Cumulative Prospect Theory, Option Prices, and the Variance Premium*

Lieven Baele Joost Driessen Sebastian Ebert Juan M. Londono

Oliver Spalt

December 2016

Abstract

The variance premium and the pricing of out-of-the-money (OTM) equity index options are major challenges to standard asset pricing models. We develop a tractable equilibrium model with Cumulative Prospect Theory (CPT) preferences that can overcome both challenges. The key insight is that the variance premium can be written as the expected return on a portfolio of OTM call and put options, and the probability weighting feature of CPT can explain the puzzlingly low returns observed for these options. Using GMM on a sample of U.S. index option returns between 1996 and 2010, we show that the CPT model fits well observed option prices and, therefore, the variance premium. In a dynamic setting, probability weighting and time-varying equity return volatility combine to match the observed time-series pattern of the variance premium.

JEL Classification: C15, G11, G13
Keywords: Cumulative prospect theory, variance risk premium, probability weighting

*Juan M. Londono is at the Division of International Finance at the Federal Reserve Board of Governors. All other authors are at the Department of Finance at Tilburg University. The authors greatly benefited from comments and discussions with Nicholas Barberis, Ing-Haw Cheng (discussant), Lei Sun (discussant), David Veredas (discussant), and seminar participants at Tilburg University, the 2012 EFA meeting in Copenhagen, the 2015 FIRS conference in Reykjavik, the 2016 Belgian Financial Research Forum in Brussels, and the 2016 Research in Behavioral Finance Conference in Amsterdam. Please send correspondence to: J.J.A.G.Driessen@uvt.nl.

Cumulative Prospect Theory, Option Prices, and the Variance Premium

December 2016

Abstract

The variance premium and the pricing of out-of-the-money (OTM) equity index options are major challenges to standard asset pricing models. We develop a tractable equilibrium model with Cumulative Prospect Theory (CPT) preferences that can overcome both challenges. The key insight is that the variance premium can be written as the expected return on a portfolio of OTM call and put options, and the probability weighting feature of CPT can explain the puzzlingly low returns observed for these options. Using GMM on a sample of U.S. index option returns between 1996 and 2010, we show that the CPT model fits well observed option prices and, therefore, the variance premium. In a dynamic setting, probability weighting and time-varying equity return volatility combine to match the observed time-series pattern of the variance premium.

1 Introduction

A central empirical success of prospect theory (Kahneman and Tversky (1979); Tversky and Kahneman (1992)) is its ability to explain major puzzles in financial economics. For example, Benartzi and Thaler (1995) show that prospect theory can help in understanding the equity premium puzzle, and Barberis, Huang, and Santos (2001) demonstrate how a dynamic prospect theory framework can simultaneously generate a high equity premium, predictability of equity returns, and excess volatility. In these and other settings, prospect theory offers a rigorous and testable alternative framework to more traditional asset pricing theories (see, e.g., Barberis (2013) for a recent survey).

In this paper, we show that prospect theory can also help us understand one of the most important recent asset pricing puzzles: the variance premium. The variance premium, defined as the difference between the option-implied and the expected realized variance of stock returns, is strongly positive, on average, and time varying. Theoretically, the variance premium is a major puzzle because the standard consumption-based model with constant relative risk aversion (CRRA) preferences cannot generate a nonzero variance premium, irrespective of the risk aversion level and even when consumption variance varies over time (see, e.g., Drechsler and Yaron (2011)).

Empirically, the variance premium is a first-order phenomenon. For example, Coval and Shumway (2001), Driessen and Maenhout (2007), and Eraker (2013) find that the Sharpe ratio for volatility-selling strategies, such as shorting straddles, which effectively bet on the variance premium, is at least twice the Sharpe ratio of the underlying equity index.[1] If the equity premium is a puzzle, then, by this metric, the variance premium is an even bigger puzzle.

In this paper we show that cumulative prospect theory (CPT) can explain the variance premium puzzle. The key insight we use is that the variance premium can be written as the

[1]See also Bakshi and Kapadia (2003), Jiang and Tian (2005), Bakshi and Madam (2006), Carr and Wu (2009), Bollerslev et al. (2011), and Bollerslev et al. (2009), among others, for further evidence on the variance premium.

expected return on a portfolio of calls and puts with different strike prices. As a consequence, pricing the underlying options accurately is a sufficient condition for explaining the variance premium. Our central result is that a calibrated equilibrium model with CPT investors can generate option prices close to those observed in the data. We therefore show that CPT can explain the variance premium by solving the more general problem of accurately pricing options, and, in particular, out-of-the-money (OTM) call and put options.

The central intuition, formalized in a theoretical CPT model below, can be easily explained. Start with the standard pricing equation as in Cochrane (2005) applied to investors with CRRA preferences:

$$s_0 = E(x)/R_f + cov\left(m^{CRRA}, x\right), \tag{1.1}$$

which says that the price s_0 of an asset with payoff x is equal to its expected payoff discounted at the risk-free rate R_f, plus a risk correction given by the well-known covariance between the asset's payoff and the pricing kernel m^{CRRA}. A negative covariance corresponds to a positive risk premium. In case of CRRA preferences, the pricing kernel depends positively on the marginal utility of the investor.

Even though the CRRA pricing kernel is widely used, it does not explain observed option prices. To illustrate the disconnect between model and data, Figure 1 plots observed average option returns for different maturities (black squares) against option returns implied by a standard CRRA-model that we formalize below (blue x-es). Panel A shows that the CRRA model is a complete failure when pricing calls. First, because call options pay off when investors' marginal utility is low, the model predicts that calls should have positive expected returns. In the data, however, calls across all strike prices have negative returns. Second, Panel A of Figure 1 shows that, in the data, returns from investing in call options decrease for higher strike prices, whereas equation (1.1) implies that OTM call prices should increase because they pay off in the states with the lowest marginal utility. The CRRA model thus fails on two fronts: it predicts the wrong sign on call option returns and, on top, it predicts the wrong sign on the relation between call option returns and moneyness.

Panel B of Figure 1 presents analogous results for put returns. Qualitatively, the standard CRRA model does better on put returns than it does on calls. In particular, the model is consistent with the observed negative returns on puts, and it correctly predicts that OTM puts have the lowest returns. Quantitatively, however, the model fit is as bad as for call options. For reasonable parametrizations, standard CRRA models predict dramatically higher returns for puts, and especially OTM puts, than those observed. In Figure 1, this is reflected by the vertical difference between the data (black squares) and CRRA-implied values (blue x-es).

While the put option challenge has received a fair amount of attention in the existing literature (see, amongst many others, Bondarenko (2003a,b), Jones (2006), and Santa-Clara and Saretto (2009)), there is comparatively little work on the call option challenge (e.g., Bakshi et al. (2010) and Constantinides et al. (2013)).[2] This lack of attention is notable because, from a conceptual standpoint, the CRRA model does not even get the direction of call prices as a function of moneyness right.

In the theory part of the paper we show that when investors have CPT preferences the pricing equation (1.1) can be written as

$$s_0 = E(x)/R_f + cov\left(m^{CRRA}, x\right) + cov\left(m^{CPT}, x\right), \tag{1.2}$$

which says that the price of an asset is the discounted expected value, plus the CRRA risk correction, *plus* an additional covariance term. This additional covariance term depends on m^{CPT}, which we label the "CPT pricing kernel." Figure 2 plots m^{CPT} as a function of equity index returns. The crucial feature of the CPT pricing kernel is that it has a U-shape, which implies that the price of an asset is high if it pays off in extreme states of the world, irrespective of whether those states are good or bad. The CPT model therefore predicts simultaneously

[2] These two articles document the return patterns for both put and call options and propose reduced-form models to explain these patterns, but do not explore preference-based explanations for these patterns.

higher prices and lower returns for OTM call options and OTM put options. For comparison, Figure 2 also plots m^{CRRA}, which is monotonically decreasing, and can therefore not explain higher prices for OTM calls. Figure 1 illustrates our CPT model's fit. The CPT model-implied option returns (red crosses) line up remarkably well with the actual data for both puts and calls, which stands in sharp contrast to the poor fit of the CRRA model. Because pricing returns of puts and calls is a sufficient condition for explaining the variance premium, the results in Figure 1 also shows that the CPT model explains the variance premium puzzle.

An attractive feature of the framework captured by equation (1.2) is that it allows us to explain two separate puzzles—why returns of OTM puts are so low and why returns of OTM calls are so low—using one unifying mechanism. A second advantage of equation (1.2) in terms of theory building is that it nests the standard CRRA model as a special case, which allows us to directly compare the standard model and our proposed extension. Finally, it is important to note that equation (1.2) is not an ad hoc pricing kernel devised to explain a particular puzzle. Rather, as we formally show below, it is the pricing kernel that arises in an equilibrium with CPT preferences.

The key driver of the second covariance term is the probability weighting feature of CPT. Probability weighting is a non-linear transformation of objective probabilities whose main implication for our setting is that the tails of a distribution are overweighted when evaluating its attractiveness. Intuitively, probability weighting is a modeling device that captures demand for lottery tickets and insurance through the same underlying mechanism. Probability weighting is relevant in the context of option pricing because OTM calls and OTM puts, like lottery tickets and insurance, pay off in unlikely states of the world.

An extensive literature in decision science documents that probability weighting is a pervasive trait of human decision making (see, e.g., Epper and Fehr-Duda (2012)). Barberis (2013) surveys a growing literature in finance and economics on probability weighting, and writes that *"in risk-related fields of economics, such as finance, insurance, and gambling, there is now more empirical support for probability weighting than for loss aversion, an ar-*

guably better-known component of prospect theory." Introducing probability weighting into the discussion on the variance premium and the overpricing of OTM options, as well as documenting its quantitative impact on both, is new and a main contribution of this study.

Our paper proceeds as follows. We first introduce an equilibrium asset pricing model with CPT preferences in Section 2. The model builds on the work of Barberis et al. (2001), who consider a representative agent model in which the agent's preferences are the sum of a CRRA utility function and a gain-loss prospect theory utility function. Because our focus is on pricing options, we do not explicitly model consumption and dividends but directly depart from a lognormal distribution of stock market returns. We show that the resulting model yields the equilibrium pricing equation (1.2) above, which allows us to derive the expected return on the stock market index and on a range of options on the index. We can then derive the model-implied variance premium from these equilibrium returns.

A key advantage of the model is that it is very tractable. In addition, the model-implied portfolio weights are unique, finite, and wealth-independent. Exploding portfolio weights are a well-known shortcoming of existing CPT portfolio choice models (see, e.g., Barberis and Huang (2008); Bernard and Ghossoub (2010)). Even more importantly, we can prove that if, and only if, the CRRA component of preferences is given by log-utility, the CPT investor behaves like a one-period myopic investor, even when returns are not iid. This result, which generalizes results by Merton (1969) and Samuelson (1969) to a CPT setting, is interesting because it clarifies when insights obtained for a one-period setting extend to a multi-period setting. For example, this theoretical result allows us to study a setting with non-i.i.d. returns using a conditional one-period equilibrium model.

In Section 3, we bring the model to the data. Based on S&P 500 equity returns and S&P option prices from 1996 to 2010, we construct monthly call and put option returns for 13 different strike levels. Using GMM, we show that the model yields a very good fit for the cross-section of expected option returns, and, therefore, for the variance premium. All our tests show that probability weighting is the central determinant of the model's fit, and our tests formally reject the null hypothesis of no probability weighting. In our benchmark setup, the estimate of the Tversky-Kahneman probability distortion parameter is 0.67, and thus very similar to values found for subjects in lab experiments.

When we allow for separate distortions for gains and losses, we find that distorting probabilities in both tails is important. Distorting only gains yields an acceptable fit for calls, but the model then fails to explain puts well. Distorting only the probability of losses yields the opposite result. As we explain in greater detail below, the ability to distort also the right tail of the distribution distinguishes our CPT model from well-known alternative approaches to explain the variance premium in the literature that focus on the left tail (e.g., rare disasters, long-run risks, etc.). It also highlights why our strategy of solving the variance premium puzzle by solving the option pricing puzzle (the sufficient condition for the variance premium) is informative: any model that focuses on distorting either marginal utility or probabilities only on the downside is likely to fail to price call options well, *even though* the model may fit the variance premium.

In Section 4, we show that the model can also generate meaningful time variation in the variance premium. We focus on three potential drivers: time-varying equity return volatility, time-varying probability distortion, and time-varying loss aversion. The estimated model yields intuitive time-series patterns in loss aversion and probability weighting. The main finding, however, is that, with time-varying volatility, the CPT model successfully captures the dynamics of the variance premium, even when probability distortion is kept fixed at the benchmark value. Hence, we do not need time variation in investor preferences to explain the time variation in the variance premium. The intuition for this result is straightforward: the power of probability weighting comes from overweighting the tails of the wealth distribution. Higher return volatility translates into thicker tails, which, in turn, are then overweighted more. As a result, the variance premium is positively correlated with the volatility level, in line with empirical observations.

1.1 Related Literature

Our paper contributes to the literature that studies the usefulness of prospect theory in finance and to a growing number of recent papers exploring the usefulness of probability weighting in asset pricing contexts (see Barberis (2013) for a recent survey). Related papers that use a prospect theory framework with probability weighting include Driessen and Maenhout (2007), who empirically investigate portfolio choice of CPT-style investors when options are part of the asset menu; Barberis and Huang (2008), who theoretically investigate the pricing of assets with positively skewed returns; and De Giorgi and Legg (2012), who examine dynamic portfolio choice with narrow framing and probability weighting. Several empirical asset pricing studies provide evidence that is consistent with a role for CPT's probability weighting feature (for example, Green and Hwang (2012) for IPOs, Boyer and Vorking (2014) for stock options, and Ilmanen (2012) for a recent survey). Polkovnichenko and Zhao (2013) and Kliger and Levy (2009) are related papers that also use index options prices to identify CPT parameters. However, none of the above papers focus on the variance premium. To the best of our knowledge, our paper is the first to apply prospect theory to understand the variance premium and the related puzzle of the overpricing of OTM put and call options.

We also contribute to the recent literature that studies potential explanations for the variance premium puzzle. Relative to the CRRA-lognormal benchmark model, most work in this literature modifies either the data generating process, which makes extreme events more likely; investor preferences, which makes investors care more about extreme states of the world; or investor beliefs, which makes investors perceive extreme states as more likely than they are. For example, Gabaix (2012) shows how a model with time-varying disaster risk generates a positive variance premium even under CRRA preferences. Shaliastovich (2015) develops a long-run risk model in which consumption growth remains conditionally gaussian but whose expected growth rate has to be learned via a "recency"-biased updating procedure. In his model, expected consumption growth and its uncertainty are time-varying, while uncertainty is subject to jumps. The model predicts OTM put options to be expensive relative to at-the-money (ATM) options, as the former hedge confidence jump risks. Other papers make adjustments to both beliefs and preferences. Several authors combine the long-run risk approach of Bansal and Yaron (2004), which features recursive preferences and persistent time variation in consumption and dividend volatility, with nonlinear dynamics in fundamentals, such as volatility-of-volatility (Bollerslev et al. (2009), Londono (2014)) or jumps (Drechsler and Yaron (2011)). Instead, Du (2011) combines rare jumps in fundamentals with habit preferences as in Campbell and Cochrane (1999). Bekaert and Engstrom (2016) show how a positive variance premium can emerge through a combination of preferences with habit formation and Bad Environment-Good Environment (BEGE) dynamics for consumption. Drechsler (2013), building on the model of Liu et al. (2005), shows how time-varying model uncertainty can amplify the impact of jumps on equity and variance risk premia. Schreindorfer (2014) proposes a model with generalized disappointment aversion and regime-switches in volatility of consumption and dividends. This model generates substantial equity and variance premiums and high put option prices. Finally, Jin (2015) develops a dynamic equilibrium model in which beliefs about crash risk vary over time, with periods where investors overestimate the likelihood of a crash. This model generates variation in (crash) risk premiums over time, and predicts that the crash risk premium increases after a crash. The latter prediction is consistent with our empirical findings in the time-varying setting in Section 4.

There are a number of key differences between all these approaches and our paper. First, our model with CPT preferences generates a substantial variance premium even when asset returns are iid lognormal. As such, rare disaster models, such as the one in Gabaix (2012), present the polar opposite to our approach. These models maintain the CRRA nature of preferences, but change the return generating process. Second, previous work predominantly focuses on the left tail of the distribution, both in terms of the data generating process and preferences, and on the overpricing of OTM put options. Instead, our CPT model, via the probability weighting feature, can speak to both tails of the distribution, and as such explains the overpricing not only of OTM put options but also of OTM call options. Third, an important and potentially testable difference is that the above models are necessarily dynamic, whereas the CPT model can generate a variance premium even in a one-shot game with known probabilities. It is therefore possible that some of the variance premium is driven by a fundamentally different mechanism than suggested by the existing literature.

The remainder of this paper is organized as follows. Section 2 introduces a one-period version of our model. In Section 3, we perform a comparative static exercise that conveys the intuition for most of the results that follow. We then use a more formal GMM approach to match our CPT model to actual data in an unconditional setting. In Section 4, we explore the ability of CPT to characterize the time-varying nature of the variance premium. Section 5 concludes.

2 The Model

We start our analysis by considering a simple representative agent model with preferences given by a mixture of terminal wealth expected utility and gain-loss cumulative prospect theory (CPT) utility, the most parsimonious model that allows us to derive our key insights. The next subsection defines the investor's preferences. Subsection 2.2 solves a standard portfolio problem for these preferences in a static setting and discusses dynamic extensions, which are detailed in the appendix. Subsection 2.3 employs these results to derive the pricing kernel of the economy. Subsection 2.4 explains the relationship between option prices and the variance premium, in general and as predicted by our model.

