Erschienen in: Annals of Finance ; 14 (2018), 3. - S. 289-329.
https://dx.doi.org/10.1007/s10436-017-0317-9

The pricing kernel puzzle: survey and outlook

Horatio Cuesdeanu[1] · Jens Carsten Jackwerth[1]

Abstract It has been a while since the literature on the pricing kernel puzzle was summarized in Jackwerth (Option-implied risk-neutral distributions and risk-aversion, The Research Foundation of AIMR, Charlotteville, 2004). That older survey also covered the topic of risk-neutral distributions, which was itself already surveyed in Jackwerth (J Deriv 2:66–82, 1999). Much has happened in those years and estimation of risk-neutral distributions has moved from new and exciting in the last half of the 1990s to becoming a well-understood technology. Thus, the present survey will focus on the pricing kernel puzzle, which was first discussed around 2000. We document the pricing kernel puzzle in several markets and present the latest evidence concerning its (non-)existence. Econometric studies are detailed which test for the pricing kernel puzzle. The present work adds much breadth in terms of economic explanations of the puzzle. New challenges for the field are described in the process.

Keywords Pricing kernel puzzle · Stochastic discount factor · Options · S&P 500

We would like to thank George Constantinides, Maria Grith, and Alexandros Kostakis for helpful comments.

✉ Jens Carsten Jackwerth
jens.jackwerth@uni-konstanz.de

Horatio Cuesdeanu
horatio.cuesdeanu@uni-konstanz.de

1 University of Konstanz, PO Box 134, 78457 Constance, Germany

1 Introduction and a simple model of the pricing kernel

The pricing kernel m is of fundamental concern to all of modern finance as it is the basis for all pricing:

$$E[mR] = 1 \tag{1}$$

where E[] is the expectation under the physical (true) probabilities p across states and R is the return in each state. The pricing kernel m is the ratio of state prices π and physical probabilities p or, alternatively, of discounted risk-neutral probabilities (q/R_f) and physical probabilities:

$$m = \frac{\pi}{p} = \frac{q}{R_f p} \tag{2}$$

The pricing kernel informs us on how we need to adjust payoffs X such that we can take simple expectations in order to obtain the price of the security. It thus contains important information about the investor's assessment of different states: payoffs in states associated with low wealth/consumption are valued highly (m is large).

We can appreciate the link between the pricing kernel and preferences in a simple one-period economy. The representative investor maximizes the expected value of end of period utilities of consumption $U(C_i)$ in states i, where the investor is endowed with an initial wealth of w_0 and the utility function is concave. The investor can choose to collect h_i units of wealth in state i and, in equilibrium, consumption needs to equal collected wealth in each state. The optimization problem, including the budget constraint, reads then as:

$$\begin{aligned} &\max_{C_i,h_i} E\left[U\left(C\right)\right] = \max_{C_i,h_i} \sum_{i=1}^{N} p_i U\left(C_i\right) \\ &\text{s.t.} \sum_{i=1}^{N} h_i \pi_i \leq w_0 \\ &\text{and } C_i = h_i \quad \text{for } i = 1, \ldots, N, \end{aligned} \tag{3}$$

where p_i are the physical probabilities in states i and π_i are the state prices. Assuming an interior solution, we write the first order conditions as:

$$\pi_i = \frac{p_i U'(h_i)}{\lambda} \quad \text{for } i = 1, \ldots, N, \tag{4}$$

where λ is the Lagrange multiplier associated with the budget constraint. Note that neither the pricing kernel ($m_i = \pi_i / p_i = U'(h_i)/\lambda$) nor the risk-free rate are so far uniquely identified. To achieve such identification, we assume that the representative agent needs to hold all securities in equilibrium so that the collected wealth h_i needs to equal $w_0 R_i$, where R_i is the return on the market in state i. Then,

$$\pi_i = \frac{p_i U'(w_0 R_i)}{\lambda} \quad \text{for } i = 1, \ldots, N. \tag{5}$$

Summing up across states and further assuming the existence of a risk-free security means that the sum $\sum_{i=1}^{N} \pi_i = (1/R_f) = \frac{\sum_{i=1}^{N} p_i U'(w_0 R_i)}{\lambda}$, where R_f is one plus the risk-free rate. This allows us to identify λ and to express the pricing kernel, while substituting for λ, as:

$$m_i = \frac{\pi_i}{p_i} = \frac{U'(w_0 R_i)}{R_f \sum_{k=1}^{N} p_k U'(w_0 R_k)} \quad \text{for } i = 1, \ldots, N. \tag{6}$$

Equation (6) informs us that the pricing kernel is proportional to marginal utility. Any insight into the pricing kernel thus translates into knowledge about aggregate investor preferences in our economy. In particular, standard concave utility functions, such as power and exponential utility, lead to positive and monotonically decreasing pricing kernels.

Empirically, it emerges that estimated ratios of risk-neutral and physical probabilities often exhibit non-decreasing parts. Such findings constitute the pricing kernel puzzle, and we will return to it promptly in the next section.

A second line of research uses market data to infer the (parametric) utility function of a representative investor. A starting point is the equity premium puzzle of Mehra and Prescott (1985). Here, a stylized economy with a representative investor economy with power utility is being calibrated to market data. The resulting risk aversion coefficients tend to be much too high when compared to survey based estimates. This literature has been continued in Kocherlakota (1996) and Mehra (2008), with international evidence added in Pozzi et al. (2010) and Dimson et al. (2012). Closely related is the work by Bartunek and Chowdhury (1997) who use power utility and Benth et al. (2010) who use exponential utility instead; both papers calibrate to options data.

2 The pricing kernel puzzle

We are more interested in a third approach, the direct estimation of the pricing kernel m via Eq. (2), which uses as inputs the physical distribution and the option-implied risk-neutral distribution. Bates (1996a, b) points out that the two stochastic processes seem to be incompatible. Finally, Jackwerth (2000), Ait-Sahalia and Lo (2000), and Rosenberg and Engle (2002) estimate the empirical pricing kernel by dividing the risk-neutral distribution by the physical distribution. They document the surprising result that the pricing kernel is locally increasing while a simple model such as the one in Eq. (6) suggests a monotonically decreasing pricing kernel. The pricing kernel puzzle was born.

We now document the pricing kernel puzzle in a worked example, which broadly follows Jackwerth (2000, 2004) and draws on additional papers to make basic points about the empirical implementation of the pricing kernel puzzle. The full range of papers on the topic will be surveyed in Sect. 3.

To start our investigation, we need an asset, which is highly correlated with overall wealth in the economy. For the US, the S&P 500 index is the asset of choice as it is seen as a reasonable proxy for the market return even though it does not cover all investment opportunities of a representative investor. It also comes with a deep and liquid market for options on the index, which we will need momentarily. Studying Eq. (2), we require three quantities, (one plus) the risk-free rate R_f, the risk-neutral probabilities p, and the physical probabilities q. Estimation of the interest rate is an easy task as the discounting effect is small over the typical horizons of 30–60 days.

2.1 How to estimate the risk-neutral distribution q?

Rubinstein's (1994) seminal article allowed for the first time to recover risk-neutral, option-implied distributions. Jackwerth and Rubinstein (1996) extended and applied that technique to the S&P 500 index options. Taking the last step of finding the empirical pricing kernel through dividing the risk-neutral probability distribution by the physical distribution seems obvious in retrospect but was not quite so clear at the time.

Estimation of the risk-neutral distribution is by now a well-established field of research and a large literature covers it, from which we summarize some papers in Table 1.

Given a sufficiently large cross section (more than 10 option strike prices), most methods perform relatively similarly and yield the desired risk-neutral distributions where one typically uses the SPX options on the S&P 500 index with maturities of 30–60 days.

Table 1 We list some papers on the extraction of risk-neutral densities from option prices and the use of such densities

Name of paper	Comments
Papers covered in the text	
Rubinstein (1994)	First risk-neutral density fit
Jackwerth and Rubinstein (1996)	Extended fit based on smoothness criterion
Jackwerth (2004)	Survey and new "fast and stable" method
Haerdle et al. (2015)	Analyzes the effect of errors in the raw data; kernel based method
Papers not covered in the text	
Jackwerth (1999)	Survey
Bahra (1997)	Survey, mixture of two lognormals
Melick and Thomas (1997)	Mixture of three lognormals
Ait-Sahalia and Lo (1998)	Kernel regression of implied volatilities
Bliss and Panigirtzoglou (2002, 2004)	Splines, based on Shimko (1993)
Figlewski (2010)	
Fengler and Hin (2015)	B-splines
Ludwig (2015)	Neural networks
EZB (2011)	Applications using risk-neutral densities
Carr et al. (2002)	Based on CGMY Levy-process
David and Veronesi (2014)	Risk-neutral volatilities and macro variables
Martin (2017)	Expected market return and risk-neutral volatility
Christoffersen et al. (2013)	Survey on forecasting with risk-neutral information

In the recent literature, curve fitting of the implied volatility has become the most popular starting point for backing out risk-neutral distributions. To illustrate, we review the fast and stable curve fitting approach of Jackwerth (2004). First, the option prices observed in the market are converted to implied volatilities $\{\bar{\sigma_i}\}_{i=1}^{I}$, where $\bar{\sigma_i}$ denotes the Black-Scholes implied volatility of an option with strike price K_i. Second, one chooses a smoothness parameter λ and solves the following optimization problem to obtain an implied volatility curve $\{\sigma\}_{j=1}^{J}$ on a fine grid:

$$\min_{\sigma_i} \frac{\Delta^4}{2\,(J+1)} \sum_{j=0}^{J} \left(\sigma_j''\right)^2 + \frac{\lambda}{2I} \sum_{i=0}^{I} (\sigma_i - \bar{\sigma_i})^2 \tag{7}$$

The coarseness of the grid is given by Δ, which defines the distance between two consecutive strike prices, and σ_j'' denotes the numerical approximation to the second derivative of the implied volatility curve. Equation (7) has a straightforward closed form solution for $\{\sigma\}_{j=1}^{J}$, see Jackwerth (2004). By varying over λ one can choose a reasonable trade-off between smoothness (sum of the second derivatives over j) and fit (sum of the squared errors over i). Finally, the smooth implied volatilities curve is translated back to a call option price curve, whose second derivative is the compounded risk-neutral density:

$$q(K_i) = R_f \frac{\partial^2 Callprice(K)}{\partial K^2}|_{K=K_i} \tag{8}$$

The relationship in Eq. (8) was first established in Breeden and Litzenberger (1978).

Care needs to be taken in implementing Eq. (8), as a smooth solution depends heavily on the spacing Δ. Using a \$5 spacing as in the market data leads to jagged solutions, see for example Barone-Adesi and Dall'O (2010). Using half or a quarter of \$5 leads to much better results.

Turning attention from the impact of spacing to the impact of pricing errors, Haerdle et al. (2015) pick up on the older work of Bliss and Panigirtzoglou (2002) and analyze the impact of errors in option prices or implied volatilities on the empirical pricing kernel. Both the risk-neutral distribution and the physical distribution are being obtained through kernel based techniques, which allows Haerdle et al. (2015) to describe the uniform confidence bands around the empirical pricing kernel in statistical terms.

In Fig. 1 we present representative risk-neutral distributions for the S&P 500 in the US, the DAX 30 in Germany, the FTSE 100 in the UK, and the Nikkei 225 in Japan. We also depict the physical distributions, which we discuss next.

2.2 How to estimate the physical distribution p?

Those physical distributions in Fig. 1 are based on 38- to 45-day, non-overlapping returns of the S&P 500 index within moving, 4-year historical windows. The horizons of the returns (38–45 days) are chosen such that they match the maturity of the underlying options. The returns are then smoothed through a kernel density estimator where the bandwidth is chosen according to Silverman's (1986) rule of thumb.

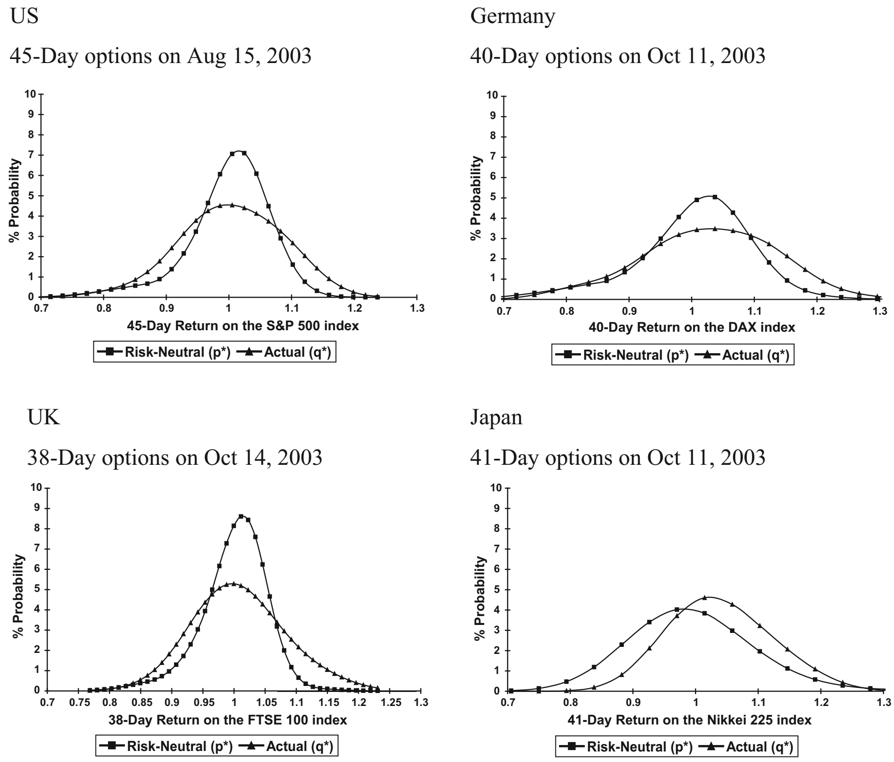

Fig. 1 Risk-neutral and actual distributions. The actual distributions are calculated with the same return horizon as the time-to-expiration of the options. For the US we used the historical sample from Sep 2, 1997 to Aug 15, 2003, for Germany from Jan 2, 1997 to Oct 9, 2003, for the UK from Jan 2, 1997 to Oct 9, 2003, and for Japan from Jan 5, 1998 to Oct 10, 2003. Returns are reported as 1 plus the rate of return

Kernel densities do not rely on any distributional assumptions except the stationarity of the returns and were used in Jackwerth (2000, 2004) and Ait-Sahalia and Lo (2000). However, in the presence of time-varying volatility and structural breaks, more recent papers have turned to GARCH models. Thus, Rosenberg and Engle (2002), Barone-Adesi et al. (2008), and Barone-Adesi and Dall'O (2010) all fit the Glosten et al. (1993) GARCH model (GJR GARCH) to historical returns.[1]

In particular, the daily log return r_t is modeled as the sum of a constant μ and an error term ε_t:

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \mu + \varepsilon_t \tag{9}$$

The error term is given by $\varepsilon_t = z_t\sigma_t$, where z_t is standard normal and the volatility process σ_t is recursively defined by:

[1] For the fit of a continuous time Levy-processes see Carr et al. (2002).

$$\sigma_t^2 = \kappa + \gamma \sigma_{t-1}{}^2 + \alpha \varepsilon_{t-1}{}^2 + \mathbb{1}_{(\varepsilon_{t-1}<0)} \zeta \varepsilon_{t-1}{}^2 \tag{10}$$

where κ, γ, α, and ζ are model parameters, which are usually estimated by maximum likelihood. In contrast to ordinary GARCH models, the GJR GARCH can account for the leverage effect by treating positive and negative shocks differently through the indicator function $\mathbb{1}_{(\varepsilon_{t-1}<0)}$.

