid.

Remember, we denote the spread of the butterfly by ΔK, which captures the distance between the strike K and the point where the payoff of the butterfly is zero. We make ΔK increasing in S_0, T and implied volatility by choosing ΔK equal to the difference in strike of call options with deltas 0.35 and 0.50 on a certain trading day. Approximately, we keep the expected payoff of a butterfly given a moneyness level constant, as the expected value would decrease in any of the aforementioned variables. To make sure that the prices of butterflies with moneyness more towards the tails of the distribution are significantly larger than zero, we scale ΔK in the tails of the return distribution by two. By doing so, we can still exploit the differences in price for butterflies with extreme moneyness levels.

As we are discretizing the state-space, we have to choose the distance between two consecutive states. We choose the amount of states on which the pricing kernel is estimated to be stable at 20 estimates each day. The spread of the butterfly is twice as large in the tails of the return distribution, similarly we choose the distance between consecutive states to be twice as large in the tails. The amount of states are distributed over the interval as follows, relatively twice as many states lay in the interval of one-standard deviation around the mode of skewed-t distribution adjusted for the skew. Calculated in the following way:

$$\left[m_t - \frac{c \cdot \sigma_t}{1+\xi^2}; m_t + \frac{c \cdot \sigma_t}{1+\frac{1}{\xi^2}} \right],$$

where m_t is the mode of the distribution and $c = 2 \cdot t^{-1}(0.84, v)$ comes from a t-distribution with v degrees of freedom. The described interval is the skewed-t equivalent of the interval $[\mu - \sigma; \mu + \sigma]$ of a normal distribution. Outside the interval with the largest probability mass, the relative amount of states goes down by a factor two or equivalently, the step-size between intermediate states goes up by a factor of two. Which is sensible, given that we scale the spread of the butterfly in the tails of the return distribution by two.

4 Results

We begin our analysis by estimating the one-month pricing kernel and we show in Section 4.1 that, in line with previous research, we find a strongly u-shaped short-term kernel. In Section 4.2 we analyze the term structure of the pricing kernel and show results for the one-month, six-month forward and twelve-month forward pricing kernel. We show that the U-shape disappears for six and twelve month pricing kernel, for the six-month forward kernel we still find a small violation of monotonicity. Finally, in Section 4.3 we report the time-series variation of the one- six- and twelve month (forward) pricing kernel. We show that the short-term pricing kernel exhibits significant time-series variation and the U-shape is more pronounced in good times. The six- and twelve-month forward pricing kernel exhibit relatively little time-series variation.

4.1 One-month Pricing Kernel

We start by estimating the one-month kernel, in order to do so we have to model the one-month return distribution using equations (4)-(5). The MLE results are represented in the first row of Table 2. The second and third rule are used to estimate the kernel with six and twelve month maturity, respectively. We estimate the parameters using MLE on an overlapping return sample adjusting the maturity accordingly.

Table 2: Maximum Likelihood Estimates of the parameters of the skewed t-distribution with volatility linear function of the VIX of equations (4)-(5). These parameters are estimated using T-month overlapping simple returns of the S&P 500 from 1990-2016.

n	μ	α	β	ξ	v
1	0.0058	−0.0066	0.8488	0.7022	14.0911
6	0.0400	−0.0060	0.8054	0.7568	6.1238
12	0.0867	0.0434	0.5977	0.8027	6.3923

The average return estimate increases with maturity as expected. Even though the VIX equals the model-free one-month risk neutral volatility, it does capture time-variation of the volatility of the six and twelve month return distribution. The skewness

related parameter goes up when the maturity is longer, indicating that the left-skew of the distribution becomes less pronounced. Lastly, for each of the maturities we find evidence of fat tails given that the degrees of freedom are rather small. Together with the volatility surface data we can construct butterflies which allows us to estimate the pricing kernel.

Figure 3 plots the time-series mean and median of the one-month pricing kernel as a function of the return. The range of return levels for which we represent the pricing kernel correspond to a 99% confidence interval of the return distribution with mean 0.58% and volatility 0.045%. We choose a 99% confidence in order to account for the fact that the empirical return distribution is skewed and heavy-tailed.

Figure 3: In the figure the time-series average (solid) and median (dashed) of the one month pricing kernel are shown. The dotted line represents the pricing kernel of the risk neutral investor. Blue triangles in the the figure indicate 95% confidence interval using Newey-West standard errors. Note, for each moneyness level 1% highest and lowest estimates are discarded.

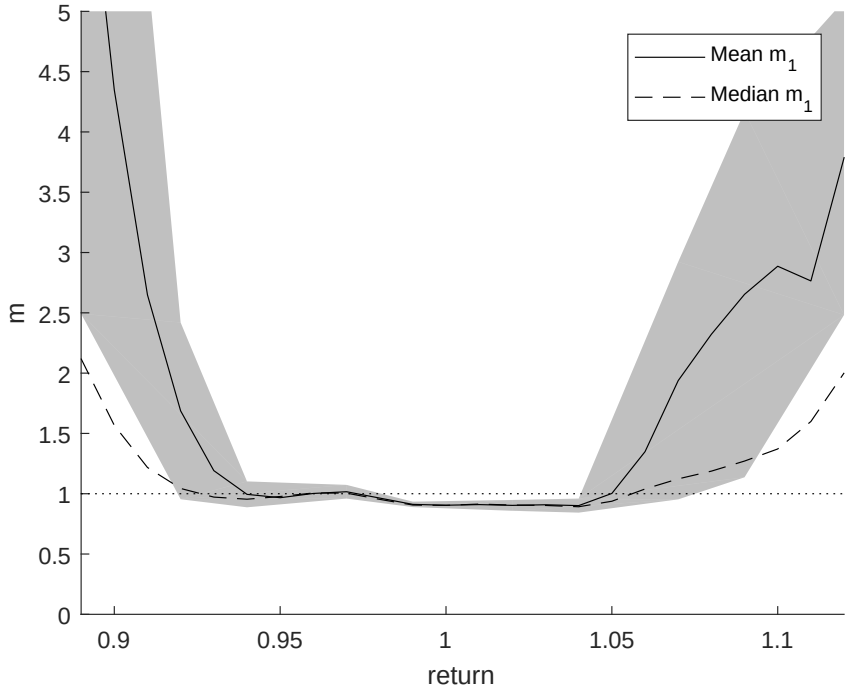


Figure 3 shows that investors are willing to pay more for instruments which pay out either in bad or very good states of the economy, indicated by a high level of the pricing kernel. Bad states are characterized by monthly returns being negative. Economically, an investor is willing to pay about 1.5 times (median) the expected value for an Arrow-Debreu security paying when the stock market looses 10% (0.90 in the

figure), which is equivalent to accepting an expected return of -33%. In the extremely good states of the economy, the pricing kernel slopes upward for monthly stock returns greater than 5%. The U-shape, or violation of monotonicity of the pricing kernel for one-month maturity has been documented before by Jackwerth (2000), Aït-Sahalia and Lo (2000), Rosenberg and Engle (2002), Chabi-Yo (2012) and Song and Xiu (2016). This result can be characterized as a preference of investors for betting on (extremely) good states of the economy and willing to accept a negative expected return. The difference between the mean and median increases more towards the tails of the return distribution, indicating that the time-series distribution of the pricing kernel for a given return level in either of the tails is right-skewed. This suggests that there is time-series variation of pricing kernel which we analyze in detail in Section 4.3.