2.1 Preferences

The representative agent's *total utility* is given by a weighted sum of the *expected utility* over terminal wealth W_T and a *CPT utility* over the gain or loss X_T. The final gain or loss is defined as $X_T = W_T - W_{Ref,T}$, where W_{Ref} denotes the *reference point* (Markowitz (1952)), which determines whether terminal wealth levels are perceived as gains ($W_T \geq W_{Ref}$) or losses ($W_T < W_{Ref}$). Formally, total utility is represented by

$$\Psi(W_T, X_T) = EU[W_T] + b_T CPT[X_T] \tag{2.1}$$

where $b_T \geq 0$ is a scaling term that governs the relative importance of the expected utility part, $EU[W_T]$, and the CPT part, $CPT[X_T]$. For the expected utility part, we assume that $EU[W_T] = E[U(W_T)]$ satisfies constant relative risk aversion (CRRA), i.e., the utility function U is given by

$$U(W_T) = \begin{cases} \frac{W_T^{1-\gamma}}{1-\gamma} & \text{for} \quad \gamma \neq 1 \\ \ln W_T & \qquad\;\; \gamma = 1 \end{cases}, \tag{2.2}$$

where γ is the risk aversion coefficient. Next, we define the CPT part. To this means, first, let v denote the value function (sometimes also called a *utility function)*, which is defined over gains and losses X_T. We assume that v takes the piece-wise linear form

$$v(X_T) = \begin{cases} X_T & \text{for} \quad X_T \geq 0 \\ \lambda X_T & \qquad\;\; X_T < 0 \end{cases}, \tag{2.3}$$

where $\lambda \geq 1$ is called the *loss aversion parameter.* For example, if $\lambda = 2$, then a decrease of terminal wealth beneath the reference level by one ($X_T = W_T - W_{Ref} = -1$) feels twice as bad as a gain of one beyond the reference level ($X_T = 1$) feels good. Second, in CPT, probabilities may be processed non-linearly. We define *probability weighting functions* $w^+$and w^-, for gains and losses respectively, by

$$w^-(p) = \frac{p^{c_1}}{\left[p^{c_1} + (1-p)^{c_1}\right]^{1/c_1}}, \tag{2.4}$$

$$w^+(p) = \frac{p^{c_2}}{\left[p^{c_2} + (1-p)^{c_2}\right]^{1/c_2}},$$

where c_1, $c_2 \in [0.28; 1]$ in equation (2.4) control the curvature of each weighting function.[3]

The weighting functions are inverse-S shaped, which means that probabilities close to zero are overweighted ($w(p) > p$), while probabilities close to 1 are underweighted ($w(p) < p$). The cumulative prospect theory of Tversky and Kahneman (1992), however, distorts cumulative and decumulative probabilities rather than marginal probabilities p to obtain *decision weights* for each state. Intuitively, using decision weights instead of distorted marginal probabilities, leads to an overweighting of unlikely and extreme states, i.e., the tails of the distribution. Formally, assume that there are N discrete states of the world at time T, each occurring with objective probability p_i, and that each state is associated with a specific final wealth level $W_{T,i}$. All states are then ordered ("ranked", see, e.g., Quiggin (1982)) from worst to best and related to the investor's reference point W_{Ref}: $W_1 \leq ... \leq W_{k-1} \leq W_{Ref} \leq W_k \leq ... \leq W_N$. Then, the decision weight π_i of state i is given by:

$$
\begin{aligned}
\pi_i &= w^-\left(p_1 + ... + p_i\right) - w^-\left(p_1 + ... + p_{i-1}\right) \text{ for } 2 \leq i \leq k \\
\pi_i &= w^+\left(p_i + ... + p_N\right) - w^+\left(p_{i+1} + ... + p_N\right) \text{ for } k+1 \leq i \leq N-1,
\end{aligned}
$$

where $\pi_1 = w^-\left(p_1\right)$ and $\pi_N = w^+\left(p_N\right)$. The CPT value of the gain or loss X_T is then computed as

$$
CPT(X_T) = \sum_{i=1}^{N} \pi_i v(X_{T,i}). \tag{2.5}
$$

We now discuss some special cases of our preference model. Like Barberis et al. (2001) (BHS; see also Kőszegi and Rabin (2006)), we assume that the investor's total utility is given by a weighted sum of expected utility and a gain loss utility term. The main departure from BHS is that our gain-loss utility term incorporates probability weighting. This additional feature will be the key driver of our results. When $c_1 = c_2 = 1$, the weighting functions in 2.4 are given by $w^+(p) = w^-(p) = p$, so that overall utility is the sum of terminal wealth and gain-loss utility, as is the case in BHS. For $b = 0$, our preference specification nests

[3] The lower bound is a technical condition that ensures that decision weights are positive (Rieger and Wang (2006) and Ingersoll (2008)).

pure expected utility. Following BHS, we make the assumption that gain-loss utility scales with wealth according to $b = \hat{b}W_T^{-\gamma}$ for some $\hat{b} \geq 0$. Moreover, following the prior literature (e.g., Barberis and Xiong (2009, 2012)), we assume that $W_{Ref} = W_0 R^f$, i.e., the investor experiences positive gain-loss utility if and only if her investment yields more than the risk-free return.

2.2 The Investor's Problem

We consider a simple market with the following assets: (i) a risk-free asset with a constant and exogenously given gross return R^f; (ii) equity that pays x_i^E in state i of the world at time T, which occurs with probability p_i; and (iii) D derivatives on equity, where derivative $d \in \{1, \ldots, D\}$ pays x_i^d in state i of the world at time T. The derivatives are in zero net supply, and we normalize the supply of equity to be one. The expected gross returns on equity, $E[R^E]$, and on each derivative d, $E[R^d]$, are to be determined in equilibrium. The prices s_0^d of the derivatives can then be obtained as $s_0^d = \frac{E[x^d]}{E[R^d]}$. The representative investor's problem is to choose a one-period optimal portfolio with positions in the risky asset, α_E, each derivative d, α_d, and the risk-free asset, such that she maximizes her total utility,

$$\max_{\alpha_d, \alpha_d} \Psi(W_T, X_T). \tag{2.6}$$

Given the investor's initial wealth W_0, her terminal wealth is given by

$$\begin{aligned} W_T &= \left[(1 - \alpha_E - \sum_d \alpha_d)R^f + \alpha_E R^E + \sum_d \alpha_d R^d\right] W_0 \\ &= \left[(R^f + \alpha_E(R^E - R^f)) + \sum_d \alpha_d (R^d - R^f)\right] W_0, \end{aligned}$$

and, accordingly, the gain or loss experienced is

$$X_T = \left[\alpha_E(R_i^E - R^f) + \sum_d \alpha_d (R^d - R^f)\right] W_0.$$

The first-order condition (FOC) that determines the optimal weight in equity (the FOCs for the derivative shares are determined analogously) is given by[4]

$$\begin{aligned} 0 &= E[U'(W_T)] + \hat{b} W_0^{-\gamma} \sum_{i=1}^{N} \pi_i v'(X_{T,i}) \\ &= \sum_i (R_i^E - R^f) \left[p_i (W_0(R^f + \alpha_E(R_i^E - R^f)))^{-\gamma} + \hat{b} W_0^{-\gamma} \pi_i (1 + (\lambda - 1) 1_{X_{T,i}<0}) \right] \\ &= \sum_i (R_i^E - R^f) \left[p_i (R^f + \alpha_E(R_i^E - R^f))^{-\gamma} + \hat{b} \pi_i (1 + (\lambda - 1) 1_{X_{T,i}<0}) \right], \end{aligned} \tag{2.7}$$

where the indicator function $1_{X_{T,i}<0}$ is equal to one if the investment in state i yields a loss and equal to zero otherwise. It is easy to verify that the second-order condition is satisfied, and therefore:

Proposition 1. *The investor's optimal portfolio weights are unique, finite, and wealth-independent.*

The finiteness of the the portfolio weights is noteworthy. In a standard portfolio problem framework with pure CPT preferences (and no terminal wealth utility) that is otherwise identical to ours, Bernard and Ghossoub (2010) show that infinite leverage is optimal whenever the expected return of the risky asset exceeds that of the riskless asset. A similar issue arises in the pure CPT model of Barberis and Huang (2008). Due to the rational terminal-wealth utility part, our model always results in a non-degenerate solution to the investor's maximization problem. Moreover, as in the model of BHS without probability weighting, our choice of b ensures that the portfolio weights are wealth independent. In other words, the wealth-independency property known from CRRA models carries over to our extended CPT model model with probability weighting.

In Appendix A, we extend the above model to a multi-period setting and prove the following result:

[4]For simplicity of presentation (i.e., to avoid case distinctions in the formulas), we ignore the technical issue of v not being differentiable at the reference point.

Proposition 2. *Consider a multi-period version of the model above where the investor has CRRA terminal wealth utility, can trade each period, and, in addition, experiences prospect theory utility over portfolio gains and losses each period. The investor's optimal portfolio weights are myopic if, and only if, the investor has log utility.*

Myopia means that the investor's portfolio choice in each period is independent of her investment horizon. In that case, the investor's portfolio weights remain unchanged compared to the one-period model. Note that the proposition does not assume that returns are independently and/or identically (iid) distributed. If returns are iid, then the investor chooses the same weights each period. The result generalizes Merton (1969) and Samuelson (1969), who find that portfolio choice is myopic for investors with log utility as well as for investors with general CRRA utility when returns are iid. When the investor, additionally, has gain-loss prospect theory utility as in our model, this latter result no longer holds true.

While we believe that Proposition 2 is of interest in its own right, it also adds some generality to the results in this paper. In the next subsection, we derive a pricing kernel for our economy. For simplicity of presentation, we do this in a one-period setting. From Proposition 2, it follows that this pricing kernel remains unchanged in a multi-period setting when the investor has log utility so that our results are not necessarily restricted to a one-period setting.

2.3 Equilibrium Pricing Kernel and Equity Risk Premium

The key to deriving the equity risk premium in our model is to note that, as derivatives are in zero net supply, market clearing requires $\alpha_E = 1$ and $\alpha_d = 0$ for all d. Upon substituting these equilibrium weights into the wealth constraint $W_T = W_0 R^E$ as well as into $X_T = R^E - R^f$, the FOC becomes:

$$
\begin{aligned}
0 &= \sum_i (R_i^E - R^f)\left[p_i\left(R_i^E\right)^{-\gamma} + \hat{b}\pi_i(1+(\lambda-1)1_{R_i^E<R^f})\right] \\
&\Longleftrightarrow 0 = \sum_i p_i(R_i^E - R^f)\left[\left(R_i^E\right)^{-\gamma} + \hat{b}\frac{\pi_i}{p_i}\left(1+(\lambda-1)\,1_{R_i^E<R^f}\right)\right] \qquad (2.8) \\
&\Longleftrightarrow 0 = E\left[(R^E - R^f)\left[m^{CRRA} + m^{CPT}\right]\right], \qquad (2.9)
\end{aligned}
$$

where m^{CRRA} and m^{CPT} refer to the random variables that, in state i, yield outcome $\left(R_i^E\right)^{-\gamma}$ and $\hat{b}\frac{\pi_i}{p_i}\left(1+(\lambda-1)\,1_{R_i^E<R^f}\right)$, respectively. The random variable $m := m^{CRRA} + m^{CPT}$ in equation (2.9) is a pricing kernel that can be used to price any asset in the economy that is in zero net supply—in equilibrium, the representative investor will not hold such assets in her portfolio. Recalling that $R^d = x^d/s_0^d$, it follows that (see, e.g., Cochrane (2005), p.13, equation (1.8)) $s_0^d = \frac{E[x^d]}{R^f} + cov(m, x^d)$ so that we have proven the following result:[5]

Proposition 3. *The equilibrium price of derivative d in the above model is given by*

$$
s_0^d = \frac{E[x^d]}{R^f} + cov(m^{CRRA}, x^d) + cov(m^{CPT}, x^d). \qquad (2.10)
$$

Figure 2 illustrates the two components of the pricing kernel. The price of each derivative is given by its discounted expected payoff plus the covariance of the pricing kernel with the payoff. If $\hat{b} = 0$, then $m^{CPT} = 0$, and the second covariance term vanishes. Since $m^{CRRA} = \left(R^E\right)^{-\gamma}$describes the investor's marginal utility from equity returns, m^{CRRA} is large when equity payout is low. Accordingly, the prices of derivatives whose payoffs correlate with the market $cov(m^{CRRA}, x^d) > 0$ are lower. As such, calls should have low prices while puts should have high prices. When $\hat{b} > 0$, derivative prices are influenced by a second covariance

[5]Proposition 3 holds when the pricing kernel is used to also price the risk-free asset. In the empirical analysis, we do not price the risk-free asset. Instead, we take the (observed) risk-free rate as given and focus directly on excess returns as in equation (2.9). Note that, to match any given risk-free rate level, our model can be generalized by scaling the total utility function (and hence the pricing kernel$m^{CRRA} + m^{CPT}$) with an additional free parameter. Since this would not affect equation (2.9), this would not change any of our results.

term $cov(m^{CPT}, x^d)$. The prospect theory part of the pricing kernel, m^{CPT}, is relatively larger when $\frac{\pi_i}{p_i} > 1$ and/or when the equity payoff is experienced as a loss (i.e., when the equity return is smaller than R^f). First, as regards the effect of probability weighting, it can be shown that $\frac{\pi_i}{p_i}$ is approximately equal to the derivative of the weighting function used in the computation of π_i. Recalling the inverse-S-shaped form of the weighting functions, the derivatives are large for both small and large probabilities. Taking into account the ranking of states according to the investor's terminal wealth when computing the decision weights π_i, this means that the pricing kernel is increased in states of extreme equity payouts, both large and small. Second, the factor $\left(1 + (\lambda - 1)\, 1_{R_i^E < R^f}\right)$ increases the CPT contribution to marginal utility by the constant loss aversion parameter λ if and only if state i is a loss state, thus leading to a discontinuity in the pricing kernel. Overall, the novel pricing kernel derived above results in lower expected returns of assets that pay in extremely bad and/or extremely good states (while, under standard CRRA, expected returns are larger the worse the state), and expected returns are lower for loss states than for gain states.

2.4 The Model-Implied Variance Premium

In this subsection, we define the variance premium and derive a new intuitive expression for the variance premium in terms of expected option returns.

As is common in the literature, we define the variance premium as the difference between the risk-neutral and the actual expected variance of the equity market return (see, e.g., Bollerslev et al. (2009)). Specifically, define the actual probability measure P as the measure under which the actual equity payoffs x_i^E are generated, and Q as the risk-neutral measure under which equity and derivatives are priced. The variance premium is then defined by

$$\begin{aligned} VP &\equiv V^Q - V^P \\ &= \sum_i q_i \left[ln\left(R_i^E\right) - E^Q\left[ln\left(R_i^E\right)\right]\right]^2 - \sum_i p_i \left[ln\left(R_i^E\right) - E^P\left[ln\left(R_i^E\right)\right]\right]^2, \quad (2.11) \end{aligned}$$

where p_i and q_i denote the actual and risk-neutral probability of state i, respectively. If investors were risk-neutral, P and Q would be identical and the variance premium would be equal to zero. Hence, the variance premium depends on investor preferences about variation in equity returns.

In our model, the risk-neutral probability for state i is given by

$$q_i = p_i \frac{\left[m_i^{CRRA} + m_i^{CPT}\right]}{E^P\left[m^{CRRA} + m^{CPT}\right]}. \tag{2.12}$$

Using the expressions for $m_i^{CRRA} + m_i^{CPT}$ from equation (2.9) and with an assumption on the actual distribution of returns, we can calculate the model-implied variance premium. Equation (2.12) shows how the difference between the actual and risk-neutral probability is determined by the preferences of the investor and illustrates how our setting differs from more traditional models. In addition to the standard component in traditional CRRA-based asset pricing models given by m_i^{CRRA}, the risk-neutral probability q_i depends on the CPT part of the pricing kernel, m_i^{CPT}.

The following proposition shows that the variance premium and expected option returns are closely related:

Proposition 4. *Define $R^c(K)$ and $R^p(K)$ as the return on a call and put option with strike K, respectively. If there are no arbitrage opportunities, we have*

$$VP \approx \int_0^{s_0^E} v^p(K)\, E^P\left[R^p(K) - R^f\right] dK + \int_{s_0^E}^{\infty} v^c(K)\, E^P\left[R^c(K) - R^f\right] dK, \tag{2.13}$$

where $v^c(K)$ and $v^p(K)$ are scalars whose expressions are given in Appendix B.

Proposition 4 shows that the variance premium is a weighted average of the expected returns on OTM put and call options across strikes, since equation (2.13) integrates over put options for strikes $K < s_0^E$ and over call options for strikes $K > s_0^E$. In Appendix B, we show that the weights $v^c(K)$ and $v^p(K)$ are effectively always negative. Hence, one can "earn" the variance premium by selling put and call options with various strikes. The resulting strategy forms a portfolio of short positions in so-called "strangles," where each strangle is a combination of an OTM put and an OTM call option. The variance premium, $V^Q - V^P$, thus represents the expected profit on this portfolio, which effectively sells both insurance against bad states of the world and provides a lottery-like payoff in the good states of the world. Hence, the variance premium reflects information about the compensation received in extreme states of the underlying asset (the equity market in case of index options), as these very negative and positive outcomes determine the value of OTM puts and calls, respectively. Equation (2.13) thus shows that the pricing of both tails of the return distribution affects the variance premium.

