CFS Working Paper Series

No. 566

Luiz Felix, Roman Kräussl, and Philip Stork

Single stock call options as lottery tickets

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

The CFS Working Paper Series

presents ongoing research on selected topics in the fields of money, banking and finance. The papers are circulated to encourage discussion and comment. Any opinions expressed in CFS Working Papers are those of the author(s) and not of the CFS.

The Center for Financial Studies, located in Goethe University Frankfurt's House of Finance, conducts independent and internationally oriented research in important areas of Finance. It serves as a forum for dialogue between academia, policy-making institutions and the financial industry. It offers a platform for top-level fundamental research as well as applied research relevant for the financial sector in Europe. CFS is funded by the non-profit-organization Gesellschaft für Kapitalmarktforschung e.V. (GfK). Established in 1967 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia. GfK members comprise major players in Germany's financial industry. The funding institutions do not give prior review to CFS publications, nor do they necessarily share the views expressed therein.

Single stock call options as lottery tickets

Luiz Félix[a,b], Roman Kräussl[c] and Philip Stork[a,♣]

a) VU University Amsterdam

b) APG Asset Management

c) Luxembourg School of Finance, University of Luxembourg

This version: February 2016

Abstract

This paper investigates whether the overpricing of out-of-the money single stock calls can be explained by Tversky and Kahneman's (1992) cumulative prospect theory (CPT). We argue that these options are overpriced because investors overweight small probability events and overpay for such positively skewed securities, i.e., characteristics of lottery tickets. We match a set of subjective density functions derived from risk-neutral densities, including CPT with the empirical probability distribution of U.S. equity returns. We find that overweighting of small probabilities embedded in CPT explains on average the richness of out-of-the money single stock calls better than other utility functions. The degree that agents overweight small probability events is, however, strongly time-varying and has a horizon effect, which implies that it is less pronounced in options of longer maturity. We also find that time-variation in overweighting of small probabilities is strongly explained by market sentiment, as in Baker and Wurgler (2006).

Keywords: Cumulative prospect theory; Market sentiment; Risk-neutral densities; Call options

JEL Classification Codes: G02, G12

♣ Philip Stork (corresponding author) is Professor at VU University Amsterdam, FEWEB, School of Finance and Risk Management, de Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Tel.: + 31 20 59 83651. E-mail address: p.a.stork@vu.nl.

1. Introduction

The most distinctive contribution of Tversky and Kahneman's (1992) cumulative prospect theory (CPT) is that individuals overweight small probability events when making decisions under risk. Barberis and Huang (2008), however, are the ones to hypothesize that the CPT's overweighting of small probability events may explain a number of seemingly unrelated pricing puzzles. Differently from an earlier literature which concentrates on the CPT's value function (see Benartzi and Thaler, 1995; Barberis and Huang, 2001; Barberis and Huang, 2008) focus on the probability weighting functions of the model. They conclude that the CPT's overweighting of small probability events helps explaining why investors prefer positively skewed returns, or "lottery ticket" type of securities. Due to such preference, investors overpay for positively skewed securities, turning them expensive and causing them to yield low forward returns. This overpricing is the reason for the low long-term average return of IPO stocks, the private equity premium puzzle, distressed stocks, and the overpricing of deep out-of-the money (OTM) single stock calls, among other irrational pricing phenomena.

The proposition by Barberis and Huang (2008) that deep OTM single stock calls resemble overpriced, lottery tickets-type securities has not yet been verified empirically. Empirical studies on probability weighting functions implied by option prices are offered by Dierkes (2009), Kliger and Levy (2009), and Polkovnichenko and Zhao (2013)[1]. The evidence in these papers is, however, based on the index put options market, which is very different from the single stock option market. The main buyers of OTM index puts are institutional investors, which use them for portfolio insurance (Bates, 2003; Bollen and Whaley, 2004; Lakonishok et al., 2007; Barberis and Huang 2008). Because institutional investors comprise around two-thirds of the total equity market capitalization (Blume and Keim, 2012), their option trading activity strongly impacts the pricing of put options (Bollen and Whaley, 2004) by making them expensive. The results of Dierkes (2009) and Polkovnichenko and Zhao (2013) reiterate this evidence and suggest that overweighting of small probabilities partially explains the pricing puzzle present in the equity index option market.

Contrary to the index put market, trading activity in single stock calls is concentrated among individual investors (Bollen and Whaley, 2004; Lakonishok et al., 2007). Moreover, individual investors' demand for single stock options is speculative in nature (Lakonishok et al., 2007), whereas Mitton and Vorking (2007) provide important support to the link between preference for skewness and individual investor trading activity. Given the very distinct clientele of these two option markets (institutional investors vs. retail investors) and the different motivation for trading (portfolio insurance vs. speculation),

[1] These studies focus on the rank-dependent expected utility (RDEU) rather than the CPT, as the RDEU is seamlessly effective in dealing with the overweighting of probability phenomena. The RDEU's probability weighting functions are strictly monotonically increasing, whereas the CPT's one is not. RDEU functions are also easier to estimate because they use one less parameter than the CPT.

we reason that the OTM single stock calls overpricing is a puzzle in itself, requiring an independent empirical proof from the index option market.

The first contribution of our study is to investigate whether the CPT can empirically explain the overpricing of OTM single stock call options. To that purpose, we empirically test whether tails of the CPT density function outperform the risk-neutral density and rational subjective probability density functions on matching tails of the distribution of realized returns. We find that our estimates for the CPT probability weighting function parameter γ do not differ much from the one predicated by Tverky and Kahneman (1992), particularly for short-term options. This analysis complements the results of Barberis and Huang (2008) and provides novel support to explain the overpricing of OTM single stock calls. Our empirical results extend the findings of Dierkes, 2009; Kliger and Levy, 2009; Polkovnichenko and Zhao, 2013, because we show that investors' overweighting of small probabilities[2] is not restricted to the pricing of index puts but also applies to single stock calls.

Secondly, we provide evidence that overweighting of small probabilities is strongly time-varying and connected to the Baker and Wurgler (2006) investors' sentiment factor. These findings contrast the CPT model, where the probability weighting parameter for gains (γ) is constant at 0.61. In fact, our estimations suggest that the γ parameter fluctuates widely around that level, sometimes even reflecting underweighting of small probabilities. We show that overweighting of small probabilities was most acute during the dot-com bubble, which coincided with a strong rise in investor's sentiment.

Moreover, we find that overweighting of small probabilities is largely horizon-dependent, as such bias is mostly observed within short-term options prices (i.e., three- and six-months) rather than in long-term ones (i.e., twelve-months). We reason that such horizon-effects exist because individual investors may speculate using the cheapest available call at their disposal. In other words, individual investors buy the cheapest lottery tickets that they can find. As three- and six-month options have much less time-value than twelve-month ones, more pronounced overweighting of small probabilities within short-term options seems sensible. This result is consistent with individual investors being the typical buyers of OTM single stock calls and the fact that they mostly use short-term instruments to speculate on the upside of equities (Lakonishok et al., 2007).

[2] We acknowledge that it is yet fully unclear whether overpicing of OTM calls is caused by overweighting of small probablities (i.e., a matter of preferences) or rather by biased beliefs. Barberis (2013) eloquently discusses how both phenomena are distinctly different and how both (individually or jointly) may potentially explain the existence of overpriced OTM options, as well as many other puzzling facts in financial markets. In this paper we take a myopic view and use only the first explanation, for ease of exposition. Disentangling the two (beliefs and preferences) would potentially be very interesting, but we deem it to be outside the scope of this paper.

In our analysis of probability weighting functions, we focus on the outmost tails of RNDs[3]. We argue that, as distribution tails (mostly estimated from OTM options) are the sections of the distribution that reflect low probability events, we may analyze these locally, thus, isolated from the distribution's body. To this purpose, we use extreme value theory (EVT) and Kupiec's test (as a robustness check), which are especially suited for the analysis of tail probabilities and, so far, have not been employed yet to the evaluation of overpricing of OTM options. As an additional robustness check, we replace the CPT by the rank-dependent expected utility (RDEU) function of Prelec (1998). This alteration reconfirms the presence of overweighted small probabilities by investors within the OTM single stock call market but, at the same time, suggests that such bias is less pervasive than our CPT-based results indicate. Time-variation of the weighting function parameters is also observed when RDEU is applied.

The remainder of this paper is organized as follows. Section 2 describes the CPT model. Section 3 describes the data and methodology employed in our study. Section 4 presents and discusses our empirical analysis as well as robustness tests. Section 5 concludes.

2. Cumulative prospect theory

The prospect theory (PT) of Kahneman and Tversky (1979) incorporates behavioral biases into the standard utility theory (von Neumann and Morgenstern, 1947), which presumes that individuals are rational[4]. Such behavioral biases are (i) loss aversion[5], (ii) risk seeking behavior[6], and (iii) non-linear preferences[7]. The CPT is described in terms of a value function (υ) and a probability distortion function (π). The value function is analogous to the utility function in the standard utility theory and it is defined relative to a reference point zero. Therefore, positive values within the value function are considered as gains and negative values are losses, which leads to

$$\upsilon(x) = \begin{cases} x^{\alpha}, & if\ x \geq 0 \\ -\lambda(-x)^{\beta}, & if\ x < 0\,, \end{cases} \tag{1}$$

[3] Per contrast, Dierkes (2009) and Polkovnichenko and Zhao (2013) explore the relation between overweighting of small probabilities and options prices by analyzing the *full* RND from options. Dierkes (2009) applies Berkowitz's tests, whereas Polkovnichenko and Zhao (2013) estimate an empirical weighting function via polynomial regressions.

[4] The expected utility theory of von Neumann and Morgenstern (1947) is the standard economics framework on decision making under risk. Their theory assumes that decision-makers behave as if they maximize the expected value of some function defined over the potential (probabilitistic) outcomes. Individuals are assumed to have stable and rational preferences; i.e., not influenced by the context or framing.

[5] Loss aversion is the property in which people are more sensitive to (or affected by) losses than gains. For details, see Kahneman and Tversky (1979), Tversky and Kahneman (1992), and Barberis and Huang (2001).

[6] Risk-seeking behavior happens when individuals are attracted by gambles with unfair prospects. In other words, the risk-seeking individual is the one that chooses for a gamble versus a sure thing even though the two outcomes have the same expected value. For details, see Kahneman and Tversky (1979).

[7] Non-linear preferences occur when preferences between risky prospects are not linear in the probabilities, thus, equally probable prospects are more heavily weighted by agents than others. For details, see Tversky and Kahneman (1992), Fox, Rogers and Tverky (1996), Wu and Gonzales (1996), Prelec (1998), and Hsu et al. (2009).

where $\lambda \geq 1$, $0 \leq \beta \leq 1$, $0 \leq \alpha \leq 1$, and x are gains or losses. Thus, along the dominium of x, the CPT's value function is asymmetrically S-shaped (see Figure 1) with diminishing sensitivity as $x \to \pm\infty$.

[Please insert Figure 1 about here]

The value function is, thus, concave over gains and convex over losses, differently from the traditional utility function used by standard utility theory. Such a shape of the value function implies diminishing marginal values as gains or losses becomes larger, which, in other words, means that any additional unit of gain (loss) becomes less relevant when wealth increases (decreases). As α and β increases, the effect of diminishing sensitivity decreases (see Figure 2), and as λ increases the degree of loss aversion increases. We also note in Figure 1A that the value function has a kink at the reference point, which implies loss aversion, as the function is steeper for losses than for gains.

[Please insert Figure 2 about here]

The use of a probability distortion function or decision weight function is the adjustment made to the PT to address nonlinear preferences. This function takes probabilities and weights them nonlinearly, so that the difference between probabilities at high percentiles, e.g., between 99 percent and 100 percent, has more impact on preferences than the difference between probabilities at small percentiles, e.g., between 10 percent and 11 percent. This is the main advance of the CPT over the original PT. The CPT applies probability distortions to the cumulative probabilities (i.e., the CDF), whereas the PT applies them to individual probabilities (i.e., the PDF). The enhancement brought by this new formulation satisfies stochastic dominance conditions not achieved by the PT, which renders the CPT applicable to a wider number of experiments. The probability distortion functions suggested by Tversky and Kahneman (1992), respectively, for gains (π_n^+) and losses (π_{-m}^-) are:

$$\pi_n^+ = w^+(p_n) \tag{2a}$$

$$\pi_i^+ = w^+(p_i + \cdots + p_n) - w^+(p_{i+1} + \cdots + p_n), \quad \text{for } 0 \leq i \leq n-1 \tag{2b}$$

$$\pi_{-m}^- = w^-(p_{-m}) \tag{2c}$$

$$\pi_i^- = w^-(p_{-m} + \cdots + p_i) - w^-(p_{-m} + \cdots + p_{i-1}), \text{ for } 1-m \leq i \leq 0, \tag{2d}$$

where p are objective probabilities of outcomes, which are ranked for gains from the reference point $i = 0$ to $i = n$, the largest gain, and for losses from the largest loss $i = -m$ to $i = 0$, the reference point. Further, w^+ and w^-, the parametric form of the decision weighting functions, are given by:

$$w^+(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}} \tag{3a}$$

$$w^-(p) = \frac{p^\delta}{(p^\delta + (1-p)^\delta)^{1/\delta}}, \tag{3b}$$

where the parameters γ and δ define the curvature of the weighting function for gains and losses, which leads the probability distortion functions to assume inverse S-shapes. Figure 3 depicts how low probability events are overweighted at the cost of moderate and high probabilities within the CPT probability distortion functions. Tversky and Kahneman (1992) indicate that the weighting functions for gains are slightly more curved than for losses (i.e., $\gamma < \delta$). The parameters estimated by the authors for the CPT model, which are discussed in our empirical analysis, are λ = 2.25, β = 0.88, α = 0.88, γ = 0.61, and δ = 0.69.

[Please insert Figure 3 about here]

3. Data and methodology

In this section, we first describe the theoretical background that allows us to relate empirical density functions (EDF), RND, and subjective density functions. This is a key step for testing the hypothesis that the CPT helps explaining overpricing of OTM options because we build on the assumption that investors' subjective density estimates should correspond on average[8], to the distribution of realizations (Bliss and Panigirtzoglou, 2004). Thus, testing whether the CPT's weighting function explains the overpricing of OTM options, ultimately, relates to how the subjective density function produced by CPT's preferences matches empirical returns. Because the representative agent is not observable, subjective density functions are not estimable like EDF and RND are. As such, we build on the following theory to derive subjective density functions from RNDs.

In our empirical exercise, we first derive subjective density functions for (a) the power and (b) exponential utility functions. Because the CPT model contains not only a utility function (the value function) but also a probability weighting scheme (the weighting function), we produce two density functions resulting from such model: (c) the hereafter called partial CPT density function (PCPT), where only the value function is taken into account, and (d) the CPT density function, where the value and the weighting functions are considered. Lastly, we also calibrate γ to market data and are, then, able to compute (e) the estimated CPT density (ECPT). We provide details on estimation methods for our five subjective density functions (a) to (e) in Section 3.1, and for the RND and EDF in Section 3.5.

Once all five subjective density functions are obtained, we distinguish four main independent analyses in our methodology section: 1) the estimation of a long-term and time-varying γ weighting function parameter (from which we can produce the ECPT density); 2) EVT-based tests of consistency between tails of the EDF, the RND and our five subjective probability distributions; 3) the application of

[8] This implies that investors are somewhat rational. Such an assumption is not incoherent with the CPT assumption that the representative agent is less than fully rational. The CPT suggests that investors are biased, to the extent that its model parameters suggest, but not that decision makers are utterly irrational to the point of rejecting that their subjective density forecast should correspond on average to the realized return distribution.

Kupiec's method in order to test violation of the EDF extreme quantile returns by the RND and our subjective density functions; and 4) a regression linking the CPT time-varying probability weighting parameter (γ) to sentiment measures as well as numerous control variables.

Barclays Capital provided the single stock weighted average IV data used for the largest 50 stocks of the S&P 500 Index. The data consists of closing mid-prices from January 2, 1998 to March 19, 2013 for fixed maturities for five moneyness levels, i.e., 80, 90, 100, 110, and 120, at the three-, six- and twelve-month maturity. Continuously compounded stock market returns are calculated through our analysis from the S&P 500 index prices using a daily frequency, which is obtained via Thomson Reuters Datastream.

3.1. Subjective density functions

Standard utility theory tells us that since the representative agent does not have risk-neutral preferences, RNDs are inconsistent with subjective and physical densities[9]. Hence, if investors are risk-averse or risk-seeking, their subjective probability function should differ from the one implied by option prices. The relation between the RND, $f_Q(S_T)$, and the EDF, $f_P(S_T)$, with S_T being wealth, is accomplished by the representative investor utility function, $U(S_T)$:

$$\frac{f_P(S_T)}{f_Q(S_T)} = \lambda \frac{U'(S_T)}{U'(S_t)} \equiv \zeta(S_T), \tag{4}$$

where λ is the constant subjective discount factor and $\zeta(S_T)$ is the pricing kernel[10]. Thus, the empirical distribution equates to the RND adjusted by the pricing kernel, the subjective density function of the representative investor of utility function $U(S_T)$. By applying Eq. (4) we can estimate the subjective density function for an (rational) investor that has power and exponential utility functions, hereafter, called power and exponential density functions.