A perennial problem for estimating the physical distribution is the so-called Peso problem. What if the historical returns do not include an event (say a crash), but investors have longer memories and incorporate such fears into their subjective distributions? Luckily, Jackwerth (2000) argues that a peso problem cannot explain the pricing kernel puzzle, since for the first 4 years past the crash, the crash is "visible" in the physical distribution based on the historical returns. Still, the results do not change compared with periods where the crash is no longer visible because, on the date of the investigation, the crash lies more than 4 years into the past. In a theoretical setting, Ziegler (2007) confirms the point that a Peso problem cannot explain the pricing kernel puzzle. See Sect. 4.1 for details.

2.3 Possible shapes of the pricing kernel and statistical evidence

After dividing the risk-neutral distribution by the physical distribution, we obtain the empirical pricing kernels, which are depicted in Fig. 2. Note the tilde-shaped hump around at-the-money, which is inconsistent with Eq. (6) according to which the empirical pricing kernel is monotonically decreasing in returns since it is proportional to the marginal utility of a risk-averse investor. For such a risk-averse investor, utility is concave and marginal utility is decreasing. Moreover, equilibrium is ruled out as a non-decreasing pricing kernel implies the existence of a portfolio that stochastically dominates the market, see Sects. 5 and 6. A non-decreasing pricing kernel hence clashes with our basic intuitions and contradicts most standard market models. The violation of monotonicity has been labeled as the "pricing kernel puzzle," and we will investigate possible explanations in Sect. 5.

When looking at the empirical pricing kernels from the beginning of this research area, one observes various shapes at different points in time. Figure 3 shows a tilde-shaped pricing kernel in 1993, a u-shape in 1999, and w-shaped pricing kernels in 2004 and 2013.

To understand the different shapes, we refer to the empirical findings of Cuesdeanu (2016), who examines the S&P 500 from 1988 to 2015 and finds that (i) missing out-of-the money calls, (ii) misestimated subjective probabilities, and (iii) a time varying variance risk premium all contribute to the empirical shapes. (i) If deep out-of-the-money calls cannot be observed, one has to make assumptions about the right end of the implied volatility curve. The right end of the pricing kernel then reacts very sensitively to small changes in the implied volatility curve. However, when deep out-of-the-money calls are observed, the pricing kernel turns out to be increasing at the right end. (ii) The time-series model for estimating the subjective density matters in particular for the right end of the pricing kernel. Models with a fat right tail and thus a pricing kernel, which decreases at the right end, fit the data more poorly than models with a thin right

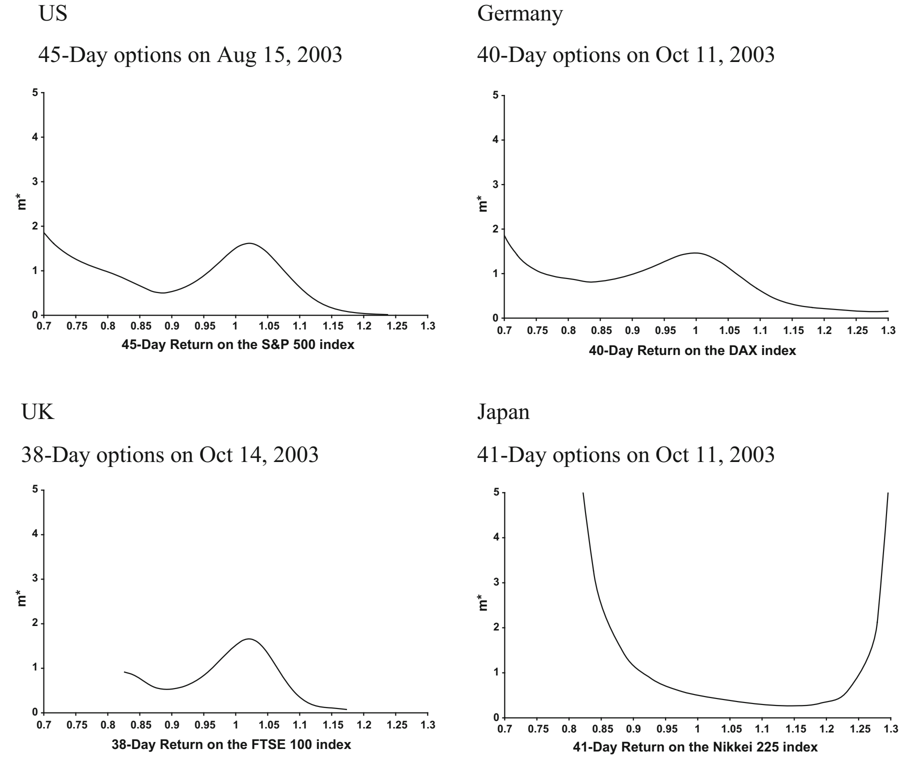

Fig. 2 Empirical pricing kernels. Typical post-1987 stock market crash implied pricing kernels. The pricing kernels are calculated as the ratio of the option implied risk-neutral distribution and the historical smoothed return distribution. Returns are reported as 1 plus the rate of return

tail and thus a pricing kernel, which increases at the right end. Hence, issues (i) and (ii) imply either w- or u-shaped pricing kernels as opposed to tilde-shaped ones. (iii) Last, obtaining sometimes w-shaped and sometimes u-shaped pricing kernels can be explained by a time varying variance risk premium. Pricing kernels tend to be u-shaped in times of high uncertainty (variance risk premium is high) and w-shaped in calm periods (variance risk premium is low). Moreover, tilde-shaped pricing kernels tend to emerge during calm periods when no out-of-the-money calls are observed, the fit to these options is poor, or the right tail of the subjective density is overestimated. Finally, monotonically decreasing pricing kernels can emerge during volatile periods when no out-of-the-money calls are observed or the fit to these options is poor.

2.4 Economics of the pricing kernel puzzle

Now that we have established the pricing kernel puzzle, we turn to the economics of the puzzle and its solutions. Taking the pricing kernel puzzle to be literally true, however, seems like a naïve interpretation. In that case, the representative investor of the simple economy in Sect. 1 would need to have a convex segment in the utility

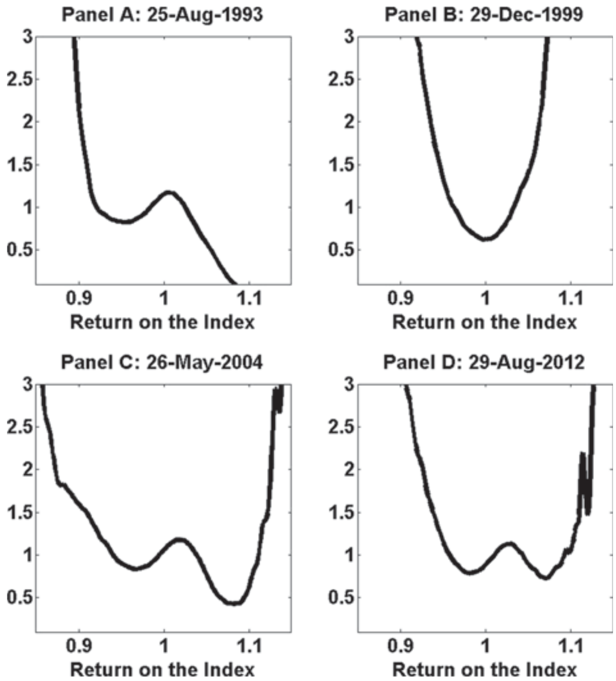

Fig. 3 Empirical pricing kernels at different points in time. The figure shows a tilde-shaped pricing kernel in 1993, panel **A**, a u-shaped pricing kernel in 1999, panel **B**, and w-shaped pricing kernels in 2004 and 2013, panel **C** and **D** the subjective distributions are estimated by a GJR-GARCH(1,1) and the risk-neutral densities are obtained by the fast and stable method of Jackwerth (2004) as introduced in Eq. (7)

function, akin to the Friedman and Savage (1948) utility function in Fig. 4.[2] For a representative investor, this is hard to reconcile with equilibrium. It would mean that the representative investor was better off by not investing into states of world, where the index pays off when the utility function is convex. Rather, the representative investor would prefer a lottery over the two adjacent states (0.9 and 1.1 on the return axis of Fig. 4) where the utility function turns concave again. But such avoidance of states jars with the notion that the representative investor needs to hold all assets by definition. Rather, security prices need to adjust so that the representative investor is willing to hold all assets in equilibrium. This point is made more rigorously in Hens and Reichlin (2013).

Looking at the equilibrium problem for a different angle, Beare (2011) works out, based on some earlier results by Dybvig (1988), measure preserving derivatives

[2] Note that Friedman and Savage (1948) introduced their utility function for individuals and not for the representative investor. In particular, their concern was with small stakes gambling such as buying a lottery ticket. Chetty and Szeidl (2007) provide a microeconomic motivation for Friedman–Savage utility via consumption commitments (e.g. housing), for which the spending cannot easily be adjusted. Again, this is a model of individual investors, and it is not obvious that the convexities would survive aggregation to a representative investor. See Ingersoll (2014) for related results on another partially convex utility function, namely cumulative prospect theory.

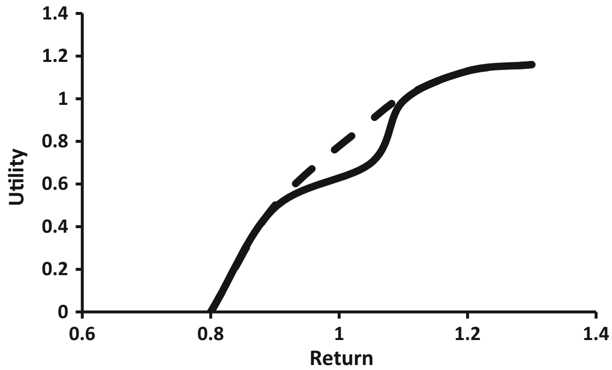

Fig. 4 Friedman and Savage (1948) utility function. We depict a utility function along the lines of Friedman and Savage (1948) with concave–convex–concave segments in the return dimension as a solid curve. We also depict the concavified version of the utility function in the center by a dashed curve

which any investor should prefer to investing into the market (see also Rieger 2011). Their prices are less than the price of the market in times where the pricing kernel puzzle exists, and Beare and Schmidt (2015) show that returns on an option portfolio exploiting this circumstance actually stochastically dominate market returns. While such disequilibrium could well exist for some period in time, it is hard to see how such a situation could persist unabated ever since the crash of 1987.

There exists a close connection between pricing kernels and the concept of stochastic dominance, which expresses dominance relations between probability distributions on which all investors of a certain class agree. Our setting of positive and decreasing pricing kernels uses the class of risk-averse investors (i.e., those with concave utility functions), and the corresponding concept is the one of second order stochastic dominance.

We look at solutions in more detail in Sect. 4. Based on the above reservations about single state variable models, much interest centers on multiple state variable models, where the pricing kernel is monotonic in several dimensions such as index return and volatility. Projecting the multivariate pricing kernel onto the index return dimension can then lead to the pricing kernel puzzle. A promising alternative approach is the demand based model of Bollen and Whaley (2004), who explicitly model the portfolio insurance demand of investors for out-of-the-money puts on the index. In particular, their model can reconcile the moderate implied volatility smiles for stocks with the steep smiles for the index.

The pricing kernel puzzle also touches on a number of related economic concepts. First, Eq. (1) governs not only index returns but also index option returns. We can thus also investigate vestiges of the pricing kernel puzzle in option returns; a discussion which we follow in Sect. 3.2. Second, a complementary aspect to the pricing kernel puzzle is the problem of bounds on option prices, detailed in Sect. 5. Here, the maximal and minimal option prices are found, which are still consistent with a class of particular pricing kernels (say, those of risk-averse investors). Rather than focusing on the pricing kernel exactly consistent with observed option prices (the focus of the pricing kernel puzzle), the interest here is on the most restrictive class of pricing kernel, which can just explain observed option prices. The classes of pricing kernels (say, those of risk-averse investors) often have alternative expressions in terms of stochastic dominance relations (here, of second order stochastic dominance).

Finally, we mention two applications of the pricing kernel puzzle. Kostakis et al. (2011) use the option-implied, risk-neutral distribution for the S&P 500 index and the assumption of an exponential or power utility to obtain forward-looking physical distributions. They basically apply the methodology of Bliss and Panigirtzoglou (2004) to a dynamic asset allocation problem. The risk aversion coefficient is iteratively estimated up to time t in order to make a prediction of the physical distribution at time t+1. They find that the forward-looking physical distributions produce better portfolios than the historical distributions, even though the approach ignores the pricing kernel puzzle by design.[3] It would be interesting to see if a more flexible pricing kernel would outperform the pricing kernels based on exponential and power utility functions.

3 Does the pricing kernel puzzle exist?

Most of the work on the pricing kernel puzzle investigates the S&P 500 index, and there are recent additions to this literature. The pricing kernel puzzle exists in the returns on the index and also in the returns on options on the index. A large number of studies have subsequently investigated if the pricing kernel puzzle also exists in other indices and have largely confirmed this finding for a number of large indices (e.g. the DAX and the FTSE). Little is known about the time-series properties of the pricing kernel puzzle. Finally, we turn to investigations of the pricing kernel puzzle in markets other than index markets. The main issue here is that the pricing kernel is now the projection of the economy-wide pricing kernel onto the space of returns investigated (say returns on gold). Depending on the correlation between the index (proxying for aggregate wealth) and gold (as a possible return under investigation), the projected pricing kernel might not exhibit any puzzling behavior.

3.1 Yes, the pricing kernel puzzle exists in index markets

Three early papers establish the pricing kernel puzzle. Using monthly S&P 500 index options from 1986 through 1995, Jackwerth (2000) suggested to approximate the risk aversion function $-U''(R_i)/U'(R_i))$ directly as $(p'/p)-(q'/q)$, which is a positive function as long as the utility function is concave and marginal utility is positive. The risk aversion function turns out to be more complicated as opposed to the more straightforward pricing kernel (Eq. 2). A number of robustness checks confirm the result that the empirical risk aversion functions are u-shaped and negative around at-the-money during the post-87-crash period, while they are mainly positive and decreasing during the pre-crash period. As a (locally) negative risk aversion function implies a (locally) increasing pricing kernel, the pricing kernel puzzle emerges.