An important fact in our setting is that equation (2.13) is "model-free" in the sense that it does not require assumptions on preferences or distributions. The only assumption is the absence of arbitrage opportunities. Hence, the approximation holds both in the data, as well as in the model. We can thus write the difference between the model-implied variance premium, VP_{CPT}, and the observed variance premium, $\widehat{VP}$, as follows

$$VP_{CPT} - \widehat{VP} \approx \int_0^{s_0^E} v^p(K)\left(E^P_{CPT}\left[R^p(K)\right] - \hat{E}\left[R^p(K)\right]\right)dK + \int_{s_0^E}^{\infty} v^c(K)\left(E^P_{CPT}\left[R^c(K)\right] - \hat{E}\left[R^c(K)\right]\right)dK, \tag{2.14}$$

where $E^P_{CPT}[R]$ denotes the expected return implied by the model and $\hat{E}[R]$ the sample average return.[6] Thus, if the model-implied option returns exactly match the option returns in the data, then the variance premium is the same for both model and actual data. Pricing the cross-section of OTM options is therefore a sufficient condition for fitting the variance premium.[7] The equation also shows that a model can exactly match the variance premium

[6] In small samples equation (2.14) is subject to sampling error in the estimates of average returns and the variance premium.

[7] There are two minor caveats to this reasoning. First, equation (2.13) is only an approximation. We

while, at the same time, generating completely counterfactual option prices. Thus, fitting the variance premium by fitting options, as we do in this paper, is a stricter test than fitting only the variance premium.

Finally, as an important point of reference for our subsequent analysis, note that the variance premium is zero for the CRRA case. If the market payoff x^E and hence the terminal wealth of the investor W_T have a lognormal distribution, and if the representative agent has pure CRRA preferences ($b = 0$), Samuelson and Merton (1969) and Rubinstein (1976) show that call and put options are priced by the Black-Scholes formula, even though the investor is not allowed to trade during time 0 and T. This directly implies that the risk-neutral variance, as implied by option prices, equals the actual equity return variance and, hence, the variance premium is equal to zero.[8]

3 Empirical Results

This section explores the ability of the CPT model described in section 2 to fit the data. We first show that the model yields a good fit for observed option prices for a plausible set of input parameter values, which implies that the model also yields a good fit for the variance premium under those parameters. We then conduct a comparative statics exercise that conveys the intuition for most of the results that follow. Finally, we use a GMM approach to match the CPT model to the data.

show in Appendix B that the approximation error is equal to the difference between central and non-central second moments. For our application, this term is numerically small, as we focus on one-month returns—it is equal to 1.3 in our benchmark model, while empirically the variance premium is equal to 157.4. Second, empirically, we only have options for a finite number of strike prices.

[8]Brennan (1979) in fact shows that CRRA preferences are the only preferences that lead to the Black-Scholes formula in a setting with discrete trading and lognormal returns. Thus, only CRRA preferences generate a zero variance risk premium. The intuition for this result is that CRRA investors, when faced with constant investment opportunities and lognormal equity returns, optimally choose a constant equity exposure even over longer horizons. Hence, these investors are not interested in dynamic trading over time. Given that options are equivalent to dynamic trading strategies, these investors thus price options in the same way, irrespective of whether they are allowed to trade continuously (the Black-Scholes world) or not. Drechsler and Yaron (2011) show that CRRA preferences generate a zero variance premium also in a multi-period setting with long-run risk.

3.1 Data

Our main dataset consists of daily closing midquotes of S&P 500 index options for various strikes and maturities obtained from the OptionMetrics database from January 1996 to October 2010. OptionMetrics creates a "surface" of interpolated option prices for fixed levels of the Black-Scholes delta and for fixed maturities (30 calendar days, 60 days, etc.) for both puts and calls. We focus on 30-day options because these are the most liquid and the ones most frequently used in the literature. On each day and for both put and call options, we construct the return on buying an option and holding it to maturity, i.e., for 30 calendar days. This yields a panel containing the returns on 26 options across deltas. Of these 26 options, 13 are calls with deltas ranging from 0.2 to 0.8 (from OTM to in-the-money (ITM)), and 13 are puts with delta ranging from -0.8 to -0.2 (from ITM to OTM).

Table 1 presents summary statistics for option returns. The observed values reported in Figure 1 (black squares) are identical to the numbers in Table 1 except for the fact that deltas are replaced by the average moneyness of the options (K/S_0) for a given delta. The observed option returns are in line with the documented stylized facts in the literature (e.g., Coval and Shumway (2001); Jones (2006); Driessen and Maenhout (2007); Bakshi et al. (2010); Constantinides et al. (2013)). First, average call option returns are negative. This is in contrast to predictions from a standard CAPM or a model with pure CRRA preferences, which imply positive call option returns, because call options have a positive exposure to the underlying equity index. Second, put options have strongly negative average returns. For example, for deep OTM put options, the average return is about -53% per month.

In addition to option returns, we also construct the corresponding S&P 500 index returns for the 30-day holding period at a daily frequency. This gives an average return of 0.58% per 30 calendar days and a standard deviation of 5.18% per month. The average 1-month T-bill rate over the sample period is 0.25%, which gives an in-sample equity premium of 0.33% per month.

The variance premium in the data is calculated as the difference between the risk-neutral expected variance, V^Q, and the actual expected variance of equity index returns, V^P. We follow the existing literature and use the square of the VIX as our measure of risk-neutral expected variance (e.g., Bollerslev et al. (2009), Carr and Wu (2009)).[9] To calculate the actual expected variance, V^P, we regress realized variance, calculated using daily returns over the last 22 trading days, on the one-month-lagged realized variance and the square of the VIX level at the beginning of the month over which the realized variance is measured. The expected realized variance is then calculated as the forecast implied by this regression model (see Londono (2014)). In our sample, the annualized average of the squared VIX equals $572.5\%^2$, and the annualized average actual expected variance equals $415.1\%^2$. This gives an average variance premium of $157.4\%^2$, which is similar to values found in the related literature (e.g, Bollerslev et al. (2009); Londono (2014)).

3.2 Benchmark Parametrization and Model-implied Variance Premium

Before we estimate the model parameters with GMM, we assess the implications of our model for a benchmark parametrization. We assume that the equity return, R^E, is lognormally distributed with parameters $\mu_E = E\left(R^E\right)$, which is to be determined endogenously, and volatility σ_E.[10] We assume a one-period constant risk-free rate of $R^f = 0.25\%$ and an unconditional volatility of equity returns of $\sigma_E = 5.18\%$, which are, respectively, the observed unconditional mean of the risk-free rate approximated by the U.S. 1-month T-bill rate and the unconditional monthly volatility of the S&P 500 index returns over our sample period.

We parametrize preferences using standard values in the literature. Specifically, following Tversky and Kahneman (1992), we use a loss aversion parameter of $\lambda = 2.25$ and, to parametrize the probability weighting function in equation (2.4), we use $c_1 = c_2 = 0.65$. We

[9] Formally, this risk-neutral variance is given by the integrated variance of continuous-time returns over a given time period. We work in discrete time. If the continuous-time returns are uncorrelated over time, this integrated variance equals the variance of the return over a discrete time period.

[10] We use a discrete approximation of the lognormal distribution, taking a grid of 500 potential outcomes for the one-month-ahead equity return ranging from -50% to 150%.

use a benchmark coefficient of relative risk aversion of $\gamma = 1$, which implies log utility for the CRRA component of the total utility function of the investor. Because the overall risk aversion of the investor is governed by both the CRRA component and the gain-loss component of equation (2.6), our model with $\gamma = 1$ is less restrictive compared to a pure CRRA model.

Finally, we need to make an assumption on the parameter b which governs the relative importance of the CRRA and gain-loss utility components. In the absence of clear guidance from the existing literature, we assume $b \simeq 0.65$, which, in combination with the other parameters in the benchmark specification, makes CPT's contribution to the value function equal to 50%.[11] We later show that our main results on the variance premium are robust to different values for b.

Based on these inputs, we can calculate the model-implied variance premium as follows. First, we solve equation (2.9) for the return on equity $E\left(R^E\right)$. Second, once $E\left(R^E\right)$ is fixed, we use the pricing kernel given by equation (2.9) to price the cross-section of options. Finally, we use the relation derived in Proposition 4 to compute the variance premium from the option prices calculated in the previous step.[12]

For the benchmark parameters in this section, we thus obtain an annualized value for the variance premium of $162.4\%^2$. This is remarkably close to the actual value of $157.4\%^2$ in the data. The CPT model can thus generate a variance premium similar to empirically observed levels for a plausible set of input parameters.

3.3 Comparative Statics

Since a pure CRRA model yields a zero variance premium, it follows that the good fit of the CPT model documented in the previous section must be due to the CPT component in the investor's total utility function. In this section, we use comparative statics to show which CPT parameters drive the model's ability to generate a variance premium.

[11] The contribution of CPT to the total utility function is calculated as $\frac{bCPT(X_T)}{\Psi(W_T,X_T)}$, see equation (2.6).

[12] Equivalently, one could also derive the VP by computing the risk neutral probabilities and then using equation (2.11).

We start with the scale parameter b, which governs the weight of the CPT component in the investor's utility function. Panel A of Figure 3 shows model-implied variance premiums for values of b ranging from zero, in which case the CPT contribution to the investor's utility is also zero, to infinity, in which case the CPT contribution approaches 100%. The black line in Panel A shows that varying b induces economically significant variation in the variance premium. As b approaches zero, the VP approaches zero. The VP increases monotonically with b and reaches a level above 300 for $b \to \infty$, which is about twice the level of the variance premium observed in the data.

The red line in Panel A of Figure 3 provides a first indication about the deeper drivers of CPT's ability to fit the variance premium. The red line is constructed identically to the black line, with the difference being that probability weighting is set to zero ($c_1 = c_2 = 1$), such that the decision weights used by the investor coincide with the actual probabilities. Once we switch probability weighting off, the model completely loses its ability to generate a substantial VP. This finding indicates that it is not loss aversion which induces the VP, but instead probability weighting.

Panel B of Figure 3 explores the importance of probability weighting in greater detail. In that panel, we fix the scaling parameter b at its benchmark value of 0.65, and vary probability weighting from 0.4 (strong weighting) to 1 (no weighting). We again find that varying the degree of probability weighting induces substantial changes in the model-implied VP, with values ranging from about $0\%^2$(for the case of no probability weighting) to $400\%^2$ (for the case of strong weighting). These results show that probability weighting is the central driver of the VP in this model.

The mechanism underlying the effect of probability weighting on the variance premium is straightforward. In the presence of probability weighting, the state prices of extreme outcomes increase as the investor attaches a higher decision weight to these states. Intuitively, investors with probability weighting find insurance and lottery tickets attractive, and are thus willing to pay higher prices for OTM put and call options. Equivalently, buying strangles is very attractive for a CPT investor, because straddles provide both lottery and insurance. By Proposition 4, higher prices for strangles, which are the result of stronger probability weighting, imply a positive variance premium.

Contrasting the strong effect of probability weighting, Panel C shows that the variance premium hardly moves with the loss aversion parameter. For the benchmark set of parameters, we find a weakly positive relation between λ and the variance premium. On the one hand, loss-averse agents are willing to pay a higher price to hedge themselves against the risk of extreme outcomes. On the other hand, as equation (2.9) shows, more loss aversion increases the equilibrium expected return on the risky asset, which makes extremely positive returns more likely and negative returns less likely. This, in turn, makes calls more expensive and puts less expensive. The overall effect of loss aversion on the variance premium will thus depend on which effect dominates. Our results in Panel C show that, qualitatively, the first effect dominates. Quantitatively, however, loss aversion has only a minor effect on the variance premium.

3.4 GMM Estimation

We now provide more formal evidence on the ability of our CPT model to match the data. Specifically, we use GMM to estimate optimal preference parameters $\theta = [\gamma, \lambda, b, c_1, c_2]$, under different sets of parameter restrictions.

3.4.1 GMM Specification

We estimate optimal preference parameters θ using the following set of moment conditions:

$$g(\theta) = \begin{bmatrix} \mu_E(\theta) - T^{-1}\sum_1^T R_t^E \\ VP(\theta) - T^{-1}\sum_1^T VP_t \\ \mu_{p,1}(\theta) - T^{-1}\sum_1^T R_t^{p,1} \\ .. \\ \mu_{p,13}(\theta) - T^{-1}\sum_1^T R_t^{p,13} \\ \mu_{c,1}(\theta) - T^{-1}\sum_1^T R_t^{c,1} \\ .... \\ \mu_{c,13}(\theta) - T^{-1}\sum_1^T R_t^{c,13} \end{bmatrix},$$

where $\mu_E(.)$, $\mu_p(.)$, and $\mu_c(.)$ are the model-implied equilibrium expected returns for equity, calls, and puts (indexed by their strike prices), respectively, and $VP(.)$ is the model-implied variance premium, all matched to their empirical counterparts in our sample. Then, we find $\widehat{\theta}$ that minimizes the GMM goal function $VF = g(\theta)'Wg(\theta)$, where W is the weighting matrix for the moment conditions. The weighting matrix gives an equal weight of $1/4$ to the equity, call, put, and variance premium moments. We calculate Newey-West corrected standard errors for the estimated parameters to account for potential autocorrelation and heteroskedasticity in the residuals due to overlapping returns as we construct holding-to-maturity returns each day. Thus, we can do inference on the estimated parameters by exploiting that

$$\widehat{\theta} \sim N(\theta, V/T),$$

where $V = [G'WG]^{-1}G'WSWG[G'WG]^{-1}$, $G = \delta g/\delta\theta'$, and S is the standard Newey-West-corrected covariance matrix.

3.4.2 GMM Baseline Results

Table 2 presents our baseline GMM estimates. As a point of reference, specification (1) sets $b = 0$ and $\gamma = 1$ which means our model collapses to the pure CRRA case with log utility.

Not surprisingly, the statistics shown in Table 2 reflect the inability of the CRRA model to fit the option data. The average put option return estimated in this case is -6.2%, while the first column shows that the return in the data is -24.7%. The model does even worse on calls. While the model predicts a positive 7.1% average call return, the data yield a negative -12.7%. The blue x-es in Figure 1, which we have already discussed in the introduction, show the model-implied returns from specification (1) for each of our 26 options.

The variance premium is zero in specification (1), as it should be for the CRRA model. Finally, the CRRA model slightly undershoots the equity premium, but note that we do not model consumption, so fitting the equity premium is less of an achievement than it otherwise would be. The real test in our setting is pricing the cross-section of option returns, and the CRRA model fails this test.

Because the VP is guaranteed to be zero in a CRRA model regardless of the coefficient of risk aversion, using alternative values of γ would not meaningfully alter any conclusions. We will thus restrict γ to be equal to one in the remainder of the paper.[13] This assumption has two advantages. First, by virtue of Proposition 2, the results we derive for a one-period CPT setting remain valid for a multi-period version of the model in which the investor is allowed to adjust portfolio weights each period. Second, because γ captures the curvature of the CRRA part of the investor's utility function, and because λ captures the concavity of the gain-loss part of the utility function, identifying γ and λ separately is empirically challenging. By restricting $\gamma = 1$, we can get better estimates of loss aversion λ.

Specification (2) estimates the CPT model. We start by restricting probability weighting to be the same for losses and gains ($c_1 = c_2 = c$), a restriction we relax below. The CPT model matches the variance premium almost perfectly with a value of $157.28\%^2$, relative to $157.38\%^2$ in the data. At the same time, it also fits the underlying option returns very well. For example, the average call return is -13.3% in the model, compared to -12.7% in the data, and the average put return in the model is -25.4%, compared to -24.7% in the data. The

[13]To fit the equity premium $\gamma = 1$ is also a reasonable assumption. This is because we find that a pure CRRA model with a value of $\gamma = 1.22$ exactly fits the equity premium in our sample.

red +es in Figure 1 show the model-implied returns from specification (2) for each of our 26 options and confirm that the model is successful at pricing all options across the spectrum of strike prices. Finally, comparing the value of the GMM goal function VF, which can be thought of as a weighted average of the error for each of the moment conditions, shows that the CPT model is orders of magnitude more accurate than the CRRA model ($VF = 1.68$ and $VF = 6,846.5$, respectively).

Next we discuss the parameter estimates. We estimate loss aversion and probability distortion in specification (2), while we fix the scale parameter b so that the contribution of CPT to the utility function is 50%. The key result is that probability weighting, c, is very precisely estimated at 0.67. This number is interesting for two reasons. First, the p-value implies that the estimated 0.67 is significantly different from the benchmark case of no probability weighting ($c = 1$). We can thus formally reject the null hypothesis of no probability weighting. We show, in Subsection 3.4.3, that this conclusion is essentially independent of the value we pick for b. Second, the degree of probability distortion of 0.67 is remarkably close to the benchmark value originally reported by Tversky and Kahneman (1992), as well as to prominent values used in the literature (for example, Barberis and Huang (2008) use 0.65).

When estimating specification (2), we fixed the scale parameter b. The reason is that it is empirically hard to jointly identify b and c with sufficient precision, even though the two parameters are conceptually distinct. Intuitively, the problem arises because "too little" probability weighting can be compensated for by giving a higher weight to the CPT part in the utility function. Conversely, even extreme levels of probability weighting do not matter much if the CPT component enters the investor's total utility function with only little weight. In Appendix C, we document this identification issue more formally using Gentzkow and Shapiro (2015) sensitivities. We caution that, because we have fixed the scale parameter, our benchmark estimate for c is not a fully independent estimate of probability weighting. But we nevertheless argue that these estimates constitute strong evidence suggesting that probability weighting can help us understand observed option prices. To the extent that a CPT contribution of 50% is considered a reasonable number, the 0.67 estimate suggests that levels of probability weighting measured from the lab are externally valid for actual option investors. Conversely, to the extend that the lab estimates of probability weighting are considered valid starting points for examining field data, the results in specification (2) show that assuming a 50% CPT contribution would make the model consistent with the data.