4.2 Term Structure

In Section 2.2 we discuss how we estimate the one-month forward pricing kernel from the two-month and one-month pricing kernel. We present results on the one-month pricing kernel described before, the six-month expected forward pricing kernel and the twelve-month expected forward pricing kernel. The latter two are derived from the five- and six-month pricing kernel and the eleven- and twelve-month pricing kernel, respectively. Later on in this Section we show that the results are robust when we use different maturities.

We calculate the six-month expected forward kernel with in the following way:

$$\mathbb{E}_t\big(m_{t+5,t+6}(r_{t+5,t+6})|r_{t+5,t+6}=r\big)=\int_0^\infty \frac{m_{t,t+6}(r_{t,t+5}\cdot r)}{m_{t,t+5}(r_{t,t+5})}f_t(r_{t,t+5})dr_{t,t+5}.$$

The expected forward kernel with a forward horizon of five months is dependent on the pricing kernel with six-months horizon and pricing kernel with five-months horizon. Therefore, we first estimate the five- and six-month pricing kernel in the same way as for the one-month pricing kernel. In order to calculate the expected forward kernel, we need to take the expectation of the forward pricing kernel with respect to the five-month distribution conditional on the return in the sixth month. For the five-month return distribution we use the model (4)-(5) and for the return in the sixth month the model

(10)-(12). In order to estimate the twelve-month forward pricing kernel the maturities have to adjusted accordingly.

We use the one-month estimates of Table 2 for the distribution of the return in the sixth month and use the forecast of the VIX_{t+n} to derive the volatility. The estimates of the five-month return and the five-month forward VIX process are given in Table 3.

Table 3: In Panel A the Maximum Likelihood Estimates of the parameters of the skewed t-distribution with volatility linear function of the VIX, equations (4)-(5). These parameters are estimated using n-month overlapping simple returns of the S&P 500 from 1990-2016. In Panel B the estimates for the VIX process with n periods forward, equations (10)-(12).

Panel A

n	μ	α	β	ξ	v
5	0.0337	−0.0090	0.8055	0.7751	7.6392
11	0.0790	0.0296	0.6483	0.8026	6.5666

Panel B

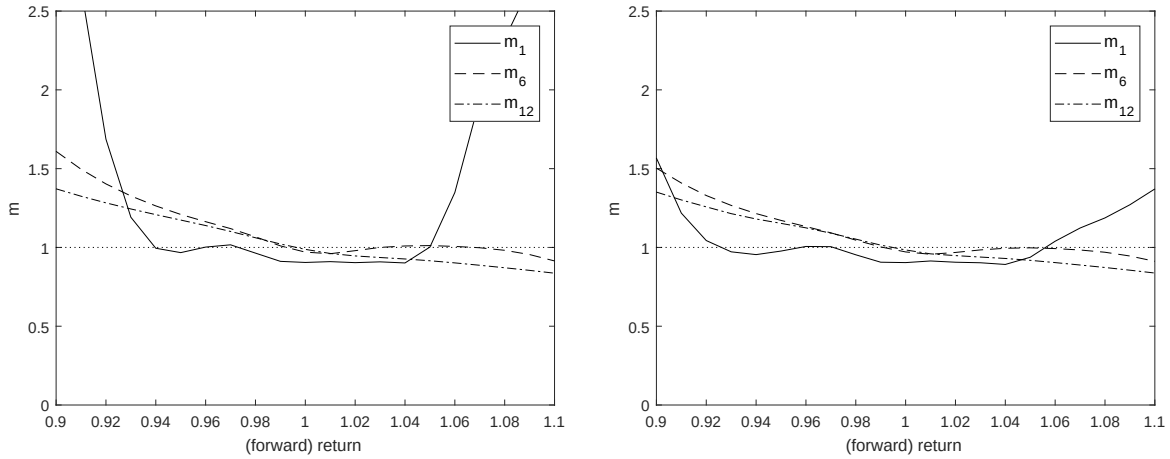
n	μ_σ	λ	ρ_1	ρ_2	σ_u
5	0.0517	0.5105	−0.0094	0.0049	0.0120
11	0.0505	0.3522	−0.0084	0.0060	0.0157

As expected the results for the five- and eleven-month return distribution are very similar to the results of Table 2. In Panel B of Table 3, μ_σ is the average monthly VIX. As λ is positive it indicates some persistence in the VIX process, ρ_2 indicates that if the volatility of the return is large the VIX is also large and ρ_1 indicates that if returns are negative the VIX increases, known as the leverage effect. In order to estimate the expected twelve-month forward kernel, we have to adjust the maturities accordingly.

We compare the means and medians of the one-month kernel, the six-month forward kernel and twelve-month forward kernel.

Figure 4: The figure on the left (right) represents the time-series mean (median) of the one-month (forward) pricing kernel. The solid, dashed and dot-dashed line represent the one, six and twelve month forward pricing kernel, respectively. Forward pricing kernels are estimated using the methodology described in section 2.2.

(a) Mean Pricing Kernel. (b) Median Pricing Kernel.



As reported in the previous section, the one-month kernel is strongly U-shaped. Interestingly, this U-shape disappears for longer horizons and the pricing kernels approach to be monotonically decreasing. For the six months forward kernel we find a small violation of monotonicity, whereas the twelve month forward kernel slopes downward monotonically. Investors are sensitive towards extreme bad and good returns for one-month horizon, whereas for longer horizons the investors are only sensitive towards large negative returns. We show that the pricing kernel puzzle, which has been discussed by Jackwerth (2000), Aït-Sahalia and Lo (2000), Rosenberg and Engle (2002), Chabi-Yo (2012) and Song and Xiu (2016), vanishes once you look at longer horizons.

Table 4: The mean differences between seven points of the (forward) pricing kernel are tested. $m_1(r)$ corresponds to the one-month pricing kernel, $m_6(r)$ corresponds to six months forward kernel and $m_{12}(r)$ corresponds to the twelve months forward kernel. In brackets the hac t-statistics are given, with appropriate amount of lags according to the ACF of the residuals.