[3] Zdorovenin and Pezier (2011) use a close variant, too, and are subject to the same critique as Kostakis et al. (2011).

Ait-Sahalia and Lo (2000) derived the pricing kernel independently of Jackwerth (2000) as the ratio of the risk-neutral distribution (obtained via the method of Ait-Sahalia and Lo 1998) and the physical distribution obtained through a kernel based estimator. Based on half-yearly returns during the year 1993 they can document the pricing kernel puzzle. The authors very graciously delayed publication so that their paper would not appear in print before the publication of Jackwerth (2000) which was started earlier but was long delayed at the journal.

The third of the canonical models, which are typically cited to establish the pricing kernel puzzle, is Rosenberg and Engle (2002). Using monthly data from 1991 to 1995 on the S&P 500 index options, they start by obtaining the physical distribution from the parametric GJR GARCH model of Glosten et al. (1993) fitted to historical returns. They next specify the pricing kernel parametrically, which allows them to obtain the risk-neutral distribution and thus derive model-implied option prices. The parameters of the pricing kernel are optimized such that the sum of squared option pricing errors is being minimized. A monotonically decreasing pricing kernel is being fitted, but mispricing can be much reduced when more flexible functional forms for the pricing kernel are allowed, leading to the pricing kernel puzzle yet again.

3.1.1 Testing the pricing kernel puzzle

The canonical models provide bounds around the pricing kernel estimates simply based on the sample variation of the inputs, namely, historical returns and option prices, and those bounds do not constitute formal tests of monotonicity. Using the bounds suggests that the estimated pricing kernels exhibit local increases exceeding those bounds. The main finding of Jackwerth (2000) is presented in his Fig. 3 where the risk aversion functions are negative by more than two standard deviations. Ait-Sahalia and Lo (2000) provide the 5 and 95% quantiles around their pricing kernel, and, by visual inspection, the upper quantile at an index value of 400 is very close to the lower quantile at an index value of 435. While this argument is not a formal statistical test, it is still highly suggestive of the presence of the pricing kernel puzzle. Rosenberg and Engle (2002) document the pricing kernel puzzle in their Fig. 6, which shows a clear local increase in the pricing kernel beyond the two standard deviation bounds.

A careful study of small sample noise in both the physical and the risk-neutral distribution is Leisen (2014).[4] He finds that spurious non-monotonicities can arise for simulations of power utility pricing kernels. The problem is particularly relevant if the physical distribution is based on historical samples of only 48 monthly returns, and the situation improves much once a GARCH(1,1) model is estimated. Also, the risk-neutral distributions are based on Ait-Sahalia and Lo's (2000) kernel-based method, which is noisier than other methods for backing out risk-neutral distributions from option prices.

The complicated issue of formally testing for locally increasing segments of the estimated pricing kernel has been taken up in Golubev et al. (2014) under the strong assumption of iid realized returns. The idea is to map the problem to an exponential

[4] See also Lioui and Malka (2004) for reported differences due to using either only call or only put options.

model and check for pricing kernel monotonicity between any two realized returns in the sample. The fairly complicated test then considers the joint distribution of monotonicity violations across all possible combinations of observed returns. Applying their test to the DAX index during the summers of 2000, 2002, and 2004, monotonicity could be rejected at the 10% significance level in 2002, but not for the years 2000 and 2004.

Another test is Haerdle et al. (2014), which uses the market model of Grith et al. (2017). Here, the pricing kernel is parameterized as two decreasing segments with some breakpoint where the pricing kernel jumps up or down. Comparing GMM estimates of the restricted model (the two segments join smoothly in a decreasing manner) versus the unrestricted model, the authors employ a so-called D-test and reject pricing kernel monotonicity in typically four out of five cases.

A further attempt at designing a formal statistical test is Beare and Schmidt (2014) who base their test on the equivalence of the monotonicity of the pricing kernel and the concavity of the ordinal dominance function. The latter function is the cumulative risk-neutral distribution of the quantile function of the physical distribution. They find that in about half the months from 1997 to 2009, the pricing kernel puzzle can be detected at the 5% significance level. We collect tests in Table 2.

Cuesdeanu and Jackwerth (2017) suggest a simpler test based on risk-neutral distributions, which have been divided by some pricing kernel to find the physical probability distributions at each observation date. Working out the quantiles of the observed market returns under the physical cumulative distribution function, the quantiles throughout the sample should be standard uniformly distributed, see Bliss and

Table 2 We list statistical tests of pricing kernel monotonicity

Name of paper	Comments
Papers covered in the text	
Jackwerth (2000)	Using in-sample two standard deviation bounds
Ait-Sahalia and Lo (2000)	5 and 95% quantiles
Rosenberg and Engle (2002)	Using in-sample two standard deviation bounds
Leisen (2014)	Theoretical and simulated aspects of noise in pricing kernel estimation
Golubev et al. (2014)	Assumes iid returns; formal test of monotonicity violations
Haerdle et al. (2014)	Tests for increasing breakpoint between two decreasing segments of the pricing kernel
Beare and Schmidt (2014)	Concavity test of the ordinal dominance function
Cuesdeanu and Jackwerth (2017)	Test for uniformity of the percentiles of observed returns under the physical cumulative distribution function
Papers not covered in the text	
Shive (2003)	Bootstrap test
Shive and Shumway (2004)	Bootstrap test
Patton and Timmermann (2010)	Monotonic relation test for asset returns. Has not been used empirically

Panigirtzoglou (2004), Diebold et al. (1998) and Diebold et al. (1999). The authors then optimize several test statistics of uniformity while either restricting or not restricting the pricing kernel to be monotonically decreasing. The discrepancy in optimized test statistics can then be tested against its simulated distribution. Cuesdeanu and Jackwerth (2017) confirm the presence of the pricing kernel puzzle in the S&P 500 index options data from 1987 to 2015.[5] Note that the paper, as opposed to the earlier canonical studies (which mix backward-looking estimates of the physical distribution with forward-looking risk-neutral distributions in order to finally find the empirical pricing kernel as the ratio q/p), uses only forward-looking data, namely, the physical returns are forward-looking and no longer based on historical samples.

3.1.2 Further studies on the S&P 500, other indices, and time series properties of the pricing kernel puzzle

Most of the initial studies use the S&P 500 index. We summarize a number of follow-up studies on the S&P 500 and other indices (DAX 30, FTSE 100, and others) in Table 3.

Song and Xiu (2016) add information about the VIX level when estimating empirical pricing kernels for the S&P 500 using kernel based methods akin to Ait-Sahalia and Lo (2000). They confirm the pricing kernel puzzle unconditionally, but cannot establish it conditionally on high or low VIX levels. Thus, they speculate that stochastic volatility could be driving the pricing kernel puzzle but find that standard stochastic volatility option pricing models cannot generate the observed patterns.[6] See also Sect. 4.2 which suggests solutions to the pricing kernel puzzle based on volatility as a second state variable.

Very interesting are the following two studies on the DAX which try to explain the time series properties of the pricing kernel puzzle. First, Giacomini et al. (2008) use tick data for the DAX from January 1999 to April 2002 and fit a GARCH model in order to obtain the physical distribution. The risk-neutral distribution estimation follows Ait-Sahalia and Lo (2000). Then, time series of simple statistics of the pricing kernel plus the absolute and relative risk aversion functions at different maturities are being calculated and subjected to a principle component analysis. The principle components are finally regressed on returns on the DAX and on changes in at-the-money implied volatility. The main result seems to be the rather obvious finding that large changes in implied volatility lead to more volatile and time-varying pricing kernels.

Similarly, but using a slightly different technique, Grith et al. (2013) use DAX data between April 2003 and June 2006. They fit a smoothing polynomial to the implied volatilities, translate those into option prices, and use Breeden and Litzenberger (1978)

---

[5] Compare Linn et al. (2014), who can only establish the pricing kernel puzzle for the FTSE 100 but not for the S&P 500. Cuesdeanu and Jackwerth (2017) attribute this result to (i) a lack of scaling so that the physical distributions of Linn et al. (2014) are not integrating to one and (ii) a mismatch in their optimization (based on moments of the uniform distribution via GMM) and their measurement of fit (based on the Cramer van Mises statistic).

[6] In particular, they find that the empirical volatility pricing kernel is u-shaped; a fact that is not captured by any option pricing model so far. A related observation by Boes et al. (2007) is that that the risk-neutral distribution, conditional on a low spot volatility, does not exhibit negative skewness.

Table 3 We list empirical papers, which document the pricing kernel puzzle in index markets

Name of paper	Market studied	Comments
Papers covered in the text		
Jackwerth (2000)	S&P 500	Locally increasing pricing kernel
Ait-Sahalia and Lo (2000)	S&P 500	Locally increasing pricing kernel
Rosenberg and Engle (2002)	S&P 500	Locally increasing pricing kernel
Song and Xiu (2016)	S&P 500	Also use information on VIX
Giacomini and Haerdle (2008)	DAX	Locally increasing pricing kernel
Grith et al. (2013)	DAX	Locally increasing pricing kernel
Coval and Shumway (2001)	S&P 500	Use of option returns and not prices
Broadie et al. (2009)	S&P 500	Use of option returns and not prices
Chaudhuri and Schroder (2015)	S&P 500	Use of option returns and not prices
Bali et al. (2017)	S&P 500	Use of option returns and not prices
Bakshi and Madan (2007)	S&P 500	Use of option returns and not prices
Bakshi et al. (2010)	S&P 500	Use of option returns and not prices
Papers not covered in the text		
Figlewski and Malik (2014)	S&P 500	Locally increasing pricing kernel
Hill (2013)	S&P 500	Locally increasing pricing kernel
Yang (2009)	S&P 500	
Audrino and Meier (2012)	S&P 500	Uses B-splines for the pricing kernel
Carr et al. (2002)	13 stocks and 8 indices, including S&P 500	Fits Levy-processes, u-shaped pricing kernel
Wu (2006)	S&P 500	Extends Carr et al. (2002)
Belomestny et al. (2017)	DAX	
Dittmar (2002)	20 industry portfolios	u-shaped pricing kernel
Schweri (2010)	30 industry portfolios	u-shaped pricing kernel
Shive (2003)	S&P 500, DAX, FTSE	Locally increasing pricing kernel
Shive and Shumway (2004)	S&P 500, DAX, FTSE, OMX Sweden	Locally increasing pricing kernel in the unconstrained version

Table 3 continued

Name of paper	Market studied	Comments
Shive and Shumway (2009)	S&P 500, DAX, AMEX Japan	Subsumes Shive (2003) and Shive and Shumway (2004), same results
Fengler and Hin (2015)	S&P 500	Locally increasing pricing kernels on one day for several maturities
Golubev et al. (2014)	DAX	Locally increasing pricing kernel in 6/2002, but not in 6/2000 or 6/2004
Detlefsen et al. (2010)	DAX	Locally increasing pricing kernels in bear, "sideways," but not bull market
Haerdle et al. (2014)	International cross sections of 20 stocks each	Locally increasing pricing kernel
Liu et al. (2009)	FTSE	Locally increasing pricing kernel
Perignon and Villa (2002)	CAC 40 France	Locally negative risk aversion function, implying the puzzle
Coutant (1999)	CAC 40 France	u-shaped risk aversion function

to obtain risk-neutral distributions. The physical distributions are based on 2 years' worth of historical returns via kernel density estimation. Finally, power utility functions are extended with four additional parameters (additive and multiplicative parameters inside and outside the power function) to allow for non-monotonic pricing kernels. Changes in these parameters and the location of the peak of the pricing kernel are being regressed on changes in the credit spread, the yield curve slope, and the short interest rate, as well as the underlying return. The authors conclude from the correlations between those macro variables and the additional shape parameters that the locally risk loving behavior is pro-cyclical as the hump of the empirical pricing kernel seems to be more pronounced in calm periods.

3.2 Yes, the pricing kernel puzzle exists in index option returns

So far we studied the pricing kernel puzzle in terms of returns of the underlying security, often a broad index such as the S&P 500. But Eq. (1), which we repeat here, also holds for option returns:[7]

$$E[mR] = 1 \tag{11}$$

[7] For a study on forecasting option returns, see Israelov and Kelly (2017).

The return on a call option (R_{call}) with strike price K is the payoff $(S - K)^+$ divided by the price of the option, which is $E[m(S - K)^+]$. We thus start our discussion by looking at the expected return on a call option under the physical measure:

$$E\left[R_{call}\right] = \frac{E\left[(S-K)^+\right]}{E\left[m\,(S-K)^+\right]} \tag{12}$$

Under the assumption of a monotonically decreasing pricing kernel, call returns should be positive and increasing in moneyness as, intuitively speaking, the pricing kernel in the denominator shifts mass to the region where the call payoff is zero. A stronger result is presented in Coval and Shumway (2001): the expected return on a call should be greater than the expected return on the underlying, which broadly holds in the data.[8] Coval and Shumway (2001) derive the relation between the derivative of the expected call price with respect to the strike price and the covariance between the pricing kernel and the asset price:

$$\frac{\partial E[R_{call}]}{\partial K} = \frac{-Cov[E[m|s],\ S-K|S>K]}{c}, \tag{13}$$

where c is a positive term that depends on the pricing kernel and the underlying but does not influence the result. From Eq. (13) and assuming that the pricing kernel is negatively correlated with the underlying, it becomes clear that expected call returns will be increasing in the strike price. This monotonicity result, and the fact that a call with a strike price of zero corresponds to the underlying, explains why the expected call return should exceed the return on the underlying.

The authors then investigate returns on option straddles and find evidence of priced volatility risk, which they cannot reconcile with power utility for the representative investor. This evidence is consistent with the pricing kernel puzzle but does not outright prove the case. Broadie et al. (2009) caution using unscaled option returns, which tend to be so noisy that one cannot even reject the assumption that the returns were being generated by the Black-Scholes model. Such findings strongly suggest scaling option returns in a suitable way (e.g. straddles as above or by standardizing betas as in Constantinides et al. 2013).

Chaudhuri and Schroder (2015) extend the results of Coval and Shumway (2001) by showing that the pricing kernel is only monotonically decreasing if (conditional) expected returns on certain option positions (called "log-concave" and encompassing long calls, puts, butterfly spreads, and others) increase in the strike price. They confirm the pricing kernel puzzle based on data for the S&P 500 index but fail for individual stock options. This is expected due to the much flatter implied volatility smiles of the individual stock options. Another extension in Bali et al. (2017) looks at the higher risk-neutral moments of option returns. Song (2012) applies the ideas of Coval and

[8] Branger et al. (2011) do not confirm their result in more recent data, thus documenting the presence of the pricing kernel puzzle in the data. They further argue that stochastic volatility, stochastic jump option pricing models, which also have jumps in the volatility process, can explain those call option returns.

Shumway (2001) to returns on options on volatility in the case of u-shaped pricing kernels.