Since CPT-biased investors price options as if the data-generating process has a cumulative distribution $F_{\tilde{P}}(S_T) = w(F_P(S_T))$[11], its density function becomes $f_{\tilde{P}}(S_T) = w'(F_P(S_T)) \cdot f_P(S_T)$ (Dierkes (2009) and Polkovnichenko and Zhao (2013)). Thus, CPT-biased agents assess probability distributions as if their tails would contain more weight than in reality they do. In other words, CPT-biased

[9] Anagnou et al. (2002) and Bliss and Panigirtzoglou (2004) have tested the consistency between RNDs and physical densities estimated from historical data and found such distributions are inconsistent, i.e., RNDs are poor forecasters of the distribution of realizations.
[10] The condition necessary for Eq. (8) to hold is that markets are complete and frictionless and a single risky asset is traded.
[11] Similarly, if investors are rational, their subjective density functions should be consistent, on average, with the empirical density function. Bliss and Panigirtzoglou (2004) find that subjective density functions, produced from RND adjusted by two types of representative investors' utility functions (power and exponential) with plausible relative risk aversion parameters, outperform RND on forecasting density functions.

agents 'see' fat tails where, in fact, they are not. Consequently, evaluating whether the CPT's propositions apply is equivalent to testing whether Eq. (4) still holds if $f_P(S_T)$ is replaced by $f_{\tilde{P}}(S_T)$, thus:

$$\frac{w'(F_P(S_T)) \cdot f_P(S_T)}{f_Q(S_T)} = \zeta(S_T), \tag{5}$$

which, re-arranged into Eq. (7), demonstrates that for the CPT to hold, the subjective density function should be consistent with the probability weighted EDF.

$$\underbrace{f_Q(S_T)}_{RND} = \underbrace{w'(F_P(S_T))}_{probability\ weigthing} \cdot \underbrace{f_P(S_T)}_{EDF} \cdot \underbrace{\zeta(S_T)}_{pricing\ kernel} \tag{6}$$

$$\underbrace{f_Q(S_T)}_{RND} = \underbrace{f_{\tilde{P}}(S_T)}_{probability\ weigthed\ EDF} \cdot \underbrace{\zeta(S_T)}_{pricing\ kernel}$$

$$\underbrace{\frac{f_Q(S_T)}{\lambda\frac{U'(S_T)}{U'(S_t)}} = \frac{f_Q(S_T)}{\zeta(S_T)}}_{Subjective\ density} = \underbrace{f_{\tilde{P}}(S_T)}_{probability\ weigthed\ EDF} \tag{7}$$

Following Ait-Sahalia and Lo (2000) and Bliss and Panigirtzoglou (2004), Eq. (7) can be manipulated so that the constant λ of the pricing kernel vanishes, producing Eq. (8), which directly relates the probability weighted EDF, the RND, and the marginal utility, $U'(S_T)$:

$$\underbrace{f_{\tilde{P}}(S_T)}_{probability\ weigthed\ EDF} = \underbrace{\frac{\lambda\frac{U'(S_T)}{U'(S_t)}Q(S_T)}{\int\frac{U'(S_t)}{U'(x)}Q(x)dx} = \frac{\frac{f_Q(S_T)}{U'(S_T)}}{\int\frac{f_Q(x)}{U'(x)}dx}}_{Generic\ Subjective\ density\ function}, \tag{8}$$

where $\int\frac{Q(x)}{U'(x)}dx$ normalizes the resulting subjective density function to integrate to one. Once the utility function is estimated, Eq. (8) allows us to convert RND into the probability weighted EDF. As we hypothesize that the representative investor has a CPT utility function, its marginal utility function is $U'(S_T) = v'(S_T)$, and, thus, $v'(S_T) = \alpha\, {S_T}^{\alpha-1}$ for $S_T \geq 0$, and $v'(S_T) = \lambda\beta(-S_T)^{\beta-1}$ for $S_T < 0$, leading to Eq. (9):

$$f_{\tilde{P}}(S_T) = \frac{\frac{f_Q(S_T)}{\alpha\, {S_T}^{\alpha-1}}}{\int\frac{f_Q(x)}{\alpha x^{\alpha-1}}dx} \quad \text{for} \quad S_T \geq 0\text{, and} \tag{9}$$

$$\underbrace{f_{\tilde{P}}(S_T)}_{probability\ weigthed\ EDF} = \underbrace{\frac{\frac{f_Q(S_T)}{\lambda\beta(-S_T)^{\beta-1}}}{\int\frac{f_Q(x)}{\lambda\beta(-x)^{\beta-1}}dx}}_{Partial\ CPT\ density\ function} \quad \text{for} \quad S_T < 0 \tag{10}$$

Eqs. (9) and (10), hence, relate the EDF where probabilities are weighted according to the CPT probability distortion functions, on the LHS, to the subjective density function derived from the CPT value function, on the RHS, separately for gains and losses, i.e., the PCPT density function. The relationships specified by

Eqs. (9) and (10) fully state the relation we would like to depict, although one additional manipulation is convenient for our argumentation.

Assuming that the function $w(F_P(S_T))$ is strictly increasing over the domain [0,1], there is a one-to-one relationship between $w(F_P(S_T))$ and a unique inverse $w^{-1}(F_P(S_T))$. So, the result $f_{\tilde{P}}(S_T) = w'(F_P(S_T)) \cdot f_P(S_T)$ also implies $f_{\tilde{P}}(S_T)\,(w^{-1})'(F_P(S_T)) = f_P(S_T)$[12]. This outcome allows us to directly relate the original EDF to the CPT subjective density function, by 'undoing' the effect of CPT probability distortion functions within the PCPT density function:

$$\underbrace{f_P(S_T)}_{EDF} = \underbrace{\frac{\frac{f_Q(S_T)}{v'(S_T)}}{\int \frac{f_Q(x)}{v'(x)}dx}(w^{-1})'(F_P(S_T))}_{CPT\ density\ function} \tag{11}$$

Thus, once the relation between the probability weighting function of EDF and the PCPT density is established, as in Eqs. (9) and (10), one can eliminate the weighting scheme affecting returns by applying the inverse of such weightings to the subjective density function without endangering such equalities, as in Eq. (11). This result allows us to obtain a clear representation of the CPT subjective density function, thus, where the value and the weighting function are simultaneously taken into account. At this stage, as we can produce RND and the set of subjective densities of our interest, including the CPT density, one can evaluate how consistent with realizations their tails are.

3.2. Estimating CPT parameters

We start evaluating the empirical validity of the CPT for single stock call options by comparing EDF to the CPT density function parameterized by Tversky and Kahneman (1992). Subsequently, we estimate CPT weighting function parameters λ, β, α, and γ with the same goal. We only estimate γ within the probability weighting function, and not δ, because we are interested in the gains-side of the distribution, which is extracted from call options. We estimate these parameters non-parametrically, by sequentially minimizing the squared distance between physical distribution and the partial CPT density function for every bin of the two distributions, as follows:

$$v(\alpha = \beta, \lambda^*) = Min \sum_{b=1}^{B} \left(EDF_{prob}^{b} - CPT_{prob}^{b}\right)^2, \tag{12a}$$

$$v(\lambda, \alpha^*, \beta^*) = Min \sum_{b=1}^{B} \left(EDF_{prob}^{b} - CPT_{prob}^{b}\right)^2, \tag{12b}$$

[12] One drawback of the CPT model is that it enables for non-strictly increasing functions, which would not allow invertibility. This is the reason why a newer literature on probability distortions functions favors other strictly monotonic functions, such as Prelec's (1998), $w(p) = e^{-(-\ln(p))^{\delta}}$ as their weighting functions. Nevertheless, because the CPT parameters of our interest (γ = 0.61; δ = 0.69) impose strict monotonicity, we can obtain the inverse of the probability function, $w^{-1}(p)$ numerically.

where, $\alpha = \beta$, the value function diminishing sensitivity parameters, in Eq. (12a) are constrained to be between [0 and 1.6]; and λ, the risk aversion parameter, in Eq. (12b) is constrained to the [0.5,3] interval and B is the total number of bins. Once the optimal λ, β, and α are known, we minimize

$$w^{\mp}(\gamma, \delta^*) = Min \sum_{b=1}^{B} (EDF_{prob}^{b} - CPT_{prob}^{b})^2 , \quad (13)$$

where, γ, the probability weighting parameter for gains, is constrained to the (0.28, 1.2] interval.

Our non-linear bounded optimization is a single parameter one, where we first estimate optimal (constraint) intervals for the, $\alpha = \beta$, and γ, and subsequently we estimate them as suggested by the sequence of optimizations described by Eqs. (12a), (12b), and (13). This method resembles the one of Dierkes (2009), Kliger and Levy (2009), Chabi-Yo and Song (2013), and Polkovnichenko and Zhao (2013). Once optimal parameters λ, $\alpha = \beta$, and γ are estimated, we can produce another subjective density function: the ECPT, which stands for estimated CPT, where we apply the optimal γ's for the characterization of its probability weighting function.

3.3. Density function tails' consistency test

We check for tail consistency of our set of five subjective density functions (CPT, PCPT, ECPT, power and exponential), RND, and the EDF by applying extreme value theory (EVT). EVT allows us to estimate the shape of the tails of these eight PDFs and to extract the returns implied by an extreme quantile within our PDFs. We estimate the tail shape estimator (φ) by means of the Hill (1975) estimator:

$$\hat{\varphi} = \frac{1}{\hat{\theta}} = \frac{1}{k} \sum_{j=1}^{k} ln\left(\frac{x_j}{x_{k+1}}\right) , \quad (14)$$

where k is the number of extreme returns used in the tail estimation, and x_{k+1} is the tail cut-off point. The tail shape estimator φ measures the curvature, i.e., the fatness of the tails of the return distribution: a high (low) φ indicates that the tail is fat (thin). The inverse of φ is the tail index (θ), which determines the tail probability's rate of decay. A high (low) θ indicates that the tail decays quickly (slowly) and, therefore, is thin (fat). Such tail shape estimator and tail index give us a good representation of the curvature of the tails, but since tails may have the same shape while estimating diverse extreme observations, we also employ the semi-parametric extreme quantile estimator from de Haan et al. (1994):

$$\hat{q}_p = x_{k+1} \left(\frac{k}{pn}\right)^{\frac{1}{\hat{\theta}}}, \quad (15)$$

where n is the sample size, p is the corresponding exceedance probability, which means the likelihood that a return x_j exceeds the tail value q, and $x_{k+1}$is the tail cut-off point. We note that $\hat{q}_p$ has as one of its inputs the tail shape estimator φ. Similar to value-at-risk (VaR) modeling, the $\hat{q}_p^-$ statistic indicates the level of the worst return occurring with probability p, which is small. This is the reason why we call $\hat{q}_p$ extreme quantile return (EQR). As we are interested only in the upside returns with a p probability estimated from calls, we only compute $\hat{q}_p^+$ by applying the same methodology to the right side of the RND obtained from the single stock option market[13].

In addition to the EQR, we also evaluate the density function tails' using expected shortfall (ES), which captures the average loss beyond the tail cut-off point. As we are interested in the upside of the distribution, we call such measure expected upside (EU) as the average gain beyond the tail cut-off point. We evaluate the EU following Danielsson et al. (2006) formulae for the ES, which relates the EQR (i.e., the VaR) to the ES (i.e., the CVaR) as described below:

$$\widehat{EU}_{q(p)} = \frac{\hat{\theta}}{\hat{\theta}-1} * x_{k+1} \left(\frac{k}{pn}\right)^{\frac{1}{\hat{\theta}}}, \quad (16)$$

where θ is the tail index.

De Haan et al. (1994) show that the tail shape estimator statistic $\sqrt{k}(\hat{\varphi}(k) - \varphi)$ and the tail quantile statistic $\frac{\sqrt{k}}{ln\left(\frac{k}{pk}\right)}\left[ln\frac{\hat{q}\,(p)}{q\,(p)}\right]$ are asymptotically normally distributed. Hence, according to Hartmann et al. (2004) and Straetmans et al. (2008), the *t*-statistics for such estimators are given by:

$$T_{\varphi} = \frac{\hat{\varphi}_1 - \hat{\varphi}_2}{\sigma[\hat{\varphi}_1 - \hat{\varphi}_2]} \quad \sim N(0,1), \quad (17a)$$

and

$$T_{q} = \frac{\hat{q}_1 - \hat{q}_2}{\sigma[\hat{q}_1 - \hat{q}_2]} \quad \sim N(0,1), \quad (17b)$$

where the denominators are calculated as the bootstrapped difference between the estimated shape parameters $\hat{\varphi}$ and the quantile parameters $\hat{q}_p$ using 1,000 bootstraps. The null hypothesis of this test is that the $\hat{\varphi}$ and $\hat{q}_p$ parameters do not come from independent samples of normal distributions, therefore, $\hat{\varphi}_1 = \hat{\varphi}_2$ and $\hat{q}_1 = \hat{q}_2$. The alternative hypothesis is that the $\hat{\varphi}$ and $\hat{q}_p$ parameters have unequal means. Such *t*-test is also applied to our EU analysis, as the distribution of the EU follows the same distribution of the tail quantile statistic $\frac{\sqrt{k}}{ln\left(\frac{k}{pk}\right)}\left[ln\frac{\hat{q}\,(p)}{q\,(p)}\right]$, given that the EU is the extreme quantile estimator multiplied by a constant.

3.4. Estimating RND and EDF

For the estimation of the RND, the first step taken is the application of the Black-Scholes model to our IV data to obtain options prices (*C*) for the S&P 500 index. Once our data is normalized, so strikes are expressed in terms of percentage moneyness, the instantaneous price level of the S&P 500 index (S_0) equals 100 for every period for which we would like to obtain implied returns. Contemporaneous dividend

[13] Our EQR measure is closely connected to the risk-neutral tail loss measure of Vilkov and Xiao (2013).

yields for the S&P 500 index are used for the calculation of *P* as well as the risk-free rate from three-, six-, and twelve-months T-bills. Because we have IV data for five levels of moneyness, we implement a modified Figlewski (2010) method for extracting our RND structure, as in Felix et al. (2015).

The Figlewski (2010) method is close to the one by Bliss and Panigirtzoglou (2004), where body and tails are also extracted separately. Bliss and Panigirtzoglou (2004) use a weighed natural spline algorithm for interpolation, which has the same decreasing-noise effect in RNDs of using splines in the absence of knots, as done in Figlewski (2010). The extrapolation in Bliss and Panigirtzoglou (2004) is done by the introduction of a pseudo-data point, which has the effect of pasting lognormal tails into the RND. One advantage of these two approaches is that the extrapolation does not result in negative probabilities, which is possible when spline interpolation is applied in such case. Nevertheless, we favor Figlewski's (2010) approach as the lognormal tails employed by Bliss and Panigirtzoglou (2004) assume that IV is constant beyond the observable strikes, resembling the Black-Scholes model. The modification made to the Fligleswki (2010) method by Felix et al. (2015) entailed having flexible inner anchor points (as opposed to having fixed anchor points) for fitting tails to the risk neutral density. The aim of this modification is to prevent the method to estimate distribution density functions with implausible shapes.

We estimate the EDF in two different ways. First, using the entire sample of realized returns (r) we estimate the 'long-term' EDFs non-parametrically, where $r = ln(S_T/S_t)$ and S_t is the S&P 500 index at time t in the sample, and S_T is the forward level of the same index three, six or twelve months forward. Because of overlapping periods, we estimate our empirical distribution of returns for these three maturities using multiple samples and distinct starting points.

In a second step, we estimate time-varying EDFs, built from an invariant component, the *standardized innovation density* and a time-varying part, the conditional variance ($\sigma^2_{t|t-1}$) produced by an EGARCH model (see Nelson, 1991). We first define the *standardized innovation,* being the ratio of empirical returns and their conditional standard deviation ($\ln(S_t/S_{t-1})/\sigma_{t|t-1}$) produced by the EGARCH model. From the set of standardized innovations produced, we can then estimate a density shape, i.e., the *standardized innovation density*. The advantage of such a density shape versus a parametric one is that it may include typically observed fat tails and negative skewness, which are not incorporated in simple parametric models, e.g., the normal. As mentioned, such density shape is invariant and it is turned time-varying by multiplication of each standardized innovation by the EGARCH conditional standard deviation at time t, which is specified as follows:

$$\ln(S_t/S_{t-1}) = \mu + \varepsilon_t, \qquad \varepsilon_t \sim f(0, \sigma^2_{t|t-1})$$

and (18)

$$\sigma^2_{t|t-1} = \omega_1 + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{t-1|t-2} + \vartheta Max[0, -\varepsilon_{t-1}]^2,$$

where α captures the sensitivity of conditional variance to the lagged squared innovations (ε_{t-1}^2), β captures the sensitivity of the conditional variance to the conditional variance ($\sigma_{t-1|t-2}^2$), and ϑ allows for the asymmetric impact of lagged returns ($\vartheta Max[0, -\varepsilon_{t-1}]^2$). The model is estimated using maximum log-likelihood where innovations are assumed to be normally distributed.