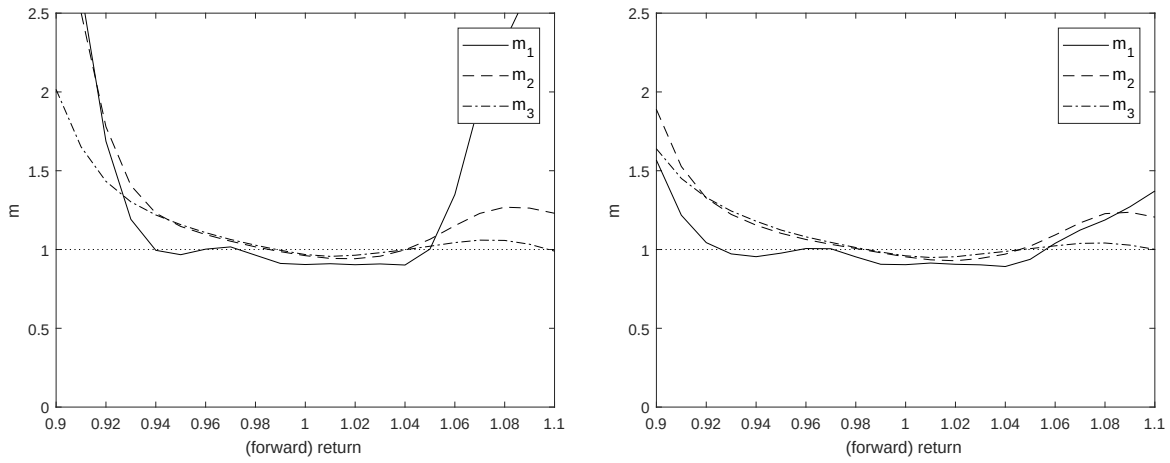
r	0.90	0.95	0.98	1.00	1.02	1.05	1.10
$m_1(r) - m_6(r)$							
Mean	2.75 (2.10)	−0.24 (−7.40)	−0.10 (−4.75)	−0.07 (−3.74)	−0.08 (−4.56)	−0.01 (−0.18)	1.91 (2.35)
$m_1(r) - m_{12}(r)$							
Mean	2.98 (2.20)	−0.21 (−5.08)	−0.10 (−5.03)	−0.08 (−8.31)	−0.04 (−1.23)	0.09 (1.50)	1.98 (2.43)
$m_6(r) - m_{12}(r)$							
Mean	0.24 (6.31)	0.04 (1.14)	0.01 (0.28)	−0.02 (−1.06)	0.03 (2.13)	0.10 (7.90)	0.08 (9.13)

The difference between the one-month kernel and the forward kernels is large, especially in the tails, indicating that the one-month kernels is more U-shaped than the forward kernels. Furthermore, the difference between the forward kernel of with forward horizon six and twelve months is much smaller than its respective difference with the one-month kernel. Only in the tails, the average six-months forward pricing kernel is a bit larger than the twelve-month forward pricing kernel. For positive returns, the difference between the six- and twelve-month forward pricing kernel is largest for 5% return, indicating the hump shape of the six-month forward pricing kernel. The evidence suggests that the representative agent is highly sensitive towards large good or bad returns for short-horizons, whereas for longer horizons the investor is only sensitive towards large negative returns.

To show the robustness of our results, we consider the full term structure of the pricing kernel up to twelve months and show that the u-shape disappears with the maturity. In the next figures we compare the short-end of the pricing kernel, namely the first three months.

Figure 5: The figure on the left (right) represents the time-series mean (median) of the (forward) pricing kernel. The solid, dashed and dot-dashed line represent the one, two and three month forward pricing kernel, respectively. Forward pricing kernels are estimated using the methodology described in section 2.2.

(a) Mean Pricing Kernel. (b) Median Pricing Kernel.

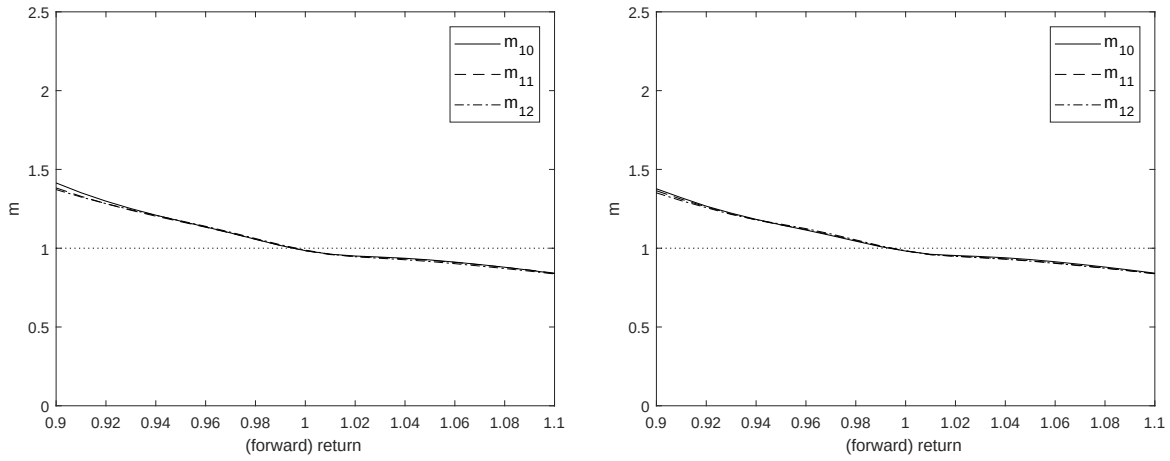


From the left graph of Figure 5 it is clear that the violation of monotonicity is strongest for the one-month pricing kernel. The two- and three-month forward pricing kernel violate monotonicity, but the strong U-shape has dissappeared. Also level of the forward pricing kernel for low returns goes down quite substantially moving from the two- to three-month forward pricing kernel.

Similarly, we compare the longest maturity forward pricing kernels, which are the ten-, eleven- and twelve-month forward pricing kernel.

Figure 6: The figure on the left (right) represents the time-series mean (median) of the (forward) pricing kernel. The solid, dashed and dot-dashed line represent the ten, eleven and twelve month forward pricing kernel, respectively. Forward pricing kernels are estimated using the methodology described in section 2.2.

(a) Mean Pricing Kernel. (b) Median Pricing Kernel.



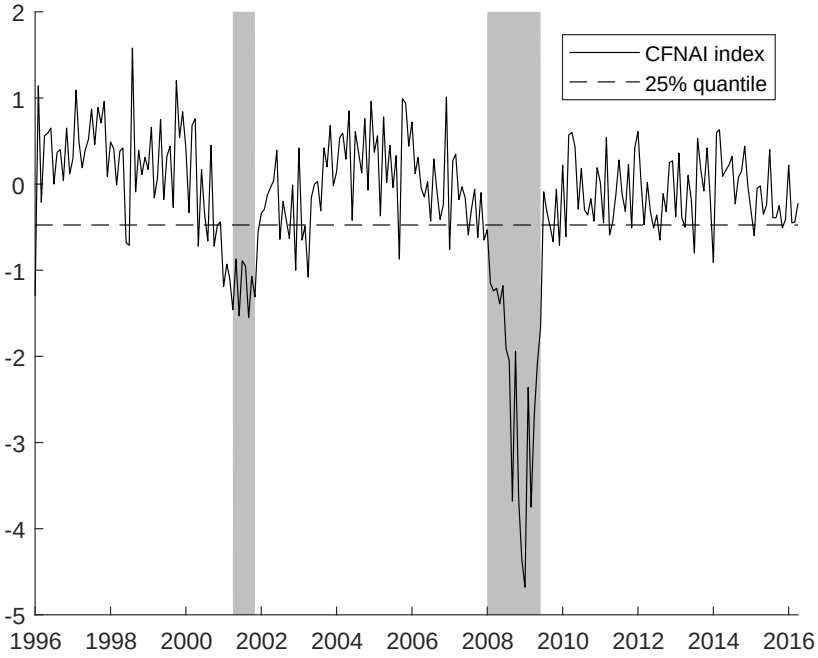
The forward pricing kernels of Figure 6 are very similar and all monotonically decreasing in the market return. As we will see later, such a pricing kernel is in line with standard asset pricing models.