Bakshi and Madan (2007) present a market model where the pricing kernel is u-shaped since a group of pessimistic investors are shorting the market index. In addition, these investors buy call options as an insurance against a rising index and, hence, are willing to pay a premium for the calls; for more details, see Sect. 4.1. Consistent with such market model, Bakshi et al. (2010) find evidence for a positive dependence between short-selling activity and expected call returns. Looking again at Eqs. (12 and 13), it is clear that a u-shaped pricing kernel directly implies that expected returns of call options with a strike above a certain threshold are negative and decreasing in the strike price. Bakshi et al. (2010) document evidence for such a u-shaped pricing kernel.[9]

3.3 No, the pricing kernel puzzle does not exist with overly restricted pricing kernels

While the canonical early papers backed out the pricing kernel, other researchers tried to find the forward looking physical probabilities by assuming a functional form for the pricing kernel. However, imposing severe restrictions on the pricing kernel can lead to estimates which will than no longer exhibit the pricing kernel puzzle despite its presence in the data. E.g., Chernov and Ghysels (2000) fitted the Heston (1993) model to S&P500 index returns and option prices. While the paper provides expressions for the pricing kernel, it is not immediately clear that the pricing kernel puzzle can be generated altogether, given the restrictive choice of only two constant risk premia (one for the market and one for volatility), which account for the parameter differences between the physical and the risk-neutral versions of the model. We collect papers, which overly restrict the pricing kernels in Table 4.

A second line of investigation, which specifies the utility function to be of power or exponential type, is also inherently not able to document the pricing kernel puzzle. The leading exponents are Bliss and Panigirtzoglou (2004) who start out with the risk-neutral distribution obtained from option prices, which they change into the physical distribution through division by the pricing kernel, which is given by the marginal utility of either a power or exponential utility function. As the parametric utility functions lead to monotonically decreasing pricing kernels, Bliss and Panigirtzoglou (2004) could not document the pricing kernel puzzle even if it were present in the data.

The prevailing thought is that only one of the three quantities, namely risk-neutral probabilities, physical probabilities, and the pricing kernel, can be backed out from the other two. Yet Ross (2015) argues that it would be preferable to use only risk-neutral information, as that is well estimated, and infer both the forward looking physical distribution and the pricing kernel. His insight is that this can be achieved if all risk-neutral transition probabilities are known, as opposed to only the risk-neutral

[9] The empirical evidence is consistent with Branger et al. (2011); compare for the theoretical results also Chaudhuri and Schroder (2015).

Table 4 Models with overly restricted pricing kernels

Name of paper	Type of constraint	Comments
Papers covered in the text		
Chernov and Ghysels (2000)	SVSJ	
Bliss and Panigirtzoglou (2004)	POWER/EXP	
Ross (2015)	ROSS RECOVERY	
Jackwerth and Menner (2015)	ROSS RECOVERY	Tests Ross (2015)
Jensen et al. (2016)	ROSS RECOVERY	
Papers not covered in the text		
Pan (2002)	SVSJ	Extends Bates (2000)
Bates (2008)	SCSJ	
Santa-Clara and Yan (2010)	SVSJ	
Duan and Zhang (2014)	POWER/EXP	
Weber (2006)	POWER/EXP	Uses collateralized debt obligations
Backus et al. (2011)	POWER/EXP	Assumes that the Merton (1976) model holds
Kang and Kim (2006)	POWER/EXP	
Benth et al. (2010)	POWER/EXP	
Bates (2012)	POWER/EXP	
Coutant (2000)	POWER/EXP	Based on the CAC 40 France
Lioui and Malka (2004)	POWER/EXP	Based on the TA-25 Israel
Stutzer (1996)	MAX ENTROPY	Extended by Alcock and Smith (2014) using Haley and Walker (2010)
Barone-Adesi et al. (2008)	RND SHAPE	Contradicted by the very similar paper Barone-Adesi et al. (2013), which finds the puzzle
Barone-Adesi and Dall'O (2010)	RND SHAPE	
Sala (2016)	RND SHAPE	
Sala and Barone-Adesi (2016)	RND SHAPE	
Audrino et al. (2015)	ROSS RECOVERY	
Jensen et al. (2016)	ROSS RECOVERY	

Types of constraint are stochastic volatility, stochastic jump models (SVSJ), power and exponential utility (POWER/EXP), maximum entropy approaches (MAX ENTROPY), shape restrictions on the risk-neutral density (RND SHAPE), and Ross (2015) recovery based approaches (ROSS RECOVERY)

distribution. The difference is that the risk-neutral distribution is one single distribution emanating from the initial (known) state and indicating the (risk-neutral) probability of moving to a future state. The risk-neutral transition probabilities are richer and also indicate the risk-neutral probabilities of moving from all hypothetical initial states to all future states, see Fig. 5.

While the approach is theoretically very appealing, Ross (2015) requires some strong assumptions, which severely restrict possible pricing kernels, even though those

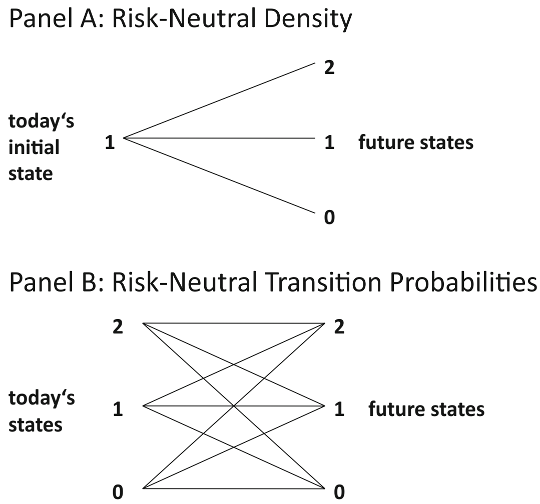

Fig. 5 Risk-neutral probabilities versus risk-neutral transition probabilities. In panel **A** we depict the typical situation of a tree emanating from today's initial state (1) and moving to several future states (0, 1, and 2). In panel **B**, we depict the data requirements of Ross (2015) where, in addition, one also needs to know the (hypothetical) transition probabilities from alternative states today (0 and 2) to all future states

assumptions do not outright preclude the existence of the pricing kernel puzzle.[10] Jackwerth and Menner (2015) study the empirical implementation of the Ross (2015) recovery and find a number of intractable problems. Such problems lead to poorly estimated pricing kernels and physical probability distributions. Jackwerth and Menner (2015) test these physical distributions based on the realized returns, which supposedly stem from them, and strongly reject the proposed physical distributions,[11] whereas the assumption that physical distributions can be estimated by using historical return distributions cannot be rejected.

Jensen et al. (2016) develop a recovery framework that makes no assumption on the underlying probability distribution and allows for a closed-form solution. Practical implementation relies only on current option prices for different maturities, and, hence, there is no need for a full matrix of transition distributions as in the Ross (2015) model. Empirically, they find that their recovered physical return distribution has some predictive power, although they stress that their empirical implementation primarily has an illustrative purpose. Applying the Berkowitz (2001) test to the realized returns, they reject the hypothesis that the recovered distribution is equal to the true physical distribution.

[10] Carr and Yu (2012) replace the assumptions on the utility function of a representative investor by assuming that the dynamics of the numeraire portfolio under the physical measure are being driven by a bounded diffusion. Walden (2017) extends Ross (2015) recovery to unbounded diffusion processes and Huang and Shaliastovich (2014) to the state dependent, recursive preferences of Epstein and Zin (1989). Schneider and Trojani (2015) suggest recovery based on assumptions on the signs of risk premia on different moments of market returns.

[11] This point is also made in Borovicka et al. (2015) who attribute these problems to "misspecified recovery," which happens when the pricing kernel has non-trivial martingale components.

3.4 No, the pricing kernel puzzle does not exist in non-index asset markets

First a word of caution on computing the empirical pricing kernel for non-index assets altogether. If one adheres to some notion of preferences over consumption, then a concentration on the index makes much sense. After all, consumption should be correlated with wealth and that in turn is driven to a large extent by the evolution of large indices such as the S&P 500. But considering an asset such as gold makes much less sense. As always, one investigates the projection of the economy-wide pricing kernel onto a particular return dimension (here gold). But as gold has a low correlation with the stock market and thus with consumption and wealth, we have no clear prediction of the shape of such projected pricing kernel in the gold dimension: a low gold price is not related to low stock market prices (poor state of the world, low consumption, high risk aversion) nor is the opposite true for high gold prices. Thus, pricing kernels on non-index assets might well turn out to be disappointingly flat and with little room for interpretation. We summarize such approaches in Table 5.

The situation would be different for asset classes more highly correlated with the index. Moreover, for a careful, bivariate analysis of the pricing kernel puzzle, one

Table 5 We list papers, which present projected pricing kernels in non-index markets

Name of paper	Market studied	Comments
Papers covered in the text		
Jackwerth and Vilkov (2017)	S&P 500 and Volatility (VIX)	Bivariate model
Ni (2009)	Equities	Monotonically decreasing pricing kernels
Chaudhuri and Schroder (2015)	Equities	Locally increasing pricing kernels
Figlewski and Malik (2014)	Exchange traded funds	Locally increasing pricing kernels
Papers not covered in the text		
Shive and Shumway (2009)	Commodities	u-shaped projected pricing kernels
Haas et al. (2012)	Foreign exchange	Use transformation of Liu et al. (2007)
Li and Zhao (2009)	Interest rates (Libor)	Use estimator of Ait-Sahalia and Duarte (2003)
Liu et al. (2015)	Interest rates (Libor)	
Kitsul and Wright (2013)	Inflation (TIPS)	u-shaped projected pricing kernels
Song and Xiu (2016)	Volatility (VIX)	u-shaped projected pricing kernels
Bakshi et al. (2015)	Volatility (VIX)	u-shaped projected pricing kernels
Chernov (2003)	Index, equities, gold, interest rates (T-bills)	Estimates the pricing kernel in several dimensions

would need to estimate bivariate risk-neutral distributions, which is exceedingly difficult as there are few options written on both assets at the same time (knowing only options on one asset and options on the other asset separately is typically not enough), and bivariate physical distributions. Jackwerth and Vilkov (2017) have recently made inroads here in estimating the bivariate risk-neutral distribution on the S&P 500 and the VIX, using longer-dated options to circumvent the above problem in this special set-up.

Considering non-index asset classes, the individual stocks take up a halfway position as they are the constituents of the index. Ni (2009) and Chaudhuri and Schroder (2015) analyze individual stock options within the S&P 500. Chaudhuri and Schroder (2015) find evidence of return patterns compatible with the pricing kernel puzzle and criticize the earlier paper of Ni (2009), which cannot find such evidence, for methodological reasons. Details can be found in Sect. 3.4. Similarly, the work of Figlewski and Malik (2014) is based on option data on exchange traded funds having the S&P 500 as an underlying. Due to the high correlation with the S&P 500, we do not really view this exchange traded fund as a non-index asset. Not surprisingly, their work finds non-monotonic pricing kernels. By considering exchange traded funds that aim to provide (i) twice the return on a long position in the S&P 500 and (ii) twice the return on a short position, they also contribute to the literature on heterogeneous investors and the pricing kernel puzzle, see Sect. 4.1.

4 Solutions

Considering the empirical evidence and the statistical tests so far, it emerges that the pricing kernel puzzle seems to be present in the data. We will now investigate models which try to explain the pricing kernel puzzle. We start with models using only a single state variable, then the important class of models with more than one state variable, before turning to behavioral and sentiment models, and finally to ambiguity aversion models. We collect such models in Table 6.

4.1 Models with a single state variable

Another way of extending the simple setting of Sect. 1 is to replace the representative investor with several (classes of) heterogeneous investors. In the simplest case, there are two groups of investors, pessimists believing that the mean return will turn out to be low and optimists believing that the mean return will turn out to be high. We depict such a situation in Fig. 6, where the aggregated subjective distribution then turns out to be bi-modal. Given a typical risk-neutral distribution, the pricing kernel puzzle obtains. Note that the pricing kernel puzzle critically depends on the bi-modality of the subjective distribution. Adding more moderate groups of investors in the center would undo the bimodality and could result in monotonically decreasing pricing kernels. Hence, such extremely bimodal subjective distributions seem to be unrealistic, and the main challenge in this strand of the literature is to formally derive the three objects of interest (m, p, and q) under aggregation.

Table 6 Models of the pricing kernel puzzle

Name of paper	Type of model	Comments
Papers covered in the text		
Brown and Jackwerth (2012)	SINGLE STATE	
Bakshi and Madan (2007)	SINGLE STATE	
Ziegler (2002)	SINGLE STATE	
Ziegler (2007)	SINGLE STATE	
Hens and Reichlin (2013)	SINGLE STATE	
Figlewski and Malik (2014)	SINGLE STATE	
Grith et al. (2017)	SINGLE STATE	
Christoffersen et al. (2013)	SEVERAL STATES	
Chabi-Yo (2012)	SEVERAL STATES	
Bakshi et al. (2015)	SEVERAL STATES	
Garcia et al. (2003)	SEVERAL STATES	
Chabi-Yo et al. (2008)	SEVERAL STATES	
Bollen and Whaley (2004)	BEHAVIORAL	
Garleanu et al. (2009)	BEHAVIORAL	
Kliger and Levy (2009)	BEHAVIORAL	
Polkovnichenko and Zhao (2013)	BEHAVIORAL	
Dierkes (2013)	BEHAVIORAL	
Chabi-Yo and Song (2013)	BEHAVIORAL	
Gollier (2011)	AMBIGUITY	
Kang et al. (2014)	AMBIGUITY	
Drechsler (2013)	AMBIGUITY	
Cuesdeanu (2016)	AMBIGUITY	
Papers not covered in the text		
Bakshi et al. (2010)	SINGLE STATE	Based on Bakshi and Madan (2007), u-shaped pricing kernel
Ziegler (2002)	SINGLE STATE	Similar to Ziegler (2007) but with only two extreme investors
Siddiqi and Quiggin (2016)	SINGLE STATE	
Haerdle et al. (2009)	SINGLE STATE	Early version of Grith et al. (2017)
Detlefsen et al. (2010)	SINGLE STATE	
Babaoglu et al. (2016)	SEVERAL STATES	Based on Christoffersen et al. (2009), u-shaped pricing kernel
Bollerslev and Todorov (2011)	SEVERAL STATES	u-shaped pricing kernel
Dittmar (2002)	SEVERAL STATES	

Table 6 continued

Name of paper	Type of model	Comments
Kiesel and Rahe (2017)	SEVERAL STATES	
Yamazaki (2017)	SEVERAL STATES	u-shaped pricing kernel
Han and Turvey (2010)	SEVERAL STATES	
Lundtofte (2010)	SEVERAL STATES	Locally increasing pricing kernel
Andreou et al. (2014)	BEHAVIORAL	Based on Han (2008)
Hodges et al. (2008)	BEHAVIORAL	
Gemmill and Shackleton (2005)	BEHAVIORAL	

Type of model refers to models with a single state variable (SINGLE STATE), models with several state variables (SEVERAL STATES), behavioral models (BEHAVIORAL), and ambiguity aversion models (AMBIGUITY)

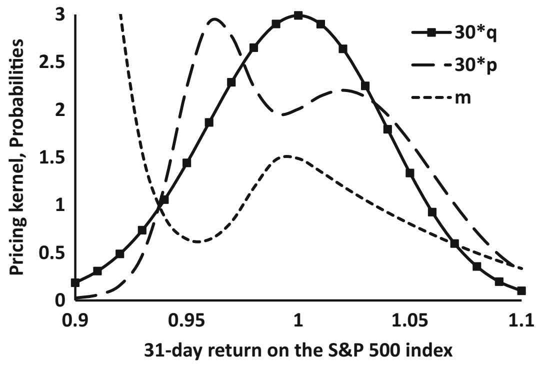

Fig. 6 Hypothetical pricing kernels with investor heterogeneity. We depict a risk-neutral distribution (q, scaled up 30 times for better readability) and a subjective distribution (p, also scaled up 30 times for better readability). The subjective distribution is a mixture of the beliefs of the pessimists (low expected mean return) and the optimists (high expected mean return). The pricing kernel m obtains as the ratio of risk-neutral by subjective probabilities. For simplicity and ease of depiction, we assume a zero interest rate

Bakshi and Madan (2007) assume heterogeneity in beliefs in a complete market. Investors have different subjective distributions (instead of homogeneous belief in the physical distribution); consequently, investors expecting positive returns are long in the market, while investors expecting negative returns are short. The aggregation of both groups of investors can lead to a u-shaped pricing kernel.