Up to this point, we managed to produce a one-day horizon EDF for every day in our sample but we still lack time-varying EDFs for the three-, six-, and twelve-months horizons. Thus, we use bootstrapping to draw 1,000 paths towards these desired horizons by randomly selecting single innovations (ε_{t+1}) from the one-day horizon EDFs available for each day in our sample. We note that once the first return is drawn, the conditional variance is updated ($\sigma_{t-1|t-2}^2$) affecting the subsequent innovation drawings of a path. This sequential exercise continues through time until the desired horizon is reached. In order to account for drift in the simulated paths, we add the daily drift estimated from the long-term EDF plus the risk-free rate to drawn innovations, thus the one-period simulated returns are given as $\varepsilon_{t+1} + \mu + Rf$. The density functions produced by the collection of returns implied by the terminal values of every path and their starting points are our three-, six-, and twelve-months EDFs. These simulated paths contain, respectively, 63, 126, and 252 daily returns. We note that by drawing returns from stylized distributions with fat-tails and excess skewness, our EDFs for the three relevant horizons also imbed such features. Finally, once these three time-varying EDFs are estimated for all days in our sample, we estimate λ and γ for each of these days using Eq (12b) and (13) [14].

Our approach for estimating both the long-term EDF and the time-varying EDF is closely connected to the method applied by Polkovnichenko and Zhao (2013). The time-varying method used by these authors is based on Engel and Rosenberg (2002). The choice for an EGARCH approach versus the standard GARCH approach is due to the asymmetric feature of the former model that imbeds the "leverage effect".[15]

4. Empirical analysis and results

In this section, we present our results of the empirical analysis described in Sections 3. We note that since we estimate EDF in the two ways described (the "long-term" and time-varying EDFs), we are able to estimate time-varying γ*'s* and long-term γ*'s* by minimizing Eq (13). We use our long-term γ estimates to compute the ECPT with the aim to compare it to the other subjective density functions using the tests

---

[14] Due to drift, time-varying EDF's mode occasionally does not match the PCPT model's one (for estimation of γ). This difference has been challenging for the optimization that estimates γ (Eq. (13)) for the twelve-month horizon, as a larger amount of γ estimates produce unreasonable PDFs such as non-monotonic CDFs. Therefore, to perform such optimizations we neutralize the drift's impact by forcing the mode of the simulated EDF to match the one of the PCPT.

[15] The leverage effect is the negative correlation between an asset's returns and changes in its volatility. For a comparison between alternative GARCH approaches, see Bollerslev (2009).

described in Section 3.3. The time-varying estimates of λ and γ are analyzed in Sections 4.3 and 4.4, respectively, with the use of an ordinary least squares (OLS) regression model. We describe this regression together with its results in section 4.4. Finally, in Section 4.5, we perform robustness tests on our results by using an alternative weighting function to the CPT; the one imbedded in the Prelec (1998) model.

4.1. Estimated CPT long-term parameters

We report the estimated CPT parameters ($\lambda,\ \alpha = \beta$, and γ) extracted from long-term density functions in Table 1, Panel A. Our first finding is that the λ parameter of risk aversion, which is calibrated to 2.25 in the CPT, does indeed fluctuate around that number. Our estimation of λ from three-month options is 1.98, whereas for the six- and twelve-months option it is 2.49 and 1.85, respectively. This finding suggests that risk aversion is higher at the six-months maturity, even higher than suggested by the CPT. The α and β parameters obtained are all close to unity, suggesting that the diminishing sensitivity to gains and losses is higher than suggested by the CPT (i.e., 0.88).

The estimated probability weighting function parameters γ matches the one suggested by the CPT (i.e., 0.61) at the three-month horizon but overshoots the CPT ones at longer horizons. At the six-month option, γ is around 0.7, whereas at the twelve-month horizon, it is close to unity, 0.97. These results suggest that overweighting of small probabilities is generally present within the average pricing of single stock options. Such lottery tickets buying occurs essentially in the short-term option markets (up to six-months), while the twelve-month option market seems to be more rational. These findings provide initial support of our hypothesis that individual investors do behave as buying lottery tickets (i.e. overweighting small probability events) when purchasing single stock call options, as suggested by Barberis and Huang (2008).

4.2. Density functions tails' consistent test results

As specified in Section 3.3, we test the empirical consistency of density function tails among a set of five subjective distributions (CPT, PCPT, ECPT, power, exponential), the RND, and the EDF. We perform such tests by employing EVT through the application of Eqs. (14) to (17). In order to apply such methods, we require return streams (x_j), which are only available for the long-term EDF. Thus, we apply an inversion transform sampling technique to our other PDFs to gather sampled returns for them. Such method, also known as the Smirnov method, entails drawing n random numbers from a uniformly distributed variable $U = (u_1, u_2, \dots, u_n)$ bounded at interval [0,1] and, subsequently, computing $x_j \leftarrow F^{-1}(u_j)$, where F are the CDFs of interest (see Devroye, 1986, p. 28). Hence, the Smirnov method simulates returns that resemble the ones of the inverse CDF by randomly drawing probabilities along such function.

Once we obtain returns for all five PDFs, the next step is to set k as the optimal number of observations used for estimation of φ by Eq. (14), the Hill-estimator. For this purpose, we produce Hill-plots for the right tail of our distributions, which depict the relationship between k and φ as a curve (Straetmans et al., 2008). Picking the optimal k is done by observing the interval in such curve where the value of φ stabilizes while k changes. This area suggests a stable trade-off between a good approximation of the tail shape by the Pareto distribution and the uncertainty of such approximation (by the use of fewer observations). The interval that corresponds to roughly four to seven percent of observations seems to be a stable region across the Hill-plots of the tails of the EDF and the CPT. As an increase in k increases the statistical power of the estimator but may distort the shape of the tail, we decide to set k as chosen from the Hill-plots for EDF and CPT tails, equal to four percent.

We examine whether the tail shape parameter (φ), computed via the Hill (1975) estimator, for the RND and for our subjective density functions (i.e., power, exponential, PCPT, CPT and ECPT) match the one for the EDF. The outcomes from the statistical tests performed to compare tail shape parameters (Eq. (17a)) are provided in Table 1, Panel B. The results suggest that for the three-month maturity options, φ for the CPT and ECPT (at 0.25) are very close and statistically equal to the EDF parameter (at 0.24). The φ estimate for the RND (0.2) does also statistically equal the one from the physical returns, despite being slightly off the EDF parameter. The φ estimate for the power, exponential, and PCPT density functions do not match the one for the EDF, as they are all around 0.17 and, thus, exhibit fatter tails than the EDF.

We observe that the results for the six- and twelve-months options are very similar to the ones obtained for the three-months expiry. The parameter estimate φ of the EDF is statistically equal to the CPT and ECPT. Parameter φ ranges from 0.18 to 0.24 for the CPT, ECPT, and RND for the six- and twelve-months maturities, whereas it is 0.23 for the EDF. The estimate of φ for the RND (0.19 and 0.22 for the six- and twelve-months maturities, respectively) also matches the one for the EDF but with less statistical power than for the CPT. The parameter estimates φ for the power, exponential, and PCPT density functions do not match the EDF's φ in any case. Generally, the parameter estimates φ for these subjective density functions are too small in comparison to the one of the EDF. This means that such six- and twelve-month maturity subjective density functions have fatter tails than the EDF and the other subjective densities (CPT and ECPT), and the RND. These results suggest that the shape of the CPT density function seems to be a good match to the shape of realized tails, supporting our hypothesis that individual investors behave as lottery buyers when trading in single stock options.

After k is chosen and the shape estimator φ for the EDF, RND, power, exponential, PCPT, CPT, and ECPT is computed, extreme quantile returns (EQR) can also be estimated via Eq. (15). Subsequently, the t-test in Eq. (17b) is applied using the one, the five and the ten percent statistical significance levels. It aims to evaluate whether the EQRs estimated from a set of two distributions being compared (RND, power, exponential, PCPT, and CPT versus EDF) have equal means (the null hypothesis). The results of this test are shown in Table 2, Panel A.

[Please insert Table 2 about here]

Analyzing the density functions derived from the three-months option maturity, we find that the EQR implied by the CPT matches very closely the realized ones: 11 percent at the tenth quantile, 13 percent for the fifth quantile, and around 20 percent for the first quantile. The EQR implied by the ECPT is the same as implied by the CPT, thus it also matches the EDF. This result was expected as the estimated long-term parameter γ used in the ECPT matches the one of the CPT (0.61). Contrarily, the EQRs for the RND, power, exponential, and PCPT densities always overshoot the one for the EDF. All comparisons between these distributions' EQR at the three-months maturity reject the null hypothesis that returns at the same quantile are equal. This pattern is observed across all quantiles analyzed, i.e., at the tenth, the fifth, and the first quantiles. This empirical finding indicates that the equity market upside implied in option markets the (i.e., the RND) and the power, exponential and PCPT densities are always higher than the ones realized by the equity market. The results for the PCPT are somewhat similar to the ones for the RND. The exponential utility density has the farthest off EQR relative to the EDF. On average, EQR for the exponential and power utilities overshoot the one for the EDF by roughly ten percent. Thus, the EQRs from the CPT and the ECPT are by far the best matches of the EQR for the EDF.

For the six-months maturity, upside returns priced by the RND and ECPT are the ones to best match the EQR for the realized return distribution. The EQR for the EDF are roughly 18, 21, and 30 percent for the tenth, the fifth, and the first quantile of returns, respectively, whereas the EQRs for the ECPT are 17, 20, and 26 percent. For the RND, such extreme upside return estimates are 19, 22, and 30 percent. Thus, the ECPT statistically matches the realized EQR best at the tenth and fifth quantile, whereas the RND is the best match for the first quantile. No rational subjective density function consistently matches the EQR of the EDF. The power, exponential, and PCPT densities always overshoot the EQR of realized returns. Contrarily, the CPT density always undershoots the EDF's extreme returns. Despite always overshooting the EQR of the EDF, the PCPT is the only other subjective density (apart from the ECPT) that has a statistically equal EQR of the EDF, which happens only at the first quantile EQR and with only weak (5 percent) statistical significance.

In contrast to the three- and six-months maturities, the EQR from the RND for the twelve-months maturity all underestimate the EQR from realized returns. The EQR of realized returns are 25, 29, and 42 percent for the tenth, fifth and first quantiles, respectively, whereas for the RND these are 22, 26, and 37 percent, respectively. However, the EQRs of the power and of the exponential densities continue to largely overshoot the ones for the EDF. The PCPT and the ECPT, in which probabilities are neutrally (or rationally) weighted (recall that γ estimated for the twelve-month horizon is 0.97), are the models that best match the EQR of the realized returns, as they are around 25, 29, and 39 percent, respectively, for the three quantiles studied. As such, the EQR of realized returns, the PCPT and the ECPT densities are strongly statistically equal at the tenth and fifth quantiles and somewhat weaker statistically equal at the first quantile. This finding supports our earlier evidence that twelve-months single stock options are likely priced more rationally than short-term ones. Given that the CPT model will always reduce the EQR from the PCPT, as it is parameterized for overweighting small probabilities, the EQRs for these distributions undershoot the EDF ones. Such an overshooting is very large in the current case, as the EQR implied by the CPT is 15, 18, and 26 percent for the tenth, fifth and first quantiles, respectively.

The above results suggest that the CPT-related densities, despite not always matching the EQRs of the EDF, seem to be the framework with the highest consistency in matching the EQR of the EDF. The ECPT is, by far, the best performing model in matching realized extreme quantile returns. This result is not a surprise as allowing γ to vary within the CPT weighting function entails extra flexibility specified within the CPT model on matching the data versus traditional utility functions. Thus, if our findings suggest that the CPT is not the fully applicable model to explain single stock options pricing, it is important to note that the basic features of the CPT of overweighting small probabilities is still fully embraced by the empirical evidence, with exception of the twelve-month options. These findings reiterate our takeaway from Section 4.1, in which a horizon-effect seems to play a role with the pricing of these options: twelve-months options seem to be priced more rationally than the shorter term ones, which seem to be priced as a result of lottery buying by individual investors.

In line with the results for the EQR presented in the previous section, Table 2, Panel B, shows that the expected upside for the EDF is more closely matched in the three-months horizon by the expected upside for the CPT and ECPT density functions for the tenth, fifth and first quantile. The three-months horizon expected upside estimated from the realized returns is given as 15, 17, and 26 percent for the mentioned quantiles. The CPT and ECPT expected upside for this same horizon is 14, 17, and 25 percent, respectively. These estimates are statistically equal to the ones estimated from the realized returns. Once again, similarly to our analysis on the EQR, for the other subjective densities, the expected upside for all quantiles is also much larger than the EDF expected upside. The exponential density has the highest expected upside across the different quantiles, being the furthest away from the realized returns. The RND-implied expected upside is somewhat conservative and relatively closer to the realized ones, however, never statistically significant equal to it.

For the six-months maturity, the expected upsides for the CPT and ECPT density functions are no longer that close to each other nor to the realized ones. The EDF expected upside is always higher than the ones for the CPT and ECPT. Only at the tenth quantile, the expected upside of the ECPT density function is statistically significant equal to the realized one. The CPT's expected upside is always lower than the one of the ECPT density function, as the γ parameter estimated equals 0.71, implying less overweighting of small probabilities than the CPT. The densities which better match the expected upside of the EDFs are the PCPT and, specially, the RND for which the expected upside is statistically equal to the EDF ones across all quantiles.The expected upside is statistically equal for the PCPT at the one percent level and for the first quantile tail only.

Similar to the results from our EQR analysis, Table 2, Panel B, shows that the expected upside is statistically equal at the twelve-months horizon for the PCPT, ECPT, and the realized returns for the tenth and fifth quantiles. The expected upside for the realized returns is 32, 37, and 54 percent for the tenth, fifth and first quantile, respectively, whereas for the PCPT it is 32, 36, and 48 percent, respectively, and 31, 35, and 47 percent, respectively, for the EPCT. The proximity between the expected upside for the PCPT and for the ECPT is due to the fact that the estimated γ parameter at the twelve-months horizon is 0.97, thus, we observe a virtually neutral-weighting of probabilities. The only exception to the match of the realized expected upside and the one for the PCPT and ECPT density functions is at the first quantile, where the expected upside for the exponential and power densities matches the expected upside for the EDF of roughly 54 percent. At the twelve-months horizon, the expected upside coming from the exponential and power utilities, typically overshoots the ones for the EDF, whereas the RND expected upside undershoots them. The expected upside implied by the CPT at the twelve-months horizon strongly undershoots the expected upside of the realized returns. This observation suggests that a pronounced overweighting of small probabilities or lottery buying is strongly negated by the expected upside from realized returns. As such, we conclude that long-dated options seem to be priced more rationally than short-term options.

These results on the expected upside reiterate the insights gained by our earlier findings using the EQR and the estimation of long-term γ. The CPT-model densities, despite not always matching the EDF's expected upside, seem to constitute the framework that more frequently approximates the right tail of the EDF, provided that an appropriate γ is used. The extent that the original CPT matches the data seems to be related to the option horizon: the CPT model and our proposition of single stock options as lottery tickets is more applicable to short-term options than long-term options.

Figure 4 compares graphically the CDFs from six of our equity return densities: the EDF, the RND, the CPT, the PCPT, the exponential and the power density[16]. We focus on the right tails of these distributions as we are interested in how closely the RND from call options and derived subjective density functions match the tails of the EDF. The plots display the cumulative probabilities on the *y*-axis and the terminal price levels on the *x*-axis, given an initial price level of 100.

[Please insert Figure 4 about here]

We see that the tails implied by option prices (RND, in red) seem fatter than the tails from the CPT (in dark blue) and EDF (in green) density functions over the three-months horizon. The tails for the CPT and the EDF are almost identical above the 110 terminal level, i.e., at the 10 percent return. The right tail of the RND distribution is clearly much fatter than the ones of the CPT and EDF but it is still thinner than the ones of the PCPT, the power and the exponential densities. Thus, the upside risk implied from options is much higher than the one realized by the EDF, a sign of a potentially biased behavior by investors in such options. This observation is confirmed by the tail shape parameter (φ), the EQRs and the EU estimated across the different quantiles, which in all cases report higher a upside in the RND than in the EDF and the CPT. Figure 4 also suggests that the upside risk of the RND seems to be much more consistent with the one of the PCPT density, whereas the CPT tails seem very distinct from the PCPT, which is in line with our earlier findings.

