Single Stock Call Options as Lottery Tickets

Félix, Luiz; Kräussl, Roman; Stork, Philip

published in
Journal of Behavioral Finance
2019

DOI (link to publisher)
10.1080/15427560.2018.1511792

document version
Publisher's PDF, also known as Version of record

citation for published version (APA)
Félix, L., Kräussl, R., & Stork, P. (2019). Single Stock Call Options as Lottery Tickets: Overpricing and Investor Sentiment. *Journal of Behavioral Finance*, *20*(4), 385-407. https://doi.org/10.1080/15427560.2018.1511792

Journal of Behavioral Finance

ISSN: 1542-7560 (Print) 1542-7579 (Online) Journal homepage: https://www.tandfonline.com/loi/hbhf20

Single Stock Call Options as Lottery Tickets: Overpricing and Investor Sentiment

Luiz Félix, Roman Kräussl & Philip Stork

To cite this article: Luiz Félix, Roman Kräussl & Philip Stork (2019) Single Stock Call Options as Lottery Tickets: Overpricing and Investor Sentiment, Journal of Behavioral Finance, 20:4, 385-407, DOI: 10.1080/15427560.2018.1511792

To link to this article: https://doi.org/10.1080/15427560.2018.1511792

Published online: 22 Jan 2019.

JOURNAL OF BEHAVIORAL FINANCE
2019, VOL. 20, NO. 4, 385–407
https://doi.org/10.1080/15427560.2018.1511792

Single Stock Call Options as Lottery Tickets: Overpricing and Investor Sentiment

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

[a]VU University Amsterdam; [b]University of Luxembourg

ABSTRACT
The authors investigate whether the overpricing of out-of-the money single stock calls can be explained by Tversky and Kahneman's [1992] cumulative prospect theory (CPT). They hypothesize that these options are expensive because investors overweight small probability events and overpay for positively skewed securities (i.e., lottery tickets). The authors find that overweighting of small probabilities embedded in the CPT explains the richness of out-of-the money single stock calls better than other utility functions. Nevertheless, overweighting of small probabilities events is less pronounced than suggested by the CPT, is strongly time varying, and most frequent in options of short maturity. The authors find that fluctuations in overweighting of small probabilities are largely explained by the sentiment factor.

KEYWORDS
Cumulative prospect theory; Investor sentiment; Risk-neutral densities; Call options

Introduction

Barberis and Huang [2008] hypothesized that Tversky and Kahneman's [1992] cumulative prospect theory (CPT) explains a number of seemingly unrelated pricing puzzles. In contrast to previous literature, which concentrates on the CPT's value function (see Benartzi and Thaler [1995], Barberis and Huang [2001]), Barberis and Huang [2008] focused on the probability weighting functions of the model. They concluded that the CPT's overweighting of small probability events explains why investors prefer positively skewed returns, or "lottery ticket" type of securities. Because of such preference, investors overpay for positively skewed securities, turning them expensive and causing them to yield low forward returns. The authors argue that this mechanism is the reason why IPO stocks, private equity, distressed stocks, single segment firms and deep out-of-the money (OTM) single stock calls are overpriced among other irrational pricing phenomena.

The proposition made by Barberis and Huang [2008] that deep OTM single stock calls resemble overpriced lottery-like securities due to investors' overweight of tails has not yet been verified empirically.[1] Empirical studies on probability weighting functions implied by option prices are offered by Dierkes [2009], Kliger and Levy [2009], and Polkovnichenko and Zhao [2013].[2] The evidence provided by these papers is, however, restricted to the index put options market, which behaves very differently from the single stock option market as the main buyers of OTM index puts are institutional investors seeking portfolio insurance (Bates [2003], Bollen and Whaley [2004], Lakonishok et al. [2007], Barberis and Huang 2008). Despite that, the results of Dierkes [2009] and Polkovnichenko and Zhao [2013] suggest that overweighting of small probabilities partially explains the pricing puzzle in equity index options.

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]) and is speculative in nature (Lakonishok et al. [2007], Choy [2015]). Beyond that, Mitton and Vorking [2007] and Kumar [2009] provided empirical support to the link between preference for skewness and individual investor trading activity. The fact that many individual investors have a substantial portion of their portfolios tied up in low-risk investments, such as pensions, social security, 401(k)s, or IRAs, or are averse (or constrained) to borrow (Frazzini and Pedersen [2014]) encourages them to buy financial instruments with implicit leverage such call options. Hence, given the very distinct clientele of these 2

CONTACT Luiz Félix luizffelix@yahoo.com VU University Amsterdam, School of Business and Economics, de Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/HBHF.

option markets (institutional investors vs. retail investors) and the different motivation for trading (portfolio insurance vs. speculation), we reason that the OTM single stock calls overpricing is a puzzle in itself, requiring an independent empirical analysis from the index option market.

The first contribution of our study is to investigate whether the CPT can empirically explain the claimed overpricing of OTM single stock call options. We empirically test whether tails of the CPT density function outperform the risk-neutral density (RND) and rational subjective probability density functions on matching tails of the distribution of realized returns. We do not use traded options prices, but we calculate them off implied volatility data. We find that our estimates for the CPT probability weighting function parameter γ are qualitatively consistent with the ones predicated by Tversky and Kahneman [1992], particularly for short-term options. Our estimates do suggest that overweight of small probabilities is less pronounced than suggested by the CPT. 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], and Polkovnichenko and Zhao [2013] because we show that investors' overweighting of small probabilities is not restricted to the pricing of index puts but it also applies to single stock calls.

Second, we provide evidence that overweighting of small probabilities is strongly time-varying and connected to the Baker and Wurgler [2007] investor sentiment factor. These findings contrast the CPT model, in which the probability weighting parameter for gains (γ) is constant at 0.61. In fact, our estimates 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 quite strong during the dot-com bubble, which coincided with a strong rise in investor sentiment. The strong time variation in overweight of tails indicates that investors have either a "bias in beliefs" or time-varying (rather than static) skewness preferences.[3]

Finally, we find that overweighting of small probabilities is largely horizon-dependent, because this bias, captured by γ, is mostly observed within short-term options prices rather than in long-term ones (i.e., γ being smaller than 1 for 3- and 6-month options and being often higher than 1 for 12-month options). We reason that such positive term structure of tails' overweighting exists because individual investors may speculate using the cheapest available call at their disposal, that is, they buy the cheapest lottery tickets that they can find. As 3- and 6-month options have much less time value than 12-month ones do, 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]).

The remainder of this article is organized as follows. The second section describes the CPT model. The third section describes the data and methodology employed in our study. The fourth section presents our empirical analysis and the fifth section discusses our robustness tests. The sixth section concludes.

Cumulative prospect theory

The prospect theory (PT) of Kahneman and Tversky [1979] incorporates behavioral biases into the standard utility theory, which presumes that individuals are rational. Such behavioral biases are (a) loss aversion, (b) risk-seeking behavior, and (c) nonlinear preferences.[4] The CPT is described in terms of a value function (υ) and a probability distortion function (w). The value function is analogous to the utility function in the standard utility theory and it is defined relative to a reference point 0. Therefore, positive values within the value function are considered as gains and negative values are losses, which leads to

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

where the loss-aversion parameter $\lambda \geq 1$, curvature parameters α and β range between [0,1], and x are gains or losses. Thus, along the dominium of x, the CPT's value function is asymmetrically S-shaped with diminishing sensitivity as $x \to \pm\infty$.

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 increase, which means that any additional unit of gain (loss) becomes less relevant when wealth increases (decreases). As α and β increase, the effect of diminishing sensitivity decreases, and as λ increases the degree of loss aversion increases. The value function has a kink at the reference point, which implies loss aversion, as the function is steeper for losses than for gains.

Furthermore, the CPT also embeds a probability distortion function or a decision weight function. This function takes probabilities and weights them nonlinearly, so that the difference between probabilities at high percentiles (e.g., between 99% and 100%) has more impact on preferences than the difference between probabilities at small percentiles (e.g., between 10% and 11%). The CPT applies probability distortions to the cumulative probabilities (CDF), whereas the PT applies them to the probability density function (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 parametric form of the probability distortion functions suggested by Tversky and Kahneman [1992], respectively, for gains w^+ and losses w^- are:

$$w^+(p) = \frac{p^\gamma}{\left(p^\gamma + (1-p)^\gamma\right)^{1/\gamma}}, \tag{2a}$$

$$w^-(p) = \frac{p^\delta}{\left(p^\delta + (1-p)^\delta\right)^{1/\delta}}, \tag{2b}$$

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. As a result, low probability events are overweighted at the cost of moderate and high probabilities. The parameters estimated by Tversky and Kahneman [1992] for the CPT model, which are discussed in the following empirical analysis, are $\lambda = 2.25$, $\beta = 0.88$, $\alpha = 0.88$, $\gamma = 0.61$, and $\delta = 0.69$.

Data and methodology

In this section, we first describe the theoretical background that allows us to relate empirical density functions (EDFs), RND, and subjective density functions. This is a key step for testing the hypothesis that the CPT helps to explain overpricing of OTM options, because we build on the assumption that investors' subjective density estimates should correspond, on average, to the distribution of realizations (see Bliss and Panigirtzoglou [2004]). This implies that investors are somewhat rational. This assumption is not inconsistent with the CPT assumption that the representative agent is less than fully rational. The CPT suggests that investors are biased, not that decision makers are utterly irrational to the point that their subjective density forecast should not correspond, on average, to the realized return distribution. 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 the weighting function, we produce 2 density functions: (c) the hereafter called partial CPT density function (PCPT), in which only the value function is taken into account, and (d) the CPT density function, in which both the value and the weighting functions are considered. Last, we also calibrate the CPT's probability weighting function parameter γ to market data and are, then, able to compute (e) the estimated CPT density (ECPT). We provide details on estimation methods for our 5 subjective density functions, a–e, in the Subjective Density Functions section, and for the RND and EDF in the Estimating RND and EDF section.

Once all 5 subjective density functions are obtained, we distinguish 4 analyses in our empirical analysis section: (a) the estimation of long-term CPT value and weighting function parameters (from which we can produce the ECPT density) (see Estimated CPT Long-term Parameters section); (b) EVT-based tests of consistency between tails of the EDF, the RND, and our 5 subjective probability distributions (Density Functions Tails' Consistent Test Results section); (c) the estimation of time-varying γ parameter (Estimated CPT Time-Varying Parameter section); and (d) a regression linking the CPT time-varying probability weighting parameter (γ) to sentiment measures as well as numerous control variables (Time Variation in the Probability Weighting Parameter and Investor Sentiment section).

We use single stock weighted average IV data used for the largest 100 stocks of the S&P 500 index within our RND estimations. Appendix A.1 shows how single stock weighted average IV are computed. Weights applied are the S&P 500 index weights normalized by the sum of weights of stocks for which IVs are available. Following the S&P 500 index methodology and the unavailability of IV information for every stock in all days in our sample, stock weights in this basket change on a daily basis. The sum of weights is, on average, 58% of the total S&P 500 index capitalization and it fluctuates between 46 and 65%. The IV data comes from closing mid-option prices from January 2, 1998 to March 19, 2013 for fixed maturities for 5 moneyness levels (i.e., 80, 90, 100, 110, and 120) at 3-, 6-, and 12-month maturity. Continuously compounded stock market returns are calculated throughout our analysis from the basket of stocks weighted with the same daily varying loadings used for aggregating the IV data. A index is calculated from this stream of returns, which we call the realized return index. IV data and stock weights are kindly provided by Barclays.[5] Single stock returns are downloaded via Bloomberg.

We take the perspective of end-users of single-stock OTM call options. Hence, we assume that supply imbalances are minimal and do not impact IVs. We think this assumption is reasonable because (a) option markets for the largest 100 U.S. stocks are liquid, (b) any unhedged risk run by market makers can be easily hedged by purchasing the stock, and (c) unhedged risk by market makers is likely much smaller when supplying call options relative to put options. Market makers run little unhedged risk when supplying call options vis-à-vis supplying puts because stocks returns are negatively skewed, making gap and jump risk much lower on the upside than on the downside. Gârleanu et al. [2009] showed that this condition is different for the index option market, where market makers mostly provide put options for portfolio insurance programs. This implies that put sellers become more risk sensitive following equity market declines, as their un-hedged risk increases, which makes them unwilling to write additional puts to the market. Our IV data show no indication of an increase in the IV skew from 120% moneyness options, nor from at-the-money options around moments of market stress (e.g., the 2008–2009 global financial crisis). Hence, we find no evidence of the presence of supply imbalances in the OTM calls in our sample.