Ziegler (2002) uses a very similar set-up and can show that even the risk-neutral distribution can become bi-modal, if the beliefs are strongly heterogeneous.[12] He documents negative relative risk aversion functions, consistent with the pricing kernel puzzle.

[12] In such setting, Shefrin (2008a, b) coins the term sentiment for the ratio of the mixture of the different subjective distributions and the physical distribution. His ideas become clearer when one assumes that the shapes of the subjective distributions and the physical distributions remain the same but the mean is low for the pessimists, high for the optimists, and in between for the physical distribution, see Shefrin (2008b, Fig. 1).

Ziegler (2007) examines a complete market with multiple investors and assumes that the index is a good proxy for consumption. His results indicate that neither (i) aggregation of (heterogeneous) preferences, (ii) misestimation of beliefs, nor (iii) heterogeneous beliefs can lead to reasonable explanations of the pricing kernel puzzle. He shows that, given reasonable individual utility functions, aggregation of heterogeneous preferences alone cannot explain the puzzle as the economy-wide risk-aversion inherits the behavior of the individual risk-aversions. In order to deal with misestimated beliefs, the stochastic volatility, stochastic jump model of Pan (2002) is considered.[13] Fitting the model to the data and assuming that investors have homogeneous beliefs but cannot estimate them correctly, Ziegler (2007) argues that the resulting misestimation is too severe to be credible.

When allowing for heterogeneity among beliefs, Ziegler (2007) needs a large share of investors with very pessimistic beliefs to explain the puzzle. Hence, a fat left tail can only be captured if some investors expect extremely negative returns. However, a setting with three groups of investors is only capable of generating the pricing kernel puzzle if two of the groups are unrealistically pessimistic. Ziegler (2007) then already suggests that a solution of the pricing kernel puzzle needs to go beyond the rather simple setting of a complete, frictionless market with a single state variable.

In a two dates exchange economy with a finite number of states, Hens and Reichlin (2013) systematically examine violations of three basic assumptions of their model (namely, risk-averse behavior, unbiased beliefs, and complete markets). All three relaxations can then generate the pricing kernel puzzle. Quite obviously, allowing for a partially convex utility function (e.g., Friedman and Savage 1948) will generate the pricing kernel puzzle by design. However, a representative investor would not allocate wealth to states where the utility function is convex, and the relaxation is thus unrealistic.

Biased beliefs are modeled in two ways by Hens and Reichlin (2013). First, as humans tend to overweigh less probable extreme events, beliefs could be systematically distorted according to the model of Tversky and Kahneman (1992). Second, beliefs could be biased as different investors fashion different subjective forward-looking distributions based on the same historical return distribution. In isolation, both types of biased beliefs are incapable of explaining the puzzle. However, by combining both types, the authors can generate the pricing kernel puzzle, although only at the cost of assuming a negative expected mean return for the representative investor. Finally, Hens and Reichlin (2013) introduce background risk as a form of market incompleteness. In a simple four state example, two investors facing background risk individually can generate the pricing kernel puzzle.

The plausibility of heterogeneous beliefs and preferences is considered in Figlewski and Malik (2014) from an empirical point of view. The authors examine options on an

[13] Although the model captures stochastic volatility and jumps, the risk-aversion functions turn negative for high return states, implying a u-shaped pricing kernel. Such behavior contradicts the standard assumption of a risk-averse representative investor. This leads to the question, if stochastic volatility, stochastic jump models are typically incapable of fitting the historical risk-neutral and physical distribution simultaneously, or if the assumptions on the functional form of the risk-premium parameters are mis-specified in such models.

exchange traded fund replicating the S&P 500 (SPY), on one that aims to provide the return on a two-times long position in the index (SSO), and on one that aims to provide the return on a two-times short position (SDS). Presumably, optimistic investors will buy the SSO fund; pessimistic investors the SDS. The paper then studies two extreme cases: (i) pricing kernels could be the same but not subjective distributions or (ii) pricing kernels could differ but all investors share the belief in the same physical distribution. It turns out that setting (i) explains the data better. Unfortunately, the set-up does not allow for intermediate settings between the extreme cases. Last, it is suggested that preferences within each group should be constant over time and the daily change in expectations stems from a change in the risk-neutral distributions.

As opposed to many of the above papers, which use equilibrium approaches to aggregate the individual investors' utility functions to a market-wide pricing kernel, some authors use rather ad-hoc assumptions in order to aggregate utility functions. Grith et al. (2017) piece together the pricing kernel from many segments, which (between reference points) are decreasing but can jump upwards at the reference points. Investors are allowed to have different reference points. Given a sufficient number of such reference points, the authors can generate a flexible pricing kernel specification, which can exhibit increasing parts. One can study its piece-wise nature in their Figs. 2 and 3 in detail. In their empirical section they find that the local maximum of the pricing kernel near at-the-money is more pronounced when the variance risk premium is low.

In conclusion, it seems rather hard to explain the pricing kernel puzzle with only one state variable. Moreover, there is always the nagging doubt of how a locally increasing segment of the pricing kernel can be reconciled with equilibrium. A representative investor would not want to hold securities that pay off in such states, and models with several (groups of) investors need to have rather strongly diverging beliefs (very pessimistic investors vs. rather optimistic ones), while ignoring the large mass of moderate investors in the middle.

4.2 Models with several state variables

One way out is being hinted at by Brown and Jackwerth (2012) who introduced the (weighted) average historical volatility as a new variable. While it is still deterministically driven by the return process (which technically makes it a single state variable model), it opens up the perspective of introducing additional state variables. The pricing kernel would then exist across those several dimensions, and the pricing kernel projected onto the return dimension might then exhibit the pricing kernel puzzle.[14]

To illustrate the additional flexibility in modeling economies when using multiple state variables, consider the situation with two state variables in Fig. 7, which is adapted from Brown and Jackwerth (2012).

[14] A number of papers show that such additional state variables seem to be empirically needed in order to explain option prices, see for example Buraschi and Jackwerth (2001), Coval and Shumway (2001), and Constantinides et al. (2013).

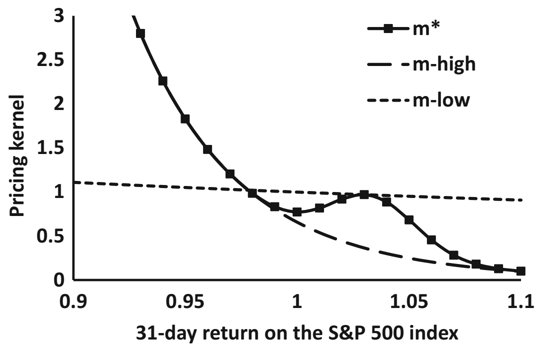

Fig. 7 Hypothetical pricing kernels depending on the second state variable volatility. We graph the simplest setting where volatility can take either of two values and we have a pricing kernel (m-high) in the high volatility state and another one (m-low) in the low volatility state. As wealth decreases or increases, the likelihood of being in the high volatility state increases, while for unchanged wealth (returns around 1) the likelihood of being in the low volatility state increases. Taking expectations of m over the two volatility states yields the desired empirical pricing kernel m*

Here it is assumed that the pricing kernel depends not only on the return (first state variable) but also on a second state variable. The most prominent choice of such a second state variable in the literature is probably volatility and therefore we stick to volatility in this example. In particular, it is assumed that volatility can either be high (m-high, with long dashes) or low (m-low, with short dashes). When returns are both very high or very low, volatility tends to be high, and the pricing kernel of the high state dominates. When returns are close to 1, volatility tends to be low, and the pricing kernel of the low state dominates. In line with the empirical evidence in Song and Xiu (2016), we model the conditional pricing kernels to be monotonically decreasing in returns. Taking the expectation over volatility yields a non-decreasing pricing kernel m* (solid line, with black squares) even though the two conditional pricing kernels were monotonically decreasing in returns.

Christoffersen et al. (2013) stay close to the above idea and extend the Heston and Nandi (2000) model by introducing a variance risk premium in addition to the equity risk premium. Similar to the setting of Fig. 7, the pricing kernel is now a function of returns and volatility. When projected onto returns only, by construction, a u-shaped pricing kernel emerges whenever the variance premium is negative. Fitting this GARCH model to the historical time series and cross sections of Wednesday options on the S&P 500 from 1996 to 2009 while allowing for a variance premium, and hence for a u-shaped pricing kernel, improves the risk-neutral and physical fit substantially. The quadratic functional form of the pricing kernel is rigidly assumed by the model and at times does not fit the empirical tilde-shaped pricing kernel in the empirical section of their paper.

Chabi-Yo (2012) shows that a recursive small-noise expansion results in a pricing kernel that incorporates stochastic volatility, stochastic skewness, and stochastic kurtosis, while an ordinary Taylor expansion would lead to a pricing kernel, which is a polynomial in the market return. Using French's 30 monthly industry portfolios, he recovers the higher moment preferences of the representative investor. His empirical pricing kernel is a function of volatility and return. Holding volatility fixed, it is monotonically decreasing in the market return. Yet, when projected onto the market return only, the empirical pricing kernel shows the puzzling behavior. For robustness, he shows that the pricing kernel projected onto the market return exhibits a similar shape if it is estimated with the S&P 500 option data rather than industry portfolio returns.

While most of the literature on heterogeneous beliefs and the pricing kernel focused on disagreement on the expected return (see e.g. Ziegler 2007; Hens and Reichlin 2013), Bakshi et al. (2015) consider heterogeneity with respect to future volatility and allow the investors with exponential utility to have different levels of risk-aversion, too. As a result, they obtain a u-shaped pricing kernel in the volatility dimension from options on VIX. In contrast, most standard models imply that the pricing kernel is monotonically increasing in volatility. Therefore, the model could potentially solve the pricing kernel puzzle as returns around zero are associated with low volatility, and low volatility on the other hand is associated with an increasing pricing kernel. Unfortunately, the paper does not explore this intriguing aspect.

Garcia et al. (2003) first introduced regime switches in the fundamental state variables of an equilibrium model and used this model to price options. Extending this work, Chabi-Yo et al. (2008) show that the pricing kernel puzzle can be explained by regime-switches in some latent state variable, which in turn drives fundamentals (the joint distribution of the pricing kernel and returns). Their model uses two preference specifications. For one a recursive Epstein and Zin (1989) utility and, alternatively, an external habit model with state dependence in the beliefs, which is based on Veronesi (2004) and Campbell and Cochrane (1999).[15] The intuition is that, conditional on the latent state variable, the pricing kernel is not violating the standard monotonicity assumption, whereas a projection of the pricing kernel onto returns leads to a locally increasing pricing kernel. Indeed, a simulation with hypothetical parameters can reproduce the desired shapes for the conditional and unconditional pricing kernels. One can note in the figures that the modeled pricing kernels often do not match the empirical pricing kernels in shape and magnitude. A more full-fledged empirical exercise might be able to improve the fit.

4.3 Behavioral and sentiment models

After first looking at demand based models, we next turn to models with probability weighting.

4.3.1 Demand based models

Bollen and Whaley (2004) come tantalizingly close to tackling the pricing kernel puzzle in their study of demand for out-of-the-money put options. They first establish that

---

[15] Benzoni et al. (2011) offer a similar model but do not show the model pricing kernel in the return dimension, and one cannot easily determine if it exhibits the pricing kernel puzzle; the pricing kernel in the dimension of consumption is monotonically decreasing by assumption.

the physical distributions for individual stocks and for the S&P 500 index are not that different. They then turn to the implied volatility smiles, which are mildly u-shaped for individual stock options and steeply skewed for the index. Their explanation is that strong investor demand for portfolio insurance exists for out-of-the-money index puts, but is weaker for individual stock option puts. The high demand for out-of-the-money index puts by institutional investors is only met with supply by the market makers at rather high prices, moving the implied volatilities up, and causing the steep smile. Having thus explained the cause of the steep index smile, they unfortunately do not connect their story to the pricing kernel puzzle, even though just one final argument is required. Namely, as the steep index smile leads to a left-skewed, leptokurtic risk-neutral distribution, the pricing kernel puzzle emerges once the risk-neutral distribution is being divided by the more normally distributed physical distribution. For the individual stock options, the mild smile leads to rather normally distributed risk-neutral distribution in the dimension of individual stock returns, and, thus, the pricing kernel puzzle does not emerge when dividing by the physical distribution.

Motivated by these empirical results, Garleanu et al. (2009) develop a demand based option pricing model by departing from no-arbitrage principles, considering the options market as being separated from the underlying, and highlighting the importance of the market maker. In the presence of jumps and stochastic volatility, market makers cannot fully hedge their exposures and will demand higher prices for options paying off in states where hedges are critical. Hence, the resulting implied volatility smile is increasing in regions where hedging is more difficult for the market maker, which mainly concerns out-of-the-money puts. Similarly to Bollen and Whaley (2004), they find that option end-users are typically long index puts and short single stock calls. Again, an explicit treatment of the pricing kernel is missing.

4.3.2 Models with probability weighting functions

Kliger and Levy (2009) revert the direction of investigation by starting with the pricing kernel puzzle, using power utility, and backing out the implied physical distribution from the risk-neutral distribution. As a result, the implied physical distribution inherits the left-skewed and leptokurtic shape of the risk-neutral distribution, which is incompatible with the physical distribution derived from bootstrapped past S&P 500 returns. Thus, they introduce a probability weighting function in order to reconcile the implied physical distribution with the bootstrapped distribution. The estimated probability weighting functions are inverse-S-shaped in their sample from 1986 to 1995.