3.4.3 GMM Extensions and Robustness

In Table 3, we present results for alternative assumptions about the CPT contribution, which is governed by the parameter b. In specifications (2) to (4), we fix b at a low value of 0.3, which implies a CPT contribution of about 18%; a value of 5, which implies a CPT contribution of 85%; and a very high value of 100, which implies a CPT contribution of 99%.

There are several insights. First, the loss aversion estimate is largely insensitive to changes in the CPT contribution. Combined with the fact that we assume $\gamma = 1$, this implies that overall risk aversion, and therefore the model-implied equity premium, remains remarkably constant.

Second, the probability weighting estimate varies across specifications in a way that is expected. If the contribution of CPT to the utility function is high, less probability weighting is needed to fit options well, and vice versa, for low CPT contributions, substantial degrees of probability weighting are needed to fit the data. Remarkably, even as the CPT contribution becomes very high, the estimated probability weighting parameter does not approach the no-probability-weighting value of 1, but, rather, approaches a value of about 0.8. Estimated probability weighting parameters in Table 3 range between 0.49 and 0.79, which is not unreasonable given estimates in the related decision science literature (e.g., Camerer and Ho (1994) estimate a value of 0.56, and Wu and Gonzalez (1996) estimate 0.71 [check Stott (2006) Table 5]). Echoing our earlier argument, the estimates in Table 3 imply that for arguably plausible levels of the CPT contribution (18 to 99%), the CPT model yields probability weighting estimates that are in the same region as estimates found in lab studies.

Third, and importantly, across all specifications, we can statistically reject the hypothesis of no probability weighting. Hence, while the value we pick for b has an influence of the degree of probability weighting, it does not change the central conclusion that probability weighting is the central ingredient to fitting the data. Finally, specifications (2) to (4) show that the CPT model yields an excellent fit of the variance premium and the underlying options in all cases, with only minor variation between the different models. In sum, from specifications (2) to (4), we conclude that our central findings are robust for reasonable values of b.

Specification (5) in Table 3 jointly estimates b, λ, and c. Based on the point estimates of these parameters, our previous conclusions remain completely unchanged. The GMM goal function indicates that the model now fits the data even better, with only minimal remaining pricing errors for both options and the variance premium. The model yields an estimated value for b of 0.44, which implies a CPT contribution of about 27%. The associated estimate for the probability weighting parameter is $c = 0.56$, and thus identical to the value estimated by Camerer and Ho (1994) in the lab. While these results are reassuring, one caveat is that it is very hard for the model to cleanly identify b and c simultaneously with high degrees of confidence. Consistent with the evidence from the Gentzkow and Shapiro (2015) sensitivities in Appendix C, the p-values are above standard levels.

In a final set of tests, in Table 4 we show that our results are not specific to lognormal equity index returns. First, we use a normal distribution of returns instead of a lognormal distribution, where the volatility is again matched to the equity return data. Specification (2) shows that we obtain very similar GMM estimates for the preference parameters and a very similar fit to the variance premium and option returns. In another specification, we consider the class of skewed student-t distributions, which are flexible enough to match the skewness and kurtosis observed in actual equity return data via a skewness parameter ξ and a kurtosis parameter ν.[14] Specification (3) shows that, despite allowing for negative skewness

[14]The normal distribution is a special case of the skewed-t distribution with $\xi = 1$ and $\nu = \infty$. In Appendix D, we describe in detail how we estimate the parameters of this distribution using Maximum Likelihood on

and fat tails in the equity return distribution, we find a level of probability distortion that is only slightly lower than in our benchmark case—a higher $c = 0.74$, compared to $c = 0.67$ in the benchmark model. This estimate is still statistically different from the no distortion case ($c = 1$) at any reasonable confidence level. The model-implied equity and variance premiums are very similar to those in the benchmark CPT specification, and the model prices calls and puts well. The GMM goal function has a low value of 2.57. Hence, we conclude that, while assuming a lognormal distribution for equity returns is a natural starting point for our analysis, nothing of substance hinges upon this assumption.

3.4.4 Distortion on the Downside versus Distortion on the Upside: $c_1 \neq c_2$

We now relax the assumption that $c_1 = c_2$ and allow for differential levels of probability weighting for losses (the "downside") and gains (the "upside"). These results are shown in Table 5, specification (2). We find that the model fits the data now only marginally better, which implies that $c_1 = c_2$ is not particularly restrictive. In terms of probability weighting estimates, the results show that the model-implied weighting on the downside becomes more pronounced, while the weighting on the upside becomes less pronounced than in the benchmark case $c_1 = c_2 = 0.67$. However, quantitatively, at $c_1 = 0.64$ and $c_2 = 0.70$, the optimal values are very close to the benchmark. Importantly, both parameters remain significantly different from the no-probability-weighting case at any conventional significance level.

Specifications (3) and (4) of Table 5 present results for the case of weighting probabilities only on the downside, or only on the upside, respectively. Both models perform well in terms of fitting the variance premium. However, the results also show that this good fit of the VP masks severe problems of the model when fitting the cross-section of options. In specification (3), when only losses are weighted, the model is far off the mark for call option returns (3.96% for the average OTM call versus 21.0% in the data). In specification (4), when only

the equity return data. We find that the estimated return distribution exhibits negative skewness ($\xi < 1$) and substantial kurtosis ($\nu = 6.33$).

gains are weighted, we see the opposite pattern, namely that the model now undershoots for put returns (24.86% for the average OTM put versus 36.08% in the data). This pattern is intuitive. By switching probability weighting off entirely, we are, effectively, pricing options as we would under the CRRA model. Hence, the restricted CPT model that only weights probabilities on the downside (upside) inherits the CRRA model's poor fit for calls (puts). Figure 4 documents this result for the individual options in our data. Distorting only on the downside does reasonably well for puts (the x-es in Panel B). For calls, however, the relation between model-implied call returns is convex and upward-sloping for deep OTM calls, which is opposite to the concave and downward-sloping relation observed in the data (black squares in Panel A). The analogous pattern emerges for weighting probabilities only on the upside.

The reason why option returns implied by our model can be strongly off even when we provide a good fit for the VP is that the VP is a weighted sum of expected option returns across strikes (equation (2.13)). Because the VP is a weighted sum, the model can still produce a good fit for the VP if the fitting errors for expected option returns cancel out across all options. These results suggest that comparing models based on their ability to fit the VP alone may not be a very informative test because many such models can fit the VP only at the expense of generating counterfactual option returns for at least a subset of the options. This reinforces our belief that implementing the stricter test of pricing the cross-section of options is much more informative about which models researchers should rely on than tests that focus on the VP alone. The GMM goal functions reflect this as well, as they are substantially higher for specifications (3) and (4), relative to specification (2)—a value of 54.80 and 39.85 versus a value of 1.36.

The results in Table 5 also provide additional insight into the workings of the model. What is notable is the fact that both specifications (3) and (4) feature rather extreme degrees of probability weighing. Because the VP can be thought of as a portfolio of strangles, and because strangles generate a positive return when there is demand for both lottery and insurance, the absence of a preference for lottery-like payoffs (induced by weighting on the upside) requires an extreme desire for insurance to still induce high enough strangle returns, which implies an extreme degree of probability weighting in the model. In the absence of a desire for insurance, an analogous pattern emerges.

Combined, the results in this section show that distorting probabilities in both tails of the equity return distribution is important. Distorting only one tail yields substantial pricing errors for individual options. The ability to price both puts and calls using only one underlying driver, probability weighting, is a particular advantage of the CPT model. In contrast, prominent alternative models which can also generate a variance premium, like rare disaster models or long-run risk models, are not well suited to pricing calls well, since they focus on the downside only.

4 Time-Varying Variance Premium

In the previous section, we showed that the CPT model can match the unconditional level of the variance premium as well as the cross-section of option returns when allowing for probability distortion. We now investigate whether the CPT model can also capture the substantial movement of the variance premium over time. We focus on time variation in three key parameters: equity return volatility, probability distortion, and loss aversion. We first explore the impact of the level of equity volatility using a comparative statics exercise. We then extend the GMM method introduced in Section 3.4 to a time-varying setting.

By introducing time-varying volatility we deviate from the assumption of iid equity returns that we used up to this point. However, Proposition 2 shows that we can still use the one-period setup even when returns are not iid, with the only difference that all expectations are conditional upon the volatility level at the beginning of the period.

4.1 Comparative Statics: The Impact of Volatility

In the unconditional analysis in Subsection 3.4, we set volatility equal to the average historical one-month volatility of the S&P 500. But, in the data, volatility varies substantially over time. Because volatility governs the thickness of the tails of the investor's wealth distribution, volatility is closely linked to the model-implied variance risk premium.

Figure 5 shows that the model-implied variance premium depends strongly on the level of equity volatility, and increases whenever volatility increases. The model-implied variance premium nearly triples when monthly equity volatility is set at 10% (about 35% annually) rather than at the benchmark level of 5.18% (about 18% annually), and drops substantially for very low levels of volatility. This pattern is intuitive, as higher volatility leads to more extreme outcomes, which are overweighted by CPT investors, and thus leads to an even higher model-implied variance premium. The figure also shows that high volatility alone is not sufficient to generate a variance premium. Only when coupled with probability weighting does higher volatility lead to a higher variance premium.

4.2 Time-Varying CPT Setting

Volatility and probability weighting govern how likely extreme events are and how much they are overweighted, respectively. Figures 3 and 5 show that both parameters play a key role in the CPT model's ability to generate a variance premium. We now explore whether allowing for time-variation in volatility and probability weighting enables the CPT model to also match the time-series pattern of the variance premium. Motivated by Barberis et al. (2001), we also allow for time variation in loss aversion.

We start by calculating the conditional monthly equity market variance, $h_t = V_t\left[r_{t+1}\right]$, as the integrated variance based on a simple GARCH(1,1) model estimated on daily returns over the 1996-2011 period.[15] The estimated GARCH(1,1) specification is

$$h_{t+1} = 1.38e - 06 + 0.91 h_t + 0.08 \left(r_t - \mu \right)^2, \tag{4.1}$$

where r_t represents daily equity returns. All estimated parameters are statistically significant at the 1% confidence level. Consistent with previous evidence, we find the daily conditional variance process to be highly persistent and stationary. To calculate the conditional variance for the next month, we simply take the sum of the model-implied predictions of the variance for all days over the next month ("integrated variance").

To model time variation in loss aversion, we closely follow Barberis et al. (2001) who posit that loss aversion varies with the recent performance of the representative agent's portfolio. Specifically, loss aversion depends on a state variable z_t as follows:

$$\lambda_t = \hat{\lambda} + \kappa_\lambda (z_t - \bar{z}_t), \tag{4.2}$$

where $\hat{\lambda}$ is the unconditional estimate the loss aversion parameter, $\bar{z}_t$ is the average of z_t, and κ_λ is the sensitivity of loss aversion to recently realized gains and losses as measured by z_t. Following BHS, we assume that the representative agent compares the current price of the risky asset with a benchmark level, where the benchmark level of the price responds sluggishly to changes in the value of equity. The sluggishness is defined in BHS as follows: "when the stock price moves up by a lot, the benchmark level also moves up, but by less. Conversely, if the stock price falls sharply, the benchmark level does not adjust downwards by as much." Formally:

$$z_{t+1} = \eta (z_t \frac{\overline{R^E}}{R^E_{t+1}}) + (1 - \eta), \tag{4.3}$$

[15] Although we believe that the integrated variance from a GARCH process is a good approximation of the actual variance, in unreported results (available on request), we show that the results in this section are robust to alternative conditional variance methods, such as an exponentially weighted moving average or, even more simply, last month's realized variance.

where $\overline{R^E}$ is a fixed parameter calibrated to guarantee that half of the time the agent has prior gains ($R_t^E > \overline{R^E}$) and the rest of the time she has prior losses ($R_t^E < \overline{R^E}$). The parameter η measures the degree of sluggishness, which can be interpreted as the agent's memory horizon. We assume $\eta = 0.5$, which implies that the half life of the representative agent's memory is 1 month.

In the absence of clear guidance from the existing literature, we model time variation in probability weighting in close analogy to time variation in loss aversion. Specifically, we posit:

$$c_t = \hat{c} + \kappa_c(z_t - \bar{z}_t), \tag{4.4}$$

where $\hat{c}$ is the unconditional estimate of probability distortion, and κ_c is the sensitivity of probability distortion to recently realized gains and losses as measured by z_t. We consider models that restrict the level of probability weighting to be the same on the downside and on the upside as well as a more flexible model in which probability weights for gains and losses can have different sensitivities to the state variable z_t.

4.3 Estimation Method and Results

We extend the moments in the GMM procedure in Section 3.4 to a time-varying setting as follows. At each point in time t, we compare the model-implied equity and variance premiums to empirical estimates of the time-t conditional equity and variance premiums. The conditional variance premium at any time t is estimated as the difference between the square of the VIX at time t and the time-t expectation of the realized variance, as explained in Section 3.4. To obtain an estimate for the conditional equity premium, we use two methods. In the first, the Price-Dividend (PD)-based method, we perform a predictive regression of monthly S&P 500 returns on the PD ratio and use the fitted values of this regression to obtain the empirically observed conditional equity premium at each point in time. In the second method, the PD-and-VP-based method, we add the variance risk premium to the predictive regression of S&P 500 returns.[16] We use the same two methods for modeling the conditional expectations of the option returns. To reduce the computational burden implied by the time-varying setting, we use monthly, instead of daily data, and consider only 10 deltas for the options—5 for the puts and 5 for the calls—instead of 26. We thus obtain a panel of errors by comparing, at each point in time (month), the model-implied and empirically observed values for the equity premium, variance premium, and option returns. The moment conditions are thus given by:

$$g_t(\kappa_c, \kappa_\lambda) = \begin{bmatrix} \mu_{E,t}\left(\kappa_c, \kappa_\lambda\right) - E_t(R^E_{t+1}) \\ VP_t\left(\kappa_c, \kappa_\lambda\right) - E_t(VP_{t+1}) \\ \mu_{p,1,t}(\kappa_c, \kappa_\lambda) - E_t(R^{p,1}_{t+1}) \\ .. \\ \mu_{p,5,t}(\kappa_c, \kappa_\lambda) - E_t(R^{p,13}_{t+1}) \\ \mu_{c,1,t}(\kappa_c, \kappa_\lambda) - E_t(R^{c,1}_{t+1}) \\ .. \\ \mu_{c,5,t}(\kappa_c, \kappa_\lambda) - E_t(R^{c,13}_{t+1}) \end{bmatrix}.$$

All parameters in the model are fixed at their unconditional GMM estimates except for κ_c and κ_λ, which we estimate by minimizing the sum of the squared errors in the above moment conditions. We sum across assets and over time using the same weighting of the equity premium, variance premium, and put and call returns as for the moment conditions defined in Section 3.4. To avoid numerical problems, we use a grid of values for κ_c and κ_λ to identify the values that minimize the weighted sum of the squared errors for each specification.

Table 6 presents the results for alternative specifications of the time-varying CPT setting. For each specification, we report parameter estimates as well as the value function divided

[16]The model in Bollerslev et al. (2009) implies that the variance premium has predictive power for equity returns, an implication for which they find empirical evidence.

by the number of months and the model's fit for the equity premium, the variance premium, and an aggregate for put and for call returns.

Our point of departure is a restricted model in which only the equity return volatility is time varying, while the CPT parameters are fixed at their unconditional estimates. That is, we set $\kappa_c = \kappa_\lambda = 0$. This model isolates the impact of time-variation in volatility on the fit of the model. Specifications (1) and (2) in Table 6 present estimation results for this simple model using the PD and PD-and-VP methods for calculating "observed" expected returns, respectively, and panels B and C of Figure 6 compare the model-implied with the observed equity premium with the observed one for the PD and the PD-and-VP methods, respectively. Panel A of Figure 6 shows that this simplest version of the conditional CPT model fits the level and dynamics of the variance premium remarkably well, although it seems to overestimate the variance premium in late 2008, around the collapse of Lehman Brothers, because the equity return volatility is unusually elevated in this period. The model also provides a reasonable fit of the equity premium, although it misses some of its larger movements, especially when observed equity premia are generated using the PD-and-VP-based equity premium.

The model also fits the time variation in expected option returns well. Specifications (1) and (2) in Table 6 show that the mean absolute error for the option returns (over time and across strikes) is 10.63% and 8.51% for calls and puts, respectively, for the PD-based equity premium, and 10.98% and 8.63% for calls and puts, respectively, for the PD-and-VP-based equity premium. In unreported results, we show that the residuals for option returns are of a similar magnitude irrespective of the degree of moneyness. However, the model seems to systemically underestimate the expected return of puts (and overestimate that of calls) in late 2008, precisely when we also overestimate the variance premium. In sum, when we keep all preference parameters constant over time, a CPT model with time-varying equity volatility generates a remarkably good fit of the time-series variation in the variance premium.

Turning to the setting with time-varying loss aversion (specification (3) in Table 6), we find that our estimate of κ_λ is positive, which implies a positive sensitivity of loss aversion to recent losses. Our estimate thus supports one of the key assumptions in BHS, namely that loss aversion increases after losses. The estimate for κ_λ generates a moderate range in loss aversion, from about 1 to 1.65. However, this variation has a minor effect on the fit to the equity and variance premiums and to the option returns moments. This small effect is not surprising, as the variance premium and option returns are not very sensitive to the degree of loss aversion to begin with (see Section 3.3).