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 EDF, thus both "real-world" probabilities. 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 "real-world" probability distributions, $f_P(S_T)$, with S_T being wealth or consumption, is described by $\zeta(S_T)$,the pricing kernel or the marginal rate of substitution (of consumption at time T for consumption at time t):

$$\frac{f_Q(S_T)}{f_P(S_T)} = \Lambda \frac{U'(S_T)}{U'(S_t)} \equiv \zeta(S_T), \tag{3}$$

where Λ is the subjective discount factor (the time-preference constant) and $U(\cdot)$ is the representative investor utility function.

As CPT-biased investors price options as if the data-generating process has a cumulative distribution $F_{\tilde{P}}(S_T) = w(F_P(S_T))$, where w is the weighting function, its density function becomes $f_{\tilde{P}}(S_T) = w'(F_P(S_T)) \cdot f_P(S_T)$ (see Dierkes [2009], Polkovnichenko and Zhao [2013]) and Equation 3 collapses into Equation 4:

$$\frac{f_Q(S_T)}{w'(F_P(S_T)) \cdot f_P(S_T) = \zeta(S_T).} \tag{4}$$

which, rearranged into Equation 6 via Equations 5a and 5b, demonstrates that for the CPT to hold, the subjective density function should be consistent with the probability-weighted EDF:

$$\underbrace{f_Q(S_T)}_{RND} = \underbrace{w'(F_P(S_T))}_{probability\ \ weigthing} \cdot \underbrace{f_P(S_T)}_{EDF} \cdot \underbrace{\zeta(S_T)}_{pricing\ \ kernel}, \tag{5a}$$

$$\underbrace{f_Q(S_T)}_{RND} = \underbrace{f\widetilde{P}(S_T)}_{probability\ \ weigthing\ EDF} \cdot \underbrace{\zeta(S_T)}_{pricing\ \ kernel}, \tag{5b}$$

$$\underbrace{\frac{f_Q(S_T)}{\Lambda\frac{U'(S_T)}{U'(S_t)}} = \frac{f_Q(S_T)}{\zeta(S_T)}}_{Subjective\ \ density} = \underbrace{f\widetilde{P}(S_T)}_{probability\ \ weigthed\ \ EDF} \tag{6}$$

Following Bliss and Panigirtzoglou [2004], Equation 6 can be manipulated so that the time-preference constant Λ of the pricing kernel vanishes, producing Equation 7, which directly relates the probability weighted EDF, the RND, and the marginal utility, $U'(S_T)$:

$$\underbrace{f\widetilde{P}(S_T)}_{probability\ \ weigthed\ \ EDF} = \underbrace{\frac{\Lambda\frac{U'(S_T)}{U'(S_t)} Q(S_T)}{\int \frac{U'(S_t)}{U'(x)} Q(x)dx} = \frac{\frac{f_Q(S_T)}{U'(S_T)}}{\int \frac{f_Q(x)}{U'(x)} dx}}_{Generic\ \ Subjective\ \ density\ \ function}, \tag{7}$$

where $\int \frac{Q(x)}{U'(x)} dx$ normalizes the resulting subjective density function to integrate to 1. Once the utility function is estimated, Equation 8 allows us to convert RND into the probability weighted EDF. As the CPT marginal marginal utility function is $U' = \upsilon(S_T) = \nu'(S_T)$, and thus, $\nu'(S_T) = \alpha {S_T}^{\alpha-1}$ for

$S_T \geq 0$, and $v'(S_T) = -\lambda\beta(-S_T)^{\beta-1}$ for $S_T < 0$, we obtain Equations 8 and 9:

$$f_{\tilde{P}}(S_T) = \frac{\frac{f_Q(S_T)}{\alpha\ S_T{}^{\alpha-1}}}{\int \frac{f_Q(x)}{\alpha x^{\alpha-1}}\ dx} \text{for } S_T \geq 0, \text{ and} \quad (8)$$

$$\underbrace{f_{\tilde{P}}(S_T)}_{\text{probability weigthed EDF}} = \underbrace{\frac{\frac{f_Q(S_T)}{-\lambda\beta(-\ S_T)^{\beta-1}}}{\int \frac{f_Q(x)}{-\lambda\beta(-x)^{\beta-1}}dx}}_{\text{Partial CPT density function}} \text{ for } S_T{<}0. \quad (9)$$

Equations 8 relates the probability-weighted EDF, which uses the CPT probability distortion functions for weighting, to the subjective density function derived from the CPT value function for gains. We call the RHS the partial CPT density function (PCPT), as it does not embed the probability function. Equation 9 is the corresponding equation for losses. As the function $w(F_P(S_T))$ is strictly increasing over the domain [0,1], there is a 1-to-1 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 $f_{\tilde{P}}(S_T)(w^{-1})'(F_P(S_T)) = f_P(S_T)$. This outcome allows us to directly relate the original EDF to the CPT subjective density function, by "undoing" the effect of the CPT probability distortion functions within the PCPT density function:

$$\underbrace{f_P(S_T)}_{EDF} = \underbrace{\frac{\frac{f_Q(S_T)}{v'(S_T)}}{\int \frac{f_Q(x)}{v'(x)}dx}(w^{-1})'(F_P(S_T))}_{CPTdensityfunction}, \quad (10)$$

where, $v'(S_T)$ is the CPT's marginal utility function.

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, we can evaluate to what extent the tails of the RND and the set of our subjective densities match the tails of the realized distribution.

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 and γ with the same goal. We estimate these parameters nonparametrically, by minimizing the weighted squared distance between physical distribution and the partial CPT density function for every bin above the median of the 2 distributions, as follows:

$$\upsilon(\lambda) = Min \sum_{b=1}^{B} W_b \left(EDF^b_{prob} - CPT^b_{prob}\right)^2, \quad (11)$$

where EDF^b_{prob} and CPT^b_{prob} are, respectively the probability within bin b in the empirical and CPT density functions and W_b are weights given by $1/\frac{1}{\sqrt{2\pi}}\int_{0.5}^{\infty} e^{-x^2/2}dx = 1$, the reciprocal of the normalized normal probability distribution (above its median), split in the same total number of bins (B) used for the EDF and CPT. The loss aversion parameter, λ, in Equation 11 is optimized using multiple constraint intervals: [0,3], [0,5], and [0,10].[6] Once the optimal λ is known, we minimize Equation 12 using its estimate and the CPT λ:

$$w^+(\gamma, \delta = \gamma) = Min \sum_{b=1}^{B} W_b \left(EDF^b_{prob} - CPT^b_{prob}\right)^2, \quad (12)$$

where γ, the probability weighting parameter for gains, is constrained to by the permutation of the following upper bounds (1.2, 1.35, 1.5, 1.75, and 2) and lower bounds (−0.25, 0 and 0.28).[7] Weights applied in these optimizations are due to the higher importance of matching probabilities tails in our analysis than the body of the distributions.

Our nonlinear bounded optimization is a single parameter one, where we first estimate optimal γ(which we impose to equal δ) across all permutations of upper and lower bounds to select the bounds that produce the lowest residual sum of square (RSS). Subsequently, we estimate λ and γ as suggested by the sequence of optimizations described by Equations 11 and 12. This method resembles the ones of Kliger and Levy [2009], Dierkes [2009], and Polkovnichenko and Zhao [2013]. Once optimal parameters λ and γ are estimated, we can produce another long-term subjective density function: the ECPT, in which we apply the optimal γ for the characterization of its probability weighting function. Finally, we also estimate time-varying γ using different assumptions of λ so to evaluate the sensitivity of γ to changes in λ.

Density Function Tails' Consistency Test

We check for tail consistency of our set of 5 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 7 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:

$$\hat{\varphi} = \frac{1}{\hat{\theta}} = \frac{1}{k}\sum_{j=1}^{k} ln\left(\frac{x_j}{x_{k+1}}\right), \quad (13)$$

where k is the number of extreme returns used in the tail estimation, and $x_{k+1}$is the tail cutoff 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). The 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 focus on the semiparametric extreme quantile estimator used by Straetmans et al. [2008]:

$$\hat{q}_p = x_{k+1}\left(\frac{k}{pn}\right)^{\frac{1}{\hat{\theta}}}, \quad (14)$$

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 cutoff point. The choice of the optimal value of k, and thus the tail cut-off point, x_{k+1}, is made through the use of Hill-plots, see Hill [1975] and the Density Functions Tails' Consistent Test Results section for empirical details. We note that, as $\hat{\varphi} = \frac{1}{\hat{\theta}}$, $\hat{q}_p$ depends on 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.

In addition to the EQR, we also evaluate the density function tails using expected shortfall (ES), which captures the average loss beyond the tail cutoff 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 in the following:

$$\hat{EU}_{q(P)} = \frac{\hat{\theta}}{\hat{\theta}-1} * x_{k+1}\left(\frac{k}{pn}\right)^{\frac{1}{\hat{\theta}}}, \quad (15)$$

where θ is the tail index. Straetmans et al. [2008] used the tail quantile statistic $\frac{\sqrt{k}}{ln\left(\frac{k}{pk}\right)}\left[ln\frac{\hat{q}(p)}{q(p)}\right]$,which is asymptotically normally distributed, in a t test as follows:

$$T_q = \frac{\hat{q}_1 - \hat{q}_2}{\sigma[\hat{q}_1 - \hat{q}_2]} \sim N(0,1), \quad (16)$$

where the denominator is calculated as the bootstrapped difference between the estimated quantile parameters $\hat{q}_p$ using 1,000 bootstraps. The null hypothesis of this test is that the $\hat{q}_p$ parameters do not come from independent samples of normal distributions, therefore, $\hat{q}_1 = \hat{q}_2$. The alternative hypothesis is that the $\hat{q}_p$ 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.

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. Note that we calculate these options prices off our IVs sample and we do not use traded options data. 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 yields for the S&P 500 index are used for the calculation of P as well as the risk-free rate from 3-, 6-, and 12-month T-bills. We implement a modified Figlewski [2010] method for extracting our RND structure, as in Felix et al. [2016].[8]

We estimate the EDF in 2 different ways. First, using the entire sample of realized returns (r), we estimate the long-term EDFs nonparametrically, where $r = ln(S_T/S_t)$ and S_t is the realized return index at time t, and S_T is the forward level of the same index 3, 6, or 12 months forward (i.e., 21, 63, and 252 days forward, respectively). Because of overlapping periods, we initially estimate our empirical distribution from nonoverlapping returns for these 3 maturities by using distinct starting points. This methodology was also applied by Jackwerth [2000] and Ait-Sahalia and Lo [2000]. However, because the length of the overlapping periods is relatively large compared with our total sample, especially for the 12-month forward returns, we average the distribution with distinct starting points to prevent that our results are sensitive to any distribution (with specific starting point) used[9].

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. 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 the typically observed fat-tails and negative skewness, which are not incorporated in simple parametric models (e.g., the normal). This 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:

$$\ln(S_t/S_{t-1}) = \mu + \varepsilon_t, \quad \varepsilon_t \tilde{f}\left(0, \sigma^2_{t|t-1}\right), \tag{17a}$$

and

$$\sigma^2_{t|t-1} = \omega_1 + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{t-1|t-2} + \vartheta Max[0, -\varepsilon_{t-1}]^2, \tag{17b}$$

where α captures the sensitivity of conditional variance to the lagged squared innovations (ε^2_{t-1}), β captures the sensitivity of the conditional variance to the conditional variance ($\sigma^2_{t-1|t-2}$), 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 in which innovations are assumed to be normally distributed.

Up to this point, we managed to produce a 1-day horizon EDF for every day in our sample but we still lack time-varying EDFs for the 3-, 6-, and 12-month horizons. Thus, we use bootstrapping to draw 1,000 paths towards these desired horizons by randomly selecting single innovations (ε_{t+1}) from the 1-day horizon EDFs available for each day in our sample. We note that once the first return is drawn, the conditional variance is updated ($\sigma^2_{t-1|t-2}$) affecting the subsequent innovation drawings of a path. This sequential exercise continues through time until the desired horizon is reached. 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 1-period simulated returns is $\varepsilon_{t+1} + \mu + Rf$. The density functions produced by the collection of returns implied by the terminal values of every path and their starting points are our 3-, 6-, and 12-month EDFs. We note that by drawing returns from stylized distributions with fat tails and excess skewness, our EDFs for the 3 relevant horizons also embed such features. Finally, once these 3 time-varying EDFs are estimated for all days in our sample, we estimate γ for each of these days using Equation 12.[10]

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], which is based on Rosenberg and Engle [2002]. The choice for an EGARCH approach versus the standard GARCH model is due to the asymmetric feature of the former model that embeds the "leverage effect," which is the negative correlation between an asset's returns and changes in its volatility (see Bollerslev et al. [2009]).