4.3 Time-series variation

As we show in Section 4.1, there was a significant difference between the mean and median of the one-month pricing kernel which is an indication of possible time-series variation. It is interesting to discuss the time-series variation as it provides a new test of leading asset pricing models. The simplest test is to compare the absolute size of the time-series variation found in the data to what is predicted by the models. The easiest way to check for time-series variation is to define subsamples and compare the means in each of the subsamples. We split the sample in two subsamples according to the level of the CFNAI, which is a business cycle indicator issued by the Chicago Fed at the beginning of each month. Periods where the CFNAI is lower than the 25% quantile over 1996-2016, we will refer to as bad times and for good times vice versa. We choose 25% cut-off, as these periods are then significantly bad. Figure 7 shows the time-series of the CFNAI of

our sample period.

Figure 7: The solid line represents the level of the CFNAI index at the beginning of every month. The dashed line corresponds to the 25% quantile over our sample period. Note, the shaded area corresponds to the NBER-recessions.



The shaded area in Figure 7 correspond to the NBER recessions. We observe that the two long recessions observed by the CFNAI coincide with the NBER recessions which suggests that our methodology is appropriate. The advantage of using the CFNAI over NBER recessions is however, that the CFNAI is available at the beginning of each month while the NBER recessions are specified ex post.

We calculate the average (median) pricing kernel for each of the maturities in good and bad times. As before, we plot the average and median of the pricing kernel over the range corresponding to a 99% confidence interval if the return distribution would have been normally distributed. Only in this case, we calculate the average volatility according to our return distribution in good and bad times separately. The volatility in good times is lower than in bad times, therefore the average and median pricing kernel in good times is plotted on a smaller range.

Figure 8: The blue (black) lines represent the time-series mean and median of the one month pricing kernel conditional on being in good (bad) times. The solid and dashed line represent the mean and median, respectively.

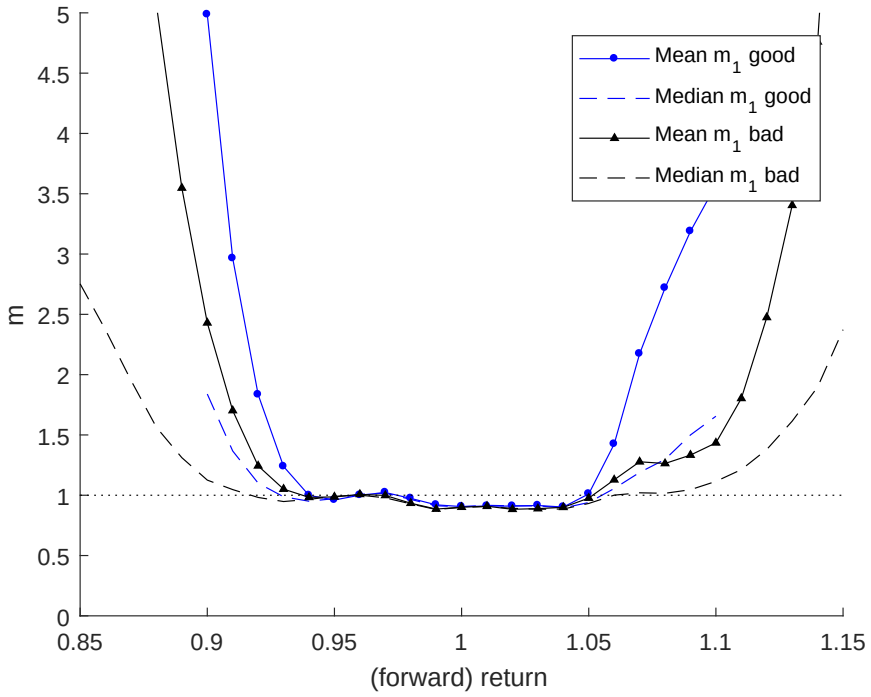
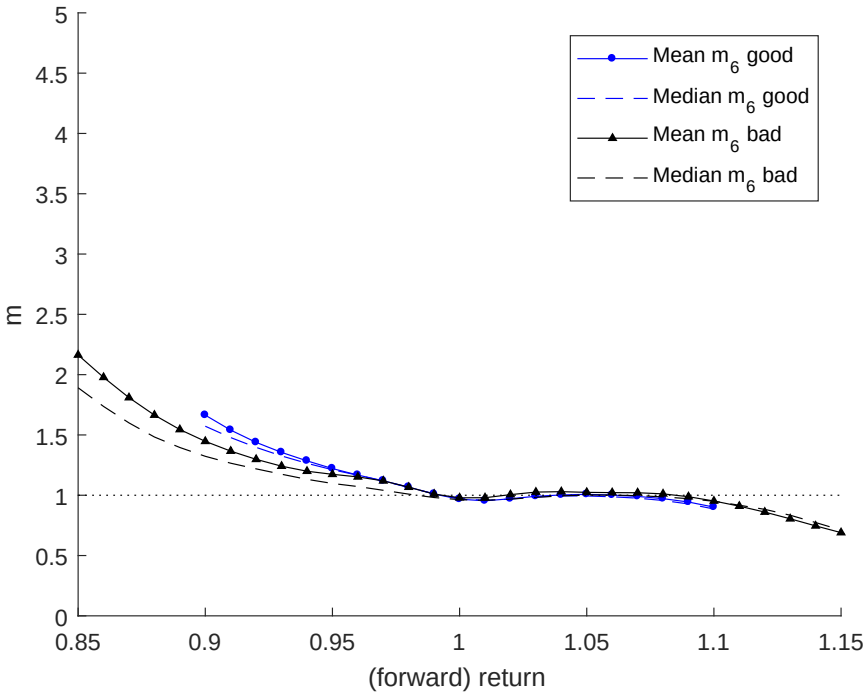