A setting with time-varying probability distortion (specification (4)) yields a negative estimate of κ_c, which implies that the probability distortion increases (lower c) following losses (high z). The negative sensitivity of c to recent realized gains and losses is in line with the model of Jin (2015), which predicts that after a crash, investors are likely to overestimate the probability of another crash, leading to higher OTM put prices and a higher variance premium. However, the economic magnitude of the estimate of κ_c is small. Given that z_t ranges from 0.95 to 1.20 in our sample, with a sample average very close to 1, $\kappa_c = -0.04$ implies that the distortion parameter ranges between 0.66 and 0.67 over the sample period. Not surprisingly, the fit of the model with time variation in the distortion is very close to the fit of the model with constant parameters.

A setting in which both λ and c are time varying (specification (5)) yields a very similar sensitivity of loss aversion and a slightly higher sensitivity of probability distortion to recent losses. However, the improvement in the fit with respect to the previous specifications is rather small.

Finally, a setting with asymmetric probability distortion in the left and right tails yields different sensitivities to recent gains and losses. In particular, the sensitivity of c_1 is negative, while that of c_2 is positive, which suggests that, following recent losses, agents distort more the left tail of the distribution and distort less (closer to 1) the right tail. This is again in line with the model of Jin (2015). Moreover, the range of variation of the distortion parameter is much wider than for the models with symmetric distortion—c_1 ranges between 0.54 and 0.67, while c_2 ranges between 0.66 and 0.87.

Overall, the main implication of the time-varying setting is that even a model with *constant* CPT parameters can generate substantial variation in the variance premium once we take into account the interplay between volatility and probability weighting. This is a very attractive feature of the model, as it means that we can make progress explaining the time-series pattern of the variance premium without needing to use any additional degrees of freedom for making CPT parameters time dependent. All we need is a positive degree of probability weighting and time-variation in volatility observed in the actual data.

5 Conclusions

This paper investigates the potential of prospect theory to capture both the level and dynamics of the variance premium. We extend the representative investor model of Barberis et al. (2001), where the investor's preferences are the sum of a CRRA utility function and a prospect theory value function, by incorporating probability weighting. Probability weighting is an integral part of cumulative prospect theory (CPT, Tversky and Kahneman (1992)). The central finding in our paper is that a CPT model with probability weighting can generate a variance premium similar to the values observed in the data for plausible parametrizations. In our benchmark specification, when we estimate the preference parameters from our theoretical model using GMM on S&P 500 equity and options data from 1996 to 2010, we obtain an estimate for the probability distortion parameter of 0.67, which is remarkably close to standard values used in the literature.

We show that a sufficient condition for fitting the variance premium is to correctly price the cross-section of options on the equity index. While the standard CRRA-lognormal model fails this stricter test, we show the CPT model is very successful with probability weighting, but not without it. An advantage of the CPT model is that it can provide a unifying explanation for two well-known option pricing puzzles—the low returns on OTM puts and the low returns on OTM calls—by incorporating only one additional modeling ingredient:

probability weighting.

We then extend the static benchmark setting and explore the model's ability to generate time variation in the variance premium. We find that the dynamic version of our model performs remarkably well once we allow for time-varying volatility, even when probability weighting is fixed. A second main insight from our paper is therefore that combining probability weighting (overweighting of extreme returns) and time-variation in volatility (the presence of extreme returns) yields a parsimonious model that matches key aspects of the time-series behavior of the variance premium.

Our findings have a number of potentially important implications for future research. In particular, relative to the standard CRRA-lognormal model, which cannot generate a variance premium, we show that one can explain the variance premium by keeping returns iid and changing only the representative agent's preference structure. This is in contrast to much of the existing literature, which changes properties of the return generating process. In contrast to approaches that rely on dynamics, the CPT model can generate a variance premium even in a one shot game with known probabilities. It is therefore possible that some of the variance premium is driven by a fundamentally different mechanism than suggested by the existing literature.

References

Bakshi, G., and N. Kapadia, 2003, Delta-hedged gains and the negative market volatility risk premium, *Review of Financial Studies* 16, 527–566.

Bakshi, G., and D. Madam, 2006, A theory of volatility spreads, *Management Science* 52, 1945–1956.

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

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

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

Barberis, 2013, Thirty years of prospect theory in economics: A review and assessment, *Journal of Economic Perspectives* 27, 173–196.

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

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

Barberis, N., and W. Xiong, 2009, What drives the disposition effect? an analysis of a long-standing preference-based explanation, *Journal of Finance* 64, 751–84.

Barberis, N., and W. Xiong, 2012, Realization utility, *Journal of Financial Economics* 104, 251–271.

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

Bekaert, G., and E. Engstrom, 2016, Asset return dynamics under bad environment-good environment fundamentals, *Journal of Political Economy, Forthcoming* .

Benartzi, S., and R. Thaler, 1995, Myopic loss aversion and the equity premium puzzle 110,, *Quarterly Journal of Economics* 110, 75–92.

Bernard, Carole, and Mario Ghossoub, 2010, Static portfolio choice under cumulative prospect theory, *Mathematics and Financial Economics* 2, 277–306.

Bollerslev, T., M. Gibson, and H. Zhou, 2011, Dynamic estimation of volatility risk premiums and investor risk aversion from option-implied and realized volatilities, *Journal of Econometrics* 160, 235–245.

Bollerslev, T., G. Tauchen, and H. Zhou, 2009, Expected stock returns and variance risk premiums, *Review of Financial Studies* 22, 4463–4492.

Bondarenko, O., 2003a, Statistical arbitrage and securities prices, *Review of Financial Studies* 16, 875–919.

Bondarenko, O., 2003b, Why are puts so expensive?, *Unpublished working paper, University of Illinois, Chicago* .

Boyer, Brian H., and Keith Vorking, 2014, Stock options as lotteries, *Journal of Finance* 69, 1485–1527.

Brennan, M. J., 1979, The pricing of contingent claims in discrete time models, *Journal of Finance* 34, 53–68.

Camerer, C. F., and T. Ho, 1994, Violations of the betweenness axiom and nonlinearity in probability, *Journal of Risk and Uncertainty* 8, 167–196.

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

Carr, P., and L. Wu, 2009, Variance risk premiums, *Review of Financial Studies* 22, 1311–1341.

Carr, Peter, and Dilip Madan, 2001, Optimal positioning in derivative securities, *Quantitative Finance* 1.

Cochrane, John, 2005, *Asset Pricing* (Princeton University Press).

Constantinides, George M., Jens Carsten Jackwerth, and Alexi Savov, 2013, The puzzle of index option returns, *Review of Asset Pricing Studies* 3, 229–257.

Coval, J., and T. Shumway, 2001, Expected option returns, *Journal of Finance* 56, 983–1009.

De Giorgi, E., and S. Legg, 2012, Dynamic portfolio choice and asset pricing with narrow framing and probability weighting, *Journal of Economic Dynamics and Control* 36, 951–972.

Drechsler, I., and A. Yaron, 2011, What's vol got to do with it?, *Review of Financial Studies* 24, 1–45.

Drechsler, Itamar, 2013, Time-varying fear, and asset prices, *Journal of Finance* 68, 1843–1889.

Driessen, J., and P. Maenhout, 2007, An empirical portfolio perspective on option pricing anomalies, *Review of Finance* 11, 561–603.

Du, Du, 2011, General equilibrium pricing of options with habit formation and event risks, *Journal of Financial Econometrics* 99, 400–426.

Epper, Thomas, and Helga Fehr-Duda, 2012, Probability and risk: Foundations and economic implications of probability-dependent risk preferences, *Annual Review of Economics* 4, 567–593.

Eraker, B., 2013, The performance of model based option trading strategies., *Review of Derivatives Research* 16, 1–23.

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

Gentzkow, Matthew, and Jesse M. Shapiro, 2015, Measuring the sensitivity of parameter estimates to sample statistics, Working Paper.

Green, Clifton, and Byoung-Hyoun Hwang, 2012, initial public offerings as lotteries: Skewness preference and first-day returns, *Management Science* 58(2), 432–444.

Ilmanen, Antti, 2012, Do financial markets reward buying or selling insurance and lottery tickets?, *Financial Analysts Journal* 68, 26–36.

Ingersoll, J., 2008, Non-monotonicity of the tversky-kahneman probability-weighting function: a cautionary note, *European Financial Management* 14, 385–390.

Jiang, G., and Y. Tian, 2005, Model-free implied volatility and its information content, *Review of Financial Studies* 18, 1305–1342.

Jin, Lawrence J., 2015, A speculative asset pricing model of financial instability, *Working Paper* .

Jones, C., 2006, A nonlinear factor analysis of s&p 500 index option returns, *Journal of Finance* 61, 2325–2363.

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

Kőszegi, Botond, and Matthew Rabin, 2006, A model of reference-dependent preferences, *The Quarterly Journal of Economics* 121, 1133–1165.

Kliger, Doron, and Ori Levy, 2009, Theories of choice under risk: Insights from financial markets, *Journal of Economic Behavior and Organization* .

Lambert, P., S. Laurent, and D. Veredas, 2012, Testing conditional asymmetry. a residual-based approach, *Journal of Economic Dynamics and Control* 36, 1229–1247.

Liu, Jun, Jun Pan, and Tan Wang, 2005, An equilibrium model of rare-event premia and its implication for option smirks, *Review of Financial Studies* 18.

Londono, J.M., 2014, The variance risk premium around the world, *Working Paper, Federal Reserve Board* .

Markowitz, Harry, 1952, Portfolio selection, *Journal of Finance* 7, 77–91.

Merton, Robert C., 1969, Lifetime portfolio selection under uncertainty: The continuous-time case, *Review of Economics and Statistics* 51, 247–257.

Polkovnichenko, Valery, and Feng Zhao, 2013, Probability weighting functions implied in option prices, *Journal of Financial Economics* 107, 580–609.

Quiggin, J., 1982, A theory of anticipated utility, *Journal of Economic Behavior and Organization* 3, 323–43.

Rieger, Marc Oliver, and Mei Wang, 2006, Cumulative prospect theory and the st.petersburg paradox, *Economic Theory* 28, 665–679.

Rubinstein, Mark, 1976, The valuation of uncertain income streams and the pricing of options, *The Bell Journal of Economics* 7, 407–425.

Samuelson, Paul A., 1969, Lifetime portfolio selection by dynamic stochastic programming, *The Review of Economics and Statistics* 51, 239–246.

Samuelson, Paul A., and Robert C. Merton, 1969, A complete model of warrant pricing that maximizes utility, *Industrial Management Review* 10, 17–46.

Santa-Clara, Pedro, and Alessio Saretto, 2009, Option strategies: Good deals and margin calls, *Journal of Financial Markets* 12, 391–417.

Schreindorfer, David, 2014, Tails, fears, and equilibrium option prices, *Working* .

Shaliastovich, Ivan, 2015, Learning, confidence, and option prices, *Working Paper* .

Stott, H. P., 2006, Cumulative prospect theory's functional menagerie, *The Journal of Risk and Uncertainty* 32, 101–130.

Tversky, A., and D. Kahneman, 1992, Advances in prospect theory, cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297–323.

Wu, George, and Richard Gonzalez, 1996, Curvature of the probability weighting function, *Management science* 42, 1676–1690.

Table 1: **Summary Statistics for Put and Call Option Returns**

This table shows the summary statistics for returns on holding S&P 500 index options with 30 calendar days to maturity for puts and calls with various strike levels as measured by their Black-Scholes delta. Data are collected from OptionMetrics, which provides daily interpolated option prices for various deltas and maturities. The sample period runs from January 1996 to October 2010, with a daily frequency and overlapping 30-day returns. The t-statistic for the average return is calculated using Newey-West with 30 lags.

Call option returns

Option delta	0.200	0.250	0.300	0.350	0.400	0.450	0.500	0.550	0.600	0.650	0.700	0.750	0.800
Average return	-0.288	-0.233	-0.190	-0.152	-0.122	-0.099	-0.078	-0.060	-0.046	-0.036	-0.027	-0.020	-0.011
t-statistic average	-2.640	-2.306	-2.017	-1.728	-1.485	-1.283	-1.082	-0.897	-0.749	-0.634	-0.518	-0.410	-0.257
Return standard deviation	2.036	1.819	1.644	1.499	1.375	1.266	1.165	1.073	0.987	0.905	0.825	0.746	0.662

Put option returns

Option delta	-0.800	-0.750	-0.700	-0.650	-0.600	-0.550	-0.500	-0.450	-0.400	-0.350	-0.300	-0.250	-0.200
Average return	-0.098	-0.116	-0.133	-0.151	-0.172	-0.194	-0.218	-0.244	-0.277	-0.320	-0.372	-0.442	-0.532
t-statistic average	-1.750	-1.884	-2.009	-2.144	-2.292	-2.456	-2.620	-2.802	-3.040	-3.370	-3.775	-4.337	-5.102
Return standard deviation	0.851	0.938	1.016	1.090	1.163	1.236	1.310	1.386	1.462	1.542	1.624	1.707	1.784

Table 2: **Unconditional Setting, GMM Estimated Parameters**

This table reports GMM estimates for the degree of loss aversion, λ, and probability distortion, c ($c_1 = c_2$), for alternative specifications. In all specifications, the reference level, W_{Ref}, is assumed equal to the risk-free rate (0.25%) and the the level of risk aversion, γ, is fixed at 1. For the benchmark CPT specification (specification (2)), the scale, b, is fixed at 1.03, which is the scale that yields a 50% contribution of the CPT value function to the utility function. The p-values to test the hypotheses $\lambda = 1$ and $c = 1$ are reported in brackets. The p-values are calculated using Newey-West standard errors. For each specification, we report CPT's contribution to the utility function, which is calculated as $\frac{bCPT(X_T)}{\Psi(W_T,X_T)}$ (see equation (2.6)). We also compare the observed variance and equity risk premiums, the average model-implied call and put returns, and the average model-implied OTM call and put returns with those implied by each alternative specification. The overall fit is summarized by the value function (VF) evaluated at the optimum. Finally, we report the level of risk aversion (RA) that would be needed in the no-CPT case (b=0) to obtain the model-implied equity premium. These parameters are estimated applying GMM to returns on the S&P 500 and on S&P 500 call and put options with different strikes and 30-days to maturity and to the variance premium. The weighting matrix gives an equal weight of 1/4 to the equity, call, put, and variance premium moments. The sample period runs from January 1996 to October 2010.

	Observed	(1) CRRA Log-Utility	(2) Benchmark CPT
γ (risk aversion)		1	1
		-	-
b (scale)		0	1.03
λ (loss aversion)			1.16
			[0.38]
c (distortion losses)			0.67
			[0.00]
CPT contribution (%)		0	50.00
Variance Premium (%^{2})	157.38	0	157.28
Equity Premium (%)	0.33	0.26	0.28
Average call return (%)	-12.67	7.06	-13.34
Average put return (%)	-24.70	-6.17	-25.44
Average OTM call return (%)	-19.61	8.44	-21.00
Average OTM put return (%)	-36.08	-7.69	-35.94
VF		6846.48	1.68
RA without CPT		1	1.10

Table 3: **Unconditional Setting, GMM Estimated Parameters, Alternative Scales**

This table reports GMM estimates for the degree of loss aversion, λ, and probability distortion, c ($c_1 = c_2$), for alternative values of b, the scale parameter controlling the contribution of the CPT component. Specification (5) reports the estimates for a specification in which the scale is also estimated. In all specifications, the level of risk aversion, γ, is fixed at 1 and the reference level, W_{Ref}, is assumed equal to the risk-free rate (0.25). For each specification, we also report CPT's contribution to the total utility function, the model-implied variance and equity risk premiums, the average model-implied call and put returns, the average model-implied OTM call and put returns, the value function (VF) evaluated at the optimum, and the level of risk aversion (RA) that would be needed in the no-CPT case to obtain the model-implied equity premium. These parameters are estimated applying GMM to returns on the S&P 500 and on S&P 500 call and put options with different strikes and 30-days to maturity and to the variance premium. The weighting matrix gives an equal weight of 1/4 to the equity, call, put, and variance premium moments. The sample period runs from January 1996 to October 2010. The p-values to test the hypotheses $\lambda = 1$, $b = 0$, and $c = 1$ are reported in brackets. These p-values are calculated using Newey-West standard errors.