Empirical analysis and results

In this section, we present our empirical results. As we estimate EDF in the 2 ways described (the long-term and time-varying EDFs), we are able to estimate long-term and time-varying γs by minimizing Equation 12. We use our long-term γ estimates to compare the ECPT to the other subjective density functions using the tests described in the section Density Function Tails' Consistency Test. The time-varying estimates of γ are analyzed in Estimated CPT Time-varying Parameters and Time Variation in the Probability Weighting Parameter and Investor Sentiment sections with the use of a regression model.

Estimated CPT long-term parameters

We report the estimated CPT parameters (λ and γ) extracted from long-term density functions in Table 1, Panel A. We find that λ, the parameter of loss aversion, which is 2.25 in the CPT shows a quite different outcome for 3-month option, 1.02, which indicates no loss aversion. For the 6- and 12-month options λ is 2.66 and 3.00, respectively. This finding suggests that loss aversion is more pronounced at longer maturities than suggested by the CPT. Apart from that, 12-month λ estimates vary strongly across the different optimization upper bounds used (i.e., 3, 5 and 10), always matching the bound value,[11] whereas estimates from 3- and 6-option maturities are very stable across upper bounds.

The estimated probability weighting function parameter γ is slightly higher than the one suggested by the CPT (i.e., 0.61) at the 3- and 6-month horizons at 0.75 and 0.81, respectively. For 12-month options, γ is

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

	Gamma (γ)		Lambda (λ),		($\gamma\|\hat{\lambda}$)	
Maturity	Estimate	RSS	Estimate	RSS	Estimate	RSS
3 months	0.75	0.02	1.02	0.12	0.54	0.01
6 months	0.81	0.02	2.66	0.30	0.87	0.02
12 months	1.09	0.06	3.00	1.64	1.12	0.07

Note. This table reports the estimated long-term CPT parameters gamma (γ), lambda (λ), and γ conditional on optimal λ ($\gamma|\hat{\lambda}$) from the single stock options as well as optimizations' residual sum of squares (RSS) of Equations 11 and 12. The parameter γ defines 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 loss aversion parameter. These parameters are long-term since their estimates are obtained by setting the average CPT density functions to match the return distribution realized within our full sample. These parameters are estimated using Equations 11 and 12.

around 1.09. These results suggest that overweighting of small probabilities occurs in short-term options (up to 6 months), while 12-month options seem to behave more rationally. These findings support our hypothesis that individual investors are, on average, biased when purchasing single stock call options, as suggested by Barberis and Huang [2008].

Density Functions Tails' Consistency Test results

As specified in Density Function Tails' Consistency Test section, we test the empirical consistency of density function tails among a set of 5 subjective distributions (CPT, PCPT, ECPT, power, exponential), the RND, and the EDF. We employ EVT through Equations 13–16. 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 obtain sampled returns[12].

Once we obtain returns for all 5 PDFs, the next step is to set k as the optimal number of observations used for estimation of φ by Equation 13, the Hill estimator. We produce Hill plots for the right tail of our distributions, which depict the relationship between k and φ as a curve (see Straetmans et al. [2008]). Picking the optimal k is done by observing the interval in this curve where the value of stabilizes while k changes. In this area there is a stable trade-off between a good approximation of the tail shape by the Pareto distribution and the uncertainty of such approximation (using fewer observations). The interval that corresponds to 4–7% of observations is the most 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 equal to 4%. Subsequently, the shape estimator φ for the EDF, RND, power, exponential, PCPT, CPT, and ECPT is computed, extreme quantile returns (EQRs) can also be estimated via Equation 14. Once k is determined, the tail cutoff point x_{k+1} follows automatically. Subsequently, the t test in Equation 16 is applied to evaluate whether the EQRs estimated from a set of 2 distributions (RND, power, exponential, PCPT, and CPT vs. EDF) have equal means (the null hypothesis). The results of this test are shown in Table 2, Panel A.

For the 3-month option maturity, we find that the EQR implied by the CPT is the only one that matches the realized EQR. The EQR implied by the ECPT is almost the same as implied by the CPT, thus, it also statistically matches the EDF. Per contrast, the EQRs for the RND, power, exponential, and PCPT densities always overshoot the one for the EDF. All other comparisons between these distributions' EQR at the 3-month maturity reject the null hypothesis that returns at the same quantile are equal. This empirical finding indicates that the equity market upside implied in option markets (i.e., the RND) and the power, exponential and PCPT densities are always higher than the ones realized by the equity market. The EQRs from the CPT and the ECPT are clearly the best matches for the EDF. For the 6-month maturity, upside returns priced by the RND and ECPT best match the EQR. No rational subjective density function consistently matches the EQR of the EDF, as they almost always overshoot it. Per contrast, the CPT density always undershoots the EDF's extreme returns. The PCPT is the only other subjective density (apart from the ECPT) that has an EQR statistically equal to the EDF. In contrast to the 3- and 6-month maturities, the EQRs from the RND for the 12-month maturity all underestimate the EQRs from realized returns. The same underestimation is documented for the PCPT, CPT and ECPT.

In line with these results for the EQR, Table 2, Panel B, shows that the expected upside (EU) for the EDF in the 3-month horizon more closely matches the EU for the CPT and ECPT density functions. Similar to our analysis on the EQR, for the other subjective densities, the EUs for all quantiles are also much larger than the EDF expected upside. For the 6-months maturity, the expected upsides for the CPT and ECPT density functions are no longer close to each other nor to the realized ones. The densities that better match the expected upside of the EDFs are the PCPT and the RND. For the 12-month horizon, in

Table 2. EVT consistency tests on tail returns.

Maturity	(1) vs. (2)	10% quantile (1)	(2) EDF	*p* value	*t* stat	5% quantile (1)	(2)	*p* value	*t* stat	1% quantile (1)	(2)	*p* value	*t* stat
Panel A - Tails extreme quantile returns (EQR)													
3 months	RND vs. EDF	0.16***	0.11	0.0%	−7.7	0.19***	0.13	0.0%	−8.2	0.26***	0.21	0.0%	−5.0
	Power vs. EDF	0.21***	0.11	0.0%	−13.9	0.23***	0.13	0.0%	−15.4	0.31***	0.21	0.0%	−10.4
	Expo vs. EDF	0.21***	0.11	0.0%	−13.9	0.24***	0.13	0.0%	−15.7	0.32***	0.21	0.0%	−11.5
	PCPT vs. EDF	0.18***	0.11	0.0%	−10.8	0.21***	0.13	0.0%	−11.6	0.27***	0.21	0.0%	−7.3
	CPT vs. EDF	0.13***	0.11	0.2%	−3.3	0.15***	0.13	0.7%	−2.9	0.21	0.21	36.3%	0.4
	ECPT vs. EDF	0.15***	0.11	0.0%	−5.4	0.17***	0.13	0.0%	−5.3	0.23*	0.21	8.1%	−1.8
6 months	RND vs. EDF	0.19	0.18	18.5%	−1.2	0.22	0.22	36.5%	−0.4	0.3*	0.32	8.9%	1.7
	Power vs. EDF	0.25***	0.18	0.0%	−9.6	0.28***	0.22	0.0%	−10.2	0.36***	0.32	0.0%	−4.3
	Expo vs. EDF	0.26***	0.18	0.0%	−10.6	0.29***	0.22	0.0%	−11.5	0.37***	0.32	0.0%	−5.5
	PCPT vs. EDF	0.22***	0.18	0.0%	−5.1	0.24***	0.22	0.0%	−4.6	0.32	0.32	36.4%	−0.4
	CPT vs. EDF	0.16***	0.18	0.0%	4.2	0.18***	0.22	0.0%	6.5	0.25***	0.32	0.0%	6.6
	ECPT vs. EDF	0.19	0.18	28.4%	−0.8	0.21	0.22	35.2%	0.5	0.28***	0.32	0.5%	2.9
12 months	RND vs. EDF	0.22***	0.32	0.0%	10.5	0.26***	0.35	0.0%	11.5	0.37***	0.44	0.0%	6.0
	Power vs. EDF	0.33	0.32	19.8%	−1.2	0.36*	0.35	9.9%	−1.7	0.45*	0.44	9.2%	−1.7
	Expo vs. EDF	0.34***	0.32	0.2%	−3.2	0.38***	0.35	0.0%	−4.0	0.47***	0.44	0.1%	−3.5
	PCPT vs. EDF	0.26***	0.32	0.0%	6.4	0.3***	0.35	0.0%	7.0	0.39***	0.44	0.0%	3.7
	CPT vs. EDF	0.18***	0.32	0.0%	23.3	0.21***	0.35	0.0%	26.9	0.3***	0.44	0.0%	12.2
	ECPT vs. EDF	0.26***	0.32	0.0%	7.1	0.29***	0.35	0.0%	7.8	0.39***	0.44	0.0%	4.1
Panel B – Tails expected upside return													
3 months	RND vs. EDF	0.2***	0.15	0.0%	−5.3	0.23***	0.19	0.0%	−5.1	0.32*	0.3	8.9%	−1.7
	Power vs. EDF	0.25***	0.15	0.0%	−10.3	0.28***	0.19	0.0%	−10.9	0.37***	0.3	0.0%	−5.9
	Expo vs. EDF	0.26***	0.15	0.0%	−10.6	0.29***	0.19	0.0%	−11.4	0.39***	0.3	0.0%	−7.2
	PCPT vs. EDF	0.22***	0.15	0.0%	−7.5	0.25***	0.19	0.0%	−7.4	0.33***	0.3	0.6%	−2.9
	CPT vs. EDF	0.16	0.15	18.6%	−1.2	0.19	0.19	39.0%	−0.2	0.26***	0.3	0.1%	3.4
	ECPT vs. EDF	0.18***	0.15	0.5%	−3.0	0.21**	0.19	3.2%	−2.2	0.28	0.3	13.3%	1.5
6 months	RND vs. EDF	0.24	0.24	36.9%	0.4	0.27	0.28	10.9%	1.6	0.37***	0.41	0.2%	3.3
	Power vs. EDF	0.3***	0.24	0.0%	−6.8	0.33***	0.28	0.0%	−6.6	0.43	0.41	14.4%	−1.4
	Expo vs. EDF	0.31***	0.24	0.0%	−7.9	0.34***	0.28	0.0%	−8.0	0.45**	0.41	1.4%	−2.6
	PCPT vs. EDF	0.26**	0.24	1.4%	−2.6	0.29	0.28	13.3%	−1.5	0.38*	0.41	5.9%	2.0
	CPT vs. EDF	0.2***	0.24	0.0%	6.0	0.22***	0.28	0.0%	8.8	0.3***	0.41	0.0%	8.1
	ECPT vs. EDF	0.23	0.24	15.3%	1.4	0.26***	0.28	0.2%	3.3	0.35***	0.41	0.0%	4.9
12 months	RND vs. EDF	0.28***	0.37	0.0%	7.6	0.33***	0.4	0.0%	7.7	0.46***	0.51	0.7%	2.8
	Power vs. EDF	0.38	0.37	14.7%	−1.4	0.42*	0.4	5.8%	−2.0	0.53*	0.51	5.8%	−2.0
	Expo vs. EDF	0.4***	0.37	0.2%	−3.2	0.44***	0.4	0.0%	−4.0	0.55***	0.51	0.1%	−3.4
	PCPT vs. EDF	0.32***	0.37	0.0%	4.8	0.36***	0.4	0.0%	4.9	0.48*	0.51	6.4%	1.9
	CPT vs. EDF	0.23***	0.37	0.0%	19.2	0.27***	0.4	0.0%	21.5	0.38***	0.51	0.0%	9.1
	ECPT vs. EDF	0.31***	0.37	0.0%	5.3	0.35***	0.4	0.0%	5.6	0.47**	0.51	3.0%	2.3

Note. This table reports the results from statistical test of the extreme quantile return, EQR (in Panel A) and tail expected upside returns, EU (in Panel B), performed according to Equations 14–16 applied to averaged density functions. Since the densities compared here are averaged for the RND and for the subjective densities or estimated using our full sample for realized returns, these tests aim to investigate the long-term consistency between the distribution tails. The null hypothesis of these tests is that the EQR and the tail expected upside returns from the distributions compared have equal means and, therefore, tails are consistent. The rejection of the null hypothesis is tested by *t* tests of Equation 16) at the 10%, 5%, and 1% statistical levels, respectively, shown by superscripts *, **, ***, assigned to the EQR or the EU for the RND, displayed in column 1, and for the same statistics for the EDF, shown in column 2.

line with the results from our EQR analysis, the power density best matches realized EUs.

In summary, across the 3 EVT tests performed (i.e., on tail shape, EQR, and EU), the 3 option maturities and the 3 quantiles evaluated, we observe that the success rate of the CPT subjective density functions on matching the EDF tails is 57%. In contrast, this success rate is 38% for the power utility, 33% for the RND and only 10% for the exponential utility density function. These results suggest that CPT-related distributions best match the EQR of the EDF, especially at the short maturities.