The plot in column B, which depicts the CDF for our studied densities at the six-months horizon, suggests that the RND and the EDF are much closer than at the three-months horizon. At the same time, the CPT density seems somewhat disconnected from the EDF, at least beyond the 105 terminal level. This finding matches our results from the EQR and the expected upside comparisons. The PCPT tail is, at this horizon, higher than the EDF, CPT, and RND ones. But, the PCPT tail is much closer to the EDF one than to the CPT one, especially at its very extreme. This finding is also confirmed by our EQR and expected upside tests, as the PCPT is statistically equal to the EDF at the one percent quantile. The exponential and power utility densities have right tails that are much fatter than the other densities, including the EDF.

Figure 4 shows that at the twelve-months horizon the CPT's CDF tails seem completely disconnected from the EDF. The EDF tails are much fatter than the CPT ones and slightly fatter than the RND ones. In fact, the RND seems to match the EDF for terminal levels above 120. This finding suggests that long-term options trade in a much less CPT-biased manner than short-term options. At this horizon, the PCPT density seems to closely match the EDF, much better than the exponential and power utilities. This finding is confirmed by the results of the EQR and the expected upside tests over such a horizon, which suggest that the tails of the EDF and PCTP are statistically equal in almost all quantile levels.

[16] We omit the ECPT for better visualization as its CDFs are very similar to the CPT ones. The similarity is caused by the ECPT left tail weighting function parameter (δ) being the same for the CPT and because its estimated long-term γ is not too different from the Tversky and Kahneman (1992) ones for the three maturities.

Overall, this visual inspection of our density function CDFs confirms our hypothesis that end-users of OTM single stock calls are likely biased and behave as buying lottery tickets when trading short-term options. The fact that the CPT and the PCPT density (as it already incorporates some elements of the CPT model) are the best matches for the right tail of the EDF suggests that individual investor overweight small probabilities. On the other hand, end-users of OTM index calls at the six-months horizon seem to be neutral to risk, as the RND is the best forecaster of the empirical returns at that time span, while options trading at the twelve-months horizon is definitely not CPT-biased.

These results strengthen the evidence provided by Barberis and Huang (2008), Ilmanen (2012), and Barberis (2013) that investors push options prices to extreme levels because they have a biased model (i.e., CPT) for estimating the distribution of equity returns. Investors seem to overweight small probabilities and to behave as buying lottery tickets when trading single stock options, especially at short-term horizons (i.e. three-months).

4.3. Estimated CPT time-varying parameters

In order to investigate any potential time-variation in the CPT's overweighting of small probabilities or lottery tickets buying reflected in single stock options, we apply Eqs. (12b) and (13) to each day in the sample to estimate the empirical γ (weighting function) and λ (risk aversion) parameters. We first evaluate the results for the estimation of λ *(*risk aversion). We report summary statistics of the estimated λ for three-, six-, and twelve-months options in Panels A and B of Table 3. Panel A reports the statistics when λ is estimated, assuming that the other two parameters of the value function, α and β, equal 0.88, i.e., the CPT calibration. In contrast, in Panel B the parameters α and β are optimized using our data following the application of Eq. (12a). Our general impression is that the λ estimates are very similar across the two calibrations. The median of λ is 1.75 for the CPT parameterization and 1.74 when the parameters α and β vary for the three-months option maturity. The median for the six-months maturities for these two methods equal 1.81 and 1.79, respectively, and for the twelve-months maturity, these values are 1.89 and 1.92, respectively. Thus, we find that the α and β parametrization virtually does not affect the λ estimates. The distributions of λ for these two types of α and β parametrization are also very close to each other, since the 25th and 75th percentile are almost the same for these different α and β parameters. We also observe that the risk aversion estimated is slightly lower than the implied one of the CPT model (at 2.25). Our findings suggest that the kink observed in the value function's reference point is less pronounced than suggested by the CPT model. The λ estimates can be highly volatile though, with a standard deviation of around 0.5 and a deviation range between 0.57 and 2.97. Such a volatility and such a wide fluctuation interval of λ confirm that this parameter is strongly time-varying.

[Please insert Table 3 about here]

We report summary statistics of the estimated CPT parameter γ for three-, six- and twelve-months options in Panels C of Table 3. We find that the median and the mean time-varying values of γ, estimated from the three-months options, roughly match the parameter value of 0.61 suggested by Tversky and Kahneman's (1992) which range between 0.66 and 0.68. This suggests that the overweighting of small probabilities is present within the pricing of three-months call options as suggested by the theory. Nevertheless, our daily estimates of γ are volatile, having a standard deviation of 0.29. The estimates for γ range from 0.28 to 1.20 (i.e., an underweighting of small probabilities) and their distribution is slightly skewed to the left as the median is just smaller than the mean.

At the six-months maturity, the overweighting of small probabilities seems less acute than suggested by the theory and by the empirical results of the three-months options. The median and the mean γ for such maturity are 0.72 and 0.75, respectively, from the time-varying estimation. The long-term γ equals 0.70 and is somewhat in line with the time-varying estimates. The distribution of γ is somewhat skewed to the left (towards a more pronounced overweighting of small probabilities), as the median is smaller than the mean. The 75th quantile of γ equals 1.05 and suggests a neutral weighting of probabilities.

The results for the twelve-months maturity tend even more towards a neutral probability weighting than the six-months ones. The median γ is 0.83, whereas the mean γ is 0.80. In contrast with other maturities, the distribution of γ is strongly skewed to the right, towards a neutral- to under-weighting of small probabilities. The estimates are highly volatile though, as the standard deviation of γ is 0.32.

In summary, statistics discussed above and reported in Table 3, Panel C, indicate that the weighting function parameters γ for the three maturities evaluated are clearly time-varying[17]. But since the risk -aversion parameter λ is also time-varying, we re-estimate γ for each day in our sample using the daily estimate of λ in Eq. (13).

We report the summary statistics of the new γ estimates in Panel D of Table 3. The new median and the new mean estimates for γ are 0.72 and 0.74, respectively, and, thus, higher than when γ was estimated under a fixed risk aversion calibration. The 75th percentile of γ also increases, from 0.87 to 0.94, respectively. At the six-months horizon, the difference between γ with a fixed or with a varying λ is even larger. The median γ for the static λ is 0.72, whereas for the time-varying it is 1.07. The means are given as 0.74 and 0.96, respectively. At the 75th percentile using the time-varying λ, γ becomes 1.2, which is the upper bound of our optimization. For the twelve-months maturity, we observe a similar effect. The median γ for the time-varying λ is 1.15, whereas for the static one it is 0.83.

[17] For the three maturities evaluated, γ estimates can be as low as zero if the lower bound is set at zero, which is unreliable as it specifies a probability weighting function for the CPT that is non-monotonic (for further details on the subject see Ingersoll, 2008). The imposition of a lower-bound of 0.28 for γ in order to avoid such non-monotonicity problem does not alter, however, our estimated mean and median γ materially.

These findings suggest that once we account for empirical risk aversion, the extent that investor overweight smaller probabilities reduces. For six- and twelve-months maturities, on average, probabilities are neutrally weighted. This implies that RND tails might be fatter than EDF ones not because investor are biasedly overweighting small probabilities, but because the risk premium charged by investors to hold call options is lower. Interestingly, an average overweighting of small probabilities remains present in three-months options, despite changes in investors' risk aversion.

Further, accounting for the time-variation of λ does not affect the time-variation in γ. The standard deviation of the γ estimates barely changes when we account for time-varying λ. This result applies to options across all maturities. Thus, the fact that six- and twelve-month options show, on average, neutral or under-weighting of tails does not preclude that overweighting (or even underweighting) of small probabilities is present at times. An alternation between periods of overweighting of small probabilities and underweighting is also observed by Chabi-Yo and Song (2013) and Polkovnichenko and Zhao (2013), with a clear prevalence of overweighting of tails. Their finding is very consistent with ours, even after we have accounted for the time-variation of λ. The fact that their conclusions are based on options with short maturities of 28 days, also connects to our findings, as our strongest results occur for options with the shortest maturity in our sample (three-months).

4.4. Time variation in the probability weighting parameter and investor sentiment

As observed in Section 4.3, the probability weighting parameter γ is clearly time-varying given the large standard deviation and the range observed. In the following we investigate which factors may explain such time-variation of γ. Our main hypothesis is that the time variation of γ is linked to investors' sentiment. The link between sentiment and overweighting of small probabilities or lottery buying in out-of-the-money (OTM) single stock calls originates from the fact that individual investors are highly influenced by market sentiment and attention-grabbing stocks (Barberis et al., 1998; Barber and Odean, 2008), and that OTM single stock calls trading is speculative in nature and mostly done by individual investors (Lakonishok et al., 2007). For instance, Lakonishok et al. (2007) find that during the IT bubble of 2000, a period of high variation of γ is linked to investors' sentiment, when the least sophisticated investors were the ones that substantially purchased calls on growth and IT stocks. Figure 5 depicts the time-variation in γ and the Baker and Wurgler (2006) sentiment factor. It provides evidence that these measures move in tandem at times. For example, during the IT bubble, the level of γ seems quite connected with the level of sentiment, especially for the six-months options. At the same time, changes in γ from three-months options seem linked to changes in sentiment, despite a relative disconnect in the levels.

[Please insert Figure 5 about here]

To formally test our hypothesis that time variation of γ is linked to investors' sentiment, we design a regression model. Within Eq. (19), the explained variables are γ for the three-, six-, and twelve-months horizons and the explanatory variables are the Baker and Wurgler (2006) sentiment measure[18], the percentage of bullish investors minus the percentage of bearish investors given by the survey of the American Association of Individual Investors (AAII), a proxy for individual investors' sentiment (see Han, 2008), and a set of control variables among the ones tested by Goyal and Welch (2008)[19] as potential forecasters of the equity market. The data frequency used in the regression is monthly as this is the highest frequency available from the sentiment data and from the Goyal and Welch (2008) data set[20]. Our regression sample starts in January 1998 and ends in December 2010[21]. Our OLS regression model is specified as follows:

$$\gamma_t = c + Sent_t + IISent + E12_t + B/M_t + Ntis_t + Rfree_t + Infl_t + Corpr_t + Svar_t + CSP_t + \varepsilon_t, \quad (19)$$

where *Sent* is the Baker and Wurgler (2006) sentiment measure, *IISent* is the AAII individual investor sentiment measure, *E12* is the twelve-months moving sum of earnings on the S&P 500 index, B/M is the book-to-market ratio, *Ntis* is the net equity expansion, *Rfree* is the risk-free rate, *Infl* is the annual inflation rate, *Corpr* is the corporate spread, *Svar* is the stock market variance, and *CSP* is the cross-sectional premium.

Additionally, we run univariate models for each explanatory factor to understand the individual relation between γ and the control variables:

$$\gamma_t = \alpha_i + \beta_i x_{i,t} + \varepsilon_{t+1}, \quad (20)$$

where x replaces the n explanatory variables earlier specified, given $i = 1 \ldots n$.

[Please insert Table 4 about here]

Table 4, Panel A presents the estimates of Eq. (19). We note the high explanatory power of the multivariate regression, ranging from 31 to 57 percent. As expected, we observe that *Sent* is consistently negative across the three different horizons studied. We find that high sentiment exacerbates overweighting of small probabilities in calls of shorter horizon. However, *Sent* is not significantly linked to γ at the twelve-months horizon. The univariate regressions of *Sent* confirm the negative link between

[18] Available at http://people.stern.nyu.edu/jwurgler/.

[19] The complete set and description of variables suggested by Goyal and Welch (2008) is provided in Appendix A. Within the complete set of variables used by Goyal and Welch (2008), we select fewer ones using the cross-correlation between such measures to avoid multicollinearity in our regression analysis. As we run a multivariate model, using the full set of variables is undesirable as some of them correlate 80 percent with each other. We exclude variables that correlate more than 40 percent with each other.

[20] Given the fact that γ is estimated on a daily basis, we average γ throughout each month.

[21] This regression sample is only possible because Goyal and Welch (2008) updated their dataset after the initial publication. The regression sample however, could not be extended further than December 2010 because the sentiment measure of Baker and Wurgler (2006) is available until that date.

sentiment and γ. But this relation is once again lost at the twelve-months horizon as the variable is no longer significant. The relationships also show up in the scatter plot in Figure 6. On the three- and six-months horizons (Plots A and B), a negative relation exists between the Baker and Wurgler (2006) sentiment measure and γ, whereas at the twelve-months horizon it does not. The explanatory power of the variable *Sent* in the univariate setting is with 11 and 23 percent, respectively, also high for the three- and six-months horizon,. The twelve-months univariate regression has, however, a R^2 of zero. These findings strengthen our hypothesis that overweighting of small probabilities increases at higher levels of sentiment and that sentiment strongly impacts the probability weighting bias of call investors. However, this conclusion applies to the three- and six-months horizons only.

[Please insert Figure 6 about here]

In contrast with the variable *Sent*, the individual investor sentiment (*IISent*) has only low explanatory power. The regression run on three- and six-months γ has zero percent explanatory power, whereas the twelve-months γ can be partially explained by *IISent* with a significance level of 5 percent. More importantly, at the twelve-months horizon, *IISent* is positively linked to γ and this relation is statistically significant, in contrast to the other univariate regression using *IISent*. The relation of *IISent* and γ within the multivariate regression is very similar to the one captured by the univariate regression: statistical significance is found only at the twelve-months horizon with a positive coefficient. This positive relationship between these variables may be attributed to potential mean-reversion in individual investor sentiment, whereas at the twelve-months horizon, an increase in sentiment relates to a more pronounced underweighting of small probabilities.

In a next step, we specify the Goyal and Welch (2008) factors as control variables in our regression in order to add potential explanation to the model. The three-, six-, and twelve-months multivariate models explain 31, 57, and 51 percent, respectively, of the level of γ. Most of these relations are stable, because the coefficient signs change only rarely. The control variables that are statistically significant in our multivariate setting are *E12, Rfree, Infl, Svar*, and *CSP* (Table 4). We observe that γ is positively linked to *E12*, the twelve-months moving sum of earnings on the S&P 500 index, as well as to *Rfree*, the risk-free rate. The positive relation between *E12* and γ could be explained by mean-reversion of earnings being linked to a greater overweighting of small probabilities, which could be justified by the higher investor sentiment outweighing earning downgrades in a rallying market. The sign of *Rfree* suggests that as interest rates rise, OTM calls become more rational, as less overweighting of small probabilities is observed. The explanatory variables *Svar*, the stock market variance, and *CSP*, the cross-sectional premium, are negatively linked to γ. Such results suggest that the higher the variance, the higher the overweighting of small probabilities is. In a univariate setting (at the six-months horizon), the explanatory power of such an univariate regression is with 17 percent relatively high. Table 4, Panel B indicates that the cross-sectional premium *CSP* is statistically significant only at the six-months horizon. The univariate regression between γ and *CSP* does not, however, capture a statistically significant relation and has a coefficient with the opposite sign to the coefficient in the multivariate setting. This proves that supportive fundamental data for equity markets do not necessarily intensify biased behavior of single stock call option investors. This is an interesting takeaway, especially if one considers that sentiment does appear to affect such behavior: single stock option investors seem to overweight small probabilities when sentiment is exuberant, not necessarily when fundamentals are exuberant.

Moreover, as we have also estimated daily γ when the risk aversion parameter λ is time-varying, we run our regression models (Eqs. (19) and (20)) using such new estimation of γ as the explained variable. Table 5 indicates that the results for *Sent* are similar to the ones obtained in our earlier regression: *Sent* is negatively linked to γ and statistically significant at the three- and six-months horizon but not at the twelve-months horizon. These results apply to both the multivariate and univariate regression models. The magnitude of the *Sent* parameters is little altered at the three- and six-months regression within both the multivariate and univariate models, suggesting a robust relationship. The explanatory power of the regressions is once again high, as R^2 ranges from 31 to 59 percent in the multivariate models. The explanatory power of *Sent* is 11 and 24 percent for the three- and six-months maturities in the univariate setting, vis-à-vis 11 and 23 percent in the previous regression setting. The explanatory power of *Sent* in the twelve-months regression is one percent, similar to our earlier univariate regression results. Table 5 shows that the explanatory variable *IISent*, the AAII individual investor sentiment measure, is not significantly linked to γ within the short-term options. However, in the longer term options with a twelve-months horizon such a relationship is positive, rather than negative, as we would have expected. Such a positive relationship means a decrease in the overweighting of tails as individual investor sentiment rises. Several signs of control variables change in the multivariate and univariate setting in comparison to the regression results that used static risk aversion parameter. The control variables that remain statistically significant are *E12* and *Rfree.* The moving sum of equity earnings *E12,* however, changes its sign from positive to negative. The default return spread *Corpr* and the cross-sectional premium *CSP* become statistically significant in the univariate regressions, whereas the relation between the γ and the book-to-market ratio *B/M* and the stock market variance factor *Svar*, which we observed as strongly statistically significant in the univariate regressions, become insignificant once we take time-variation in the risk aversion parameter into account.