		(1)	(2)	(3)	(4)	(5)
	Observed	Benchmark	Low b		High b	Estimate b
γ (risk aversion)		1	1	1	1	1
b (scale)		1.03	0.3	5	100	0.44
		-	-	-	-	[0.25]
λ (loss aversion)		1.16	1.19	1.16	1.15	1.18
		[0.38]	[0.43]	[0.29]	[0.25]	[0.44]
c (distortion)		0.67	0.49	0.76	0.79	0.56
		[0.00]	[0.00]	[0.00]	[0.00]	[0.11]
CPT contribution (%)		50.00	18.46	83.66	99.04	27.28
Implied Variance Premium (%^{2})	157.38	157.28	157.57	157.16	157.12	157.46
Implied Equity Premium (%)	0.33	0.28	0.27	0.28	0.28	0.27
Average call return (%)	-12.67	-13.34	-11.42	-13.98	-14.17	-12.30
Average put return (%)	-24.70	-25.44	-23.86	-25.96	-26.09	-24.56
Average OTM call return (%)	-19.61	-21.00	-18.20	-22.05	-22.37	-19.45
Average OTM put return (%)	-36.08	-35.94	-34.23	-36.33	-36.38	-35.03
VF		1.68	1.41	2.94	3.44	0.97
RA without CPT		1.10	1.07	1.12	1.12	1.08

Table 4: **Unconditional Setting, GMM Estimated Parameters, Alternative Equity Return Distributions**

This table reports GMM estimates for the degree of loss aversion, λ, and probability distortion, c ($c_1 = c_2$), for equity return distributions alternative to the benchmark lognormal distribution. In particular, we compare the benchmark lognormal distribution with a normal distribution (specification (2)) and with a skewed-t distribution (specification (3)). For the skewed-t distribution, the parameters driving the skewness and kurtosis—ξ and υ, respectively—are calibrated to match the observed monthly equity returns (see Appendix B). In all specifications, the level of risk aversion, γ, is fixed at 1, the scale, b, is fixed at 1.03, the scale that yields a 50% contribution of the CPT value function for the benchmark specification (specification (1)), and the reference level, W_{Ref}, is assumed equal to the risk-free rate (0.25%). For each specification, we also report CPT's contribution to the utility function, the model-implied variance and equity risk premiums, the average model-implied call and put returns, the average model-implied OTM call and put returns, the value function (VF) evaluated at the optimum, and the level of risk aversion (RA) that would be needed in the no-CPT case to obtain the model-implied equity premium. These parameters are estimated applying GMM to returns on the S&P 500 and on S&P 500 call and put options with different strikes and 30-days to maturity and to the variance premium. The weighting matrix gives an equal weight of 1/4 to the equity, call, put, and variance premium moments. The sample period runs from January 1996 to October 2010. The p-values to test the hypotheses $\lambda = 1$ and $c = 1$ are reported in brackets. These p-values are calculated using Newey-West standard errors.

	(1) Benchmark	(2) Normal	(3) Skewed-t
γ (risk aversion)	1	1	1
b (scale)	1.03	1.03	1.03
ξ			0.84
υ			6.33
λ (loss aversion)	1.16	1.12	1.11
	[0.38]	[0.44]	[0.45]
c (distortion)	0.67	0.68	0.74
	[0.00]	[0.00]	[0.00]
CPT contribution (%)	50.00	49.51	50.54
Implied Variance Premium (%^{2})	157.28	157.32	157.56
Implied Equity Premium (%)	0.28	0.55	0.29
Average put return (%)	-25.44	-25.26	-23.91
Average call return (%)	-13.34	-13.15	-11.79
Average OTM call return (%)	-21.00	-20.67	-19.17
Average OTM put return (%)	-35.94	-35.62	-33.72
VF	1.68	1.63	2.57
RA without CPT	1.10	2.12	1.14

Table 5: **Unconditional Setting, GMM Estimated Parameters, Asymmetric Distortion**

This table reports GMM estimates for the degree of loss aversion, λ, and probability distortion for gains and losses, c_1 and c_2, respectively, for alternative specifications. In all specifications, the reference level, W_{Ref}, is assumed equal to the risk-free rate (0.25), the level of risk aversion, γ, is fixed at 1, and the scale, b, is fixed at 1.03, which is the scale that yields a 50% contribution of the CPT value function to the total utility function for the benchmark specification (specification (1)). The p-values to test the hypotheses $\lambda = 1$, and $c = 1$ are reported in brackets. The p-values are calculated using Newey-West standard errors. For each specification, we also report CPT's contribution to the utility function, the model-implied variance and equity risk premiums, the average model-implied call and put returns, the average model-implied OTM call and put returns, the value function (VF) evaluated at the optimum, and the level of risk aversion (RA) that would be needed in the no-CPT case to obtain the model-implied equity premium. These parameters are estimated applying GMM to returns on the S&P 500 and on S&P 500 call and put options with different strikes and 30-days to maturity and to the variance premium. The weighting matrix gives an equal weight of 1/4 to the equity, call, put, and variance premium moments. The sample period runs from January 1996 to October 2010.

	(1)	(2)	(3)	(4)
	Benchmark		Asymmetric	
	CPT		CPT	
γ (risk aversion)	1	1	1	1
b (scale)	1.03	1.03	1.03	1.03
λ (loss aversion)	1.16	1.11	1.00	1.79
	[0.38]	[0.45]	[0.50]	[0.05]
c_1 (distortion losses)	0.67	0.64	0.46	1
	[0.00]	[0.00]	[0.00]	-
c_2 (distortion gains)		0.70	1	0.47
		[0.00]	-	[0.00]
CPT contribution (%)	50.00	49.32	46.96	54.60
Implied Variance Premium ($\%^2$)	157.28	157.30	157.61	157.33
Implied Equity Premium (%)	0.28	0.26	0.50	0.42
Average call return (%)	-13.34	-13.28	-3.13	-10.99
Average put return (%)	-25.44	-25.31	-29.68	-22.50
Average OTM call return (%)	-21.00	-20.30	-3.69	-23.07
Average OTM put return (%)	-35.94	-27.79	-45.28	-24.86
VF	1.68	1.36	54.80	39.65
RA without CPT	1.10	1.03	1.93	1.63

Table 6: **Time-Varying Setting, GMM Estimated Parameters**

This table shows the estimated parameters and the model's fit for alternative specifications of the time-varying CPT setting. Specifications (1) and (2) do not allow for time variation in the CPT parameters. Specification (1) uses the PD-based method to calculate the observed equity premium, wherein the expected equity returns are calculated using an in-sample forecast from the PD ratio. In all other specifications, the observed equity premium is calculated as an in-sample forecast from the PD ratio and the variance premium (PD-and-VP-based method). For every specification, we report the value function divided by the number of months (Avg. VF). The value function is calculated as a weighted average of the squared residuals for the variance and equity premiums as well as for the 10 options considered in the time-varying setting. We also report a decomposition of the model's fit into its variance premium, equity premium, and option return components. For each component, the fit is calculated as the mean absolute error. In specifications (1) to (5), we assume $\lambda = 1.16$, and $c = 0.67$—the parameters in the benchmark unconditional setting (specification (2) in Table 2). In specification (6), we assume $\lambda = 1.11$, $c_1 = 0.64$, and $c_2 = 0.70$—the parameters in the asymmetric unconditional setting (specification (2) in Table 2). In all specifications, we assume $\gamma = 1$ (risk aversion) and $\eta = 0.5$ (agent's memory). The time-varying volatility is calculated as the integrated volatility based on a GARCH(1,1) method (see Section 4.2).

	(1)	(2)	(3)	(4)	(5)	(6)
EP Method	PD	PD and VP				
κ_λ (loss aversion)			2.57		2.60	
κ_{c1} (distortion)				-0.04	-0.12	-0.70
κ_{c2}						1.01
Avg. VF	91.07	91.10	87.32	90.03	87.27	89.62
EP fit (%)	0.39	0.46	0.51	0.54	0.51	0.53
VP fit (%^{2})	108.42	108.42	96.69	98.94	99.93	98.17
Put fit (%)	8.51	8.63	8.63	8.74	8.63	8.68
Call fit (%)	10.63	10.98	10.99	11.08	10.99	10.98

Figure 1: **Average Option Returns and Model Fit**

This figure compares the average observed 30-day return for all options in our sample with the ones fitted by our benchmark CPT representative agent model (specification (2) in Table 2) and by a CRRA model—a restricted CPT model with $c = 1$ and $b = 0$ (specification (1) in Table 2). The sample is comprised of daily S&P 500 index option returns from 1996 to 2010 for various strike prices. Of the 26 options in the sample, 13 are calls, with deltas from 0.2 to 0.8 (from OTM to ITM). The other 13 are puts, with delta from -0.8 to -0.2 (from ITM to OTM). The figure below presents results by average moneyness (K/S_0) for a given delta, separately for calls (Panel A) and puts (Panel B).

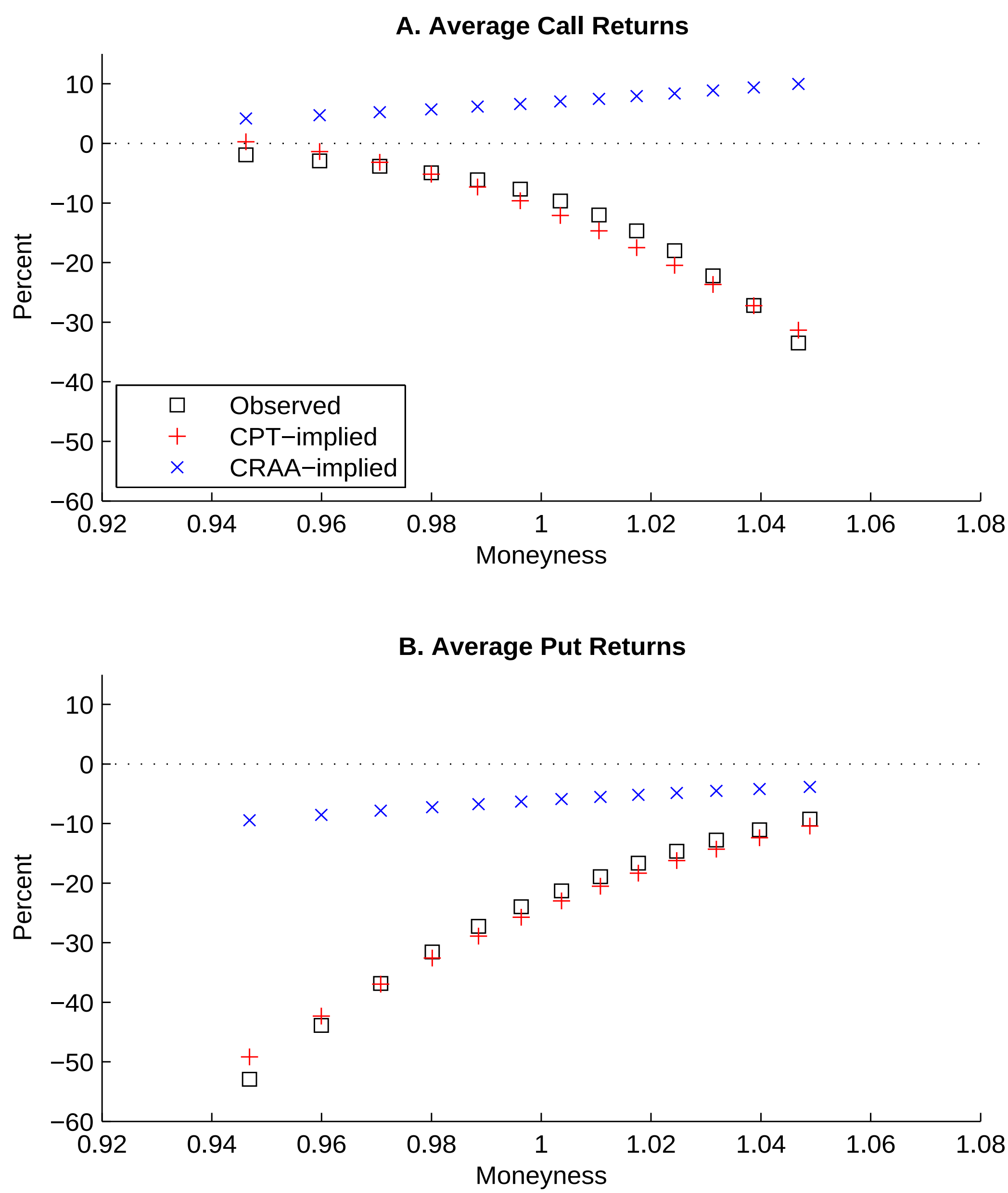


Figure 2: **Pricing Kernel**

This figure shows the CRRA and CPT components of the pricing kernel implied by our CPT setting (see equation (2.9)) as a function of market returns (R^E). To create the figure, we assume the following set of parameters. The reference level, W_{Ref}, is assumed equal to the risk-free rate (0.25%); the level of risk aversion, γ, is fixed at 1; the scale, b, is fixed at 1.03, which is the scale that yields a 50% contribution of the CPT value function to the utility function for this set of parameters; the parameter driving the probability weighting, $c_1 = c_2 = c$, is fixed at 0.67. These are the parameters obtained in the benchmark GMM estimation (see Table 2).

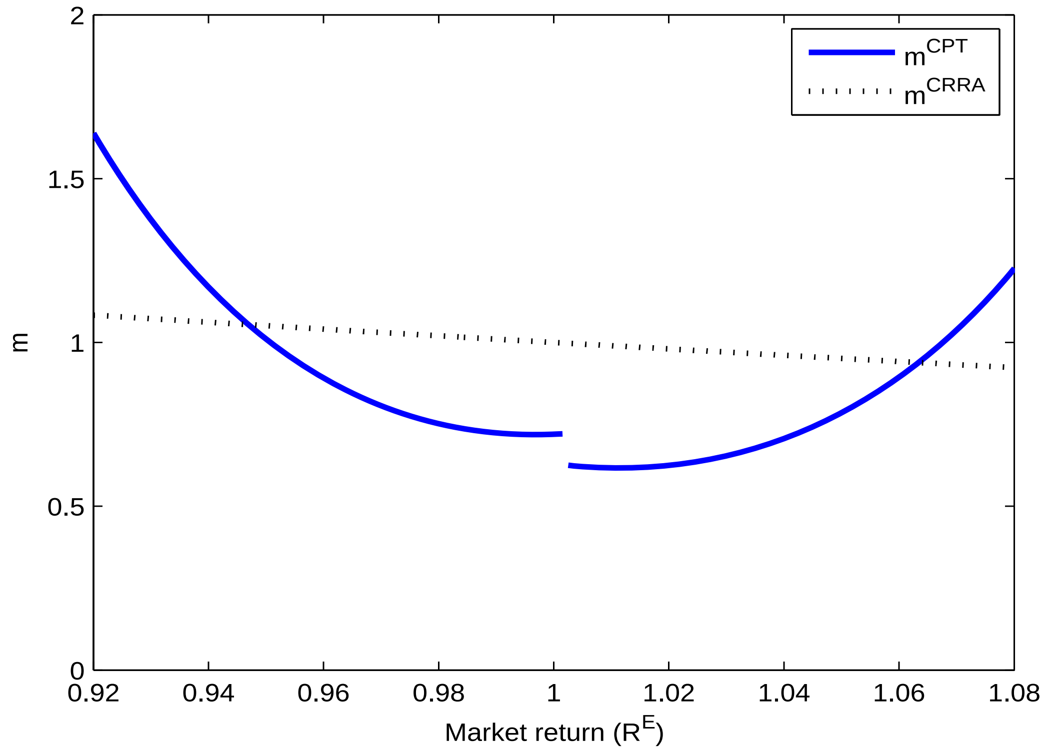


Figure 3: **CPT-Implied Variance Premium for Alternative Preference Parameter Values**

This figure shows the annualized CPT-implied variance premium for alternative values of the scale, b, probability distortion, c, and loss aversion, λ, parameters, in panels A to C, respectively. In the CPT setting, the representative agent's preference combines CRRA and CPT functions (equation (2.1)) and the probabilities are distorted, as explained in Section 2. The variance premium is defined as the difference between the risk-neutral and the expected realized variance of equity returns, as explained in Section 2.4. For the benchmark setup (the bold line), we set $c = 0.65$, $b = 0.65$, and $\lambda = 2.25$, unless otherwise indicated in the panel. We additionally assume $\gamma = 1$.

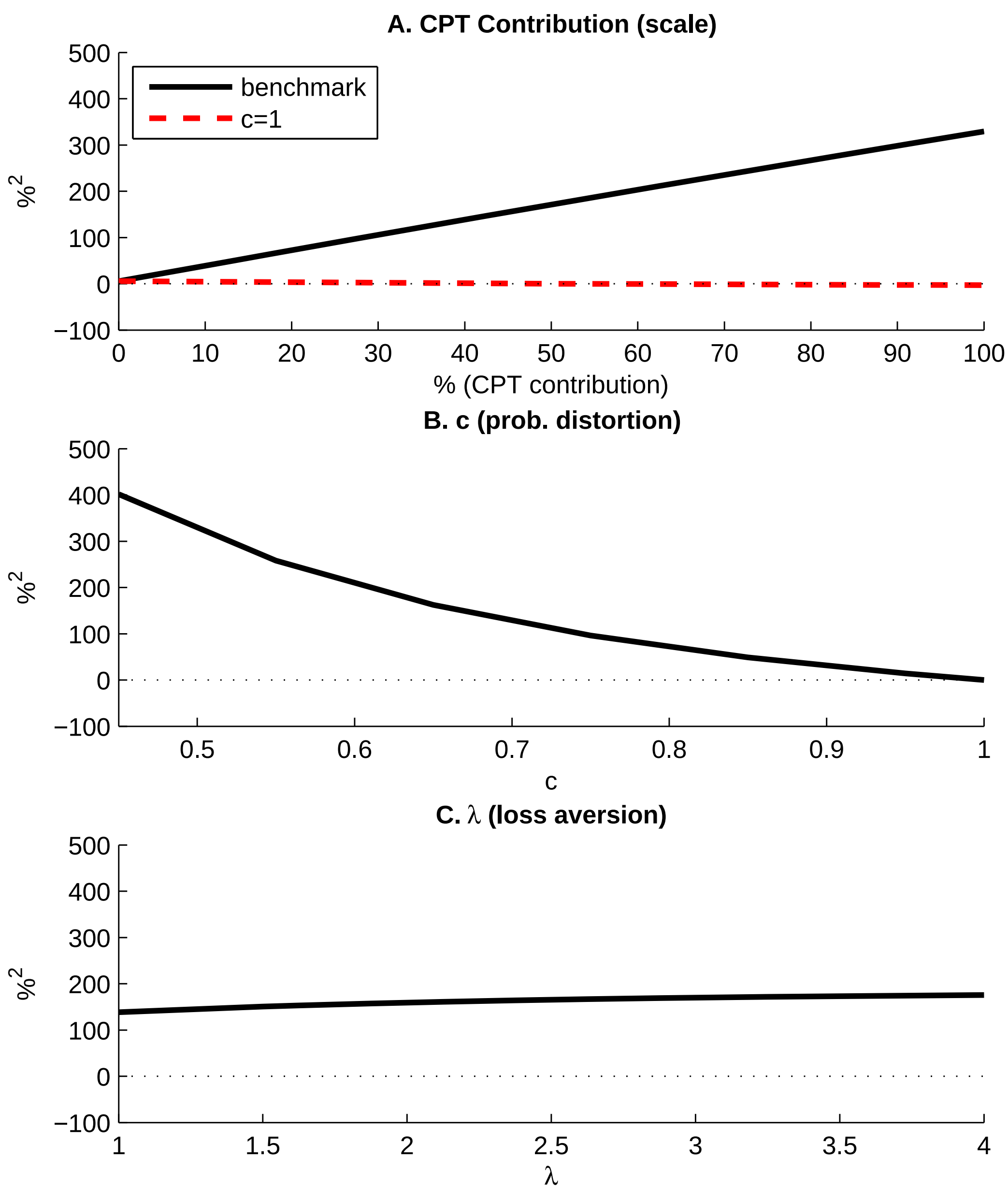


Figure 4: **Average Option Returns and Model Fit for CPT models with Asymmetric Probability Distortion**

This figure compares the average observed 30-day return for all options in our sample with the ones fitted by a CPT representative agent model with asymmetric probability distortion ($c_1 \neq c_2$), a model in which only the left tail of the distribution is distorted ($c_2 = 1$), and a model in which only the right tail of the distribution is distorted ($c_1 = 1$) (specifications (2), (3), and (4) in Table 5, respectively). Our sample includes 26 option returns grouped across strikes (or deltas). Of these 26 options, 13 are puts (panel A), with delta from -0.8 to -0.2 (from ITM to OTM). The other 13 are calls (panel B), with delta from 0.2 to 0.8 (from OTM to ITM).