These findings reiterate our takeaway from the Estimated CPT Long-Term Parameters section: 12-month options are priced more rationally than shorter-term ones, which seem to be priced as a result of lottery buying by individual investors. Figure 1 compares the CDFs from 6 of our equity return densities: the EDF, the RND, the CPT, the PCPT, the exponential-utility density, and the power-utility density.[13] 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.

In Figure 1, we see that the tails implied by option prices (RND, in red) are fatter than the tails from the CPT (in dark blue) and EDF (in green) density functions over the 3-month horizon. The right tail of the RND distribution is thinner than the ones of the PCPT, the exponential- and the power-utility 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 EQRs and the EU estimated across the different quantiles, which in all cases report higher upside in the RND than in the EDF and the CPT. Figure 1 also suggests that the upside risk of the RND is more consistent with 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 at the 6-month horizon, suggests that the RND and the EDF are closer than at the 3-month horizon. At the same time, the CPT density seems more disconnected from the EDF. These findings match 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 and closer to the EDF one than to the CPT, which matches our findings from EQR and the expected upside tests. The exponential and power utility densities have right tails that are much fatter than the other densities, including the EDF. Figure 1 shows that at the 12-month 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. The RND seems to match the EDF the best, suggesting long-term options trade more rationally than short-term ones.

Overall, Figure 1 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. These results strengthen the evidence provided by Ilmanen [2012], Barberis [2013], Conrad et al. [2013], Boyer and Vorkink [2014], and Choy [2015], who showed that investors push single stock option prices to extreme valuation levels.

Estimated CPT time-varying parameters

To further investigate time-variation in the CPT's overweighting of small probabilities in single stock

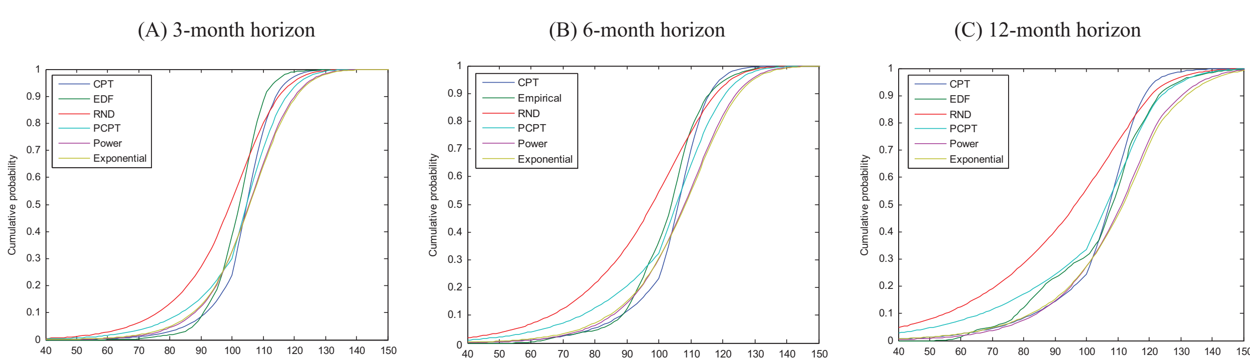

Figure 1. Cumulative density functions. This figure shows 3 plots that depict the cumulative density function (CDF) for equity returns obtained from the empirical density function (EDF), the risk-neutral density (RND), and the 4 subjective density functions: (a) the power utility density, (b) the exponential utility density, (c) the cumulative prospective theory density (CPT), and the partial CPT (PCPT). The equity returns' CDFs from these 6 sources are presented for the 3-, 6-, and 12-month horizons. The plots display cumulative probabilities on the *y*-axis and the terminal price levels on the *x*-axis, given an initial price level of 100.

options, we apply Equation 12 to each day in the sample to estimate the γ (weighting function) parameter. Lower and upper bounds of –0.25 and 1.75 are used in this optimization as they produce the lowest RSS across permutation of all bounds when using the CPT parameterization. We estimate γ under 4 different assumptions about the loss aversion parameter: () λ equals 2.25, the CPT parameterization; (b) no loss aversion, λ equals 1; (c) augmented loss aversion, λ equals 3; and (d) optimal λ, as estimated by Equation 11.

Table 3, Panel A reports the statistics when λ equals 2.25. We find that the median and mean time-varying values of γ, estimated from the 3-month options, are above its CPT value of 0.61 but still reflect overweight of small probabilities. This finding suggests that overweighting of small probabilities is present within the pricing of 3-month call options as suggested by the theory. The distribution of γ is skewed to the right and overweight of small probabilities is present 64% of times within the 3-month maturity. The 25th percentile of γ is 0.74, clearly suggesting a less pronounced overweight of small probabilities than suggested by the CPT. Interestingly, when we split the sample in 3 parts (Table 3), we observe that overweight of small probabilities is most strongly present at the beginning of our sample, in 97% of the days from January 5, 1998, to January 30, 2003, but that has it faded since 2003. This finding suggests that overpricing of single stock options is not structural, which only partially confirms our hypothesis that the CPT can empirically explain the overpricing of OTM single stock call options.

At the 6-month maturity, overweighting of small probabilities is less frequent than in the 3-month tenor. The long-term γ equals 0.81 and is somewhat out of sync with the time-varying estimates. Similarly, to the 3-month maturity, the distribution of γ is also slightly skewed to the right and the 75th quantile of γ equals 1.14, suggesting an underweighting of tail probabilities. However, probability overweighting again decreases over time as within the overall sample. Differently from the other maturities, γ estimates for the 12-month maturity tend toward underweight of tail probabilities, with median γ of 1.03 and a mean γ of 1.01.

In summary, the statistics in Table 3, Panel A, indicate that the weighting function parameters γ for all 3 maturities are time varying. Overweight of small probabilities holds for the 3-month maturity, less convincingly so for the 6-month maturity, and not at all for the 12-month maturity.

As the loss-aversion parameter λ is of high importance in the CPT model, we estimate γ under different λ parameterizations, more specifically, for (a) λ equals 1; (b) λ equals 3; (c) optimal λ, as estimated from the long-term empirical distribution (see Table 1). We report these estimates in Panel B of Table 3, when we assume λ equals 1. The estimates for γ in all 3 maturities fall strongly compared with the estimates under

Table 3. Time-varying gamma parameter.

Maturity	Min	25% Qtile	Median	Mean	75% Qtile	Max	StDev	% γ<1	% γ<1 (98-03)	% γ<1 (03-08)	% γ<1 (08-13)	RSS
Panel A – Gamma with CPT loss aversion ($\lambda = 2.25$)												
3 months	–	0.74	0.91	0.89	1.04	1.75	0.23	64%	97%	35%	59%	0.0209
6 months	–	0.81	0.99	0.96	1.14	1.75	0.28	52%	92%	18%	46%	0.0170
12 months	0.04	0.91	1.03	1.01	1.14	1.75	0.22	41%	83%	11%	29%	0.0225
Panel B – Gamma with no loss aversion ($\lambda = 1$)												
3 months	0.32	0.52	0.66	0.67	0.80	1.27	0.18	97%	100%	94%	97%	0.0253
6 months	0.32	0.55	0.71	0.72	0.87	1.75	0.21	90%	98%	79%	92%	0.0198
12 months	0.29	0.62	0.83	0.80	0.98	1.75	0.22	81%	98%	63%	83%	0.0169
Panel C – Gamma with augmented loss aversion ($\lambda = 3$)												
3 months	0.45	0.81	0.96	0.98	1.11	1.75	0.25	58%	93%	27%	53%	0.0230
6 months	0.38	0.89	1.02	1.06	1.25	1.75	0.25	45%	83%	13%	38%	0.0196
12 months	0.37	0.98	1.07	1.09	1.19	1.75	0.21	31%	66%	6%	22%	0.0265
Panel D – Gamma with optimized loss aversion												
3 months	0.33	0.52	0.66	0.67	0.81	1.75	0.18	97%	100%	93%	97%	0.0249
6 months	0.37	0.86	1.01	1.02	1.17	1.75	0.24	47%	86%	15%	40%	0.0187
12 months	0.34	0.96	1.05	1.06	1.17	1.75	0.20	35%	73%	8%	24%	0.0250

Note. This table reports the summary statistics of the estimated CPT time-varying parameter gamma (γ) from the single stock options market for each day in the full sample across different values of lambda (λ). The parameter λ is the loss 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 values close to zero denote a larger overweighting of small probabilities. The column with heading % $\gamma < 1$ reports the percentage of observations in which $\gamma < 1$, thus, the proportion of the sample in which overweight of small probabilities is observed. We report this metric for the full sample as well as for 3 equal-sized splits of our full sample, namely: (98-03), from January 5, 1998, to January 30, 2003; (03-08), from January 31, 2003, to February 21, 2008; and (08-13), from February 22, 2008, to March 19, 2013. Panel A reports the summary statistics of γ when we assume the CPT parameterization, where λ equals 2.25. Panel B reports the summary statistics of γ when we assume the loss aversion parameter λ equals 1 (no loss aversion). Panel C reports the summary statistics of γ when we assume λ equals to 3 (augmented loss aversion). Panel D reports the summary statistics of γ when we assume λ to equal its estimated (optimal) values, as reported in Table 1, Panel A.

the CPT loss aversion calibration (λ=2.25). A lower loss aversion parameter consistently gives rise to higher γ estimates, across the different options' maturities and quantiles. The opposite effect is observed when λ is increased from 2.25 to 3, as shown in Table 3, Panel C. The rise in γ estimates is observed in all 3 maturities and across both the 25 and 75% quantiles. Table 3, Panel D, which reports γ estimates when optimized λ parameters are used, shows distinct results for the 3-month maturity versus the 6- and 12-month maturity. For the 3-month maturity, we observe a downward shift to γ estimates, whereas for 6- and 12-month maturities, the estimates rise. However, this opposite effect in estimates is, in fact, qualitatively equal to the result just described when we assume λ as 1 or 3, as the optimal λ parameters estimated for the 3-, 6-, and 12-month maturities are 1.02, 2.66, and 3.00, respectively.

The reason why a lower (higher) loss aversion gives rise to a decreased (increased) γ is that it increases (decreases) the probability on the left side of distribution, influencing the probabilities and the shape of the right side of the CPT distribution. High values of λ push the CPT density to have more probability on the right side of the distribution, which is spread proportionally to the probabilities originally observed in the right-side bins, all else equal. The impact of such probability shift fades into the tail. Nevertheless, the right tail of the CPT density does turn fatter (and the γ parameter higher) as λ increases. The opposite occurs for low values of λ: the right tail of the CPT density thins, causing γ estimates to be low (which more forcefully can turn the RND right tail into such thin CPT tail). One important finding from our experimentation with different λ parameters is the existing time variation observed when λ equals 2.25 is unchanged. Overweighting of tails increases dramatically when low levels of λ are used, while showing great time variation.

We interpret these findings as a strong evidence that single stock options are not overvalued due to a structural skewness preference. We reckon that if static skewness preferences would drive overweight of small probabilities, the parameter γ would be relatively stable throughout our sample. The volatility of γ supports the view that investors experience (time varying) "bias in beliefs" or, alternatively, time-varying preferences (see Barberis [2013]). Our results are in line with Green and Hwang [2011] and Chen, Kumar and Zhang [2015], who reported similar time-varying effects in the overpricing, skewness effects, and returns for IPOs and lottery-like stocks.

Time variation in the probability weighting parameter and investor sentiment

In the following we investigate which factors may explain the reported time-variation of γ. Our main hypothesis is that it is linked to investor sentiment. The link between sentiment and overweighting of small probabilities or lottery buying in 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] argued that the IT bubble of 2000, a period of high variation of γ, is linked to elevated investor sentiment, when the least sophisticated investors were the ones most inclined to purchase calls on growth and IT stocks. Figure 2 depicts time-varying γs and the Baker and Wurgler [2007] sentiment factor, showing that these measures often move in tandem. For example, during the IT bubble, the level of γ seems quite connected with the level of sentiment, especially for the 3- and 6-month options.