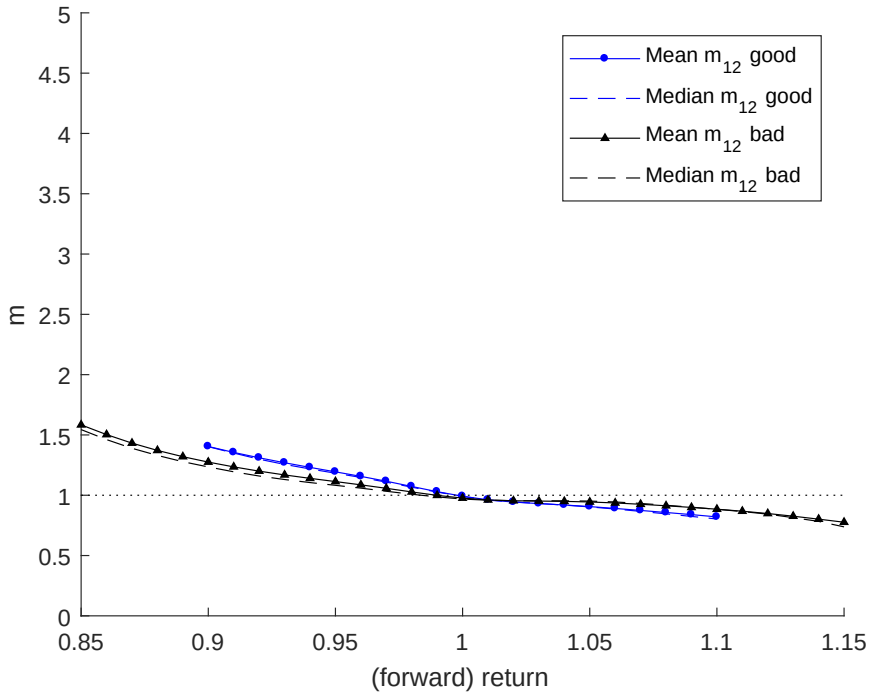
Figure 8 plots the mean and median of the one-month pricing kernel over the different subsamples. We find that the U-shape is stronger in good times. In particular, for returns larger than 5% in absolute value, the investor is willing to pay more to hedge against (or bet on) extreme events in good than in bad times. Similarly, we compute the average and median six-month forward pricing kernel over the same samples.

Figure 9: The blue (black) lines represent the time-series mean and median of the six month pricing kernel conditional on being in good (bad) times. The solid and dashed line represent the mean and median, respectively.



The results for the six-month forward kernel are much weaker than the results from the analysis of the one-month kernel. The level of the pricing kernel increases marginally in good times compared to bad times for low returns. This increased level indicates that the investor is more sensitive towards negative returns in good rather than bad times. The next graphs represent the same analysis for the twelve-month pricing kernel.

Figure 10: The blue (black) lines represent the time-series mean and median of the twelve month pricing kernel conditional on being in good (bad) times. The solid and dashed line represent the mean and median, respectively.



The pattern for level of the pricing kernel for low returns continues in the same way from the six- to the twelve-month forward kernel as it did from the one- to the six-month forward kernel. In good times the level of the forward pricing kernel is a bit higher compared to bad times. Also, the pricing kernel is monotonically decreasing in both good and bad times.

In order to test the significance of the change of curvature we do a regression of the slope of the pricing kernel in a certain region on a dummy variable equaling one in good times. We are going to take a look at four different slopes, defined in the following way:

$$\begin{aligned}
\mathrm{s}_1(T) &= m_T\big(0.90\big) - m_T\big(1.00\big),\\
\mathrm{s}_2(T) &= m_T\big(1.05\big) - m_T\big(1.00\big),\\
\mathrm{s}_3(T) &= m_T\big(1.00\big) - m_T\big(1.05\big),\\
\mathrm{s}_4(T) &= m_T\big(1.00\big) - m_T\big(1.10\big).
\end{aligned}$$

The slope is dependent on the horizon of the pricing kernel (T). As for the figures, we

distinguish a horizon of one, six or twelve months.

Table 5: Regression of the slope of the kernel on the time-series dummy indicating one when VIX is smaller than median VIX. The slopes are dependent on T, where T is 1, 6 or 12 months. Time-series regression: $\text{slope}_i(T) = \beta_0 + \beta_1 \cdot \text{good} + \epsilon$. In brackets the hac t-statistics are given, using the appropiate amount of lags according to the ACF of the residuals.

		1-month		6-month		12-month	
$\text{s}_1(T)$	β_0	-3.49 (-2.55)	-1.53 (-2.04)	-0.64 (-6.40)	-0.47 (-6.94)	-0.38 (-8.13)	-0.30 (-8.73)
	β_1	–	-2.61 (-2.06)	–	-0.23 (-2.22)	–	-0.11 (-2.89)
$\text{s}_2(T)$	β_0	-0.06 (-3.10)	-0.09 (-5.40)	-0.24 (-14.88)	-0.19 (-7.30)	-0.19 (-6.04)	-0.14 (-5.91)
	β_1	–	0.03 (1.25)	–	-0.06 (-1.79)	–	-0.07 (-3.11)
$\text{s}_3(T)$	β_0	0.10 (1.48)	0.08 (1.62)	0.04 (3.41)	0.05 (2.97)	-0.07 (-1.93)	-0.03 (-1.50)
	β_1	–	0.02 (0.49)	–	-0.00 (-0.25)	–	-0.05 (-2.14)
$\text{s}_4(T)$	β_0	1.99 (2.42)	0.53 (3.40)	-0.06 (-3.66)	-0.03 (-2.03)	-0.15 (-2.94)	-0.09 (-3.02)
	β_1	–	2.17 (1.94)	–	-0.04 (-1.68)	–	-0.08 (-2.20)

In Table 5 the first column of each T-month slope corresponds to the total sample average. The second column adds a good times dummy equaling one when the CFNAI is larger than the 25% quantile over our sample, as seen in Figure 7. The results in Table 5 show the statistical significance of the change in slope over time. For the one-month kernel, we can see in the table that the slope for negative market returns, $s_1(1)$, is driven by the highly negative slope in good times. Furthermore, the U-shape, or positive slope for positive market returns $s_4(1)$, is also driven by the large slope in good times. The pattern for the slope of negative market returns of the one-month kernel is similar for the pricing kernels with longer horizons. However, the U-shape disappears for long horizons, as we only find a positive slope for the six-month kernel in $s_3(6)$. For the twelve-month kernel, the U-shape disappeared. For the six- and twelve-month forward kernel we still find that the level for bad returns increases in good times. However, the overall magnitude is much smaller, especially for the twelve-month forward pricing kernel.