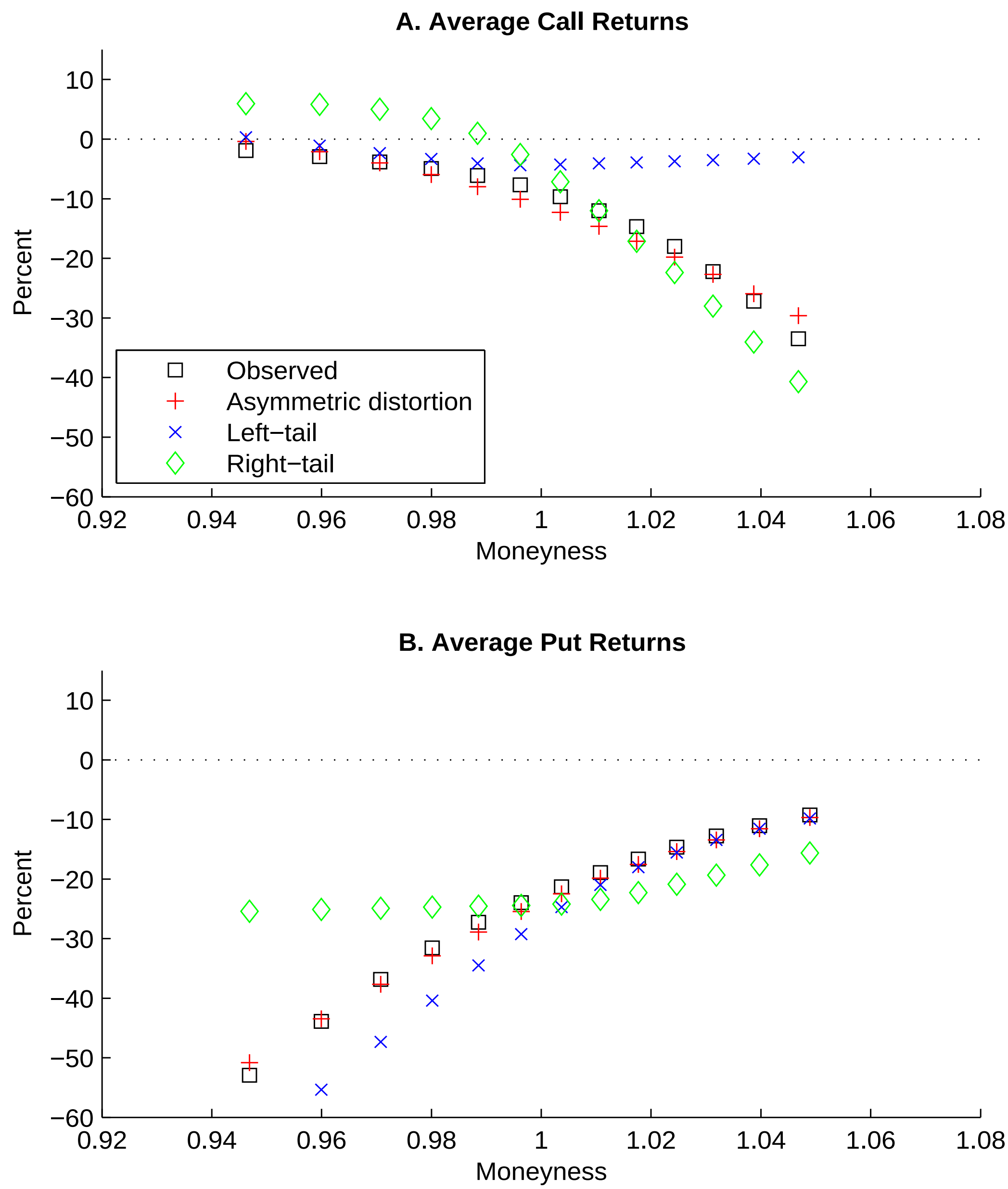


Figure 5: **CPT-Implied Variance Premium for Alternative Volatility Levels**

This figure shows the annualized CPT-implied variance premium (see Figure 3) for three different values of monthly equity volatility: 2% (low), 5.18% (benchmark), and 10% (high). Panels A to C show, respectively, the effect of volatility combined with that of the scale, b, probability distortion, c, and loss aversion, λ, for the CPT-implied variance premium. For the benchmark setup (the bold line), we set $c = 0.65$, $b = 0.65$, and $\lambda = 2.25$, unless otherwise indicated in the panel. We additionally assume $\gamma = 1$.

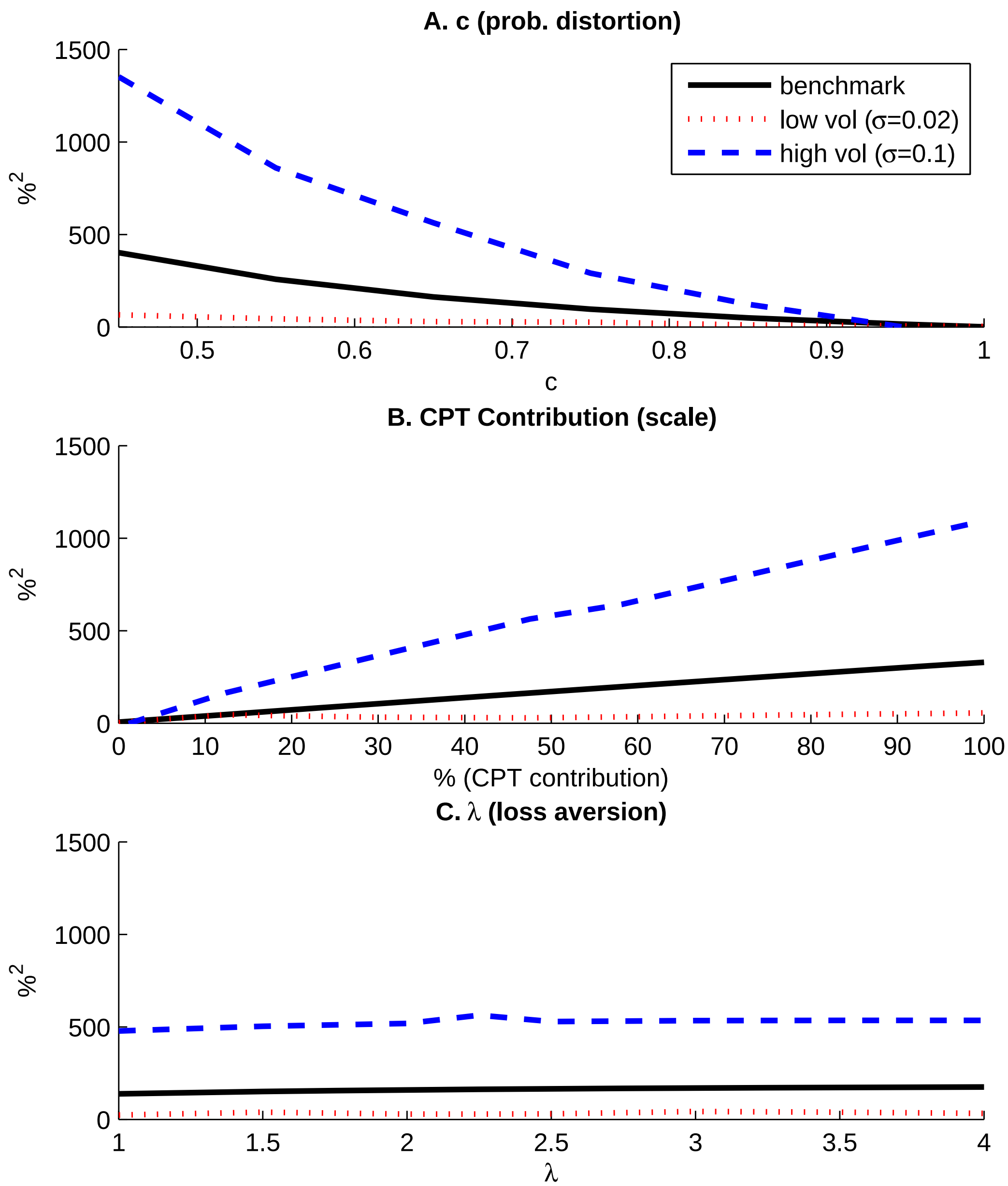
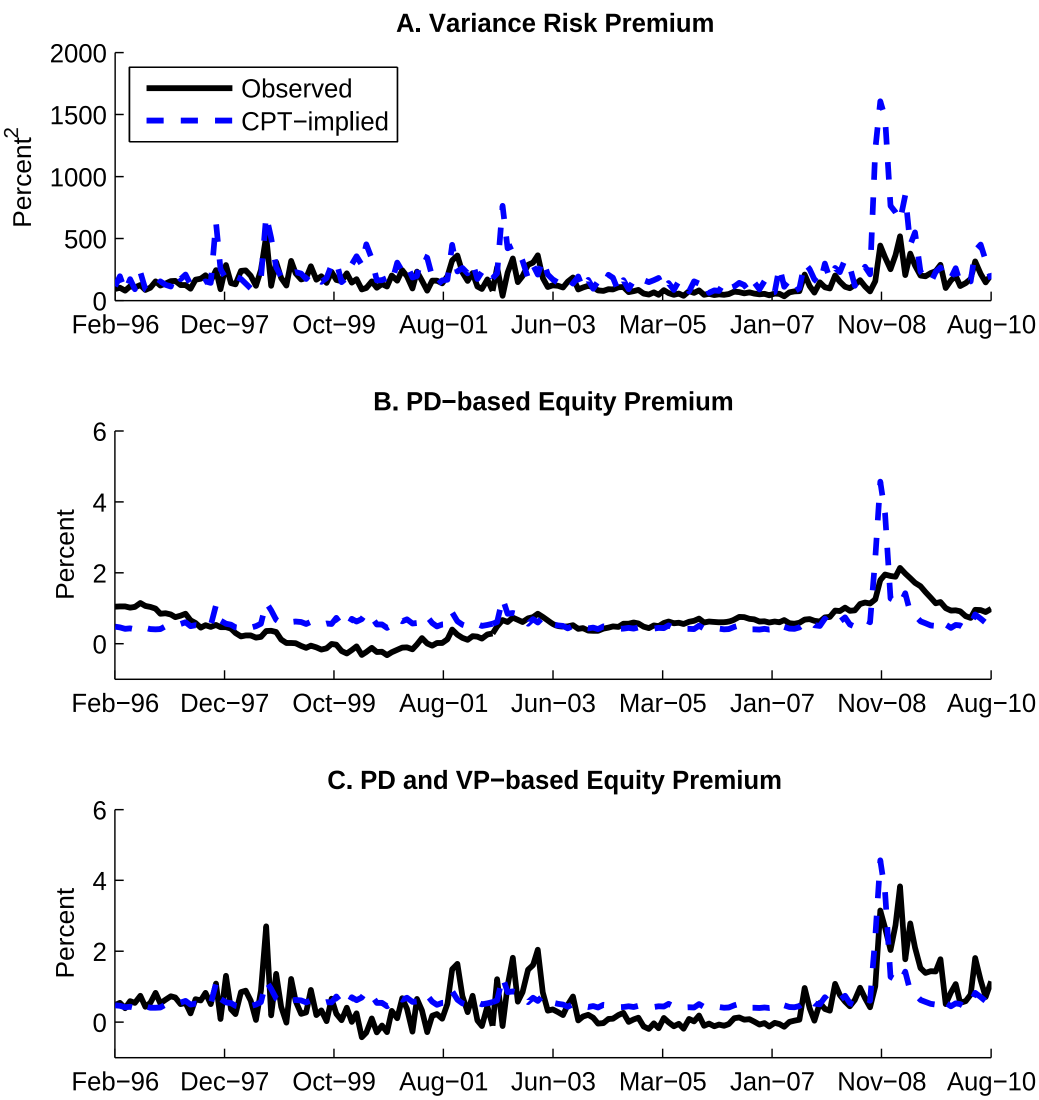


Figure 6: **Observed vs. CPT-Implied Equity and Variance Premiums**

This figure compares the observed equity and variance risk premiums with those implied by a CPT setting in which the model's CPT parameters are constant and the stock return volatility is time varying (specifications (1) and (2) in Table 6). The variance premium, in panel A, is calculated as the difference between the square of the VIX and the expected realized variance (see Section 3.4). The expected realized variance is calculated as an in-sample forecast of the stock return variance using the lagged stock return variance and the square of the VIX. The CPT-implied equity and variance risk premiums are calculated as explained in Section 2.4. In panel B, the observed equity premium is calculated as an in-sample forecast of equity returns from the price-dividend ratio (PD-based), while, in panel C, we add the variance risk premium as a predictor of equity returns (PD and VP-based).



Appendix

A Proof of Proposition 2

Without loss of generality, we focus on one risky asset (equity).[17]

The investor's problem is now given by

$$\max_{\alpha_0^E,\ldots,\alpha_{T-1}^E} V = E[U(W_T)] + \sum_{t=0}^{T-1} b_t CPT(X_{t+1})$$

where as before we follow Barberis et al. (2001) and let the weight on the prospect theory utilities be given by a multiple of current period's marginal utility of wealth, i.e., by $b_t = \hat{b} W_t^{-\gamma}$. The investor's budget constraint is

$$W_{t+1} = W_t(\alpha_t^E r_{t+1} + R^f), \quad t = 0, \ldots, T-1$$

where $r_{t+1} := R_{t+1}^E - R^f$ denotes the excess return in from period t to period $t+1$. The portfolio gain or loss each period is thus given by

$$X_{t+1} = W_{t+1} - W_t R^f = W_t \alpha_t^E r_{t+1}, \quad t = 0, \ldots, T-1.$$

Let us first consider the log-utility case ($\gamma = 1$), whose proof is simple and intuitive. In that case, the investor seeks to maximize

$$\begin{aligned} V &= E[\log((W_0 \Pi_{t=0}^{T-1}(\alpha_t^E r_{t+1} + R^f))] + \sum_{t=0}^{T-1} b_t CPT(W_t \alpha_t^E r_{t+1}) \\ &= E[\log W_0 + \sum_{t=0}^{T-1} \log(\alpha_t^E r_{t+1} + R^f)] + \sum_{t=0}^{T-1} \hat{b} W_t^{-1} CPT(W_t \alpha_t^E r_{t+1}) \end{aligned}$$

[17] Not introducing $d = 1, ..., D$ derivatives (with portfolio shares α_t^d for derivative d in period t) simplifies the notation in the following derivation significantly and does not cause any loss of economic insight.

Since the CPT functional is homogeneous of degree 1 it follows that $CPT(W_t\alpha_t^E r_{t+1}) = W_t\alpha_t^E CPT(r_{t+1})$. For this reason, the decision variables α_t^E are no longer nested within the CPT functional. It follows that

$$\begin{aligned} V &= \log W_0 + \sum_{t=0}^{T-1} E[\log(\alpha_t^E r_{t+1} + R^f)] + \sum_{t=0}^{T-1} \alpha_t^E \hat{b} CPT(r_{t+1}) \\ &= \log W_0 + \sum_{t=0}^{T-1} \Big(E[\log(\alpha_t^E r_{t+1} + R^f)] + \alpha_t^E \hat{b} CPT(r_{t+1}) \Big) . \end{aligned}$$

The objective functional is thus additively separable in the decision variables so that the FOC for each α_t^E is independent of the α_s^E, $s \neq t$ and structurally identical. In particular, the weights α_t^E do not depend on the period t itself or on the number of periods remaining – the myopia result. When returns are iid, we further have that$\alpha_0^E = \ldots = \alpha_{T-1,}^E$ so that the investor does not rebalance her portfolio.