To formally test our hypothesis that time variation of γ is linked to investor sentiment, we specify a regression model. In Equation 18, the explained variable is γ for the 3-, 6-, and 12-month horizons and the explanatory variables are the Baker and Wurgler [2007] sentiment measure[14] and the percentage of bullish investors minus the percentage of bearish investors given by the survey of the American Association of Individual Investors (AAII), used as a proxy for individual investor sentiment by Han [2008]. From Goyal and Welch [2008], we used the following control variables: *E12*, the 12-month moving sum of earnings on the S&P5000 index; *B/m*, the book-to-market ratio; *Ntis*, the net equity expansion; *Rfree*, the risk-free rate; *Infl*, the annual inflation rate; *Corpr*, the corporate spread; *Svar*, the stock market variance and; *CSP*, the cross-sectional premium.[15] Our monthly sample starts in January 1998 and ends in December 2010.

$$\begin{aligned}\gamma_t = c + \psi_1 Sent_t + \psi_2 IISent_t + \psi_3 E12_t + \psi_4 B/M_t \\ + \psi_5 Ntis_t + \psi_6 Rfree_t + \psi_7 Infl_t + \psi_8 Corpr_t \\ + \psi_9 Svar_t + \psi_{10} CSP_t + \varepsilon_t,\end{aligned} \tag{18}$$

where *Sent* is the Baker and Wurgler [2007] sentiment measure, *IISent* is the AAII individual investor sentiment measure.

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

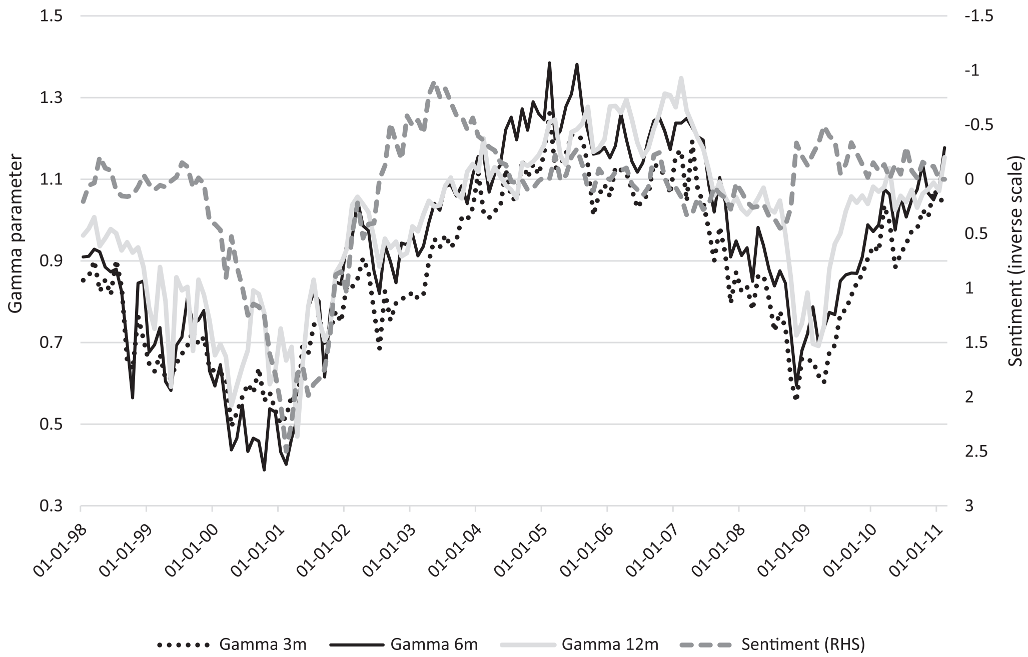

Figure 2. Time varying nature of gamma parameter in CPT. This figure depicts the time-varying nature of the gamma, γ, parameter from 3-, 6-, and 12-month single stock options estimated using the CPT parameterization as well as the sentiment factor of Baker and Wurgler [2007].

$$\gamma_t = c_i + \psi_i x_{i,t} + \varepsilon_t, \quad (19)$$

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

Table 4, Panel A presents the estimates of Equation 18. We note the high explanatory power of the multivariate regression, ranging from 68% to 71%. As expected, *Sent* is consistently negative and statistically significant across the 3 different horizons studied as well as in the univariate regressions of *Sent*. The explanatory power of the variable *Sent* in the univariate setting is high, between 22% and 29%. These findings do support our hypothesis that overweighting of small probabilities increases at higher levels of sentiment and that sentiment strongly impacts the probability weighting bias of call option investors. In contrast with the variable *Sent*, the coefficients for the individual investor sentiment (*IISent)* are positive, but not statistically significant either in the multivariate setting or in the univariate one (see Table 4).

The 9 Goyal and Welch [2008] control variables add substantial explanatory power to our multivariate regressions, as the 3-, 6-, and 12-month models explain, respectively, 71%, 68%, and 67% of γ. The control variables that are statistically significant in our multivariate setting are *E12*, *B/m*, *Rfree*, *Infl*, *Svar*, and *CSP* (Table 4). We observe that γ is positively linked to *E12*, the 12-month moving sum of earnings on the S&P 500 index, as well as to *B/m*, the book-to-market ratio, in both multivariate and univariate regressions. The positive relation between *E12, B/m* and γ could be explained by mean reversion of earnings and valuation being linked to a greater overweighting of small probabilities, which could be justified by the higher investor sentiment outweighing earning downgrades and rising valuations in a rallying market. Further, the stock market variance, *Svar*, is negatively linked to γ. Apparently a higher risk environment coincides with a higher overweighting of small probabilities.

As an extra check, we apply the Least Absolute Shrinkage and Selection Operator (Lasso) methodology to our multivariate regressions (see Tibshirani [1996] and Appendix A.2). We apply Lasso to select the regressors that are most relevant for the overall fit of γ. The coefficients that shrink to zero via Lasso are identified in Table 4 (Panel A) with a dagger (†). Model selection via the Lasso confirms that *Sent* and *IISent* are more relevant for the overall fit of γ than the fundamental factors *Ntis*, *Infl*, *Corpr*, and *CSP*.

These results indicate that single stock option investors overweight small probabilities when sentiment is exuberant, not necessarily when stock fundamentals are exuberant. More importantly, these results support our earlier findings that overweight of small probabilities is strongly time-varying and linked to sentiment. Therefore, overweight of small probabilities is unlikely to result from (static) investor preferences but from investors' bias-in-beliefs or time-varying preferences, which seem related to sentiment.

Table 4. Regression results: CPT parameterization.

	Panel A - Multivariate			Panel B - Univariate													
Horizon	3m	6m	12m	3m	6m	12m	3m	6m	12m	3m	3m	3m	3m	3m	3m	3m	3m
Intercept	0.564***	0.646***	0.769***	0.891***	0.969***	1.016***	0.864***	0.939***	0.996***	0.573***	0.513***	0.865***	0.917***	0.855***	0.869***	0.917***	0.864***
	(0.068)	(0.090)	(0.074)	(0.020)	(0.023)	(0.019)	(0.022)	(0.026)	(0.022)	(0.031)	(0.061)	(0.022)	(0.031)	(0.023)	(0.022)	(0.026)	(0.021)
Sent	−0.078***	−0.121***	−0.114***	−0.145***	−0.206***	−0.157***											
	(0.024)	(0.032)	(0.024)	(0.020)	(0.028)	(0.020)											
IISent	0.094	0.030	−0.003				0.031	−0.040	−0.067								
	(0.059)	(0.085)	(0.058)				(0.083)	(0.107)	(0.079)								
E12	0.049***	0.052***	0.044***							0.059***							
	(0.008)	(0.010)	(0.011)							(0.005)							
B/m	0.617**	0.698*	0.387								1.428***						
	(0.268)	(0.344)	(0.331)								(0.266)						
Ntis	0.357	0.201†	−0.869									0.617					
	(0.553)	(0.676)	(0.567)									(0.811)					
Rfree	−0.022**	−0.030**	−0.016										−0.022				
	(0.010)	(0.014)	(0.011)										(0.013)				
Infl	−0.580†	−3.111†	0.370†											5.867			
	(2.539)	(3.493)	(2.916)											(4.778)			
Corpr	0.100†	−0.174†	−0.037†												−0.314		
	(0.275)	(0.373)	(0.349)												(0.554)		
Svar	−9.762***	−11.792***	−8.679**													−13.280***	
	(3.179)	(4.156)	(3.686)													(4.835)	
CSP	−0.219†	−0.219†	−0.116†														0.626*
	(0.197)	(0.233)	(0.238)														(0.313)
R^2	71%	68%	67%	22%	29%	27%	0%	0%	1%	37%	31%	0%	3%	1%	0%	18%	2%
F stat	36.5	31.7	30.2	43.7	64.3	57.7	0.2	0.2	0.8	90.3	70.0	0.7	5.4	2.2	0.3	33.9	3.8
AIC	−247.1	−166.0	−233.7	−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	−213.4	−132.3	−200.0	−320.0	−179.9	40.3	0.1	0.1	0.1	0.0	0.2	0.7	0.0	3.9	0.5	2.3	0.3

Note. This table reports the regression results for Equation 18 in a multivariate setting in Panel A. The dependent variable in this regression is gamma (γ), where the explanatory variables are (a) the Baker and Wurgler [2007] sentiment measure, (b) the AAII individual investor sentiment measure, and (c) the explanatory variables used by Welch and Goyal [2008], excluding the factors that correlate with each other in excess of 40%. The regressors identified with a dagger (†) are the ones shrank to zero by the application of the Lasso. Panel B reports the regression results for Equation 19, which regresses γ and the same explanatory variables mentioned before in the univariate setting.

Furthermore, we run our regression models (Equations 18 and 19) using different assumptions about the value of λ, the loss aversion parameter. In this exercise we set λ to imply (a) no loss aversion (λ=1); (b) augmented loss aversion (λ=3); and (c) optimal loss aversion, where λ assumes the estimated value by Equation 11 and reported in Table 1, Panel A. To save space, we only report a summary of the outcome.

We find that the results for *Sent* are similar to the ones obtained in our main regressions: *Sent* is negatively linked to γ and statistically significant at all horizons but with less statistical significance, explanatory power and magnitude at the 12-month horizon. Across all options maturities, the *Sent* coefficient increases when λ equals 3 and it shrinks when λ equals 1, which is intuitive: as λ increases, the probabilities on the left side of the CPT distribution increase, favoring a thinner tail on the right side of the PCPT distribution, which then requires less overweight of tail adjustment (through a higher γ) for the PCPT to match the EDF. As a higher γ is obtained by such increase in λ, the coefficient of γ with the given sentiment factor also increases in magnitude.

The robustness of the relation between γ and *Sent* suggests that levels of loss aversion do not drive investors to overweight upside tail events, as one could hypothesize when associating upside speculation with a state of low loss aversion. Our results suggest that overweighting of small probabilities is consistently linked to sentiment, rather than positive fundamentals or loss aversion levels. Our results tie in closely with the findings of Green and Hwang [2011], who find that in IPOs the skewness effect is stronger during period of high investor sentiment. In the same line, Chen et al. [2015] concluded that when gambling sentiment is high, stocks with lottery-like characteristics earn positive abnormal returns in the short-run, followed by underperformance in the long run.

Robustness tests

In the following, we perform 4 robustness tests on our results: (a) by using the Prelec [1998] weighting function rather than the CPT one; (b) by applying Kupiec's test to probability tails; (c) by estimating time-varying γ under different assumptions of δ, α, and β; and (d) by evaluating whether the overweight of tails priced in the IV of single stock options is, in fact, not caused by a high IV priced in the index option market.

Kupiec's test for tail comparison

We employ Kupiec's [1995] test to compare the tails of the EDF with those of the subjective density functions and of the RND as a robustness test 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. Kupiec's method computes a proportion of failure (POF) statistic that evaluates 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 or consistency of tails is challenged. Kupiec's test is a log-likelihood ratio test, defined as

$$LR_{POF} = -2\ \log\left[(1-p^*)^{n-\upsilon}(p^*)^{\upsilon}\right] + 2\ \log\left[(1-[\upsilon/n])^{n-\upsilon}\ {}^{(\upsilon/n)\upsilon}\right] \sim \chi^2(1), \qquad (20)$$

where p^* is the POF under the null hypothesis, n is the sample size, and υ is the number of violations in the sample. The null hypothesis of such test is $\upsilon/n = p^*$ (i.e., the realized probability of failure matches the predicted one. If the *LR* exceeds the critical value, $\chi^2(1)$ =3.841, the hypothesis is rejected at the 5% level. In 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 υ/n is the realized number of violations. Because we 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 is outlining the expected percentage of failure (p^*) between the EQR from the EDF and from the subjective and risk-neutral densities. We assume p^* as being 5% and 10%. 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 a fatter tail reflects an overweighting of small probabilities, we expect that densities that do not adjust for the CPT weighting function will deliver a higher frequency of failures than the CPT density function. The Kupiec's test results are reported in Table 5.