[Please insert Table 5 about here]

We conjecture that the profound impact that changing the parameter λ has on the control variables in our regression is due to the fact that some of these variables are either directly linked to uncertainty or their time variation may be conditional to risk-states. For example, the stock variance factor *Svar* and the default return spread *Corpr* can be interpreted as ex-post variants of λ, which is a forward-looking risk measure. Therefore, the introduction of a time-varying λ plausibly changes the relationship between *Svar* and γ from strong and negative to statistically weak and positive, whereas the regression estimates of the default return spread changed from negative and insignificant to positive and significant.

The robust results, in which the relation between γ and *Sent* is little modified once λ varies, suggests that changes in the overweighting of tails are not conditional on the level of the investors risk aversion parameter only, as section 4.3 suggests for the six- and twelve-month options. In fact, earlier results only suggest that, on average, overweight of small probabilities has diminished for these two maturities, but fluctuations in γ may still be linked to the Baker and Wurgler (2006) sentiment measure. In other words, levels of risk aversion do not fully drive investors to overweight upside tail events, as one could hypothesize when associating upside speculation with a state of low risk aversion. Our results suggest that overweighting of small probabilities seems to be a much more stable phenomenon, linked mostly to sentiment than to fundamental factors or risk states.

4.5. Robustness tests

We employ the Kupiec's (1995) test to compare the tails of the EDF with the ones of the subjective density functions and of the RND as a triangulation to the EVT methods applied. Kupiec's test was originally designed to evaluate the accuracy of value-at-risk (VaR) models, where the estimated VaR were compared with realized ones. Because the VaR is no different from the EQR on the downside, i.e., the $\hat{q}_p^-$ statistic, we can also make use of Kupiec's method to test the accuracy of the $\hat{q}_p^+$ statistic for subjective densities and the RND on matching realized EQRs. More specifically, Kupiec's method computes a proportion of failure (POF) statistic that is used to evaluate how often a VaR level is violated over a specified time span. Thus, if the number of realized violations is significantly higher than the number of violations implied by the level of confidence of the VaR, then such a risk model is challenged. Kupiec's POF test, which is designed as a log-likelihood ratio test, is defined as:

$$LR_{POF} = -2\log[(1-p^*)^{n-b}(p^*)^b] + 2\log[(1-[b/n])^{n-b}(b/n)^b] \quad \sim \chi^2(1)\,, \qquad (21)$$

where p^* is the POF under the null hypothesis, n is the sample size, and b is the number of violations in the sample. The null hypothesis of such test is $b/n = p^*$, i.e., the realized probability of failure matches the predicted one. Thus if the *LR* exceeds the critical value, $\chi^2(1)$ =3.841, the hypothesis is rejected at the 5 percent level.

By translating this methodology to our empirical problem, p^* equals the assumed probability that the EQR of the subjective and risk-neutral densities will violate the EQR of the realized returns, whereas

b/n is the realized number of violations. We note that because we intend to apply Kupiec's test to upside returns, violations mean that returns are higher than a positive threshold.

The first step in applying Kupiec's test to our data set is outlining the expected percentage of failure (p^*) between the EQR from the EDF and from the subjective and risk-neutral densities. We pick p^* as being 5 and 10 percent. The percentages can be seen as the expected frequency that the tails of the subjective and of the RND distributions overstate the tails of the distribution of the realized returns. As fatter tails are symptoms of an overweighting of small probabilities, we expect that densities that do not adjust for the CPT weighting function may deliver a higher frequency of failures than the CPT density function. The results of our application of the Kupiec's test are reported in Table 6.

[Please insert Table 6 about here]

Panel A in Table 6 suggests that the probability of failure for the RND, power, exponential, and PCPT densities is at the three-months horizon particularly high, with more than 97 percent for the EQR at 90 and 95 percent and for p^* equal to 5 and 10 percent. The *LR*-statistics for these densities tend to go to infinity in all of these cases[22], at the same time that the *p*-values approach zero. These results suggest that the EQR for these densities usually violates (overstates) the EQR recorded for the EDF. In other words, such densities contain often fatter tails than the EDF. Among these distributions, the RND density presents POFs that are smaller than for the other densities, roughly 97.4 percent vs. 99.9 percent, though such difference is still marginal. For the CPT density, the POF is much lower across the two values of p^* used and the 90 and 95 percent EQR. The POF for the 90 percent EQR is roughly 52 percent for the CPT, irrespective of p^*. At the 95 percent EQR, the POF is 40.6 percent for the CPT. These findings suggest that at the 90 and 95 percent EQR, the CPT and ECPT densities overstate less frequently the EDF tails. The violations of the EDF tails are, however, still significant as they occur between 40.6 and 52 percent of times. Nevertheless, when we analyze the 99 percent EQR, we find that the POF for all densities decreases considerably and, for the CPT, it becomes with 8.4 percent particularly low. Using tail violation criteria, these results indicate that the CPT and the EDF tails are statistically significant at the 99 percent EQR.

Panel B of Table 6 depicts a very similar pattern of the POF for the probability densities derived from the six-months options as we find for the three-months options. The POF is very close to 100 percent for all densities apart from the CPT at the 90 percent EQR, while at the 95 and 99 percent EQR violations fall substantially, even more than what we observed for the three-months options. Nevertheless, the CPT remains the best approximation for the EDF, as its POF is the lowest. The Kupiec's test result suggests that the CPT density is statistically equal to the EDF, whereas the RND also equals the empirical returns at the 10 percent level. The results for p^* equal to 5 or 10 percent are very similar. Panel C presents the POF

[22] The calculated *LR*-statistic for several Kupiec's tests is infinite as the number of failures is very high and, especially, because the data set used is large.

for the twelve-months maturity. We find once again that the CPT tails are the ones which violate the EDF tails the least. The POF for these densities are about 20 percent for the 90 percent EQR, 7 percent for the 95 percent EQR, and 5 percent for the 99 percent EQR. This finding suggests that the tails of the CPT closely match the EDF ones, especially far out in the tail, i.e., at the 95 and 99 percent EQR. The RND, power, exponential, and PCPT densities record POFs that are much smaller than for the three- and six-months maturities but that are still high in comparison to the CPT.

We note that results for the PCPT and the CPT are quite distinct, whereas results for the PCPT are somewhat closer to the ones of the RND. This suggests that the weighting function is the component within the CPT density function that more forcefully causes the RND to approximate the EDF, so that it is not the value function. Such a finding highlights the fact that a weighting function within a utility framework is important in understanding options pricing and investors' behavior. Overall, our analysis using Kupiec's test leads to similar results as the ones reached within our EVT analysis and adds to the evidence that the CPT model is superior to others in matching realized returns.

As another robustness check, we estimate the weighting function parameter α of the RDEU model suggested by Prelec (1998) in order to test whether our conclusions are robust to other weighted functions formulations[23]. We note that according to Prelec (1998) the standard α parameter value equals 0.65. Our findings are presented in Table 7.

[Please insert Table 7 about here]

The long-term estimates of α are somewhat in line with the one suggested by the RDEU but less so for the twelve-months horizon: α estimated from the three-, six-, and twelve-months is 0.61, 0.75, and 0.99, respectively. This is quite consistent with our long-term estimates for γ being, 0.71, 0.70, and 0.97 (see Table 1), as it suggests an overweighting of small probabilities that fades with the increase in the option horizon. Nevertheless, the time-varying estimates of α differ substantially from the long-term ones. We find the mean (0.93) and median (0.93) for time-varying estimates of α from three-months options to be much higher than the ones suggested by Prelec (1998). This means that an overweighting of small probabilities within the single stock option markets is far less than our results suggested by RDEU. For the six- and twelve-months maturities, an underweighting of small probabilities is even more frequent than an overweighting. The average α for the six-month options is 1.01 (median being 1.03), and for the twelve-months options it is 1.05 (median being 1.11). The fact that investors tend to overweight small probabilities to a much lesser extent in the short-term and that estimates are higher than suggested by their respective lab-based estimates confirm our main findings. The absence of an overweighting of small

[23] A major advance of Prelec's (1998) weighting function vis-a-vis the CPT is that it is monotonic for any value of α, which is different from the CPT with a non-monotonic weighting for low levels of γ. Because of that, the optimization used to estimate α uses a lower bound of zero instead of 0.28 as in our CPT optimizations.

probabilities in six- and twelve-months maturities is in line with our results when the risk aversion parameter is also time-varying but disconnected from our results when such a parameter is fixed.

The dispersion among α estimates (ranging from 0.14 to 0.18) is much smaller than the ones for γ, which is always above 0.28 and reaches 0.32 for the twelve-months horizon. This suggests that α estimates may be more reliable than γ estimates. The 25th quantile for α estimates, which varies from 0.83 to 1.02, is already much closer to their median than for γ, confirming that γ has much more dispersed estimates. The maximum α for the three-months maturity is 1.20, which suggests that the overweighting of small probabilities does not hold through the entire sample.

Overall, the robustness checks following Prelec (1998) confirm our main findings regarding long-term estimates for overweighting of small probabilities, and they reiterate our conclusion that the overweighting weakens with the increase of the horizon.

5. Conclusions

Single stock OTM call options are deemed overpriced because investors overpay for positively skewed securities, so-called lottery tickets (Mitton and Vorking, 2007; Barberis and Huang, 2008). According to Barberis and Huang (2008), the CPT's probability weighting function of Tversky and Kahneman (1992) provides an appealing explanation why such options are expensive: investors' preferences for positively skewed securities.

We find empirically that the CPT subjective density function of stock returns outperforms the RND and two rational densities functions (from the power and exponential utilities) on matching tails of realized equity returns. We estimate the CPT probability weighting function parameter γ and find that it does not differ much from the one predicated by Tverky and Kahneman (1992), particularly for short-term options. This outcome confirms our hypothesis that investors in single stock call options are CPT-biased. This is a strong finding since explaining the overpricing of single stock call options via the CPT weighting function had not yet been accomplished empirically. It adds to the recent advances in research that explain the overpricing of OTM index puts by investors' overweighting of small probabilities (Dierkes, 2009; Kliger and Levy, 2009; Polkovnichenko and Zhao, 2013).

However, the estimated γ's indicate that such overweighting of small probabilities is less pronounced than suggested by the CPT and has a horizon effect. This observed horizon effect implies that overweighting of small probabilities becomes less pronounced as the option maturity increases. This finding suggests that investors in single stock calls are more biased when trading short-term contracts, whereas they seem to be more rational (less biased) when trading long-term calls. This result is consistent with individual investors being the typical buyers of OTM single stock calls and the fact that they mostly use short-term instruments (cheaper lottery tickets) to speculate on the upside of equities (Lakonishok et al., 2007). The fact that longer maturity options are less positively skewed than short-term ones is also consistent with our results.

We also find that an investors' overweighting of small probabilities is largely time-varying, as γ varies from an extreme overweighting to an underweighting of small probabilities. Such a time-variation in γ's remains strong even when we account for time-varying risk aversion. Nevertheless, when we allow the risk aversion to vary, the average overweighting of tails weakens for all options of different maturities. As such, we argue that such an overweighting of small probabilities might be partially linked to priced-in risk-premium. We find that the Baker and Wurgler (2006) sentiment measure explains up to 23 percent of the time-variation in the investors' overweighting of small probabilities. However, the strong link between market sentiment and investors' overweighting of small probabilities is only present at the shorter maturities (three- and six-months). The horizon effect earlier observed for γ seems to be a corollary of how market sentiment connects to investors' overweighting of small probabilities: when an overweighting of small probabilities is pervasive, sentiment seems to explain their time-variation.

Our findings have several important practical implications. First, the understanding of time-variation in investors' overweighting of small probabilities should be a strong pillar in the development of behavioral option pricing models, which remains in its infancy. To the extent that overweighting of small probabilities is a latent variable or, simply, not trivial to estimate, we contemplate that future option pricing models should be more sentiment-aware than current ones. Second, of importance for such next generation option-pricing models is the inclusion of the horizon-effect found by us. Such potential modifications on options' pricing have large and direct consequences to risk-management, hedging and arbitrage activities. Third, from a financial stability point of view, investors' overweighting of small probabilities in single stock options could be of use to regulators for triangulating the presence of equity markets bubbles. Finally, as behavioral-led overpricing of OTM options is now uncovered in the single stock call market (beyond the index put market), further research connecting this finding across these two markets is warranted. The study of time-variation in overweighting of tails across these two markets is of particular interest. We believe that the understanding of equity sentiment can be substantially expanded via this route, leading to a better comprehension of expected stock returns and volatility.

Appendix A – Equity market predictors

The complete set and summarized descriptions of variables provided by Welch and Goyal (2008)[24] is:

1. ***Dividend–price ratio*** **(*log*)**, ***D/P***: difference between the log of dividends paid on the S&P 500 index and the log of stock prices (S&P 500 index).
2. ***Dividend yield*** **(*log*)**, ***D/Y*:** difference between the log of dividends and the log of lagged stock prices.
3. ***Earnings, E12:*** 12-month moving sum of earnings on the S&P 500.
4. ***Earnings–price ratio*** **(*log*)**, ***E/P***: difference between the log of earnings on the S&P 500 index and the log of stock prices.
5. ***Dividend–payout ratio*** **(*log*)**, ***D/E***: difference between the log of dividends and the log of earnings.
6. ***Stock variance*, *SVAR***: sum of squared daily returns on the S&P 500 index.
7. ***Book-to-market ratio*, *B/M***: ratio of book value to market value for the Dow Jones Industrial Average.
8. ***Net equity expansion*, *NTIS***: ratio of twelve-month moving sums of net issues by NYSE-listed stocks to total end-of-year market capitalization of NYSE stocks.
9. ***Treasury bill rate*, *TBL***: interest rate on a three-month Treasury bill.
10. ***Long-term yield*, *LTY***: long-term government bond yield.
11. ***Long-term return*, *LTR***: return on long-term government bonds.
12. ***Term spread*, *TMS***: difference between the long-term yield and the Treasury bill rate.
13. ***Default yield spread*, *DFY***: difference between BAA- and AAA-rated corporate bond yields.
14. ***Default return spread*, *Corpr***: difference between returns of long-term corporate and government bonds.
15. ***Cross-sectional premium, CSP:*** measures the relative valuation of high- and low-beta stocks and is proposed in Polk et al. (2006).
16. ***Inflation, INFL****:* calculated from the CPI (all urban consumers) using $x_{i,t-1}$ n Eq.(1) for inflation due to the publication lag of inflation numbers.
17. ***Investment-to-capital ratio, I/K****:* ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the entire economy (Cochrane, 1991).

[24] Available at http://www.hec.unil.ch/agoyal/.

References

Ait-Sahalia, Y., Lo, A.W., 2000. Nonparametric risk management and implied risk aversion. *Journal of Econometrics* 94(1-2), 9–51.

Anagnou, I., Bedendo, M., Hodges, S., Tompkins, R., 2002. The relationship between implied and realised probability density function. *University of Warwick and the University of Technology WP, Vienna.*

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-sectional of stock returns. *Journal of Finance* 61(4), 1645–1680.

Barber, B., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21(2), 785–818.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment, *Journal of Financial Economics* 49(3), 307–343.

Barberis, N., 2013. The psychology of tail events: Progress and challenges. *American Economic Review*, 103(3), 611–616.

Barberis, N., Huang, M., 2001. Mental accounting, loss aversion and individual stock returns. *Journal of Finance* 56(4), 1247–1292.

Barberis, N., Huang, M., 2008. Stocks as lotteries: The implication of probability weighting for security prices. *American Economic Review* 98(5), 2066–2100.

Barberis, N., Huang, M., Santos, T., 2001. Prospect theory and asset prices. *Quarterly Journal of Economics* 56 (1), 1–53.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49 (3), 307–343.

Bates, D., 2003. Empirical option pricing: A retrospection. *Journal of Econometrics* 116(1-2), 387–404.

Benartzi, S., Thaler, R., 1995. Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics* 110 (1), 73–92.

Bliss, R., Panigirtzoglou, N., 2004. Option-implied risk aversion estimates. *Journal of Finance* 59(1), 407–446.

Blume, M., Keim, D., 2012. Institutional investors and stock market liquidity: Trends and relationships. SSRN WP#2147757.

Bollen, N., Whaley, R., 2004. Does net buying pressure affect the shape of implied volatility function? *Journal of Finance* 59(2), 711-754.

Bollerslev, T., 2009. Glossary to ARCH (GARCH), *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*. Tim Bollerslev, Jeffrey R. Russell and Mark Watson (eds), Oxford University Press, Oxford, chapter 8, 137-163.

Chabi-Yo, F., Song, Z., 2013. Recovering the probability weights of tail events with volatility risk from option prices. Working Paper Ohio State University.

Cochrane, J., 1991. Production-based asset pricing and the link between stock returns and economics fluctuations. *Journal of Finance* 46(1), 209-237.

de Haan, L., Jansen, D., Koedijk, K., de Vries, C., 1994. Safety first portfolio selection, extreme value theory and long run asset risks. In *Proceedings from a Conference on Extreme Value Theory and Applications*, Galambos J (ed.), Kluwer Academic: Boston, MA, chapter 11, 471–487.