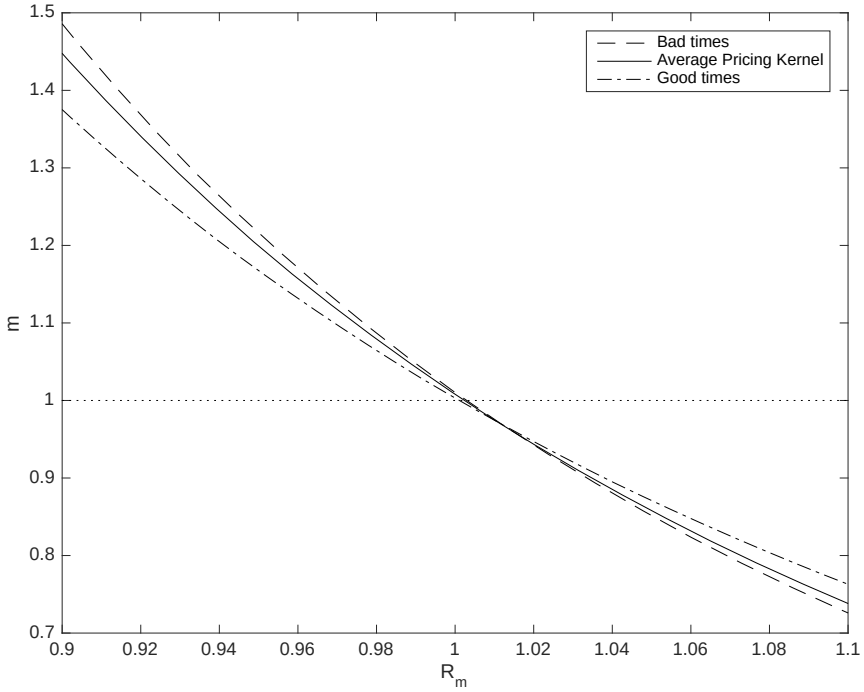
5 Theoretical Models

In this section we discuss what leading asset pricing models predict with respect to the average kernel, time-series variation and the term structure. We incorporate the following models: habit model by Campbell and Cochrane (1999), rare disaster model by Wachter (2013), Long run risk model by Bansal and Yaron (2004) and time-varying recovery by Gabaix (2012). All simulation we carry out for the models are based on a monthly calibration. We show that all the models produce a downward sloping pricing kernel as a function of the market return. Furthermore, the models have for common calibrations very little time-series variation. Therefore, the leading asset pricing models are able to capture the data on the long-term pricing kernel quite well. Given that we have shown in Section 4 that the long-term pricing kernel is downward sloping and exhibits little time-series variation. As we will show, the direction of the time-series variation in the models is dependent on the calibration, therefore it is not a strong test for the models. To compare the predictions of the models to the data, we focus on the size of the time-series variation.

5.1 Habit model

We use the calibration from Campbell and Cochrane (1999), see Appendix A.1 for the details of the models.. Given that in the Habit model the only stochastic variation comes from consumption growth, for which we have the distribution, we can calculate the pricing kernel as a function of market return. Mapping the pricing kernel in the model to the market return is the theoretical analogue of our empirical analysis. The time-series variation is obtained by conditioning on the surplus-ratio.

Figure 11: The average pricing kernel, and its time-series variation of the Habit model of Campbell and Cochrane (1999).

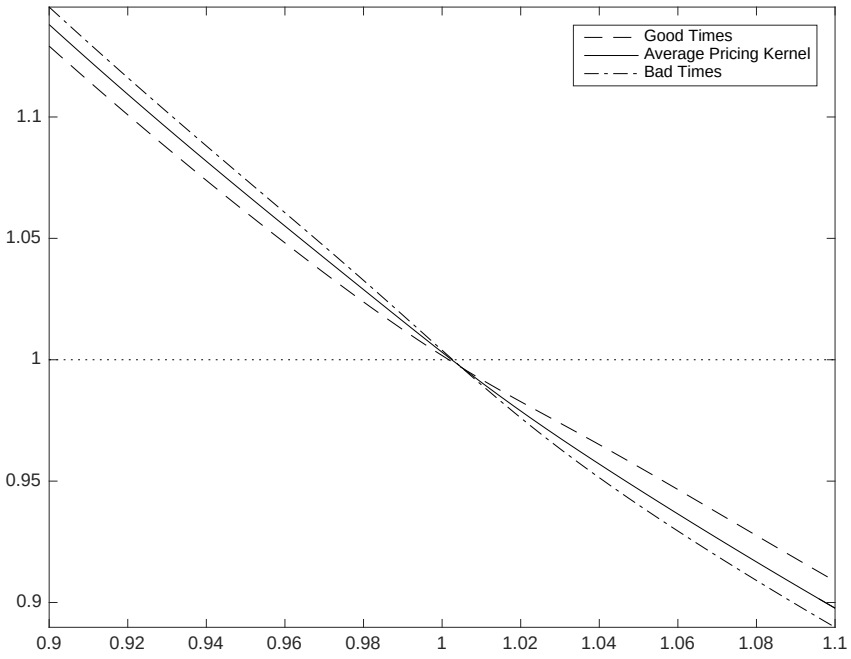


As expected, the average pricing kernel is monotonically decreasing in the market return. If the surplus ratio is larger than the long term average, current consumption exceeds the habit level more, implied risk aversion is lower which decreases the curvature. If the surplus ratio is larger than the average, it is an indication of good times and vice versa. Quantitatively, the time-series variation in the habit model is very small compared to the data on the one-month pricing kernel. Given that the time-series variation is small, there is not a lot of variation in the term structure either.

5.2 Rare disaster model

The next model we discuss is the rare disaster model by Wachter (2013), we use the discrete-time equivalent of Dew-Becker et al. (2017) see Appendix A.2 for the details. We are interested in the average and time-series variation of the pricing kernel as a function of the market return. The result is given in the following graph. The time-series variation is obtain by conditioning on the probability of disaster.

Figure 12: Average pricing kernel and its time-series variation from the disaster model of Wachter (2013) with the calibration from Table 6.

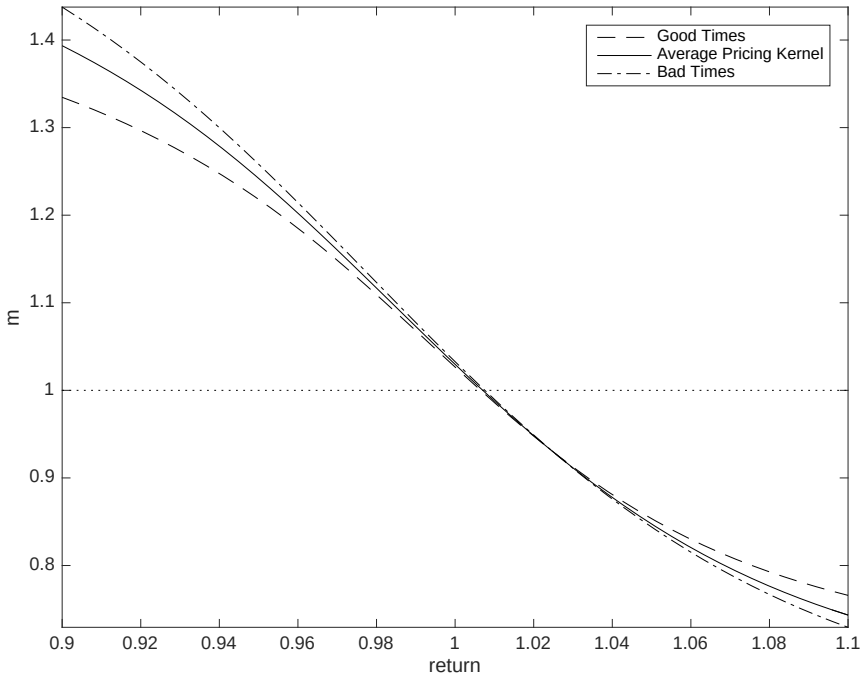


Again, we find that the pricing kernel is monotonically decreasing in market returns. In the model the pricing kernel is marginally affected by an increase in disaster probability, yielding a flat term structure. If the probability is larger than the long term average, indicating bad times, the curvature increases marginally. We are able to show that the level of the pricing kernel in a log-linearized version of the model, is increasing in the disaster probabillity for the current calibration. Furthermore, also in this model the time-series variation is minimal, therefore the term-structure does not vary much either.

5.3 Long run risk model

The next model we consider is the long run risk model introduced by Bansal and Yaron (2004), see Appendix A.3 for the details. We represent the average pricing kernel as a function of the market return. The time-series variation is obtained by conditioning at the same time on mean consumpion growth and stochastic volatility.