Panel A in Table 5 suggests that the probability of failure for the RND, power, exponential, and PCPT densities is particularly high at the 3-month horizon, with more than 99% for the EQR at 90% and 95% and for p^*equal to 5% and 10%. These densities often contain fatter tails than the EDF. For the CPT density, the POF is much lower across the 2 values of p^*used and the 90% and 95% EQR. These findings suggest that at the 90% and 95% EQR, the CPT densities overstate less frequently the EDF tails than other densities. The violations of the EDF tails are, however, still substantial as they occur between 41% and 52% of times. Nevertheless, for the 99% EQR, the POF for all densities decreases considerably and, for the CPT, it becomes 16%.

Panel B of Table 5 depicts a very similar pattern of the POF for the probability densities derived from the 6-month options as we find for the 3-month options. The POF is close to 100% for all densities apart from the CPT at the 90% and 95% EQR, while at the 99% EQR violations fall substantially. Nevertheless, the CPT remains the best approximation for the EDF, with the lowest POF. The result suggests that the CPT density is statistically equal to the EDF. The results for p^* at the 5 or 10% level are very similar. Panel C presents the POF for the 12-month maturity. Again, the CPT violates the EDF the least, especially far out in the tail. The RND, power, exponential, and PCPT densities record POFs that are much smaller than for the 3- and 6-month maturities but that are still high in comparison to the CPT.

Overall, the Kupiec's test confirms the results reached in our EVT analysis, which further evidences that the CPT model is superior in matching realized returns.

Table 5. Robustness checks: Kupiec's test.

		EQR 90%			EQR 95%			EQR 99%	
$p^* = 10\%$	POF	p value	LR stat	POF	p value	LR stat	POF	p value	LR stat
Panel A. 3-month call options									
RND vs. EDF	99.9%	0.0000	∞	99.2%	0.0000	∞	50.5%	0.0000	414.8
Power vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	84.7%	0.0000	∞
Expo vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	86.8%	0.0000	∞
PCPT vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	67.2%	0.0000	752.0
CPT vs. EDF	58.2%	0.0000	559.6	45.7%	0.0000	333.3	16.0%	0.0002	13.6
$p^* = 5\%$	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	99.9%	0.0000	∞	99.2%	0.0000	∞	50.5%	0.0000	671.5
Power vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	84.7%	0.0000	∞
Expo vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	86.8%	0.0000	∞
PCPT vs. EDF	100.0%	0.0000	∞	100.0%	0.0000	∞	67.2%	0.0000	∞
CPT vs. EDF	58.2%	0.0000	861.9	45.7%	0.0000	561.3	16.0%	0.0000	65.5
Panel B. 6-month call options									
$p^* = 10\%$	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	99.9%	0.0000	∞	93.3%	0.0000	∞	13.8%	0.0160	5.8
Power vs. EDF	99.9%	0.0000	∞	97.7%	0.0000	∞	22.1%	0.0000	49.7
Expo vs. EDF	99.9%	0.0000	∞	97.8%	0.0000	∞	23.0%	0.0000	56.4
PCPT vs. EDF	99.9%	0.0000	∞	97.3%	0.0000	∞	17.0%	0.0000	18.2
CPT vs. EDF	62.4%	0.0000	647.0	36.3%	0.0000	197.3	5.7%	0.0019	9.6
$p^* = 5\%$	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	99.9%	0.0000	∞	93.3%	0.0000	∞	13.8%	0.0000	44.8
Power vs. EDF	99.9%	0.0000	∞	97.7%	0.0000	∞	22.1%	0.0000	137.7
Expo vs. EDF	99.9%	0.0000	∞	97.8%	0.0000	∞	23.0%	0.0000	149.7
PCPT vs. EDF	99.9%	0.0000	∞	97.3%	0.0000	∞	17.0%	0.0000	76.0
CPT vs. EDF	62.4%	0.0000	∞	36.3%	0.0000	369.9	5.7%	0.5474	0.4
Panel C. 12-month call options									
$p^* = 10\%$	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	62.8%	0.0000	655.1	25.0%	0.0000	72.9	20.3%	0.0000	37.0
Power vs. EDF	93.5%	0.0000	∞	42.5%	0.0000	283.5	29.3%	0.0000	114.7
Expo vs. EDF	94.6%	0.0000	∞	43.1%	0.0000	292.7	30.4%	0.0000	126.2
PCPT vs. EDF	79.5%	0.0000	1067.2	36.1%	0.0000	194.7	24.4%	0.0000	68.3
CPT vs. EDF	29.4%	0.0000	115.2	7.2%	0.0480	3.9	8.4%	0.2666	1.2
$p^* = 5\%$	EQR 90%			EQR 95%			EQR 99%		
RND vs. EDF	62.8%	0.0000	∞	25.0%	0.0000	177.9	20.3%	0.0000	114.2
Power vs. EDF	93.5%	0.0000	∞	42.5%	0.0000	492.6	29.3%	0.0000	245.6
Expo vs. EDF	94.6%	0.0000	∞	43.1%	0.0000	505.3	30.4%	0.0000	263.5
PCPT vs. EDF	79.5%	0.0000	∞	36.1%	0.0000	366.1	24.4%	0.0000	170.2
CPT vs. EDF	29.4%	0.0000	246.4	7.2%	0.0631	3.5	8.4%	0.0048	8.0

Note. 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. The test is performed as a robustness check to the extreme value theory (EVT)–based tests performed on the EQR and on the expected upside returns. The null hypothesis, which is designed as a log-likelihood ratio test (Equation 20)), is that the realized probability of failure (v/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% 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, where v/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.

Prelec's weighting function parameter

As a second robustness check, we estimate the weighting function parameter ω of the RDEU model suggested by Prelec [1998] to test whether our conclusions are robust to other weighted functions formulations[16]. The Prelec weighting function (w_P^{+-}) is given by Equation 21:

$$w_P^{\pm}(p) = \exp(-(-\log(\mathrm{p}))^{\omega}), \quad (21)$$

where the parameter ωdefines the curvature of the weighting function for both gains and losses, which also leads to *S*-shaped probability distortion functions. In Prelec [1998], the standard ω parameter value equals 0.65. Our time-varying and long-term (LT) estimates for ω are presented in Table 6, Panel A.

The long-term estimates of ω are somewhat in line with those in the RDEU but less so for the 12-month horizon: ω estimated from the 3-, 6-, and 12-months are 0.46, 0.67, and 1.11, respectively. These parameters are somewhat consistent with our long-term estimates for γ being, 0.75, 0.81, and 1.09 (see Table 1), as they suggest overweighting of small probabilities that fades with the increase in the option horizon. Similarly, time-varying estimates of ω also indicate more overweight of small probabilities than suggested by γ estimates. We find the mean (0.95) and median (0.93) for time-varying estimates of ω from 3-month options to be higher than the ones in Prelec [1998]. This outcome means that overweighting of small probabilities within the single stock option markets is less than suggested by RDEU and that estimated Prelec parameters imply a less pronounced overweight of tails than suggested by our CPT parameter estimations. In line with our results for the CPT, for the 6- and 12-month maturities, underweight of small probabilities is, however, more frequent than overweight. The average ω for the 6-month options is 1.02, and 1.05 for the 12-month options.

The time-variation observed in our main results is thus confirmed by the usage of Prelec's weighting function, as overweight of tails is pervasive mostly in the 1998–2003 sample.

Estimating Time-Varying γ Under Different Assumptions for δ, α, and β

As a third robustness test to our time-varying estimates of γ, we run optimizations where we fix the parameter δ instead of jointly optimizing it with γ. We impose $\delta = 1$ (no overweight of small probabilities on the left side of the distribution) or 0.69, the value of δ within the CPT. Table 6, Panels B and C, shows that results from optimizations with different values for δ are qualitatively the same as our main results (i.e., a positive term structure and sample dependency of overweight of small probabilities. Unreported results also indicate a negative correlation between γ

Table 6. Robustness checks: Time-varying weighting function parameters.

Maturity	Min	25% Qtile	Median	Mean	75% Qtile	Max	StDev	% γ<1	% γ<1 (98-03)	% γ<1 (03-08)	% γ<1 (08-13)	RSS	LT
Panel A - Prelec omega (ω)													
3 months	0.42	0.76	0.93	0.95	1.07	1.75	0.27	64%	95%	36%	61%	0.0204	0.46
6 months	0.37	0.84	0.99	1.02	1.17	1.75	0.26	51%	88%	21%	45%	0.0170	0.68
12 months	0.44	0.94	1.07	1.05	1.18	1.75	0.21	39%	79%	10%	28%	0.0201	1.14
Panel B – Gamma with overweight of small probabilities on the right tail (δ=0.69)													
3 months	0.44	0.70	0.86	0.97	1.21	1.75	0.34	58%	99%	23%	53%	0.0231	
6 months	0.40	0.75	1.01	1.04	1.27	1.75	0.31	49%	93%	13%	43%	0.0198	
12 months	0.40	0.83	1.05	1.04	1.24	1.75	0.25	43%	87%	11%	32%	0.0238	
Panel C – Gamma with neutral probability weighting on the right tail (δ=1)													
3 months	0.48	0.73	0.88	0.97	1.15	1.75	0.30	62%	98%	30%	58%	0.0230	
6 months	0.43	0.80	0.99	1.03	1.24	1.75	0.30	51%	92%	16%	45%	0.0191	
12 months	0.50	0.87	1.03	1.02	1.13	1.75	0.22	44%	84%	12%	35%	0.0233	
Panel D– Gamma with pronounced diminishing sensitivities to gains and losses (α and β =0.75)													
3 months	0.45	0.81	0.96	0.98	1.09	1.75	0.25	57%	93%	27%	52%	0.0230	
6 months	0.38	0.90	1.02	1.06	1.23	1.75	0.25	44%	82%	13%	36%	0.0196	
12 months	0.32	0.99	1.07	1.09	1.18	1.75	0.20	30%	65%	5%	21%	0.0276	
Panel E – Gamma with no diminishing sensitivities to gains and losses (α and β =1)													
3 months	0.00	0.72	0.88	0.87	1.03	1.75	0.24	67%	98%	40%	64%	0.0204	
6 months	0.00	0.78	0.98	0.93	1.14	1.75	0.30	55%	94%	21%	49%	0.0163	
12 months	0.04	0.86	1.02	0.98	1.14	1.75	0.25	44%	88%	11%	33%	0.0207	

Note. This table reports robustness checks of our time-varying estimates of overweight of small probabilities. Panel A reports the summary statistics of the estimated omega (ω) parameter, which is the parameter used in the Prelec's [1998] probability weighting function (see Equation 21). Similarly to the CPT, the parameter ω defines the curvature of the weighting function for gains and losses, which leads the probability weighting functions to assume inverse *S*-shapes. An ω parameter equal to one means a weighting function with un-weighted (neutral) probabilities, whereas $\omega < 1$ denotes overweighting of small probabilities. Similarly to γ, we estimate long-term $\omega's$ (reported for γ in Table 1, Panel A) as well as time-varying $\omega's$ (reported for γ in Table 4, Panel C). Panels B and C report γ estimates when the CPT's probability weighting parameter for left side of the distribution (δ) is assumed to be, respectively, 0.69 (the CPT parameterization) and 1 (neutral probability weighting). Panels C and D report γ estimates when the CPT's value weighting parameters α and β for diminishing sensitivity to gains and losses are assumed to be, respectively, 0.75 (increased diminishing sensitivity) and 1 (no diminishing sensitivity). We assume in these robustness tests that the loss aversion parameter λ equals 2.25.

and sentiment and high explanatory power of regressions.

Similarly, we estimate γ under different specifications for α and β. We assume $\alpha=\beta=1$ (no diminishing sensitivity to gains and losses) and $\alpha=\beta=0.75$ (more pronounced diminishing sensitivity to gains and losses) instead of the CPT parameterization $\alpha=\beta=0.88$. Our results, reported in Table 6, Panels D and E, suggest that lower sensitivity to gains and losses leads to a decrease in overweight of small probabilities (higher γ), whereas higher sensitivity to gains and losses leads to an increase in overweight of tails (lower γ). This effect is similar to the one observed by changes in λ (see Estimated CPT Time-Varying Parameters section).

As indicated in the Estimated CPT Time-Varying Parameters section, we also estimate time-varying γ using different lower (−0.25, 0 and 0.28) and upper bounds (1.2, 1.35, 1.5, 1.75, and 2). We find that higher bounds produce upward shifts in the estimated γ across all quantiles, median and averages to the extent that overweight of small probabilities becomes less pronounced but remains present. The time-variation pattern observed in Figure 2 and, more importantly, the strong negative relationship with sentiment reported in Table 4 are, though, extremely robust to changes in lower and upper optimization bounds.