To illustrate how probability weighting influences the pricing kernel, we extend the one period pricing kernel from Eq. (6) along the lines of Polkovnichenko and Zhao (2013):

$$m_i = \frac{q_i}{R_f p_i} = \frac{U'(w_0 R_i)\, Z(P_i)}{R_f \sum_{k=1}^{N} p_k U'(w_0 R_k)\, Z(P_k)} \quad \text{for } i = 1, \ldots, N. \tag{14}$$

Here, Z is a probability weighting function, which applies on the cumulative probability P_i of return R_i. Note that again the numerator is state dependent while the denominator is constant. Hence, the shape of the pricing kernel is no longer proportional to marginal utility but to the product of marginal utility and the probability weighting function Z. The probability weighting function in Polkovnichenko and Zhao (2013) stems from Prelec (1998) and is fully described by the parameters β and α:

$$Z(P) = \exp\left(-\left(-\beta \log(P)\right)^{\alpha}\right) \tag{15}$$

Polkovnichenko and Zhao (2013) repeat the study of Kliger and Levy (2009) on more recent data, using power utility with a risk aversion coefficient of two, and, for the physical distribution, using an EGARCH model based on past returns. Their probability weighting functions can be S-shaped (2004–2006) or inverse-S-shaped (during the remaining years from 1996 to 2008). The former suggests that investors overweigh probabilities in the center of the distribution and underweigh the tails, while the pattern reverses for the latter. It is somewhat puzzling that the pricing kernel puzzle tends to be rather stable through time but yields in this setting very different probability weighting functions. The model also does not account for learning; investors do not pay attention to the fact that the physical distribution, as it is being revealed in realized returns, looks different from the reweighted distribution.

Dierkes (2013) makes a nice point about the lack of identification in Polkovnichenko and Zhao (2013), as the utility function cannot be derived separately from the weighting function. He suggests an intriguing solution by fitting several maturities at the same time. That allows the utility function to be the same for all maturities but the weighting function scales with maturity. Empirically, Dierkes (2013) then finds the weighting function to be inverse-S-shaped and the utility function to be convex–concave around the zero percent return.

Chabi-Yo and Song (2013) confirm the findings of Polkovnichenko and Zhao (2009) and document that the probability weighting functions are heavily time-varying, even if they use the VIX as a conditioning variable. They thus extend the model and apply probability weighting to both the return and volatility dimensions of the index in a two period setting. Using S&P 500 and VIX options, they find inverse-S-shaped probability weighting functions, which are now much more stable in comparison with the single state variable model.

4.4 Ambiguity aversion models

Here, we propose a novel approach based on the smooth ambiguity aversion model of Klibanoff et al. (2005). The model nests on the one hand the traditional expected utility setting as the ambiguity aversion approaches ambiguity neutrality and on the other hand the maximin utility approach as the ambiguity aversion goes to infinity.[16] Gollier (2011) already mentions that the pricing kernel puzzle can emerge in a smooth ambiguity aversion setting, although without explicitly deriving the formulas and without detailed examples, which we are providing here. Kang et al. (2014) achieve similar results with a model where the representative investor is worried that some worst case

---

[16] For a survey of ambiguity aversion and its relevance for asset pricing, see Epstein and Schneider (2010). For an alternative formulation of ambiguity aversion through Choquet expected utility, see Bassett et al. (2004).

stock price process with lower drift might be realized. Drechsler (2013) extends the model of Liu et al. (2005) where a representative agent faces uncertainty aversion regarding jumps in the endowment process. It would be interesting to explicitly calibrate these models to option data and see if such economies imply a non-monotonic pricing kernel.

4.4.1 The theoretical pricing kernel under ambiguity aversion

We re-derive our simple economy from Sect. 1, Eqs. (3–6) in the setting of Klibanoff et al. (2005). They assume that there are M of the above economies (called an ambiguity setting), each with a probability p_j of occurring for j = 1,…,M. Our representative investor is thus solving the following problem:

$$\max_{C_{ij},h_{ij}} E\left[U\left(C\right)\right] = \max_{C_{ij},h_{ij}} \sum_{j=1}^{M} p_j \phi\left(\sum_{i=1}^{N} p_{ij} U\left(C_{ij}\right)\right)$$
$$\text{s.t.} \sum_{j=1}^{M}\sum_{i=1}^{N} h_{ij}\pi_{ij} \leq w_0$$
$$\text{and } C_{ij} = h_{ij} \text{ for } i = 1, \ldots, N \text{and for } j = 1, \ldots, M, \tag{16}$$

where C_{ij} is the consumption in state i of ambiguity setting j, h_{ij} is the chosen wealth, ϕ is a utility function across ambiguity settings which operates on the expected utility achieved in each ambiguity setting, p_{ij} is physical probability of state i occurring in ambiguity setting j, and w_0 is the initial wealth. Note that the physical probability of being in state i is the sum of (p_{ij} p_j) across ambiguity settings j. We defer the derivation to Internet Appendix.KMM and only state the resulting pricing kernel:

$$m_i = \frac{1}{\sum_{j=1}^{M} p_j p_{ij}} \times \frac{\sum_{j=1}^{M} p_j \phi'\left(\sum_{k=1}^{N} p_{kj} U\left(w_0 R_k\right)\right) p_{ij} U'(w_0 R_i)}{R_f \sum_{j=1}^{M}\sum_{s=1}^{N} p_j \phi'\left(\sum_{k=1}^{N} p_{kj} U(w_0 R_k)\right) p_{sj} U'(w_0 R_s)} \quad \text{for i} = 1, \ldots, \text{N} \tag{17}$$

We can readily interpret the pricing kernel formula in comparison to the simple case without ambiguity.[17] There, the pricing kernel is the ratio of marginal utility and expected marginal utility. In the setting with ambiguity aversion, the pricing kernel is the scaled marginal utility in each state divided by a modified expected marginal utility. We explain the modification of expected marginal utility first and then the scaling of the pricing kernel. For the modified expectation, the probabilities of the expectation

[17] Unfortunately, we cannot easily analyze the derivative of the pricing kernel with respect to returns. The resulting expressions are intractable and cannot be nicely segregated into, say, an income and a substitution effect.

(p_{ij} p_j) are being distorted by the marginal ambiguity utility $\phi'(\sum_{k=1}^{N} p_{kj} U(w_0 R_k))$. The resulting quantities are no longer probabilities, i.e. they will not add to one. Thus, the pricing kernel needs to be scaled in order to correct for the modification. The scaling factor is the fraction in front of the marginal utility term in Eq. (17). It turns out to be the ratio of the sum of the probabilities (p_{ij} p_j), which are again being distorted by $\phi'(\sum_{k=1}^{N} p_{kj} U(w_0 R_k))$ and the sum of the probabilities themselves (p_{ij} p_j).

4.4.2 The pricing kernel puzzle in a model of ambiguity aversion

Here we use Eq. (17) with power utilities and parameters η for the ambiguity aversion and γ for the risk aversion, respectively. We set w_0 to 1. The following choice for U(x) satisfies the assumption of Klibanoff et al. (2005) that two utility values need to be independent of γ, here, $U(1) = 0$ and $U(2) = 1$.[18] The investors are ambiguity averse if $\eta > \gamma$.

$$\phi(x) = \frac{x^{1-\eta} - 1}{1 - \eta}; \quad U(x) = \frac{x^{1-\gamma} - 1}{2^{1-\gamma} - 1} \tag{18}$$

Further, we model the 30-day return being lognormally distributed with an annualized mean of 0.10. Numerical details are relegated to the Internet Appendix.KMM. The investors are ambiguous with respect to annualized volatility, which we assume to be lognormally distributed with mean log 0.19 and standard deviation 0.10.

We depict the resulting pricing kernel with $\eta = 6$ and $\gamma = 4$ in Fig. 8 and it matches quite nicely the empirically observed u-shaped pricing kernels, see for example Fig. 3, panel B. The physical probability distribution (sum of the probabilities p_{ij} p_j) has, at an annual horizon, a mean of 0.10, standard deviation of 0.19, skewness of 0.00, and kurtosis of 3.12.

The next extension is to introduce large negative jumps (− 0.20 annualized mean and 0.30 standard deviation) where the investor exhibits ambiguity aversion across the probability of such jumps occurring. The return distribution without crashes is modeled being lognormally distributed with an annualized mean return of 0.12 and a volatility of 0.19. Finally, the conditional probabilities p_{ij} are obtained by mixing the return distribution without the crashes with the jump distribution. The probabilities for the occurrence of a jump then determine the appropriate weights for the two distributions such that the p_{ij} add up to 1 for a fixed j.

The pricing kernel in Fig. 9 with $\eta = 7$ and $\gamma = 6$ exhibits now a tilde shape. In comparison to the pricing kernels observed in the empirical literature, however, see for example Fig. 2, the hump in the center is shifted slightly to the left.

Thus, a simple one-period ambiguity aversion model can exhibit the pricing kernel puzzle. It turns out that ambiguity aversion over volatility generates u-shaped pricing kernels. Ambiguity aversion over the probability of large crashes generates tilde-shaped pricing kernels and can explain the hump of the empirical pricing kernel puzzle

[18] Note that alternatively, one could also use $U(x) = \frac{x^{1-\gamma}-1}{1-\gamma}$ with $\gamma \in (0, 1)$ but the above formulation allows for a great range of risk aversion coefficients.

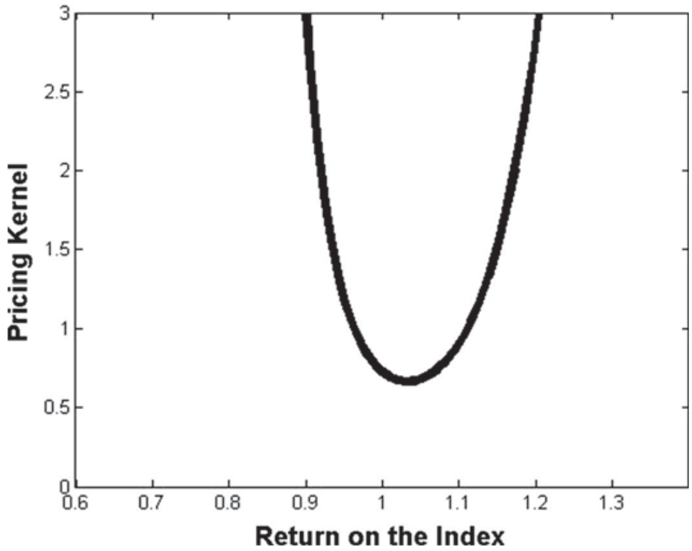

Fig. 8 The pricing kernel with ambiguity over volatilities. The pricing kernel based on the Klibanoff et al. (2005) model with ambiguity over volatilities, projected onto returns

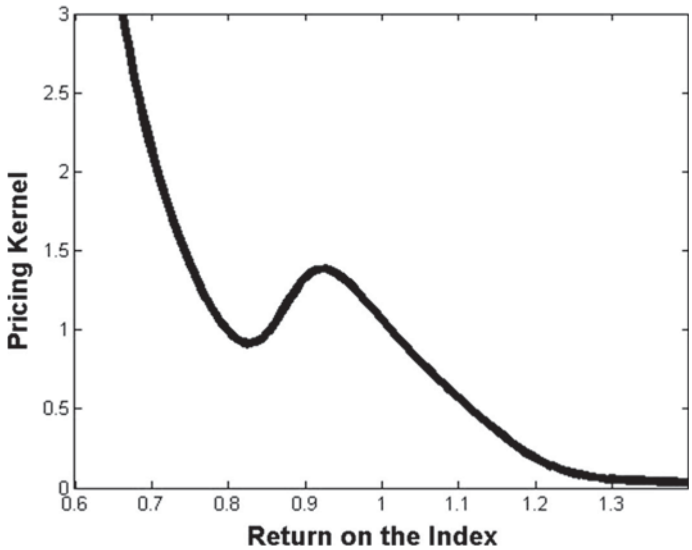

Fig. 9 The pricing kernel with ambiguity over market crashes. The pricing kernel based on the Klibanoff et al. (2005) model with ambiguity over market crashes, projected onto returns

at the center. Cuesdeanu (2016) extends this ambiguity aversion model by introducing ambiguity over volatility and jumps simultaneously. He finds that this allows for w-shaped pricing kernels as well.

5 Bounds on option prices

The literature on bounds on option prices takes a different perspective on the pricing kernel puzzle. The pricing kernel puzzle is about analyzing the empirical pricing kernel, given risk-neutral and physical distributions, where the pricing kernel turns out to be non-decreasing in returns. Turning the problem around, one can ask what are the highest and lowest option prices still compatible with a monotonically decreasing pricing kernel? This approach was developed in Perrakis and Ryan (1984) with the restrictions that the pricing kernel has to be positive and decreasing, and that it prices the stock and the bond and one reference option traded in the market. The resulting linear program then looks (for a call option with given strike price K) as follows:

$$\begin{aligned} & \text{Max/Min } \mathrm{E}[\mathrm{m}\,(\mathrm{R}\cdot\mathrm{S}-\mathrm{K})^{+}] \\ \text{s.t. } & \mathrm{E}[\mathrm{m}\,1]=\mathrm{B} \\ & \mathrm{E}[\mathrm{m}\,\mathrm{R}\cdot\mathrm{S}]=\mathrm{S} \\ & \mathrm{E}[\mathrm{m}\,(\mathrm{R}\cdot\mathrm{S}-\mathrm{K}^{\circ})^{+}]=\mathrm{C}(\mathrm{K}^{\circ}) \\ & \mathrm{m}>0,\ \text{m decreasing in wealth } \mathrm{R}\cdot\mathrm{S}, \end{aligned} \tag{19}$$

where S is the initial stock price, B the unit bond price, and K° the strike price of the observed reference option. Dividends are assumed to be zero for ease of exposition. A long literature ensued which extends the above linear program approach, adding e.g. bid/ask spreads and transaction costs, see the survey of Constantinides et al. (2008).

The resulting bounds are driven by pricing kernels which tend to be extreme, exhibiting steep drops after almost flat sections. Cochrane and Saa-Requejo (2000) address the problem of such unrealistic pricing kernels. They essentially work within the above set-up while also restricting the volatility of the pricing kernel, which leads to smoother pricing kernels and tighter bounds. Bernardo and Ledoit (2000) offer an alternative restriction by limiting the ratio of expected gains and expected losses of a security; ruling out that securities are priced much too low or high compared to their fair value.[19] Pyo (2011) achieves this goal by the ad-hoc restriction that price deviations are limited by deviations of observed prices from model prices based on a predetermined (power) pricing kernel.