Figure 13: Average pricing kernel and its time-series variation from the long run risk model of Bansal and Yaron (2004) with the calibration from Table 7.

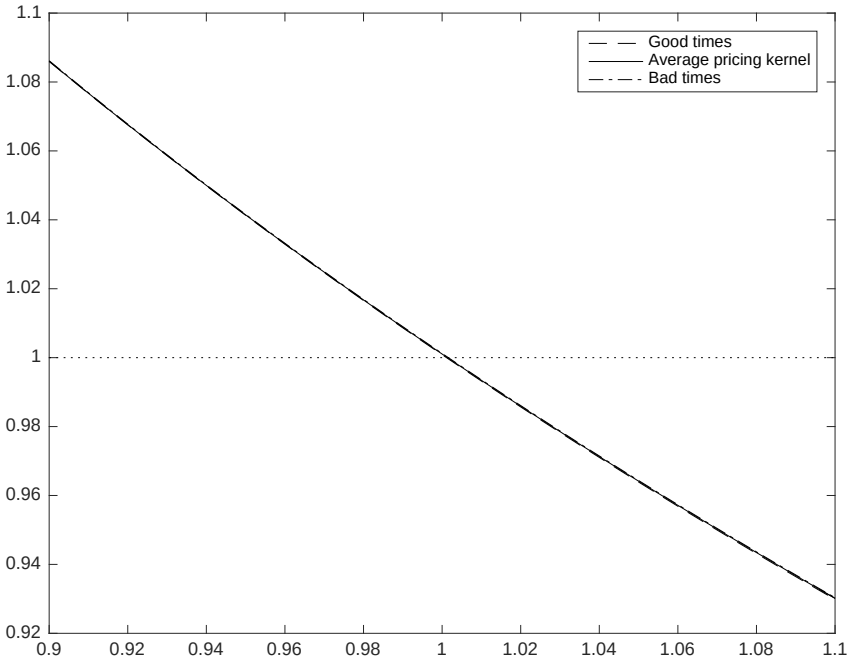


For the long-run risk model the pricing kernel is monotonically decreasing in market returns. The curvature of the pricing kernel incerases in bad times, when volatility is low and consumption growth large. We show in a simple version of the model that the level of the pricing kernel is decreasing in the level of the consumption growth in the calibration by Bansal and Yaron (2004). Only the opposite holds when we use the calibration of Bansal et al. (2012), the correlation of consumption and dividend growth is driving the differences. Furthermore, given the persistence of the state variables and the little time-series variation the term structure is flat.

5.4 Time-varying recovery model

The last model we discuss is the time-varying recovery model by Gabaix (2012) see Appendix A.4. The pricing kernel in the time-varying recovery model is standard power utility pricing kernel. Therefore, there is no time-series variation as the state-variables do not enter the pricing kernel equation as in the previously discussed models. The average kernel of the model are given in the next Figure.

Figure 14: The average pricing kernel and its time-series variation of the time-varying recovery model using the parameters of table 8.



Indeed, the model does not exhibit any time-series variation as the pricing kernel is i.i.d. due to the power utility assumption. The term structure in this model will be flat as well due to the power utility preferences.

6 Conclusion

In line with previous work, we find that the short-term pricing kernel is U-shaped. Our result that the U-shape is stronger in good times is new to the literature. We show that stylized facts of the short-term pricing kernel are incomprehensible with leading asset pricing models. However, the long-term pricing kernel slopes downward monotonically and exhibits much less time-series variation. Theoretical asset pricing models are able to match the data on long-term pricing kernls well.

7 Bibliography

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

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

Bakshi, G., Madan, D. and Panayotov, G. (2010). Returns of claims on the upside and the viability of u-shaped pricing kernels. *Journal of Financial Economics*, 97(1):130 – 154.

Bansal, R. and Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4):1481 – 1509.

Bansal, R., Kiku, D. and Yaron, A. (2012). An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review*, 1(1):183–221.

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

Bollerslev, T., Tauchen, G. and Zhou, H. (2009). Expected stock returns and variance risk premia. *The Review of Financial Studies*, 22(11):4463–4492.

Breeden, D. and Litzenberger, R. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621–651.

Britten-Jones, M. and Neuberger, A. (2000). Option prices, implied price processes, and stochastic volatility. *The Journal of Finance*, 55(2):839–866.

Campbell, J. and Cochrane, J. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205 – 251.

Chabi-Yo, F. (2012). Pricing kernels with stochastic skewness and volatility risk. *Management Science*, 58(3):624 – 640.

Chabi-Yo, F., Garcia, R. and Renault, E. (2007). State dependence can explain the risk aversion puzzle. *The review of Financial studies*, 21(2):973–1011.

Christoffersen, P., Heston, S. and Jacobs, K. (2013). Capturing option anomalies with a variance-dependent pricing kernel. *Review of Financial Studies*, 26(8):1963 – 2006.

Cuesdeanu, H. and Jackwerth, J. (2018a). The pricing kernel puzzle: Survey and outlook. *Annals of Finance*, 14(3):1–41.

Cuesdeanu, H. and Jackwerth, J. (2018b). The pricing kernel puzzle in forward looking data. *Review of Derivatives Research*, 21(3):1–24.

Dew-Becker, I., Giglio, S., Le, A. and Rodriguez, M. (2017). The price of variance risk. *Journal of Financial Economics*, 123(2):225–250.

Epstein, L. and Zin, S. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *The Quarterly Journal of Economics*, 127(2):645–700.

Jackwerth, J. (2000). Recovering risk aversion from option prices and realized returns. *The Review of Financial Studies*, 13(2):433–451.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, 47:263–292.

Linn, M., Shive, S. and Shumway, T. (2017). Pricing kernel monotonicity and conditional information. *The Review of Financial Studies*, 31(2):493–531.

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

Rosenberg, J. and Engle, R. (2002). Empirical pricing kernels. *Journal of Financial Economics*, 64(3):341–372.

Song, Z. and Xiu, D. (2016). A tale of two option markets: Pricing kernels and volatility risk. *Journal of Econometrics*, 190(1):176–196.

van Binsbergen, J. and Koijen, R. (2017). The term structure of returns: Facts and theory. *Journal of Financial Economics*, 124(1):1 – 21.

van Binsbergen, J., Brandt, M. and Koijen, R. (2012). On the timing and pricing of dividends. *American Economic Review*, 102(4):1596–1618.

Wachter, J. (2013). Can time varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, 68(3):987 – 1035.