Overweight of (right) tails driven by IV of single stock options

Finally, given that overweight of small probabilities by single stock call investors was most evident during the IT bubble period (as Table 3 suggests), we hereby evaluate whether this finding may have been driven by movements in the IV of index options rather than changes in the IV of single stock options. We perform such analysis because our methodology for calculation of average weighted stock IV volatilities partly relies on the IV on index options (as it depends on implied correlations), as Equations A.1j and A.1l in Appendix A.1 suggest. Essentially, we want to ensure that the overweight of small probabilities observed from our single stock options data is not caused by a rise in index options' IV. As overweight of small probabilities is a corollary of high IV skew[17], we examine the IV skews (120% moneyness vs. at the money [ATM]) from both index options and from single stock options within our sample using a k-nearest-neighbors (KNN) algorithm (see Appendix A.3 for detail). Figure 3 depicts a scatter plot that relates single stock IV skews (on the y-axis) with index option IV skew (on the x-axis) overlaid with the decision boundary between overweight of tails (in red) and its absence (in blue), produced by the application of the KNN algorithm to our full data sample. The picture

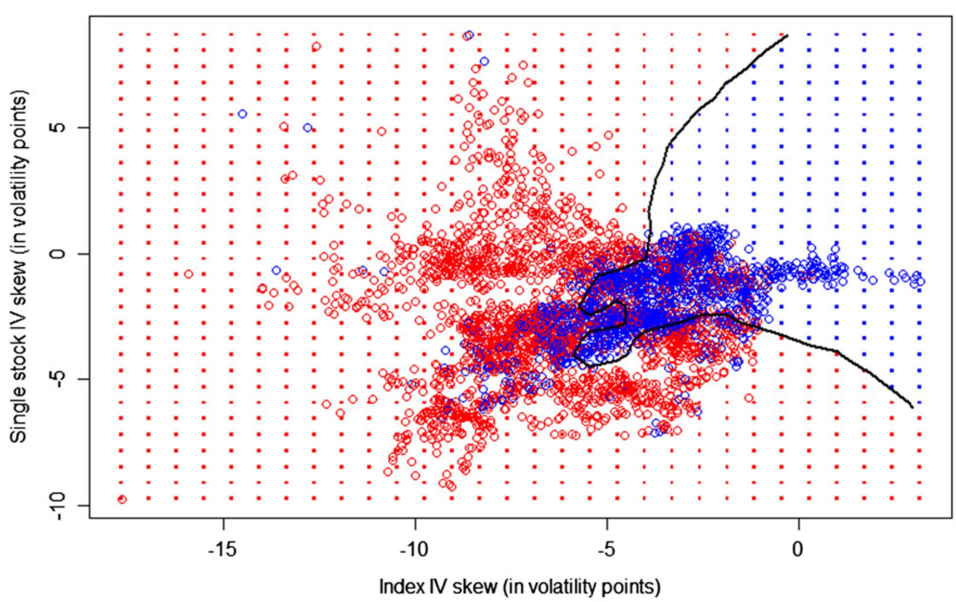

Figure 3. *k*-Nearest-neighbors Plot for IV Skews. This figure shows a scatter plots depicting the relation between single stock (120% minus ATM) IV skews (on the *y*-axis) and index (120% minus ATM) IV skews (on the *x*-axis). Observations colored in red imply the presence of overweight of small probabilities on the right side of the distribution ($\gamma < 1$), whereas observations colored in blue imply either neutral probability weighting ($\gamma = 1$) or underweight of small probabilities ($\gamma > 1$). The decision boundary is produced by a *k*-Nearest-Neighbors algorithm ($k = 41$, estimated via cross-validation) and delimits the region in which a new observation (of paired IV skews, such as the solid dotes) will be assigned to the overweight of small probabilities class (in red) or the alternative class (in blue).

suggests that that overweight of small probabilities is almost never caused by positive index IV skews, whereas positive single stock IV skews very often produce overweight of tails rather than underweights. Overweight of tails are mostly caused by situation where single stock IV skew are higher than index IV skew, which suggest that either high single stock IV skews or low implied correlation are responsible for overweight of tails, not index options' IV. These conditions can be anecdotally confirmed by our observation of IV skews during the 2000's IT bubble. During that period, when overweight of tails was pervasive, IV skew from single options was quite high, close to +10 volatility points, whereas the same IV skew from index options reached extreme low levels such as –15 volatility points. This disconnect between the 2 IV markets, which drove the implied correlation to 2.8% (an extreme low level), suggests that the index options' IV was not the driver for overweight of tails during the IT bubble. These findings reiterate our suggestions that overweight of small probabilities observed in our sample is caused by trading in single stock options by retail investors, rather than activity in the index option market.

Conclusion

Single-stock OTM call options are deemed overpriced because investors overpay for positively skewed securities, resembling lottery tickets. The CPT's probability weighting function of the Tversky and Kahneman [1992] theoretical model provides an appealing explanation why these options are expensive: investors' preferences for positively skewed securities. In our empirical analysis, we find that the CPT subjective density function implied by single stock options outperforms the RND and 2 rational densities functions (from the power and exponential utilities) in matching the tails of realized equity returns. We estimate the CPT probability weighting function parameter γ and find that it is qualitatively consistent with the one predicated by Tversky and Kahneman [1992], particularly for short-term options. This outcome endorses our hypothesis that investors in single stock call options are biased.

Our analysis provides detailed insights into the behavior of single stock option investors. Our empirical findings suggest that overweight of small probabilities is less pronounced than proposed by the CPT. We find that the overweighting of tail becomes less pronounced as the option maturity increases. Investors in single stock calls are more biased when trading short-term contracts, whereas they seem to be more rational 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 options to speculate on the upside of equities.

We also find that investors overweighting of small probabilities is strongly time-varying and sample dependent. Time variation in γs remains strong even when we account for different levels of loss aversion, different diminishing sensitivities to gains and losses, different degrees of overweighting of the left tail and an alternative (Prelec's) weighting function. The strong time-variation and sample dependency of γ suggest that investors do not have a static preference for skewness, but rather time-varying preferences or a "bias in beliefs" (see Barberis [2013]).

Such time variation in γ is also confirmed by the pronounced overweighting of tails during periods in which sentiment is high, for instance, the IT bubble period of 2000. This finding is consistent with the Baker and Wurgler [2007] sentiment measure being the main explanatory variable of overweighting of small probabilities. Our results challenge the view that single stock call options are structurally overpriced and offer the insight that overweight of tail events implied in these options are conditional on sentiment levels and option maturity rather than positive stock fundamentals, loss aversion levels, or investor preferences for skewness.

Our findings have several important practical implications. First, the understanding of time variation in investors' overweighting of small probabilities could be used in the development of behavioral option pricing models, which still remains in its infancy. To the extent that overweighting of small probabilities is a latent variable or, simply, not trivial to estimate, we suggest that future option pricing models should be made more sentiment-aware. Second, of importance for such next generation option-pricing models is the inclusion of a positive term structure of tails' overweighting. 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 identifying the presence of speculative equity markets bubbles.

Notes

1. Boyer and Vorkink [2014] provided evidence that lottery-like single stock options do deliver lower

forward returns than options with lower ex-ante skewness. However, their paper does not test why these options are overvalued, nor does it analyze the potential time-variation in ex-ante skewness and forward returns. Conrad et al. [2013] found similar results for ex ante skewness and subsequent stock returns.

2. 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 one is not. RDEU functions are also easier to estimate because they use 1 less parameter than the CPT.
3. We acknowledge that it is unclear whether overpricing of OTM calls is caused by overweighting of small probabilities (i.e., a matter of preferences), or rather by biased beliefs. Barberis [2013] discussed how both phenomena are distinctly different and how both (individually or jointly) may explain the overpricing in OTM options. We take a myopic view and use only the first explanation, for ease of exposition.
4. Loss aversion is the property in which people are more sensitive to losses than gains. For details, see Tversky and Kahneman [1992] and Barberis and Huang [2001]. Risk-seeking behavior happens when individuals are attracted by gambles with unfair prospects. The risk-seeking individual chooses a gamble over a sure thing even though the 2 outcomes have the same expected value. Nonlinear preferences occur when equally probable prospects are more heavily weighted by agents than others. For details, see Tversky and Kahneman [1992] and Prelec [1998].
5. We thank Barclays for providing the implied volatility data. Barclays disclaimer: "Any analysis that utilizes any data of Barclays, including all opinions and/or hypotheses therein, is solely the opinion of the author and not of Barclays. Barclays has not sponsored, approved or otherwise been involved in the making or preparation of this Report, nor in any analysis or conclusions presented herein. Any use of any data of Barclays used herein is pursuant to a license."
6. Zero is used as the lower bound for the estimation of λ to characterize loss aversion. The different upper bounds follow the maximum mean and median of λ across different experiments as estimated by Abdellaoni et al. [2007].
7. These upper and lower bounds were set to allow the optimizer to explore a wide range of parameters around the neutral (unweighted) probability function characterized by $\gamma = 1$ and, at the same time, allow for discerning levels of residual sum of squares to select the final bounds to be used in our experiment.
8. The main advantage of this method over other techniques is that it extracts the body and tails of the distribution separately, thereby allowing for fat tails. The Figlewski [2010] method is close to the one employed by and Bliss and Panigirtzoglou [2004], where body and tails are also extracted separately. We favor Figlewski's [2010] approach as the Bliss and Panigirtzoglou [2004] approach assumes that IV is constant beyond the observable strikes, resembling the Black-Scholes model (with lognormal tails).
9. As a robustness check to this approach, we compare our 3-, 6-, and 12-month empirical distributions with the ones calculated from nonoverlapping returns. We use data since 1871 for the U.S. equity price index, made available by Welch and Goyal [2008], who used S&P 500 data since 1926, and data from Robert Shiller's website for the preceding period. Our empirical distributions are quite similar to the ones estimated from the longer data set, suggesting that they are indeed suitable as long-term distributions.
10. Due to drift, the model of time-varying EDF for the 12-month horizon occasionally does not match the one of the PCPT model. This difference is challenging to estimation of γ (Equation 12), as a larger amount of γ estimates produce unreasonable PDFs such as nonmonotonic CDFs. Therefore, to perform the optimizations given in Equation 12, we set the mode of the simulated EDF equal to the one of the PCPT.
11. We do not find the high variability of the λ parameter from 12-month options to be a concern in our exercise as γ estimates conditional on $\hat{\lambda}$estimates using all upper bounds are not volatile and qualitatively the same as when $\lambda = 2.25$.
12. This technique is known as the Smirnov method. It entails drawing n random numbers from a uniformly distributed variable $U = (u_1, u_2, \ldots, u_n)$ 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]). Hence, the Smirnov method simulates returns that resemble the ones of the inverse CDF by randomly drawing probabilities along such function.
13. 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 the estimated long-term γ for the 3 maturities are close to the Tversky and Kahneman [1992] one.
14. Available at http://people.stern.nyu.edu/jwurgler/
15. The complete set and description of variables suggested by Goyal and Welch (2008) is available at http://www.hec.unil.ch/agoyal/. We select a smaller set to avoid multicollinearity, excluding variables that correlate more than 40 percent with each other.
16. A major advance of Prelec's [1998] weighting function vis-a-vis the CPT is that it is monotonic for any value of ω, whereas the CPT can have a nonmonotonic probability weighting for low levels of γ.
17. While this relation is widely acknowledged, Jarrow and Rudd [1982], Corrado and Su [1996], and Longstaff [1995] provided a formal theorem for the link between IV skew and risk-neutral skewness and kurtosis.
18. $\sum_{i=1}^{n} w_i^2 \sigma_i^2$ is typically of order 10^{-4} as a multiplication between powers of decimal numbers produces very small numbers, in the current case of order 10^{-5}. Distinctively, as $\left(\sum_{i=1}^{n} w_i \sigma_i\right)^2$applies the power after the sum is computed, its order of magnitude is much larger, of 10^{-2} in our data set, the same applying for σ_I^2.

19. Note that the CBOE S&P500 Implied Correlation index does not match our implied correlation metric also because it uses IVs for the largest 50 stocks of the S&P 500 index, whereas we use 100 stocks, and fixed option expirations "rotated" to proxy for 1- and 2-years maturities, whereas we use fixed maturities of 3-, 6- and 12-month maturities.

Acknowledgments

The authors are grateful to the anonymous referee at *Journal of Behavioral Finance* for their useful comments and suggestions. The authors thank seminar participants at the IFABS 2016 Barcelona Conference, VU University Amsterdam, APG Asset Management Quant Roundtable, at the 2016 Research in Behavioral Finance Conference in Amsterdam and at the Board of Governors of the Federal Reserve System for their helpful comments. The authors thank APG Asset Management for financial support and for making available part of the data.

References

Abdellaoui, M., H. Blichrodt, and C. Paraschiv. "Loss Aversion Under Prospect Theory: A Parameter-free Measurement." *Management Science*, 53, (2007), pp. 1659–1674.