Danielsson, J., Jorgensen, B., Sarma, M., de Vries, C., 2006. Comparing downside risk measures for heavy tailed distributions. *Economics Letters* 92(2), 202-208.

Devroye, L., 1986. Non-Uniform Random Variate Generation. Springer-Verlag, New York.

Dierkes, M., 2009. Option-implied risk attitude under rank-dependent utility. Unpublished working paper. University of Münster, Münster, Germany.

Felix, L., Kräussl, R., Stork, P., 2015. The 2011 European Short Sale Ban: A Cure or a Curse? *Journal of Financial Stability*, forthcoming.

Figlewski, S., 2010. Estimating the implied risk neutral density for the U.S. market portfolio. In: *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, Tim Bollerslev, Jeffrey R. Russell and Mark Watson (eds), Oxford University Press, Oxford, chapter 15, 323–353.

Fox, C., Rogers, B., Tversky, A., 1996. Options traders exhibit subadditive decision weights. *Journal of Risk and Uncertainty* 13, 5-17.

Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21(4), 1455–1508.

Hartmann, P., Straetmans, S., de Vries, C., 2004. Asset market linkages in crisis periods. *Review of Economics and Statistics* 86(1), 313–326.

Han, B., 2008. Investor sentiment and option prices. *Review of Financial Studies* 21 (1), 387–414.

Hill, B., 1975. A simple general approach to inference about the tail of a distribution. *Annals of Statistics* 3(5), 1163–1173.

Hsu, M., Krajbich, I., Zhao, C., Camerer, C., 2009. Neural response to reward anticipation under risk is nonlinear in probabilities. *The Journal of Neuroscience* 29(7), 2231-2237.

Ilmanen, A., 2012. Do financial markets reward buying or selling insurance and lottery tickets? *Financial Analyst Journal* 68(5), 26–36.

Ingersoll, J., 2008. Non-monotonicity of the Tversky-Kahneman probability-weighting function: A cautionary note. *European Financial Management* 14(3), 385-390.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47(2), 263–291.

Kliger, D., Levy, O., 2009. Theories of choice under risk: Insights from financial markets. *Journal of Economic Behavior & Organization* 71(2), 330-346.

Kupiec, P., 1995. Techniques for verifying the accuracy of risk management models. *Journal of Derivatives* 3, 73-84.

Lakonishok, J., Lee, I., Pearson, N., Poteshman, A., 2007. Option market activity. *Review of Financial Studies* 20(3), 813–857.

Mitton, T., Vorkink, K., 2007. Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies* 20(4), 1255–1288.

Nelson, D.,1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59(2), 347-370.

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

Prelec, D., 1998. The probability weighting function. *Econometrica* 66(3), 497-527.

Rosenberg, J., Engle, R., 2002. Empirical pricing kernels. *Journal of Financial Economics* 64, 341–372

Straetmans, S., Verschoor, W., Wolff, C., 2008. Extreme US stock market fluctuations in the wake of 9/11. *Journal of Applied Econometrics* 23(1), 17–42.

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

Vilkov, G., Xiao, Y., 2013. Option-implied information and predictability of extreme returns. SAFE (Goethe University Frankfurt) Working Paper Series no. 5, 1-36.

Von Neumann, J., Morgenstern, O., 1947. Theory of games and economic behavior, 2nd edition, Princeton University Press, Princeton.

Wu, G., Gonzalez, R., 1996. Curvature of the probability weighting function. *Management Science* 42 (12) , 1676-1690.

Table 1 – Long-term CPT parameters and consistency test on tail shape

Panel A of this table reports the estimated "long-term" CPT parameters gamma (γ), lambda (λ), and alpha and beta (α, β) from the single stock options market. The parameters γ define the curvature of the weighting function for gains, which leads the probability distortion functions to assume inverse *S*-shapes. Estimated parameters close to unity lead to weighting functions that are close to un-weighted probabilities, whereas parameters close to zero denotes a larger overweighting of small probabilities. The parameter λ is the risk aversion parameter. The α and β parameters are the value function diminishing sensitivity parameters. These parameters are long-term since the estimates of such parameters are obtained by setting the average CPT density functions to match the return distribution realized within our sample. These parameters are estimated using Eq. (12a), (12b), and (13). Panel B reports the results from the statistical test of the tail shape parameter phi (φ), according to Eqs (14) and (17a) applied to the averaged probability density functions. As densities compared here are averaged, for the RND and subjective densities, including the CPT, or estimated using our full sample for realized returns, such test aims to test for the long-term consistency between distribution tail shapes. The null hypothesis of these two tests is that φ from the two distributions being compared have equal means and, therefore, tail shapes are consistent. The rejection of the null hypothesis is tested by t-tests of Eq. (17a) at the ten, five, and one percent statistical levels, respectively, shown by superscripts *, **, ***, assigned to φ for the RND, displayed in column (1), and φ for the EDF, shown in column (2).

Panel A

	Gamma	Lambda	Alpha/Beta
Maturity	Long-term	Long-term	Long-term
3 months	0.61	1.98	0.94
6 months	0.70	2.49	0.98
12 months	0.97	1.85	0.99

Panel B

		Phi			
Maturity	**(1) vs. (2)**	**(1)**	**(2) EDF**	***p*-value**	***t*-stat**
3 months	RND vs. EDF	0.20*	0.24	6.8%	-1.9
	Power vs. EDF	0.17***	0.24	0.3%	-3.2
	Expo vs. EDF	0.17***	0.24	0.5%	-2.9
	PCPT vs. EDF	0.17***	0.24	0.6%	-2.9
	CPT vs. EDF	0.25	0.24	36.7%	0.4
	ECPT vs. EDF	0.25	0.24	36.7%	0.4
6 months	RND vs. EDF	0.19*	0.23	9.0%	-1.7
	Power vs. EDF	0.16***	0.23	0.1%	-3.4
	Expo vs. EDF	0.16***	0.23	0.1%	-3.3
	PCPT vs. EDF	0.17***	0.23	0.7%	-2.9
	CPT vs. EDF	0.23	0.23	37.5%	0.3
	ECPT vs. EDF	0.18**	0.23	3.5%	-2.2
12 months	RND vs. EDF	0.22	0.23	33.0%	-0.6
	Power vs. EDF	0.14***	0.23	0.0%	-4.7
	Expo vs. EDF	0.14***	0.23	0.0%	-5.1
	PCPT vs. EDF	0.18**	0.23	1.6%	-2.5
	CPT vs. EDF	0.24	0.23	34.9%	0.5
	ECPT vs. EDF	0.18**	0.23	3.2%	-2.2

Table 2 – EVT consistency tests on tail returns

This table reports the results from our statistical test of the extreme quantile return, EQR, or quantile estimator, q, (in Panel A) and tail expected upside returns (in Panel B) performed according to Eqs (15), (16) and (17b) applied to the averaged probability density functions, PDFs. Since the densities compared here are averaged, for the RND and for the subjective densities, including the CPT, or estimated using our full sample for realized returns, such a test aims to investigate the long-term consistency between the distribution tails. The null hypothesis of these two tests is that the EQR or the tail expected upside returns from the two distributions being compared have equal means and, therefore, tails are consistent. The rejection of the null hypothesis, which indicates that tails are not consistent, are tested by t-tests of Eq. (17b) at the ten, five, and one percent statistical levels, respectively, shown by superscripts *, **, ***, assigned to the EQR or the tail expected upside return for the RND, displayed in column (1), and for the same statistics for the EDF, shown in column (2).

Panel A - Tails extreme quantile returns (EQR)

		10% quantile				*5% quantile*				*1% quantile*			
Maturity	**(1) vs. (2)**	**(1)**	**(2) EDF**	***p*-value**	***t*-stat**	**(1)**	**(2)**	***p*-value**	***t*-stat**	**(1)**	**(2)**	***p*-value**	***t*-stat**
3 months	RND vs. EDF	0.16***	0.11	0.0%	-8.9	0.19***	0.13	0.0%	-10.6	0.26***	0.2	0.0%	-7.5
	Power vs. EDF	0.21***	0.11	0.0%	-16.7	0.23***	0.13	0.0%	-19.9	0.31***	0.2	0.0%	-12.7
	Expo vs. EDF	0.21***	0.11	0.0%	-17.2	0.24***	0.13	0.0%	-20.7	0.32***	0.2	0.0%	-13.6
	PCPT vs. EDF	0.18***	0.11	0.0%	-12.9	0.21***	0.13	0.0%	-15.3	0.28***	0.2	0.0%	-9.6
	CPT vs. EDF	0.11	0.11	16.4%	1.3	0.13	0.13	12.2%	1.5	0.19	0.2	26.9%	0.9
	ECPT vs. EDF	0.11	0.11	16.4%	1.3	0.13	0.13	12.2%	1.5	0.19	0.2	26.9%	0.9
6 months	RND vs. EDF	0.19*	0.18	6.1%	-1.9	0.22*	0.21	8.6%	-1.8	0.3	0.3	39.8%	-0.1
	Power vs. EDF	0.25***	0.18	0.0%	-9.2	0.28***	0.21	0.0%	-10.3	0.36***	0.3	0.0%	-6.8
	Expo vs. EDF	0.26***	0.18	0.0%	-10.3	0.29***	0.21	0.0%	-11.7	0.37***	0.3	0.0%	-8.0
	PCPT vs. EDF	0.22***	0.18	0.0%	-5.1	0.24***	0.21	0.0%	-5.3	0.32**	0.3	2.3%	-2.4
	CPT vs. EDF	0.13***	0.18	0.0%	6.4	0.15***	0.21	0.0%	8.0	0.22***	0.3	0.0%	7.9
	ECPT vs. EDF	0.17	0.18	36.5%	0.4	0.2	0.21	13.1%	1.5	0.26***	0.3	0.1%	3.6
12 months	RND vs. EDF	0.22**	0.25	3.5%	2.2	0.26***	0.29	0.4%	3.1	0.37***	0.42	0.0%	3.9
	Power vs. EDF	0.33***	0.25	0.0%	-7.6	0.36***	0.29	0.0%	-7.7	0.45***	0.42	0.5%	-3.0
	Expo vs. EDF	0.34***	0.25	0.0%	-9.1	0.38***	0.29	0.0%	-9.3	0.47***	0.42	0.0%	-4.3
	PCPT vs. EDF	0.26	0.25	14.0%	-1.4	0.3	0.29	29.8%	-0.8	0.39*	0.42	8.7%	1.7
	CPT vs. EDF	0.15***	0.25	0.0%	11.6	0.18***	0.29	0.0%	14.4	0.26***	0.42	0.0%	12.5
	ECPT vs. EDF	0.25	0.25	29.6%	-0.8	0.29	0.29	39.9%	0.0	0.39**	0.42	2.8%	2.3

Panel B – Tails expected upside return

		10% quantile				*5% quantile*				*1% quantile*			
Maturity	**(1) vs. (2)**	**(1)**	**(2) EDF**	***p*-value**	***t*-stat**	**(1)**	**(2)**	***p*-value**	***t*-stat**	**(1)**	**(2)**	***p*-value**	***t*-stat**
3 months	RND vs. EDF	0.2***	0.15	0.0%	-7.6	0.23***	0.17	0.0%	-9.0	0.32***	0.26	0.0%	-6.0
	Power vs. EDF	0.25***	0.15	0.0%	-14.4	0.28***	0.17	0.0%	-16.9	0.37***	0.26	0.0%	-10.2
	Expo vs. EDF	0.26***	0.15	0.0%	-15.1	0.29***	0.17	0.0%	-18.0	0.39***	0.26	0.0%	-11.3
	PCPT vs. EDF	0.22***	0.15	0.0%	-10.7	0.25***	0.17	0.0%	-12.5	0.33***	0.26	0.0%	-7.2
	CPT vs. EDF	0.14	0.15	24.1%	1.0	0.17	0.17	21.6%	1.1	0.25	0.26	34.7%	0.5
	ECPT vs. EDF	0.14	0.15	24.1%	1.0	0.17	0.17	21.6%	1.1	0.25	0.26	34.7%	0.5
6 months	RND vs. EDF	0.24	0.23	26.2%	-0.9	0.27	0.27	36.1%	-0.4	0.37	0.38	16.7%	1.3
	Power vs. EDF	0.3***	0.23	0.0%	-7.1	0.33***	0.27	0.0%	-7.5	0.43***	0.38	0.0%	-4.1
	Expo vs. EDF	0.31***	0.23	0.0%	-8.1	0.34***	0.27	0.0%	-8.8	0.44***	0.38	0.0%	-5.2
	PCPT vs. EDF	0.26***	0.23	0.1%	-3.3	0.29***	0.27	0.4%	-3.0	0.38	0.38	39.4%	-0.1
	CPT vs. EDF	0.17***	0.23	0.0%	6.2	0.2***	0.27	0.0%	7.7	0.29***	0.38	0.0%	7.7
	ECPT vs. EDF	0.21*	0.23	8.5%	1.8	0.24***	0.27	0.3%	3.2	0.32***	0.38	0.0%	5.3
12 months	RND vs. EDF	0.28**	0.32	1.4%	2.6	0.33***	0.37	0.1%	3.5	0.47***	0.54	0.0%	4.4
	Power vs. EDF	0.38***	0.32	0.0%	-4.8	0.42***	0.37	0.0%	-4.0	0.53	0.54	31.9%	0.7
	Expo vs. EDF	0.4***	0.32	0.0%	-6.1	0.44***	0.37	0.0%	-5.5	0.55	0.54	36.5%	-0.4
	PCPT vs. EDF	0.32	0.32	39.8%	0.1	0.36	0.37	20.3%	1.2	0.48***	0.54	0.0%	3.7
	CPT vs. EDF	0.19***	0.32	0.0%	11.4	0.23***	0.37	0.0%	14.0	0.34***	0.54	0.0%	12.2
	ECPT vs. EDF	0.31	0.32	33.4%	0.6	0.35*	0.37	8.7%	1.7	0.47***	0.54	0.0%	4.1

Table 3– Time-varying parameters

This table reports the summary statistics of the estimated CPT time-varying parameters lambda, λ, and gamma, γ, from the single stock options market for each day in our sample. The parameter λ is the risk aversion parameter and the parameter γ defines the curvature of the weighting function for gains, which leads the probability distortion functions to assume inverse *S*-shapes. An estimated γ parameter close to unity leads to a weighting function that is close to the unweighted probabilities, whereas parameters close to zero denote a larger overweighting of small probabilities. Panel A reports the summary statistics of λ when we assume the CPT parameterization where the alpha and beta parameters equal 0.88. Panel B reports the summary statistics of λ when alpha and beta are estimated via Eq. (12a) (results reported in Table 1). Panel C reports the summary statistics of γ when we assume the risk aversion parameter λ equal to 2.25, i.e., the CPT parameterization. Panel D reports the summary statistics of γ when we assume that the risk aversion parameter λ is time-varying. In this case, γ is estimated subsequent to the estimation of λ for each day in our sample.

Panel A – Lambda (λ) with alpha and beta = 0.88

	Lambda						
Maturity	Min	25% Qtile	Median	Mean	75% Qtile	Max	SD
3 months	0.64	1.43	1.75	1.76	2.10	2.90	0.49
6 months	0.59	1.45	1.81	1.78	2.13	2.96	0.49
12 months	0.64	1.45	1.89	1.82	2.18	2.95	0.50

Panel B – Lambda (λ) with optimized alpha and beta

	Lambda						
Maturity	Min	25% Qtile	Median	Mean	75% Qtile	Max	SD
3 months	0.64	1.44	1.74	1.77	2.10	2.95	0.48
6 months	0.60	1.45	1.79	1.79	2.15	2.96	0.49
12 months	0.57	1.45	1.92	1.84	2.23	2.97	0.51

Panel C – Gamma with static risk aversion (λ = 2.25)

	Gamma						
Maturity	Min	25% Qtile	Median	Mean	75% Qtile	Max	SD
3 months	0.28	0.42	0.66	0.68	0.87	1.20	0.29
6 months	0.28	0.48	0.72	0.75	1.05	1.20	0.31
12 months	0.28	0.57	0.83	0.80	1.11	1.20	0.32

Panel D – Gamma with time-varying risk aversion (λ)

	Gamma						
Maturity	Min	25% Qtile	Median	Mean	75% Qtile	Max	SD
3 months	0.28	0.55	0.72	0.74	0.94	1.20	0.30
6 months	0.28	0.74	1.07	0.96	1.20	1.20	0.28
12 months	0.28	0.78	1.15	0.97	1.20	1.20	0.30

Table 4 – Regression results: CPT parameterization

This table reports the the regression results for Eq. (19) in a multivariate setting in Panel A. The dependent variable in this regression is γ, where the explanatory variables are 1) the Baker and Wurgler (2006) sentiment measure, 2) a dummy for the earnings season, and 3) the explanatory variables used by Goyal and Welch (2008), excluding the factors that correlate with each other in excess of 40 percent. Panel B reports the regression results for Eq. (20), which regresses γ and the same explanatory variables mentioned before in the univariate setting.