The earlier papers solved the linear program explicitly and were thus limited in the complexity of the linear program, e.g., they could only handle one reference option and extending it to two was already a difficult task. Relying simply on computer solutions to the linear program, Constantinides et al. (2009) can compute bounds for S&P 500 index options while taking into account all observed options as reference assets and even formulating the linear program over two steps instead of one. Further, they use the analytical bounds from Constantinides and Perrakis (2002) for continuous, intermediate trading and proportional transaction costs. Empirically, they find a substantial number of options to be located outside their bounds, consistent with the pricing kernel puzzle. Wallmeier (2015) replicates their work and finds far fewer violations. This

[19] Marroquin-Martinez and Moreno (2013) extend Cochrane and Saa-Requejo (2000) and Bernardo and Ledoit (2000) to settings with stochastic volatility and find the resulting bounds to be tighter than in the original papers.

difference comes about as Wallmeier (2015) uses option implied information from just hours ago to adjust the physical distribution, while Constantinides et al. (2009) rely on information further into the past. The concern is that using more recent option implied information will eventually move the physical distribution so close to the risk-neutral, that one can no longer detect bound violations.

More interesting is the question if, using the earlier information of Constantinides et al. (2009), one can profitably trade based on bound violations. That exercise can be found in Constantinides et al. (2011), now using options on futures on S&P 500 and employing the analytical bounds of Constantinides and Perrakis (2007), which are suitable for these American options. The results suggest that trading strategies involving out-of-bounds options are superior to pure stock-and-bond strategies for all risk-averse investors.

6 Conclusion and outlook

In our survey of the pricing kernel puzzle, we recount the history, starting with the canonical papers which around the year 2000 divided risk-neutral distributions of S&P 500 returns by the physical distributions. These empirical pricing kernels exhibited increasing sections, which are inconsistent with simple representative investor models with a single state variable. Evidence from indexes in other countries and other periods finds the same puzzling behavior. We also discuss the (sparse) literature, which cannot detect the pricing kernel puzzle in the data and try to understand the reasons.

A number of statistical tests suggest the presence of the pricing kernel puzzle. What is still missing is a critical analysis and comparison of the several tests which so far exist. Are their assumptions realistic? Are certain tests better than others? There is still no agreement on which test to use as the standard test of the pricing kernel puzzle.

Much room is given to the potential explanations of the pricing kernel puzzle, starting with simple one-state-variable formulations and then moving to more complex settings. Similarly to the tests, many of the solutions are stand-alone model with little empirical validation. Mostly, they concentrate on a calibration, which, using some stylized facts, exhibits the pricing kernel puzzle. Much work is still needed in sorting through the alternative models and grading them according to their compatibility with the data. Ideally, some of the solutions might be joined in a nested model, allowing for a proper test of the different features. It would be interesting to know more about the true mechanism of what drives the pricing kernel puzzle. Interesting research along those lines is trying to explain the time-series patterns of the pricing kernel puzzle (e.g., its severity) using explanatory variables. This challenging work is still in its infancy and, as of now, still underwhelming.

For a glimpse into the future of the pricing kernel puzzle, one might want to consider the bivariate estimation of risk-neutral and physical distributions in Jackwerth and Vilkov (2017). Those bivariate risk-neutral distributions can normally only be obtained with the help of options written on both assets simultaneously, but Jackwerth and Vilkov (2017) were able to achieve this feat in the dimensions of index returns and volatility employing longer-dated options on the S&P 500. Dividing the two distributions into each other allows one to extract for the first time a bivariate pricing kernel.

References

Ait-Sahalia, Y., Lo, A.W.: Nonparametric estimation of state-price densities implicit in financial asset prices. J Finance **53**, 499–547 (1998)

Ait-Sahalia, Y., Lo, A.W.: Nonparametric risk management and implied risk aversion. J Econom **94**, 9–51 (2000)

Ait-Sahalia, Y., Duarte, J.: Nonparametric option pricing under shape restrictions. J Econom **116**, 9–47 (2003)

Alcock, J., Smith, G.: Testing alternative measure changes in nonparametric pricing and hedging of European options. J Futures Mark **34**, 320–345 (2014)

Andreou, P.C., Kagkadis, A., Philip, D.: Investor sentiments, rational beliefs, and option prices. Working paper, Lancaster University (2014)

Audrino, F., Meier, P.: Empirical pricing kernel estimation using a functional gradient descent algorithm based on splines. Working paper, University of St. Gallen (2012)

Audrino, F., Huitema, R., Ludwig, M.: An empirical analysis of the Ross recovery. Working paper, University of St. Gallen (2015)

Babaoglu, K., Christoffersen, P., Heston, S., Jacobs, K.: Option valuation with volatility components, fat tails, and non-monotonic pricing kernels. Working paper, University of Toronto (2016)

Backus, D., Chernov, M., Martin, I.: Disasters implied by equity index options. J Finance **66**, 1969–2012 (2011)

Bahra, B.: Implied risk-neutral probability density functions from option prices: theory and application. Working paper, Bank of England, no. 66 (1997)

Bakshi, G., Madan, D.: Investor heterogeneity, aggregation, and the non-monotonicity of the aggregate marginal rate of substitution in the price of market-equity. Working paper, University of Maryland (2007)

Bakshi, G., Madan, D.B., Panayotov, G.: Returns of claims on the upside and the viability of u-shaped pricing kernels. J Financ Econ **97**, 130–154 (2010)

Bakshi, G., Madan, D.B., Panayotov, G.: Heterogeneity in beliefs and volatility tail behavior. J Financ Quant Anal **50**, 1389–1414 (2015)

Bali, T.G., Cakici, N., Chabi-Yo, F., Murray, M.: The Risk-Neutral Distribution of Option Returns. Georgetown University (2017)

Barone-Adesi, G., Dall'O, H.: Is the price kernel monotone? Working paper, Swiss Finance Institute (2010)

Barone-Adesi, G., Mancini, L., Shefrin, H.: A tale of two investors: estimating optimism and overconfidence. Working paper, Swiss Finance Institute (2013)

Barone-Adesi, G., Engle, R., Mancini, L.: A GARCH option pricing model with filtered historical simulation. Rev Financ Stud **21**, 1223–1258 (2008)

Bartunek, K.S., Chowdhury, M.: Implied risk aversion parameter from option prices. Financ Rev **32**, 107–124 (1997)

Bassett, G.W., Koenker, R., Kordas, G.: Pessimistic portfolio allocation and Choquet expected utility. J Financ Econom **2**, 477–492 (2004)

Bates, D.: Post-'87 crash fears in S&P 500 futures options. J Econom **94**, 181–238 (2000)

Bates, D.: U.S. stock market crash risk, 1926–2010. J Financ Econ **105**, 229–259 (2012)

Bates, D.: Testing option pricing models. In: Maddala, G.S., Rao, C. R. (eds.) Statistical Methods in Finance (Handbook of Statistics, vol. 14, Chapter 20), pp. 567–611. Elsevier Science, Amsterdam (1996a)

Bates, D.: Jumps and stochastic volatility: exchange rate processes implicit in Deutsche mark options. Rev Financ Stud **9**, 69–107 (1996b)

Bates, D.: The market for crash risk. J Econ Dyn Control **32**, 2291–2321 (2008)

Beare, B.K.: Measure preserving derivatives and the pricing kernel puzzle. J Math Econ **47**, 689–697 (2011)

Beare, B.K., Schmidt, L.: An empirical test of pricing kernel monotonicity. J Appl Econom **31**, 338–356 (2014)

Beare, B.K., Schmidt, L.: Empirical implications of the pricing kernel puzzle for the return on contingent claims. Working paper, UC San Diego (2015)

Belomestny, D., Haerdle, W., Krymova, E.: Sieve estimation of the minimal entropy martingale marginal density with application to pricing kernel estimation. Int J Theor Appl Finance **20**, 1750041 (2017)
Benth, F.E., Groth, M., Lindberg, C.: The implied risk aversion from utility indifference option pricing in a stochastic volatility model. Int J Appl Math Stat **16**, 11–37 (2010)
Benzoni, L., Collin-Dufresne, P., Goldstein, R.: Explaining asset pricing puzzles associated with the 1987 market crash. J Financ Econ **101**, 552–573 (2011)
Berkowitz, J.: Testing density Forecasts with applications to risk management. J Bus Econ Stat **19**, 465–474 (2001)
Bernardo, A., Ledoit, O.: Gain, loss, and asset pricing. J Polit Econ **108**, 144–172 (2000)
Bliss, R., Panigirtzoglou, N.: Testing the stability of implied probability density functions. J Bank Finance **26**, 381–422 (2002)
Bliss, R., Panigirtzoglou, N.: Option-implied risk aversion estimates. J Finance **59**, 407–446 (2004)
Boes, M.J., Drost, F.C., Werker, B.J.M.: Nonparametric risk-neutral joint return and volatility distributions. Working paper, VU University Amsterdam and Tilburg University (2007)
Bollen, N., Whaley, R.: Does net buying pressure affect the shape of implied volatility functions? J Finance **59**, 711–754 (2004)
Bollerslev, T., Todorov, V.: Tails, fears and risk premia. J Finance **66**, 2165–2211 (2011)
Borovicka, J., Hansen, L.P., Scheinkman, J.A.: Misspecified recovery. Working paper, NBER (2015)
Branger, N., Hansis, A., Schlag, C.: The impact of risk premia on expected option returns. Working paper, University of Muenster (2011)
Breeden, D.T., Litzenberger, R.H.: Prices of state-contingent claims implicit in options prices. J Bus **51**, 621–651 (1978)
Broadie, M., Chernov, M., Johannes, M.: Understanding index option returns. Rev Financ Stud **22**, 4493–4529 (2009)
Brown, D., Jackwerth, J.C.: The pricing kernel puzzle: reconciling index option data and economic theory. In: Batten, J.A., Wagner, N. (eds.) Contemporary Studies in Economics and Financial Analysis: Derivative Securities Pricing and Modelling, vol. 94, pp. 155–183. Bingley: Emerald Group (2012)
Buraschi, A., Jackwerth, J.C.: The price of a smile: hedging and spanning in option markets. Rev Financ Stud **14**, 495–527 (2001)
Campbell, J., Cochrane, J.: By force of habit: a consumption-based explanation of aggregate stock market behavior. J Polit Econ **107**, 205–251 (1999)
Carr, P., Yu, J.: Risk, return, and Ross recovery. J Deriv **20**, 38–59 (2012)
Carr, P., Geman, H., Madan, D.P., Yor, M.: The fine structure of asset returns: an empirical investigation. J Bus **75**, 305–332 (2002)
Chabi-Yo, F.: Pricing kernels with stochastic skewness and volatility risk. Manag Sci **58**, 624–640 (2012)
Chabi-Yo, F., Song, Z.: Recovering the probability weights of tail events with volatility risk from option prices. Working paper, Ohio State (2013)
Chabi-Yo, F., Garcia, R., Renault, E.: State dependence can explain the risk aversion puzzle. Rev Financ Stud **21**, 973–1011 (2008)
Chaudhuri, R., Schroder, M.D.: Monotonicity of the stochastic discount factor and expected option returns. Rev Financ Stud **28**, 1462–1505 (2015)
Chernov, M.: Empirical reverse engineering of the pricing kernel. J Econom **116**, 329–364 (2003)
Chernov, M., Ghysels, E.: A study towards a unified approach to the joint estimation of objective and risk-neutral measures for the purpose of options valuation. J Financ Econ **56**, 207–458 (2000)
Chetty, R., Szeidl, A.: Consumption commitments and risk preferences. Q J Econ **122**, 831–877 (2007)
Christoffersen, P., Heston, S., Jacobs, K.: Capturing option anomalies with a variance-dependent pricing kernel. Rev Financ Stud **26**, 1962–2006 (2013)
Christoffersen, P., Heston, S., Jacobs, K.: The shape and term structure of the index option smirk: why multifactor stochastic volatility models work so well. Manag Sci **55**, 1914–1932 (2009)
Christoffersen, P., Jacobs, K., Chang, B.Y.: Forecasting with option-implied information. In: Elliot, G., Timmermann, A. (eds.) Handbook of Economic Forecasting, vol. 2, pp. 581–656. Amsterdam: North-Holland (2013)
Cochrane, J.H., Saa-Requejo, J.: Beyond arbitrage: good-deal asset price bounds in incomplete markets. J Polit Econ **108**, 79–119 (2000)
Constantinides, G.M., Perrakis, S.: Stochastic dominance bounds on derivative prices in a multiperiod economy with proportional transaction costs. J Econ Dyn Control **26**, 1323–1352 (2002)

Constantinides, G.M., Perrakis, S.: Stochastic dominance bounds on American option prices in markets with frictions. Rev Finance **11**, 71–115 (2007)
Constantinides, G.M., Jackwerth, J.C., Savov, A.: The puzzle of index option returns. Rev Asset Pricing Stud **3**, 229–257 (2013)
Constantinides, G.M., Jackwerth, J.C., Perrakis, S.: Option pricing: real and risk-neutral distributions. In: Birge, J.R., Linetsky, V. (eds.) Handbooks in Operations Research and Management Science: Financial Engineering, vol. 15, pp. 565–591. Amsterdam: Elsevier (2008)
Constantinides, G.M., Jackwerth, J.C., Stylianos Perrakis, S.: Mispricing of S&P 500 index options. Rev Financ Stud **22**, 1247–1277 (2009)
Constantinides, G.M., Czerwonko, M., Jackwerth, J.C., Perrakis, S.: Are options on index futures profitable for risk averse investors? Empirical evidence. J Finance **66**, 1401–1431 (2011)
Coutant, S.: Implied risk aversion in option prices using Hermite polynomials. Working paper, Banque de France (1999)
Coutant, S.: Time-varying implied risk aversion in option prices using Hermite polynomials. Working paper, Banque de France (2000)
Coval, J.D., Shumway, T.: Expected option returns. J Finance **56**, 983–1009 (2001)
Cuesdeanu, H.: Empirical pricing kernels: a tale of two tails and volatility. Working paper, University of Konstanz (2016)
Cuesdeanu, H., Jackwerth, J.C.: The pricing kernel puzzle in forward looking data. Rev. Deriv. Res. (2017). https://doi.org/10.1007/s11147-017-9140-8
David, A., Veronesi, P.: Investors' and central bank's uncertainty embedded in index options. Rev Financ Stud **27**, 1661–1716 (2014)
Detlefsen, K., Haerdle, W., Moro, R.: Empirical pricing kernels and investor preferences. J Math Methods Econ Finance **3**, 19–48 (2010)
Diebold, F.X., Tay, A.S., Wallis, K.F.: Evaluating density forecasts of inflation: the survey of professional forecasters. In: Engle, R.F., White, H. (eds.) Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger, pp. 76–90. New York: Oxford University Press (1999)
Diebold, F.X., Gunther, T.A., Tay, A.S.: Evaluating density forecasts with applications to financial risk management. Int Econ Rev **39**, 863–883 (1998)
Dierkes, M.: Probability weighting and asset prices. Working paper, University of Muenster (2013)
Dimson, E., Marsh, P., Staunton, M.: Equity premiums around the world. In: Hammond, B., Leibowitz, M., Siegel, L. (eds.) Rethinking the Equity Premium, Chapter 4, pp. 32–52. Charlottesville: Research Foundation of the CFA Institute (2012)
Dittmar, R.F.: Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. J Finance **57**, 369–403 (2002)
Drechsler, I.: Uncertainty, time-varying fear, and asset prices. J Finance **68**, 1843–1889 (2013)
Duan, J.C., Zhang, W.: Forward-looking market risk premium. Manag Sci **60**, 521–538 (2014)
Dybvig, P.H.: Distributional analysis of portfolio choice. J Bus **61**, 369–93 (1988)
Epstein, L.G., Schneider, M.: Ambiguity and asset markets. Annu Rev Financ Econ **2**, 315–346 (2010)
Epstein, L.G., Zin, S.: Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica **57**, 937–969 (1989)
EZB: Der Informationsgehalt von Optionspreisen während der Finanzkrise. Monatsbericht, EZB Februar (2011)
Fengler, M., Hin, L.Y.: Semi-nonparametric estimation of the call price surface under no-arbitrage constraints. J Econom **184**, 242–261 (2015)
Figlewski, S.: Estimating the implied risk-neutral density for the U.S. market portfolio. In: Bollerslev, T., Russell, J., Watson, M. (eds.) Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle, pp. 323–353. Oxford: Oxford University Press (2010)
Figlewski, S., Malik, F.M.: Options on leveraged ETFs: a window on investor heterogeneity. Working paper, New York University (2014)
Friedman, M., Savage, L.J.: Utility analysis of choices involving risk. J Polit Econ **56**, 279–304 (1948)
Garcia, R., Luger, R., Renault, E.: Empirical assessment of an intertemporal option pricing model with latent variables. J Econom **116**, 49–83 (2003)
Garleanu, N., Pedersen, L.H., Poteshman, A.M.: Demand-based option pricing. Rev Financ Stud **22**, 4259–4299 (2009)
Gemmill, G., Shackleton, M.B.: Prospect theory and option prices: evidence from S&P 500 index options. Working paper, Warwick University (2005)