Ait-Sahalia, Y., and A. W. Lo. "Nonparametric Risk Management and Implied Risk Aversion." *Journal of Econometrics*, 94, (2000), pp. 9–51.

Baker, M., and J. Wurgler. "Investor Sentiment in the Stock Market." *Journal of Economic Perspectives*, 21, (2007), pp. 129–157.

Barber, B., and T. Odean. "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." *Review of Financial Studies*, 21, (2008), pp. 785–818.

Barberis, N. "The Psychology of Tail Events: Progress and Challenges." *American Economic Review*, 103, (2013), pp. 611–616.

Barberis, N., and M. Huang. "Mental Accounting, Loss Aversion and Individual Stock Returns." *Journal of Finance*, 56, (2001), pp. 1247–1292.

Barberis, N., and M. Huang. "Stocks as Lotteries: The Implication of Probability Weighting for Security Prices." *American Economic Review*, 98, (2008), pp. 2066–2100.

Barberis, N., A. Shleifer, and R. Vishny. "A Model of Investor Sentiment." *Journal of Financial Economics*, 49, (1998), pp. 307–343.

Bates, D. "Empirical Option Pricing: A Retrospection." *Journal of Econometrics*, 116, (2003), pp. 387–404.

Benartzi, S., and R. Thaler. "Myopic Loss Aversion and the Equity Premium Puzzle." *The Quarterly Journal of Economics*, 110, (1995), pp. 73–92.

Bliss, R., and N. Panigirtzoglou. "Option-implied Risk Aversion Estimates." *Journal of Finance*, 59, (2004), pp. 407–446.

Bollen, N., and R. Whaley. "Does Net Buying Pressure Affect the Shape of Implied Volatility Function?" *Journal of Finance*, 59, (2004), pp. 711–754.

Bollerslev, T., G. Tauchen, and H. Zhou. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22, (2009), pp. 4463–4492.

Boyer, B., and K. Vorkink. "Stock Options as Lotteries." *Journal of Finance*, 69, (2014), pp. 1484–1527.

Chen, Y., A. Kumar, and C. Zhang. "Searching For Gambles: Investor Attention, Gambling Sentiment, and Stock Market Outcomes." SSRN, Working Paper (2015).

Choy, S-K. "Retail Clientele and Option Returns." *Journal of Banking & Finance*, 51, (2015), pp. 141–159.

Conrad, J., R. Dittmar, and E. Ghysels. "Ex-ante Skewness and Expected Stock Returns." *Journal of Finance*, 68, (2013), pp. 85–124.

Corrado, C., and T. Su. "Implied Volatility Skews and Stock Index Skewness and Kurtosis Implied by S&P 500 Index Option Prices." *Journal of Derivatives*, 4, (1997), pp. 8–19.

Danielsson, J., B. Jorgensen, M. Sarma, and C. de Vries. "Comparing Downside Risk Measures for Heavy Tailed Distributions." *Economics Letters*, 92, (2006), pp. 202–208.

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

Dierkes, M. *"Option-implied Risk Attitude Under Rank-dependent Utility*. University of Münster, Working Paper (2009).

Felix, L., R. Kräussl, and P. Stork. "The 2011 European Short Sale Ban: A Cure or a Curse?" *Journal of Financial Stability*, 25, (2016), pp. 115–131.

Figlewski, S. "Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio." In T. Bollerslev, J. Russell and M. Watson (eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, (pp. 323–353.) Oxford: Oxford University Press, 2010.

Frazzini, A., and L. Pedersen. "Betting Against Beta." *Journal of Financial Economics*, 111, (2014), pp. 1–25.

Gârleanu, N., L. Pedersen, and A. Poteshman. "Demand-based Option Pricing." *Review of Financial Studies*, 22, (2009), pp. 4259–4299.

Green, T., and B-H Hwang. "Initial Public Offering as Lotteries: Skewness Preferences and First-day Returns." *Management Science*, 58, (2011), pp. 432–444.

Han, B. "Investor Sentiment and Option Prices." *Review of Financial Studies*, 21, (2008), pp. 387–414.

Hastie, T., R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd Edition. Springer-Verlag, 2008.

Hill, B. "A Simple General Approach to Inference about the Tail of a Distribution." *Annals of Statistics*, 3, (1975), pp. 1163–1173.

Ilmanen, A. "Do Financial Markets Reward Buying or Selling Insurance and Lottery Tickets?" *Financial Analyst Journal*, 68, (2012), pp. 26–36.

Jackwerth, J. "Recovering Risk Aversion from Option Prices and Realized Returns." *Review of Financial Studies*, 13, (2000), pp. 433–451.

Kahneman, D., and A. Tversky. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica*, 47, (1979), pp. 263–291.

Kliger, D., and O. Levy. "Theories of Choice Under Risk: Insights from Financial Markets." *Journal of Economic Behavior & Organization*, 71, (2009), pp. 330–346.

Kumar, A. “Who Gambles in the Stock Market?” *Journal of Finance*, 64, (2009), pp. 1889–1933.
Kupiec, P. “Techniques for Verifying the Accuracy of Risk Management Models.” *Journal of Derivatives*, 3, (1995), pp. 73–84.
Lakonishok, J., I. Lee, N. Pearson, and A. Poteshman. “Option Market Activity.” *Review of Financial Studies*, 20, (2007), pp. 813–857.
Longstaff, F. “Option Pricing and the Martingale Restriction.” *Review of Financial Studies*, 8, (1995), pp. 1091–1124.
Mitton, T., and K. Vorkink. “Equilibrium Underdiversification and the Preference for Skewness.” *Review of Financial Studies*, 20, (2007), pp. 1255–1288.
Polkovnichenko, V. and F. Zhao. “Probability Weighting Functions Implied in Option Prices.” *Journal of Financial Economics*, 107, (2013), pp. 580–609.
Prelec, D. “The Probability Weighting Function.” *Econometrica*, 66, (1998), pp. 497–527.
Rosenberg, J., and R. Engle. “Empirical Pricing Kernels.” *Journal of Financial Economics*, 64, (2002), pp. 341–372
Straetmans, S., W. Verschoor, and C. Wolff. “Extreme US Stock Market Fluctuations in the Wake of 9/11.” *Journal of Applied Econometrics*, 23, (2008), pp. 17–42.
Tibshirani, R. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society*, 58, (1996), pp. 267–288.
Tversky, A., and D. Kahneman. “Advances in Prospect Theory: Cumulative Representation of Uncertainty.” *Journal of Risk and Uncertainty*, 5, (1992), pp. 297–323.
Welch, I., and A. Goyal. “A Comprehensive Look at the Empirical Performance of Equity Premium Prediction.” *Review of Financial Studies*, 21, (2008), pp. 1455–1508.

Appendix

A.1 - Single stock weighted average implied volatilities

Starting from the portfolio variance σ_P^2formula, Equation A.1a, we derive the single stock weighted-average IV, given in Equation A.1l:

$$\sigma_P^2 = \sum_{i=1}^{n} w_i^2\sigma_i^2 + \sum_{i\neq j}^{n} w_i w_j \rho_{ij}\sigma_i\sigma_j \tag{A.1a}$$

Where i and j index for the portfolio constituents. We re-write Equation A.1a for and index as:

$$\sigma_I^2 = \sum_{i,j=1}^{n} w_i w_j \rho_{ij}\sigma_i\sigma_j \tag{A.1b}$$

implying that,

$$\sum_{i\neq j}^{n} w_i w_j \rho_{ij}\sigma_i\sigma_j \tag{A.1c}$$

$$= \sum_{i,j=1}^{n} w_i w_j \rho_{ij}\sigma_i\sigma_j \tag{A.1c}$$

$$-\sum_{i=1}^{n} w_i^2\sigma_i^2 \tag{A.1c}$$

Where

$$\rho_{ij}(x) = \begin{cases} \overline{\rho} & if \;\; i\neq j \\ 1 & if \;\; i=j \end{cases},$$

and where σ_I is the index option IV. We use the S&P500 index and $n=100$, as we use IVs for the largest 100 stocks of this index. Then, assuming $\overline{\rho}$ as the estimator for average stock correlation we have:

$$\sigma_I^2 = \overline{\rho}\sum_{i\neq j}^{n} w_i w_j \sigma_i\sigma_j + \sum_{i=1}^{n} w_i^2\sigma_i^2 \tag{A.1e}$$

which, given equality A.1c, can be re-written as:

$$\sigma_I^2 = \overline{\rho}\sum_{i,j=1}^{n} w_i w_j \sigma_i\sigma_j - \overline{\rho}\sum_{i=1}^{n} w_i^2\sigma_i^2 + \sum_{i=1}^{n} w_i^2\sigma_i^2 \tag{A.1f}$$

$$\sigma_I^2 = \overline{\rho}\left(\sum_{i=1}^{n} w_i\sigma_i\right)^2 - \overline{\rho}\sum_{i=1}^{n} w_i^2\sigma_i^2 + \sum_{i=1}^{n} w_i^2\sigma_i^2 \tag{A.1g}$$

$$\sigma_I^2 = \overline{\rho}\left(\left(\sum_{i=1}^{n} w_i\sigma_i\right)^2 - \sum_{i=1}^{n} w_i^2\sigma_i^2\right) + \sum_{i=1}^{n} w_i^2\sigma_i^2 \tag{A.1h}$$

$$\overline{\rho} = \frac{\sigma_I^2 - \sum_{i=1}^{n} w_i^2\sigma_i^2}{\left(\sum_{i=1}^{n} w_i\sigma_i\right)^2 - \sum_{i=1}^{n} w_i^2\sigma_i^2}. \tag{A.1i}$$

As $\sum_{i=1}^{n} w_i^2\sigma_i^2$ is relatively small[18], we simplify (A.1i) into (A.1j), the implied correlation:

$$\overline{\rho} \approx \frac{\sigma_I^2}{\left(\sum_{i=1}^{n} w_i\sigma_i\right)^2} \tag{A.1j}$$

To obtain the single stock weighted average implied volatility (Equation A.1l), we square root both sides of the approximation and re-arrange its terms:

$$sqrt(\overline{\rho}) \approx \frac{\sigma_I}{\left(\sum_{i=1}^{n} w_i\sigma_i\right)}, \tag{A.1k}$$

$$\sum_{i=1}^{n} w_i\sigma_i \approx \frac{\sigma_I}{sqrt(\overline{\rho})}. \tag{A.1l}$$

Lastly, note that, given equality in Equation A.1c, Equation A.1i can be re-written as:

$$\overline{\rho} = \frac{\sigma_I^2 - \sum_{i=1}^{n} w_i^2\sigma_i^2}{\sum_{i\neq j}^{n} w_i w_j \sigma_i\sigma_j} = \frac{\sigma_I^2 - \sum_{i=1}^{n} w_i^2\sigma_i^2}{\sum_{i=1}^{n}\sum_{i\neq j} w_i w_j \sigma_i\sigma_j} \tag{A.1m}$$

which is the implied correlation (IC) measure employed by Driessen et al. [2013] and the CBOE S&P500 Implied correlation index[19].

A.2 – Least Absolute Shrinkage and Selection Operator (Lasso)

The regression coefficients in the Lasso methodology are estimated by minimizing the quantity:

$$\sum_{i=1}^{n}\left(y_1-\beta_0-\sum_{j=1}^{p}\left(\beta_j x_{ij}\right)^2\right)+\kappa\sum_{i=1}^{n}\beta_j=RSS+\kappa\sum_{i=1}^{n}\beta_j \tag{A.2}$$

Where κ is the tuning parameter, which is estimated via cross-validation. The cross-validation applied by us uses then equal-size splits of our overall data set.

A.3 – k-Nearest-Neighbor classifier

The k-Nearest-Neighbor (KNN) classifier is one of the approaches in machine learning that attempts to estimate the conditional distribution of the explained variable (Y) given the explanatory variables (X) and, subsequently, classify new observations to the class with highest estimated probability. The KNN classifier uses the Euclidean distance to first identify the closest k^{th} observations within the training data (in-sample data) to a new test (out-of-sample) observation provided (x_0). Such neighborhood of points around the test observation x_0 is defined as $\mathbb{N}_0$. KNN, then, estimates the conditional probability of x_0 to belong to a class j as the percentage of old observations (y_i) in the neighborhood $\mathbb{N}_0$ whose class is also j :

$$\Pr\left(Y=j|X=x_0\right)=\frac{1}{k}\sum\nolimits_{i\in\mathbb{N}_0}I(y_i=j). \tag{A.3}$$

In a third step, KNN applies the Bayes rule to perform out-of-sample classification (in test data) of x_0 to the class with the largest probability. For further details, see Hastie et al. [2008].