	Panel A - Multivariate			Panel B - Univariate													
Horizon	3m	6m	12m	3m	6m	12m	3m	6m	12m	6m	6m	6m	6m	6m	6m	6m	6m
Intercept	0.718***	0.955***	0.785***	0.624***	0.486***	0.792***	0.731***	0.958***	0.975***	0.819***	0.779***	0.954***	0.999***	0.948***	0.959***	0.988***	0.955***
	(0.055)	(0.046)	(0.060)	(0.009)	(0.015)	(0.031)	(0.010)	(0.014)	(0.017)	(0.022)	(0.047)	(0.014)	(0.021)	(0.015)	(0.013)	(0.015)	(0.013)
Sent	-0.035**	-0.079***	-0.009	-0.051***	-0.097***	-0.003											
	(0.017)	(0.017)	(0.020)	(0.014)	(0.017)	(0.018)											
IISent	-0.014	-0.036	0.104**				-0.022	-0.019	0.152***								
	(0.047)	(0.048)	(0.051)				(0.040)	(0.057)	(0.054)								
E12	0.020***	0.029***	0.043***							0.028***							
	(0.005)	(0.006)	(0.008)							(0.004)							
B/m	-0.081	-0.144	-0.020								0.714***						
	(0.217)	(0.194)	(0.238)								(0.186)						
Ntis	0.106	0.248	0.843									0.574					
	(0.366)	(0.342)	(0.622)									(0.527)					
Rfree	-0.019**	-0.025***	0.004										-0.018*				
	(0.008)	(0.007)	(0.009)										(0.009)				
Infl	-0.031	-1.946	-3.186											3.890			
	(2.396)	(2.172)	(2.578)											(3.200)			
Corpr	-0.258	-0.221	0.006												-0.392		
	(0.250)	(0.262)	(0.300)												(0.386)		
Svar	-3.091*	-7.418***	-4.420*													- 8.310***	
	(1.690)	(2.116)	(2.478)													(2.773)	
CSP	-0.184	-0.309**	-0.254														0.284
	(0.120)	(0.145)	(0.248)														(0.218)
R^2	31%	57%	51%	11%	23%	0%	0%	0%	5%	19%	18%	1%	6%	1%	1%	17%	1%
F-stats	6.7	19.5	15.5	18.6	47.6	0.0	0.3	0.2	7.8	37.0	35.3	1.4	9.4	2.4	1.3	31.2	1.9
AIC	-326.8	-320.4	-260.6	-326.1	-186.0	34.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BIC	-293.1	-286.8	-226.9	-320.0	-179.9	40.3	0.0	0.0	0.1	0.0	0.1	0.5	0.0	2.5	0.3	1.5	0.2

Table 5– Regression results: Time-varying risk aversion parameter

This table reports the regression results for Eq. (19) in a multivariate setting in Panel A. The dependent variable in this regression is γ, where the explanatory variables are 1) the Baker and Wurgler (2006) sentiment measure; 2) a dummy for the earnings season and 3) the explanatory variables used by Goyal and Welch (2008) excluding the factors that correlate with each other in excess of 40 percent. Panel B reports the regression results for Eq. (20), which regresses γ and the same explanatory variables mentioned before in the univariate setting.

	Panel A - Multivariate			Panel B - Univariate													
Horizon	3m	6m	12m	3m	6m	12m	3m	6m	12m	6m	6m	6m	6m	6m	6m	6m	6m
Intercept	0.718***	0.945***	0.823***	0.738***	0.971***	0.978***	0.731***	0.956***	0.959***	0.947***	0.836***	0.823***	0.872***	0.819***	0.817***	0.803***	0.825***
	(0.056)	(0.046)	(0.062)	(0.010)	(0.012)	(0.020)	(0.010)	(0.015)	(0.020)	(0.036)	(0.045)	(0.016)	(0.026)	(0.016)	(0.016)	(0.020)	(0.015)
Sent	-0.037**	-0.081***	-0.034	-0.052***	-0.102***	-0.026											
	(0.017)	(0.016)	(0.021)	(0.014)	(0.017)	(0.021)											
IISent	-0.009	-0.036	0.087				-0.018	-0.021	0.152**								
	(0.048)	(0.047)	(0.057)				(0.040)	(0.060)	(0.066)								
E12	0.020***	0.030***	0.053***							-0.026***							
	(0.006)	(0.007)	(0.008)							(0.007)							
B/m	-0.089	-0.118	-0.249								-0.059						
	(0.218)	(0.203)	(0.237)								(0.182)						
Ntis	0.099	0.351	0.624									-0.553					
	(0.370)	(0.351)	(0.631)									(0.894)					
Rfree	-0.019**	-0.025***	-0.009										-0.022**				
	(0.008)	(0.008)	(0.009)										(0.008)				
Infl	-0.134	-1.745	-2.991											0.779			
	(2.416)	(2.176)	(2.724)											(3.325)			
Corpr	-0.249	-0.233	0.041												0.654**		
	(0.256)	(0.280)	(0.306)												(0.306)		
Svar	-3.045*	-7.782***	-7.238**													4.696	
	(1.700)	(2.144)	(2.791)													(4.070)	
CSP	-0.197	-0.334**	-0.324														-0.824***
	(0.122)	(0.145)	(0.257)														(0.276)
R^2	31%	59%	55%	11%	24%	1%	0%	0%	4%	10%	0%	1%	5%	0%	1%	3%	6%
F-stats	6.7	21.3	18.3	19.2	50.1	1.7	0.2	0.2	6.0	17.6	0.1	0.8	8.4	0.1	2.3	5.4	10.3
AIC	-324.8	-318.6	-237.1	-326.1	-186.0	34.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BIC	-291.2	-284.9	-203.4	-320.0	-179.9	40.3	0.0	0.1	0.1	0.0	0.2	0.6	0.0	3.2	0.4	2.0	0.3

Table 6 – Robustness tests

This table reports the results from Kupiec's (1995) percentage of failure (POF) test for violations of the extreme quantile returns (EQR) from the empirical density function (EDF) by the EQR of a set of RND and subjective density functions. Such tests is performed as a robustness check to the extreme value theory (EVT)-based tests performed on the EQR and on the expected upside returnsS. The null hypothesis, which is designed as a log-likelihood ratio test (Eq. (21)), is that the realized probability of failure (b/n) matches the predicted one (p^*). Thus if the LR exceeds the critical value, $\chi^2(1)$ =3.841, such a hypothesis is rejected at the 5 percent level. Translating the methodology to our empirical problem, p^* becomes the assumed probability that the EQR of the subjective and of the risk-neutral densities will violate the EQR of the realized returns, whereas b/n is the realized number of violations. We note that because we apply Kupiec's test to the upside returns, violations mean that returns are higher than a positive threshold.

Panel A. 3 months call options

	EQR 90%			EQR 95%			EQR 99%		
	POF	*p*-value	*LR*-stat	POF	*p*-value	*LR*-stat	POF	*p*-value	*LR*-stat
p^* = 10%									
RND vs. EDF	99.4%	0.0000	∞	97.4%	0.0000	∞	25.5%	0.0000	77.8
Power vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	65.4%	0.0000	711.2
Expo vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	70.1%	0.0000	821.1
PCPT vs. EDF	100.0%	0.0000	∞	99.9%	0.0000	∞	39.9%	0.0000	244.9
CPT vs. EDF	52.2%	0.0000	443.6	40.6%	0.0000	255.7	8.4%	0.2666	1.2
p^* = 5%	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	99.4%	0.0000	∞	97.4%	0.0000	∞	25.5%	0.0000	186.1
Power vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	65.4%	0.0000	∞
Expo vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	70.1%	0.0000	∞
PCPT vs. EDF	100.0%	0.0000	∞	99.9%	0.0000	∞	39.9%	0.0000	438.2
CPT vs. EDF	52.2%	0.0000	709.9	40.6%	0.0000	453.5	8.4%	0.0048	8.0

Panel B. 6 months call options

p^* = 10%	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	99.8%	0.0000	∞	67.5%	0.0000	759.6	9.4%	0.7016	0.1
Power vs. EDF	99.9%	0.0000	∞	96.2%	0.0000	∞	16.5%	0.0001	15.7
Expo vs. EDF	99.9%	0.0000	∞	96.5%	0.0000	∞	17.4%	0.0000	20.5
PCPT vs. EDF	99.9%	0.0000	∞	88.1%	0.0000	∞	12.9%	0.0613	3.5
CPT vs. EDF	59.4%	0.0000	582.5	22.8%	0.0000	55.0	3.2%	0.0000	27.0
p^* = 5%	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	99.8%	0.0000	∞	67.5%	0.0000	∞	9.4%	0.0003	13.1
Power vs. EDF	99.9%	0.0000	∞	96.2%	0.0000	∞	16.5%	0.0000	70.4
Expo vs. EDF	99.9%	0.0000	∞	96.5%	0.0000	∞	17.4%	0.0000	81.0
PCPT vs. EDF	99.9%	0.0000	∞	88.1%	0.0000	∞	12.9%	0.0000	37.3
CPT vs. EDF	59.4%	0.0000	891.6	22.8%	0.0000	147.2	3.2%	0.0793	3.1

Panel C. 12 months call options

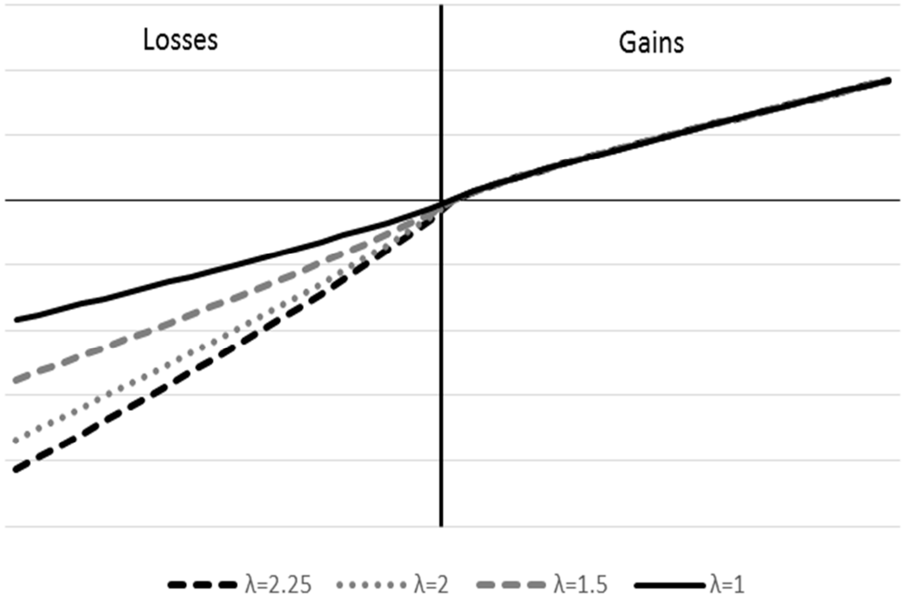
p^* = 10%	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	45.6%	0.0000	332.1	23.7%	0.0000	62.7	14.9%	0.0022	9.3
Power vs. EDF	58.2%	0.0000	558.6	42.1%	0.0000	277.5	25.0%	0.0000	73.1
Expo vs. EDF	58.9%	0.0000	572.8	42.7%	0.0000	286.9	25.8%	0.0000	80.7
PCPT vs. EDF	53.4%	0.0000	467.2	35.1%	0.0000	181.8	19.4%	0.0000	31.7
CPT vs. EDF	20.3%	0.0000	37.3	6.7%	0.0218	5.3	5.2%	0.0005	12.3
p^* = 5%	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	45.6%	0.0000	559.7	23.7%	0.0000	160.5	14.9%	0.0000	54.8
Power vs. EDF	58.2%	0.0000	860.6	42.1%	0.0000	484.1	25.0%	0.0000	178.3
Expo vs. EDF	58.9%	0.0000	879.0	42.7%	0.0000	497.3	25.8%	0.0000	190.9
PCPT vs. EDF	53.4%	0.0000	741.1	35.1%	0.0000	347.2	19.4%	0.0000	104.0
CPT vs. EDF	20.3%	0.0000	114.9	6.7%	0.1321	2.3	5.2%	0.8789	0.0

Table 7– Robustness checks: Prelec parameter

This table reports the summary statistics of the estimated Prelec's alpha (α) from the single stock option market, as a robustness check to the CPT's gamma (γ) parameter estimated in our main results. The parameter α defines the curvature of the weighting function for gains and losses, which leads the probability distortion functions to assume inverse S-shapes. An α parameter close to one means a weighting function with un-weighted (neutral) probabilities, whereas a parameter close to zero denotes a larger overweighting of small probabilities. Similarly to γ, we estimate long-term α (reported for γ in Table 1, Panel A) as well as time-varying α (reported for γ in Table 4, Panel C), as we assume for this robustness test that the risk aversion parameter λ equals 2.25.

Prelec alpha	Min	25% Qtile	Median	Mean	75% Qtile	Max	SD	Long-term
3 months	0.23	0.83	0.93	0.93	1.04	1.20	0.15	0.61
6 months	0.000	0.92	1.04	1.01	1.11	1.20	0.14	0.75
12 months	0.000	1.02	1.11	1.05	1.16	1.20	0.18	0.99

(A) **(B)**



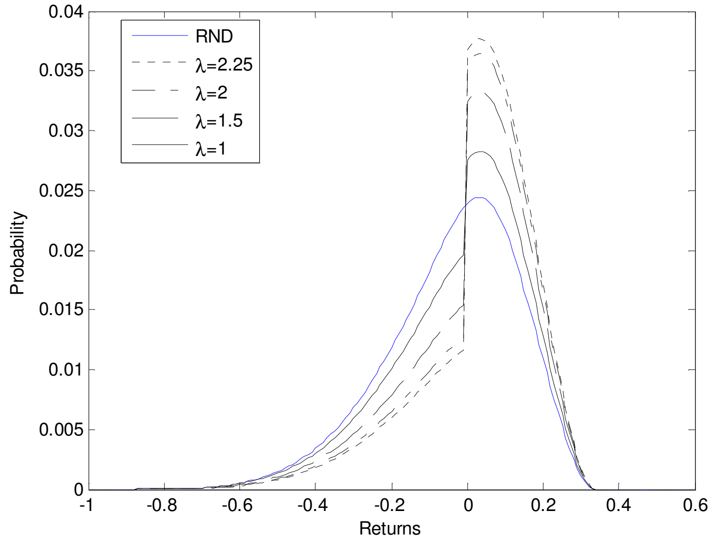


Figure 1. Impact of changes in λ. This figure shows the impact of changes in λ on the cumulative prospect theory (CPT) value function and on the CPT subjective density function. Plot A depicts the CPT value function v(x) for multiple risk aversion λ parameters. Plot B depicts the risk-neutral density (RND) extracted from the three-months option prices on July 12, 2001, and the impact of different risk aversion parameters λ on the partial CPT (subjective) density function. It is so-called "partial" because the probability weighing function, necessary for the full characterization of the CPT density function, is not yet applied. The different subjective density functions depicted are the ones required to match the RND provided that λ assumes different levels (i.e., 1, 1.5, 2, and 2.25). Thus, the higher λ is, the thinner the left tail becomes in comparison to the RND. In other words, for higher λ, a subjective density function must have thinner tails to be able to match the same (observed) RND when compared to a subjective density function with lower λ (see Eq. (9)). The α and β value function diminishing sensitivity parameter applied equal 0.88, following the CPT.