A Appendix

A.1 Habit model

Consider the Habit model of Campbell and Cochrane (1999), the k-period pricing kernel at time t is given by the following equation:

$$M_t^k = e^{-k\delta}\Big(\frac{C_{t+k}}{C_t}\Big)^{-\gamma}\Big(\frac{S_{t+k}}{S_t}\Big)^{-\gamma}, \tag{13}$$

where C_t and $S_t := (C_t - X_t)/C_t$ are the consumption and surplus ration at time t, respectively. Log consumption growth is assumed to be i.i.d. normally distributed and the surplus ratio evolved according to:

$$\log\Big(\frac{C_{t+1}}{C_t}\Big) = g + v_{t+1},$$
$$s_{t+1} = (1-\phi)\overline{s} + \phi s_t + \lambda(s_t)v_{t+1}.$$

Where v_{t+1} follows a $N\big(0, \sigma_v^2\big)$, $s_t = \log(S_t)$ and $\lambda(s_t)$ is given by:

$$\lambda(s_t) = \frac{1}{\overline{S}}\sqrt{1 - 2(s_t - \overline{s})} - 1,$$
$$\overline{S} = \sigma_v\sqrt{\frac{\gamma}{1-\phi-b/\gamma}}.$$

Note, $\lambda(s_t)$ is set to zero when $s_t > s_{max}$:

$$s_{max} = \overline{s} + \frac{1}{2}\big(1 - \overline{S}^2\big).$$

We solve the following identity given the above parameters, where P_t denotes the ex-dividend price of the claim to the future consumption stream:

$$\mathbb{E}_t\Big[M_t^1\Big(\frac{P_{t+1}}{C_{t+1}} + 1\Big)\frac{C_{t+1}}{C_t}\Big] = \frac{P_t}{C_t}.$$

If we are to assume that C_t is the dividend paid by the aggregate market, P_t/C_t is the price-dividend ratio. Note, Campbell and Cochrane (1999) also have a version of the model where the dividend process is not perfectly correlated to the consumption process.

We obtain the price dividend ratio as a function of the state variable, which is s_t for the habit model. Returns are obtain in the model via the following equation:

$$R^m_{t,t+1} = \frac{P_{t+1}/C_{t+1} + 1}{P_t/C_t} \cdot \frac{C_{t+1}}{C_t}.$$

In later models where we model dividends separately, we replace consumption with dividends.

A.2 Rare disasters model

In this section we will do the same analysis for the time-varying disaster model of Wachter (2013). We use the discrete equivalent of Dew-Becker et al. (2017). Log consumption growth in a discrete time version of this model is given by:

$$\begin{aligned}
\Delta c_t &= \mu_c + \sigma_c \epsilon_{c,t} + J_t, \\
\Delta d_t &= \eta \Delta c_t,
\end{aligned}$$

where J_t is the jump process, which we assume for simplicity to be a poisson mixture of normal distributions. It is given by:

$$\begin{aligned}
J_t &= \sum_{i=1}^{N_t} \xi_{i,t}, \\
\xi_{i,t} &\sim N\big(\mu_d, \sigma_d\big), \\
N_t &\sim Poisson(\lambda_t).
\end{aligned}$$

The intensity of the jump process follows:

$$\lambda_{t+1} = \phi\lambda_t + (1-\phi)\mu_\lambda + \sigma_\lambda \sqrt{\lambda_t}\epsilon_{\lambda,t}$$

The utility function and pricing kernel are given in this framework by:

$$v_t = (1-\beta)c_t + \frac{\beta}{1-\alpha}\log \mathbb{E}_t \exp\big(v_{t+1}(1-\alpha)\big),$$
$$M_{t+1} = \beta \exp\big(-\Delta c_{t+1}\big)\frac{\exp\big((1-\alpha)v_{t+1}\big)}{\mathbb{E}_t \exp\big((1-\alpha)v_{t+1}\big)},$$

which is a recursive utility function with EIS equal to one. Note, ϵ_c and ϵ_λ are standard normal distributed. The returns on the market in this model are given by:

$$R^m_{t,t+1} = \frac{P_{t+1}/D_{t+1}+1}{P_t/D_t}\cdot\frac{D_{t+1}}{D_t},$$

as the dividends are a leveraged claim on consumption in this model. Note, both the returns and pricing kernel are driven by shocks to consumption and shocks to the disaster probability. The montly calibration is in line with Wachter (2013), where only the disaster distribution differs. In the table the parameters are represented:

Table 6: Calibration of the Wachter (2013) model.

Parameter	Value	Parameter	Value
μ_c	$0.0252/12$	σ_c	$0.02/\sqrt{12}$
μ_d	-0.15	σ_d	0.10
μ_λ	$0.0355/12$	σ_λ	$0.067/12$
ϕ	$\exp(-0.08/12)$	β	$\exp(-0.012/12)$
η	2.6	γ	$3.0 = 1-\alpha$

A.3 Long run risk model

The next theoretical model that we consider is the long run risk model from Bansal and Yaron (2004) and we use the specification of Bansal et al. (2012) which is

given by:

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1}, \\
x_{t+1} &= \rho x_t + \varphi_e \sigma_t e_{t+1}, \\
\sigma^2_{t+1} &= \bar{\sigma}^2 + \nu\big(\sigma^2_t - \bar{\sigma}^2\big) + \sigma_w w_{t+1}, \\
\Delta d_{t+1} &= \mu_d + \phi x_t + \pi \sigma_t \eta_{t+1} + \varphi_u \sigma_t u_{t+1}.
\end{aligned}$$

As in the original specification, the representative agent is assumed to have Epstein and Zin (1989) recursive preferences and maximizes lifetime utility:

$$V_t = \Big[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta\big(\mathbb{E}_t[V_{t+1}^{1-\gamma}]\big)^{\frac{1}{\theta}}\Big]^{\frac{\theta}{1-\gamma}}.$$

The calibration we use is from Bansal and Yaron (2004) and represented in the following table. Note, $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$.

Table 7: Calibration of the long run risk model.

Parameter	Value	Parameter	Value
μ_c	0.0015	μ_d	0.0015
$\bar{\sigma}$	0.0078	σ_w	0.0000023
ρ	0.979	φ_e	0.044
ϕ	3	π	0
φ	4.5	δ	0.998
γ	10	ψ	1.5
ν	0.987		

A.4 Time-varying Recovery

The next theoretical model we consider is the time-varying recovery model of Gabaix (2012), we use the same specification as in Dew-Becker et al. (2017) and is

specified as follows:

$$\begin{aligned}
\Delta c_t &= \mu_c + \sigma_c \epsilon_{c,t} + J_{c,t}, \\
L_t &= \big(1 - \rho_L\big)\bar{L} + \rho_L L_{t-1} + \sigma_L \epsilon_{L,t}, \\
\Delta d_t &= \lambda \sigma_c \epsilon_{c,t} - L_t \cdot \mathbb{1}_{J_{c,t} \neq 0}.
\end{aligned}$$

Table 8: Calibration of Dew-Becker et al. (2017).

Parameter	Value	Parameter	Value
μ_c	$0.01/12$	σ_c	$0.02/\sqrt{12}$
μ_d	-0.3	σ_d	0.15
$\bar{L}$	$-\log(0.5)$	σ_L	0.04
ρ_L	$0.87^{1/12}$	λ	5
β	$0.96^{1/12}$	γ	7

In the model the representative agent has power utility preferences. Corresponding to a pricing kernel with the following specification.

$$M_{t+1} = \beta \cdot \Big(\frac{C_{t+1}}{C_t}\Big)^{-\gamma}.$$

As one can see from the formulas, the pricing kernel is only driven by shocks to consumption, ϵ_c or disaster J_c. Whereas returns are driven by shocks to consumption ϵ_c and shocks to the severity of the impact for consumption disasters to dividends ϵ_L.