Giacomini, E., Haerdle, W.: Dynamic semiparametric factor models in pricing kernel estimation. In: Dabo-Niang, S., Ferraty, F. (eds.) Functional and Operational Statistics, Contributions to Statistics, pp. 181–187. Heidelberg: Springer (2008)

Giacomini, E., Handel, M., Haerdle, W.: Time dependent relative risk aversion. In: Bol, G., Rachev, S.T., Würth, R. (eds.) Risk Assessment: Decisions in Banking and Finance, pp. 15–46. Heidelberg: Physica (2008)

Glosten, L.R., Jagannathan, R., Runkle, D.E.: On the relation between the expected value and the volatility of the nominal excess return on stocks. J Finance **48**, 1779–1801 (1993)

Gollier, C.: Portfolio choices and asset prices: the comparative statics of ambiguity aversion. Rev Econ Stud **78**, 1329–1344 (2011)

Golubev, Y., Haerdle, W., Timofeev, R.: Testing monotonicity of pricing kernels. AStA Adv Stat Anal **98**, 305–32 (2014)

Grith, M., Haerdle, W., Park, J.: Shape invariant modeling pricing kernels and risk aversion. J Financ Econom **11**, 370–399 (2013)

Grith, M., Haerdle, W., Kraetschmer, V.: Reference dependent preferences and the EPK puzzle. Rev Finance **21**, 269–298 (2017)

Haas, J.R., Fajardo Barbachan, J.S., Rocha de Farias, A.: Estimating relative risk aversion, risk-neutral and real-world densities using Brazilian real currency options. Econ Apl **16**, 567–577 (2012)

Haerdle, W., Grith, M., Mihoci, A.: Cross country evidence for the EPK paradox. Working paper, Humboldt University (2014)

Haerdle, W., Kraetschmer, V., Moro, R.: A microeconomic explanation of the EPK paradox. Working paper, Humboldt University (2009)

Haerdle, W., Okhrin, Y., Wang, W.: Uniform confidence bands for pricing kernels. J Financ Econom **13**, 376–413 (2015)

Haley, M.R., Walker, T.: Alternative tilts for nonparametric option pricing. J Futures Mark **30**, 983–1006 (2010)

Han, B.: Investor sentiment and option prices. Rev Financ Stud **21**, 387–414 (2008)

Han, Q., Turvey, C.G.: Equilibrium market prices for risks and market risk aversion in a complete stochastic volatility model with habit formation. Working paper, Cornell University (2010)

Hens, T., Reichlin, C.: Three solutions to the pricing kernel puzzle. Rev Finance **17**, 1065–1098 (2013)

Heston, S.L.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev Financ Stud **6**, 327–343 (1993)

Heston, S., Nandi, S.: A closed-form GARCH option pricing model. Rev Financ Stud **13**, 585–626 (2000)

Hill, S.I.: Tracking value at risk through derivative prices. J Comput Finance **16**, 79–121 (2013)

Hodges, S.D., Tompkins, R.G., Ziemba, W.T.: The favorite-longshot bias in S&P 500 and FTSE 100 index futures options: the return to bets and the cost of insurance. In: Hausch, D.B., Ziemba, W.T. (eds.) Handbook of Sports and Lottery Markets. Amsterdam: Elsevier (2008)

Huang, D., Shaliastovich, I.: Risk adjustment and the temporal resolution of uncertainty: evidence from options markets. Working paper, University of Pennsylvania (2014)

Ingersoll, J.E.: Cumulative prospect theory, aggregation, and pricing. Crit Finance Rev **4**, 1–55 (2014)

Israelov, R., Kelly, B.: Forecasting the distribution of option returns. Working paper, University of Chicago (2017)

Jackwerth, J.C.: Option-implied risk-neutral distributions and implied binomial trees: a literature review. J Deriv **2**, 66–82 (1999)

Jackwerth, J.C.: Recovering risk aversion from option prices and realized returns. Rev Financ Stud **13**, 433–467 (2000)

Jackwerth, J.C.: Option-implied risk-neutral distributions and risk-aversion. Charlotteville: The Research Foundation of AIMR (2004)

Jackwerth, J.C., Vilkov, G.: Asymmetric volatility risk: evidence from option markets. Working paper, University of Konstanz (2017)

Jackwerth, J.C., Rubinstein, M.: Recovering probability distributions from option prices. J Finance **51**, 1611–1631 (1996)

Jackwerth, J.C., Menner, M.: Does the Ross recovery theorem work empirically? Working paper, University of Konstanz (2015)

Jensen, C.S., Lando, D., Pedersen, L.H.: Generalized recovery. Working paper, Copenhagen Business School (2016)

Kang, B.J., Kim, T.S.: Option-implied risk preferences: an extension to wider classes of utility functions. J Financ Mark **9**, 180–198 (2006)
Kang, B.J., Kim, T.S., Lee, H.S.: Option-implied preferences with model uncertainty. J Futures Mark **34**, 498–515 (2014)
Kiesel, R., Rahe, F.: Option pricing under time-varying risk-aversion with applications to risk forecasting. J Bank Finance **76**, 120–138 (2017)
Kitsul, Y., Wright, J.H.: The economics of options-implied inflation probability density functions. J Financ Econ **110**, 696–711 (2013)
Klibanoff, P., Marinacci, M., Mukerji, S.: A smooth model of decision making under ambiguity. Econometrica **73**, 1849–1892 (2005)
Kliger, D., Levy, O.: Theories of choice under risk: insights from financial markets. J Econ Behav Organ **71**, 330–346 (2009)
Kocherlakota, N.R.: The equity premium: it's still a puzzle. J Econ Lit **34**, 42–71 (1996)
Kostakis, A., Panigirtzoglou, N., Skiadopoulos, G.: Market timing with option-implied distributions: a forward-looking approach. Manag Sci **57**, 1231–1249 (2011)
Leisen, D.: Small sample bias in pricing kernel estimations. Working paper, University of Mainz (2014)
Li, H., Zhao, F.: Nonparametric estimation of state-price densities implicit in interest rate cap prices. Rev Financ Stud **22**, 4335–4376 (2009)
Linn, M., Shive, S., Shumway, T.: Pricing kernel monotonicity and conditional information. Working paper, University of Michigan (2014)
Lioui, A., Malka, R.: Revealing the parameter of risk-aversion from option prices when markets are incomplete: theory and evidence. Working paper, Bar Ilan University (2004)
Liu, J., Pan, J., Wang, T.: An equilibrium model of rare-event premia and its implication for option smirks. Rev Financ Stud **18**, 131–164 (2005)
Liu, X., Kuo, J.M., Coakley, J.: A pricing kernel approach to valuing options on interest rate futures. Eur J Finance **21**, 93–110 (2015)
Liu, X., Shackleton, M.B., Taylor, S.J., Xu, X.: Empirical pricing kernels obtained from the UK index options market. Appl Econ Lett **16**, 989–993 (2009)
Liu, X., Shackleton, M.B., Taylor, S.J., Xu, X.: Closed-form transformations from risk-neutral to real-world distributions. J Bank Finance **31**, 1501–1520 (2007)
Ludwig, M.: Robust estimation of shape-constrained state price density surfaces. J Deriv **22**, 56–72 (2015)
Lundtofte, F.: Implied volatility and risk aversion in a simple model with uncertain growth. Econ Bull **30**, 182–191 (2010)
Marroquin-Martinez, N., Moreno, M.: Optimizing bounds on security prices in incomplete markets. Does stochastic volatility specification matter? Eur J Oper Res **225**, 429–442 (2013)
Martin, I.: What is the expected return on the market? Q J Econ **132**, 367–433 (2017)
Mehra, R., Prescott, E.C.: The equity premium: a puzzle. J Monet Econ **15**, 145–161 (1985)
Mehra, R.: The equity premium puzzle: a review. Found Trends Finance **2**, 1–81 (2008)
Melick, W.R., Thomas, C.P.: Recovering an asset's implied pdf from option prices: an application to crude oil during the gulf crisis. J Financ Quant Anal **32**, 91–115 (1997)
Merton, R.C.: Option pricing when underlying stock returns are discontinuous. J Financ Econ **3**, 125–144 (1976)
Ni, S.X.: Stock option returns: a puzzle. Working paper, Hong Kong University of Science and Technology (2009)
Pan, J.: The jump-risk premia implicit in options: evidence from an integrated time-series study. J Financ Econ **63**, 3–50 (2002)
Patton, A.J., Timmermann, A.: Monotonicity in asset returns: new tests with applications to the term structure, the CAPM, and portfolio sorts. J Financ Econ **98**, 605–625 (2010)
Perignon, C., Villa, C.: Extracting information from options markets: smiles, state-price densities and risk aversion. Eur Financ Manag **8**, 495–513 (2002)
Perrakis, S., Ryan, P.: Option pricing bounds in discrete time. J Finance **39**(2), 519–525 (1984)
Polkovnichenko, V., Zhao, F.: Probability weighting functions implied by options prices. J Financ Econ **107**, 580–609 (2013)
Pozzi, L.C.G., de Vries, C.G., Zenhorst, J.: World equity premium based risk aversion estimates. Working paper, Tinbergen Institute 2010-007/2 (2010)
Prelec, D.: The probability weighting function. Econometrica **66**, 497–527 (1998)
Pyo, U.: Minmax price bounds in incomplete markets. J Econ Finance **35**(3), 274–295 (2011)

Rieger, M.O.: Co-monotonicity of optimal investments and the design of structured financial products. Math Stat **15**, 27–55 (2011)
Rosenberg, J.V., Engle, R.F.: Empirical pricing kernels. J Financ Econ **64**, 341–372 (2002)
Ross, S.: The recovery theorem. J Finance **70**, 615–648 (2015)
Rubinstein, M.: Implied binomial trees. J Finance **49**, 771–818 (1994)
Sala, C., Barone-Adesi, G.: Sentiment lost: the effect of projecting the empirical pricing kernel onto a smaller filtration set. Working paper, Swiss Finance Institute (2016)
Sala, C.: Does the pricing kernel anomaly reflect forward-looking beliefs? Working paper, ESADE Business School (2016)
Santa-Clara, P., Yan, S.: Crashes, volatility, and the equity premium: lessons from S&P 500 options. Rev Econ Stat **92**, 435–451 (2010)
Schneider, P., Trojani, F.: (Almost) model free recovery. Working paper, Swiss Finance Institute (2015)
Schweri, U.: Is the pricing kernel u-shaped? Working paper, University of Zurich (2010)
Shefrin, H.: A Behavioral Approach to Asset Pricing, 2nd edn. Amsterdam: Elsevier (2008a)
Shefrin, H.: Risk and return in behavioral SDF-based asset pricing models. J Invest Manag **6**, 1–18 (2008b)
Shimko, D.C.: Bounds of probability. Risk Mag **6**, 33–37 (1993)
Shive, S.: On the shape of the option-implied stochastic discount factor. Working paper, University of Michigan (2003)
Shive, S., Shumway, T.: A non-decreasing pricing kernel: evidence and implications. Working paper, University of Michigan (2004)
Shive, S., Shumway, T.: Is the pricing kernel monotonic? Working paper, University of Notre Dame (2009)
Siddiqi, H., Quiggin, J.: Differential awareness and the pricing kernel puzzle. Working paper, University of Queensland (2016)
Silverman, B.: Density Estimation. London: Chapman & Hall (1986)
Song, Z.: Expected VIX option returns. Working paper, Federal Reserve Board (2012)
Song, Z., Xiu, D.: A tale of two option markets: state-price densities and volatility risk. J Econom **190**, 176–196 (2016)
Stutzer, M.: A simple nonparametric approach to derivative security valuation. J Finance **51**, 1633–1652 (1996)
Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncertain **5**, 297–323 (1992)
Veronesi, P.: Belief-dependent utilities, aversion to state-uncertainty, and asset prices. Working paper, University of Chicago (2004)
Walden, J.: Recovery with unbounded diffusion processes. Rev Finance **21**, 1403–1444 (2017)
Wallmeier, M.: Mispricing of index options with respect to stochastic dominance bounds? Working paper, University of Fribourg (2015)
Weber, T.: Estimating risk aversion in the European CDO market. Working paper, University of Konstanz, Germany (2006)
Wu, L.: Dampened power law: reconciling the tail behavior of financial security returns. J Bus **79**, 1445–1473 (2006)
Yamazaki, A.: A dynamic equilibrium model for u-shaped pricing kernels. Working paper, Hosei University (2017)
Yang, J.: Semiparametric estimation of asset pricing kernel. Appl Financ Econ **19**, 257–272 (2009)
Zdorovenin, V., Pezier, J.: Does information content of option prices add value for asset allocation? Working paper, University of Reading, UK (2011)
Ziegler, A.: State-price densities under heterogeneous beliefs, the smile effect, and implied risk aversion. Eur Econ Rev **46**, 1539–1557 (2002)
Ziegler, A.: Why does implied risk aversion smile? Rev Financ Stud **20**, 859–904 (2007)