(A) (B)

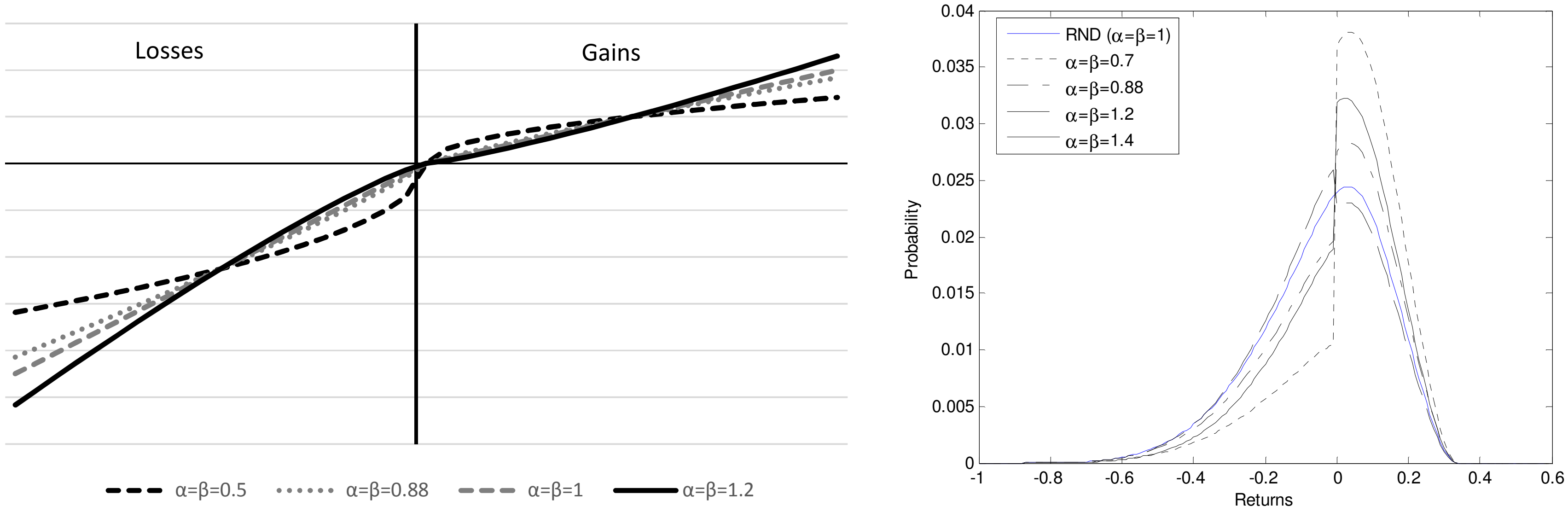
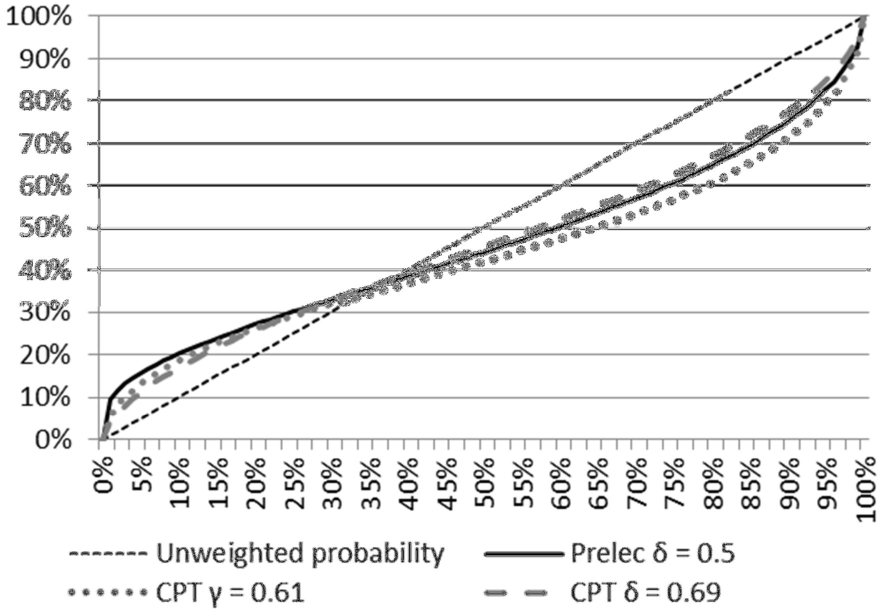


Figure 2. Impact of changes in *α* and *β*. This figure shows the impact of changes in α and β on the CPT value function and on the CPT subjective density function. Plot A depicts the CPT's value function $v(x)$ for multiple diminishing sensitivity parameters α and β for gains and losses, respectively. The higher α and β are, the higher is the sensitivity to gain or losses. When α and β is 1, the value function assumes a 45 degree inclination line (assumed that the risk aversion parameter λ is 1, as in here). An $\alpha > 1$ ($\beta > 1$) causes the value function to become convex, while an $\alpha < 1$ ($\beta < 1$) makes the value function to become concave. Plot B depicts the risk-neutral density (RND) extracted from three-months option prices on July 12, 2001, and the impact of different α and β parameters on the partial CPT (subjective) density function. The different subjective density functions depicted in Plot B are the ones required to match the RND for different levels of α and β (i.e., 0.7, 0.88, 1, 1.2, and 1.4). Taking the subjective density that has α and $\beta = 1$ as a reference (the RND) in Plot B, the smaller the values of α and β are (e.g., 0.7), the more sensitive the tails become in comparison to the RND. At first glance, this is the contrary of Plot A, where a smaller α and β suggests a lower sensitivity to relative gains or losses. This apparent disconnect is due to the fact that for smaller α and β, the subjective density function must have more sensitive tails to be able to match the same (observed) RND when compared to a subjective density function with higher α and β.

(A) **(B)**



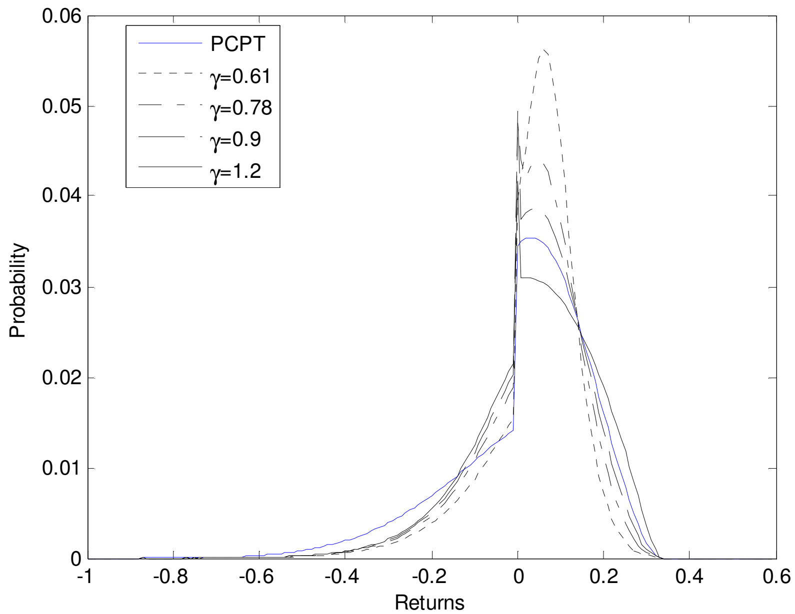


Figure 3. Impact of changes in *γ* and *δ*. This figure shows the impact of changes in γ and δ on the CPT value function and on the CPT subjective density function. Plot A depicts the CPT weighting function *w*(x) for the probability weighting parameters *γ* and *δ*, respectively, for gains and losses, as well as the weighting function of Prelec (1998) for the probability weighting parameter *δ*. Plot B depicts the partial CPT (PCPT) density function derived from the risk-neutral density (RND) extracted from three-month option prices on July 12, 2001. The PCPT is the subjective density function obtained from the RND, when only the effect of the CPT value function is considered. Plot B depicts the effect of different gamma parameters on the CPT (subjective) density function. The CPT density function is obtained from the PCPT by "undoing" the effect of the probability weighting function (see Eq. 11). We note that when $\gamma < 1$ (i.e., an investors' overweighting of small probabilities), the CPT density function has thinner tails than the PCPT model. When we observe a $\gamma > 1$ (i.e., an investors' underweighting of small probabilities), the CPT density function has fatter tails than the PCPT.

(A) 3 months horizon horizon **(B) 6 months horizon** **(C) 12 months**

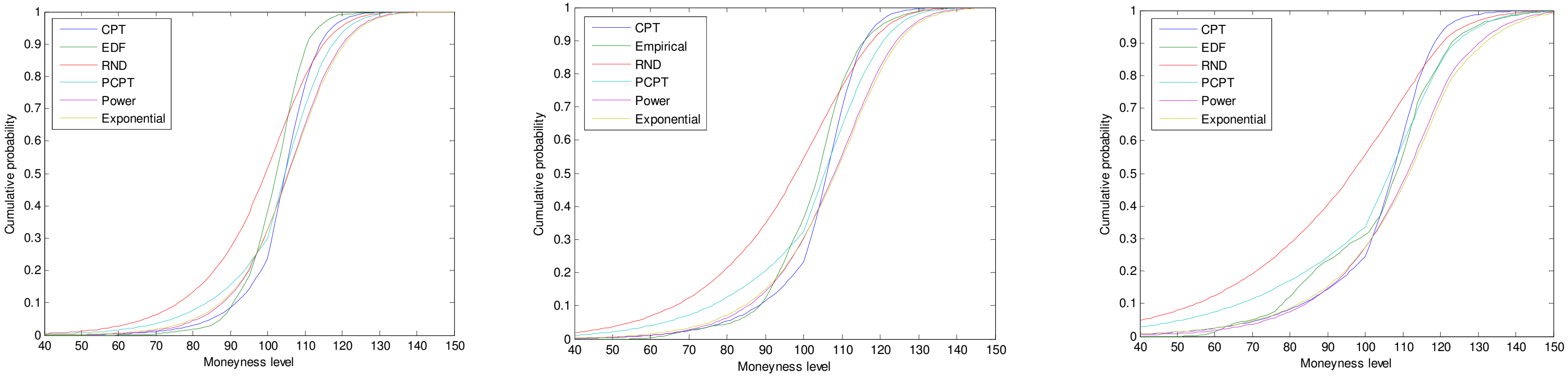


Figure 4. Cumulative density functions. This figure shows three plots that depict the cumulative density function (CDF) for equity returns obtained from the empirical density function (EDF), the risk-neutral density (RND), and the four subjective density functions: 1) the power utility density, 2) the exponential utility density, 3) the cumulative prospective theory density (CPT), and the partial CPT (PCPT). The equity returns' CDFs from these six sources are presented for three-, six-, and twelve-month horizons.

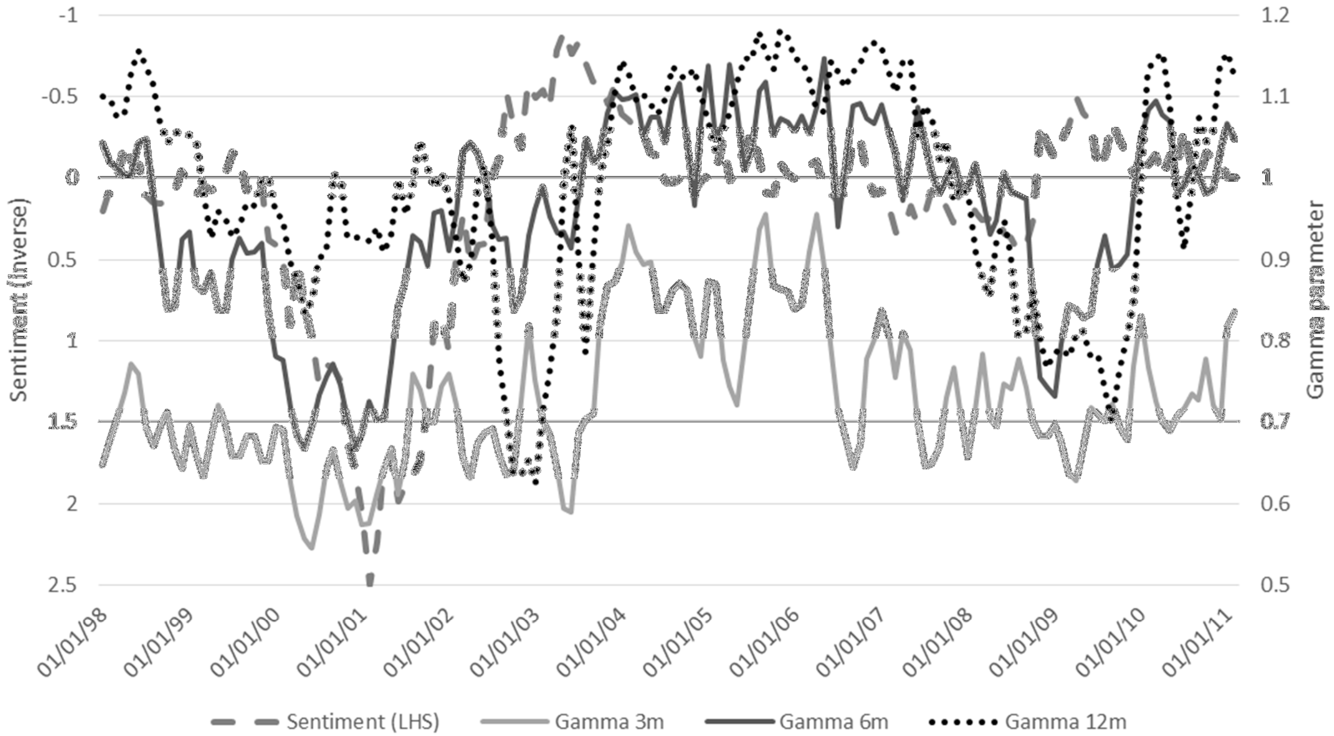


Figure 5. Time varying nature of gamma parameter in CPT. This figure depicts the time-varying nature of the gamma parameter from three-, six-, and twelve-months single stock options estimated using the CPT parameterization as well as the sentiment factor of Baker and Wurgler (2006).

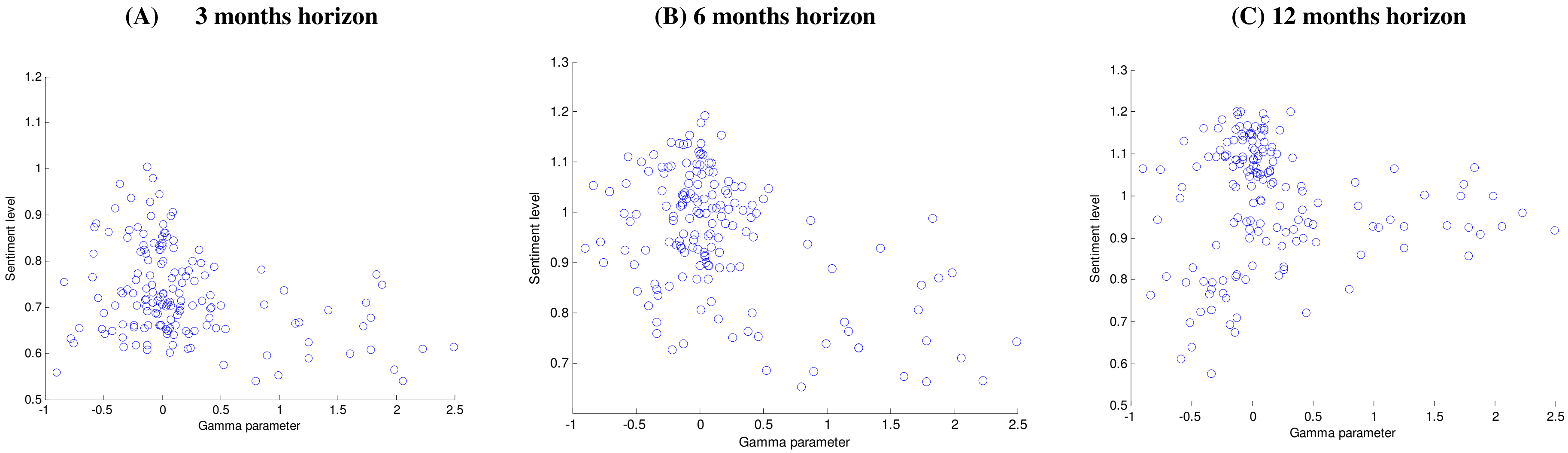


Figure 6. Scatter plot of gamma parameter and sentiment. This figure shows scatter plots depicting the relation between the estimated gamma parameter (on the *x*-axis) using the CPT parameterization and the sentiment measure of Baker and Wurgler (2006) (on the *y*-axis), using monthly data from January 1998 until February 2011. Plost A shows the relation between the estimated gamma parameter and sentiment when the gamma parameter is estimated from three-month options, whereas in plots B and C the gamma parameters are estimated from six- and twelve-month options respectively.

CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

Recent Issues

All CFS Working Papers are available at *www.ifk-cfs.de*.

No.	Authors	Title
565	Luiz Felix, Roman Kräussl, and Philip Stork	*Implied Volatility Sentiment: A Tale of Two Tails*
564	Roman Kräussl and Elizaveta Mirgorodskaya	*The Winner's Curse on Art Markets*
563	Christiane Baumeister, Reinhard Ellwanger and Lutz Kilian	*Did the Renewable Fuel Standard Shift Market Expectations of the Price of Ethanol?*
562	Thierry Mayer, Marc Melitz and Gianmarco Ottaviano	*Product Mix and Firm Productivity Responses to Trade Competition*
561	Jeronimo Carballo, Gianmarco Ottaviano and Christian Volpe Martincus	*The Buyer Margins of Firms' Exports*
560	Gianmarco Ottaviano, Giovanni Peri and Greg Wright	*Immigration, Trade and Productivity in Services: Evidence from U.K. Firms*
559	Martin Goetz	*Competition and Bank Stability*
558	Lutz Hendricks	*Accounting for Changing Returns to Experience*
557	Lutz Hendricks and Todd Schoellman	*Human Capital and Development Accounting: New Evidence from Wage Gains at Migration*
556	Richard Disney and John Gathergood	*House Prices, Wealth Effects and Labour Supply*

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de