Now consider the case of general CRRA utility, which we prove by dynamic programming. Let τ denote the number of periods that remain from current time t so that $T = t + \tau$. The value of the agent's problem at time t is given by

$$\begin{aligned} V(\tau, W_t) &= \max_{\{\alpha_s^E\}_{s=t}^{t+\tau-1}} E_t \left[\frac{W_{t+\tau}^{1-\gamma}}{1-\gamma} + \sum_{s=t}^{t+\tau-1} b_s CPT(r_{s+1}\alpha_s^E W_s) \right] \\ &= \max_{\{\alpha_s^E\}_{s=t}^{t+\tau-1}} E_t \left[\frac{W_{t+\tau}^{1-\gamma}}{1-\gamma} + \sum_{s=t}^{t+\tau-1} \hat{b}\alpha_s^E W_s^{1-\gamma} CPT(r_{s+1}) \right] \end{aligned}$$

where, similarly to the log case, we exploited the fact that the CPT preference functional is homogeneous of degree 1. The problem has become one of maximizing terminal wealth expected utility plus a sum of flow utilities, which are linear in the decision variable. In particular, letting $c_{t+1} := \hat{b}CPT(r_{t+1})$, the time-$t$ value of next period's flow utility is given by $\pi_{t+1} := c_{t+1}\alpha_t^E W_t^{1-\gamma}$. By the dynamic programming principle, the investor's problem at time t can be written as

$$V(\tau, W_t) = \max_{\alpha_t^E} E_t \left[\underbrace{\max_{\{\alpha_s^E\}_{s=t+1}^{t+\tau-1}} E_{t+1} \left[\frac{W_{t+\tau}^{1-\gamma}}{1-\gamma} + \sum_{s=t+1}^{t+\tau-1} c_{s+1} \alpha_s^E W_s^{1-\gamma} \right]}_{V(\tau-1, W_t(\alpha_t^E r_{t+1} + R_f))} + \underbrace{\alpha_t^E c_{t+1} W_t^{1-\gamma}}_{\pi_{t+1}(\alpha_t^E, W_t, c_{t+1})} \right]$$

subject to the terminal condition $V(0, W_{t+\tau}) = \frac{W_{t+\tau}^{1-\gamma}}{1-\gamma} + c_{t+\tau}\alpha_{t+\tau-1}W_{t+\tau-1}^{1-\gamma}$. This means that τ periods before the model ends, the value of the problem is determined by the problem's continuation value next period, $V(\tau-1, W_t(\alpha_t^E r_{t+1} + R^f))$, plus next periods flow utility, $\pi_{t+1}(\alpha_t^E, W_t, c_{t+1})$. The notation indicates that the continuation value is that of $\tau - 1$ periods before the end, where wealth from the previous period has changed to $W_t(\alpha_t^E r_{t+1} + R^f)$. This continuation value includes the expected flow utilities on from two periods in the future while the value of next period's flow utility is a deterministic function of the time t portfolio weight and current period wealth. Continuing,

$$\begin{aligned}
V(\tau, W_t) &= \max_{\alpha_t^E} E_t \left[\max_{\{\alpha_s^E\}_{s=t+1}^{t+\tau-1}} E_{t+1} \left[\frac{(W_t \prod_{s=t}^{t+\tau-1} (\alpha_s^E r_{s+1} + R_f))^{1-\gamma}}{1-\gamma} + \sum_{s=t+1}^{t+\tau-1} c_{s+1}\alpha_s^E W_s^{1-\gamma} \right] + \alpha_t^E c_{t+1} W_t^{1-\gamma} \right] \\
&= \max_{\alpha_t^E} E_t \Bigg[\alpha_t^E c_{t+1} W_t^{1-\gamma} + \frac{(W_t(\alpha_t^E r_{r+1} + R_f))^{1-\gamma}}{1-\gamma} \cdot \\
&\qquad \max_{\{\alpha_s^E\}_{s=t+1}^{t+\tau-1}} E_{t+1} \left[(\prod_{s=t+1}^{t+\tau-1} (\alpha_s^E r_{s+1} + R_f))^{1-\gamma} + \sum_{s=t+1}^{t+\tau-1} \frac{c_{s+1}\alpha_s^E W_s^{1-\gamma}(1-\gamma)}{W_{t+1}^{1-\gamma}} \right] \Bigg] \\
&= \max_{\alpha_t^E} W_t^{1-\gamma} E_t \Bigg[\alpha_t^E c_{t+1} + \frac{(\alpha_t^E r_{t+1} + R_f)^{1-\gamma}}{1-\gamma} \cdot \\
&\qquad \max_{\{\alpha_s^E\}_{s=t+1}^{t+\tau-1}} E_{t+1} \left[(\prod_{s=t+1}^{t+\tau-1} (\alpha_s^E r_{s+1} + R_f))^{1-\gamma} + \sum_{s=t+1}^{t+\tau-1} \frac{c_{s+1}\alpha_s^E W_s^{1-\gamma}(1-\gamma)}{W_{t+1}^{1-\gamma}} \right] \Bigg].
\end{aligned}$$

With some standard manipulations one can show that

$$\sum_{s=t+1}^{t+\tau-1} \frac{c_{s+1}\alpha_s^E W_s^{1-\gamma}(1-\gamma)}{W_{t+1}^{1-\gamma}} = (1-\gamma) \sum_{s=t+1}^{t+\tau-1} c_{s+1}\alpha_s^E \prod_{\tilde{s}=t+2}^{s} (\alpha_{\tilde{s}}^E r_{\tilde{s}+1} + R_f)^{1-\gamma},$$

where we use the convention that $\prod_{\tilde{s}=t+2}^{t+1}(x_{\tilde{s}}r_{\tilde{s}+1}+R^f)^{1-\gamma}\equiv 1$. Therefore, we obtain

$$V(\tau,W_t)=\max_{\alpha_t^E} W_t^{1-\gamma}E_t\left[\alpha_t^E c_{t+1}+\frac{(\alpha_t^E r_{t+1}+R_f)^{1-\gamma}}{1-\gamma}\cdot \right.$$
$$\left.\max_{\{\alpha_s^E\}_{s=t+1}^{t+\tau-1}} E_{t+1}\left[\prod_{s=t+1}^{t+\tau-1}(\alpha_s^E r_{s+1}+R_f)^{1-\gamma}+(1-\gamma)\sum_{s=t+1}^{t+\tau-1}c_{s+1}\alpha_s^E\prod_{\tilde{s}=t+2}^{s}(\alpha_{\tilde{s}}^E r_{\tilde{s}+1}+R_f)^{1-\gamma}\right]\right]$$

and see that the problem is homogeneous in current wealth so that, without loss of generality, we let $W_t\equiv 1$. Thus,

$$V(\tau,W_t)=\max_{\alpha_t^E} E_t\left[\underbrace{\alpha_t^E c_{t+1}}_{\pi_{t+1}}+\underbrace{\frac{(\alpha_t^E r_{t+1}+R_f)^{1-\gamma}}{1-\gamma}}_{=:u_{t+1}}\cdot \right.$$
$$\left.\underbrace{\max_{\{\alpha_s^E\}_{s=t+1}^{t+\tau-1}} E_{t+1}\left[\prod_{s=t+1}^{t+\tau-1}(\alpha_s^E r_{s+1}+R_f)^{1-\gamma}+(1-\gamma)\sum_{s=t+1}^{t+\tau-1}c_{s+1}\alpha_s^E\prod_{\tilde{s}=t+2}^{s}(\alpha_{\tilde{s}}^E r_{\tilde{s}+1}+R_f)^{1-\gamma}\right]}_{=:\psi_{t+1}(\tau-1)}\right]$$

Note that portfolio weight α_t^E depends on the investment horizon τ through the next-period term "horizon-effect term" $\psi_{t+1}(\tau-1)$, which we now study in more detail. First, observe that

$$\psi_t(\tau)=\max_{\{\alpha_s^E\}_{s=t}^{t+\tau-1}} E_{t+1}\left[\prod_{s=t}^{t+\tau-1}(\alpha_s^E r_{s+1}+R_f)^{1-\gamma}+(1-\gamma)\sum_{s=t}^{t+\tau-1}c_{s+1}\alpha_s^E\prod_{\tilde{s}=t+1}^{s}(\alpha_{\tilde{s}}^E r_{\tilde{s}+1}+R_f)^{1-\gamma}\right]\cdot\overbrace{1^{1-\gamma}}^{W_t^{1-\gamma}\equiv 1}$$

Plugging this in above leads to the recursion

$$\frac{\psi_t(\tau)}{1-\gamma}=\max_{\alpha_t^E} E_t\left[\underbrace{\alpha_t^E c_{t+1}}_{\pi_{t+1}}+\underbrace{\frac{(\alpha_t^E r_{r+1}+R_f)^{1-\gamma}}{1-\gamma}}_{u_{t+1}}\cdot \right.$$
$$\left.\underbrace{\max_{\{\alpha_s^E\}_{s=t+1}^{t+\tau-1}} E_{t+1}\left[\prod_{s=t+1}^{t+\tau-1}(\alpha_s^E r_{s+1}+R_f)^{1-\gamma}+(1-\gamma)\sum_{s=t+1}^{t+\tau-1}c_{s+1}\alpha_s^E\prod_{\tilde{s}=t+2}^{s}(x_{\tilde{s}}r_{\tilde{s}+1}+R_f)^{1-\gamma}\right]}_{\psi_{t+1}(\tau-1)}\right]$$

or likewise

$$\psi_t(\tau) = \max_{\alpha_t^E} E_t\left[\underbrace{\alpha_t^E c_{t+1}(1-\gamma)}_{\pi_{t+1}} + (\alpha_t^E r_{t+1} + R_f)^{1-\gamma} \cdot \psi_{t+1}(\tau-1)\right].$$

We see that when $\gamma \neq 1$ the optimal portfolio weight α_t^E depends on the investor's horizon τ, i.e., is non-myopic. We now show that even when returns are independent and/ or identically distributed, investment is non-myopic. When the r_{t+1} are independent, because $\psi_{t+1}(\tau-1)$ is free of α_t^E we have

$$\psi_t(\tau) = \max_{\alpha_t^E} \alpha_t^E c_{t+1}(1-\gamma) + E_t\left[\psi_{t+1}(\tau-1)\right] \cdot \max_{\alpha_t^E} E_t\left[(\alpha_t^E r_{t+1} + R_f)^{1-\gamma}\right].$$

The maximizing α_t^E satisfies

$$c_{t+1}(1-\gamma) + E_t\left[\psi_{t+1}(\tau-1)\right] \cdot E_t\left[(\alpha_t^E r_{t+1} + R_f)^{-\gamma} r\right] = 0.$$

We see now that if, and only if, $\gamma = 1$, we can divide by $E_t[\psi_{t+1}(\tau-1)]$ so that each α_t^E satisfies

$$E_t\left[(\alpha_t^E r_{t+1} + R_f)^{-\gamma} r_{t+1}\right] = 0.$$

This observation is consistent with the previous result that α_t^E is time- (and thus horizon-) independent in the log-utility case, $\gamma = 1$.[18] For all other CRRA utilities, we have shown that the α_t^E are time- and horizon-dependent.

[18]Note, however, that the proof for general γ does not nest the proof for $\gamma = 1$, because it involves expressions with $1-\gamma$ in the denominator.

B Proof of Proposition 4

Without loss of generality, we work in a setting with an infinite number of outcomes for the equity payoff $x^E > 0$ at time T. As shown by, for example, Bakshi et al. (2003) and Carr and Madan (2001), any continuous and twice-differentiable payoff function $G(x^E)$ can be written as

$$G(x^E) = G(\bar{x}) + G'(\bar{x})(x^E - \bar{x}) + \int_{\bar{x}}^{\infty} G''(K)\max(x^E - K, 0)dK + \int_{0}^{\bar{x}} G''(K)\max(K - x^E, 0)dK \tag{A.1}$$

for any value of $\bar{x}$. This equation shows that any payoff function can be replicated by an equity position and a position in an infinite number of call and put options with different strike prices. If there are no arbitrage opportunities, a risk-neutral measure exists and the risk-neutral expected value of the payoff function G can then be written as a function of call and put prices

$$E^Q\left[G(x^E)\right] = G(\bar{x}) + G'(\bar{x})(R^f s_0^E - \bar{x}) + \int_{\bar{x}}^{\infty} G''(K)R^f C(K)dK + \int_{0}^{\bar{x}} G''(K)R^f P(K)dK \tag{A.2}$$

where we use the risk-neutral pricing equations for equity, $s_0^E = \frac{1}{R^f}E^Q\left[x^E\right]$, call options with strike K, $C(K) = \frac{1}{R^f}E^Q\left[max(x^E - K, 0)\right]$, and put options with strike K, $P(K) = \frac{1}{R^f}E^Q\left[max(K - x^E, 0)\right]$. Similarly to equation (A.2) for the risk-neutral expectation, we can compute the expected value of the payoff function G under the physical measure P as

$$\begin{aligned} E^P\left[G(x^E)\right] &= G(\bar{x}) + G'(\bar{x})(E[x^E] - \bar{x}) \\ &+ \int_{\bar{x}}^{\infty} G''(K)E\left[max(x^E - K, 0)\right]dK + \int_{0}^{\bar{x}} G''(K)E\left[max(K - x^E, 0)\right]dK \end{aligned} \tag{A.3}$$

Now choose $G(x^E) = \left(ln\left(\frac{x_i^E}{s_0^E}\right)\right)^2$, and by subtracting (A.2) from (A.4) it directly follows that

$$E^Q\left[\left(ln\left(\frac{x_i^E}{s_0^E}\right)\right)^2\right] - E^P\left[\left(ln\left(\frac{x_i^E}{s_0^E}\right)\right)^2\right] = -G'(\bar{x})s_0^E\Big(\frac{E[x^E]}{s_0^E} - R^f\Big) \tag{A.4}$$
$$-\int_{\bar{x}}^{\infty} G''(K)C(K)\Big(\frac{E[max(x^E-K,0)]}{C(K)} - R^f\Big)dK$$
$$-\int_{0}^{\bar{x}} G''(K)P(K)\Big(\frac{E[max(K-x^E,0)]}{P(K)} - R^f\Big)dK$$

with $G'(x) = 2ln\left(\frac{x}{s_0^E}\right)\frac{1}{x}$ and $G''(x) = \frac{2}{x^2}\left(1 - ln\left(\frac{x}{s_0^E}\right)\right)$. We choose $\bar{x} = s_0^E$, so $G'(\bar{x}) = 0$. Defining

$$R^p(K) = E\left[max(K-x^E,0)\right]/P(K)$$

and

$$R^c(K) = E\left[max(x^E-K,0)\right]/C(K)$$

and defining weights

$$v^c(K) = -G''(K)C(K)$$

and

$$v^p(K) = -G''(K)P(K)$$

then yields the right hand side of equation (2.13) in the main text.

To see that the right hand side of equation (2.13) is an approximation to the variance premium, note that adding

$$E^P\left[ln\left(R^E\right)\right]^2 - E^Q\left[ln\left(R^E\right)\right]^2$$

to both sides of equation (A.4) above yields the variance premium exactly, as can be seen by comparing the left hand side of the resulting equation with the definition of the variance premium in equation (2.11) in the main text. This term arises due to the difference between central and non-central second moments. For our application, this term is numerically small as we focus on one-month returns: it is equal to 1.3 in our benchmark model, while empirically the variance premium is equal to 157.4. Equation (2.13) omits this term, which is why it becomes an approximation to the variance premium.

Finally, note that choosing $\bar{x} = s_0^E$ implies that equation (2.13) includes only OTM put and call options. It thus follows directly that the weights $v^c(K)$ and $v^P(K)$ are negative for any reasonable value of K. Only for deep OTM call options with strikes above $s_0^E e^1$ will the weight become positive.

C Gentzkow and Shapiro (2015) Sensitivities

To assess how well the different parameters in our CPT model are identified, we calculate the Gentzkow and Shapiro (2015) sensitivities which capture the sensitivity of the parameter estimates to the moment conditions. Formally, these are defined as the coefficients of a regression of the estimator $\hat{\theta}$ on the moment conditions $\hat{g}$. In a GMM setting, Gentzkow and Shapiro (2015) show that the sensitivities $\Lambda$are given by

$$\Lambda = -(G'WG)^{-1}G'W \tag{A.1}$$

where $G = \delta g / \delta\theta'$ and W is the GMM weighting matrix. Intuitively, the sensitivities measure how a change in a given moment condition affects the estimate of a parameter and thus capture how informative each moment condition is about a parameter. In addition, if two parameters have very similar sensitivities, it will be hard to separately identify these parameters from the given set of moment conditions. Gentzkow and Shapiro (2015) propose to standardize the sensitivities by the asymptotic variance of the estimators

$$\bar{\Lambda}_{ij} = \Lambda_{ij} \frac{Var(\hat{g}_j)}{Var(\hat{\theta}_i)}. \tag{A.2}$$

In Figure D.1 we plot these standardized sensitivities for the case where we estimate b, λ, and c (panel A), and for the benchmark case where we fix b and estimate only λ and c (panel B). We see in panel A that the sensitivities of b and λ are very similar. As discussed in Section 3.4.2 this problem arises because "too little" probability weighting can be compensated for by giving a higher weight to the CPT part in the utility function. Conversely, even extreme levels of probability weighting do not matter much if the CPT component enters the investor's total utility function with only little weight. Once we fix b, panel B shows that we can separately identify λ and c as their sensitivities differ substantially across moments. In particular, the variance premium affects the estimator for c much more than the estimator for λ, in line with the comparative statics analysis in Section 3.3.

D Alternative Stock Returns Distribution: Skewed-t Distribution

In Section 3.4.2, we investigate to what extent our baseline results are robust to using a skewed fat-tailed distribution instead of a standard lognormal distribution for stock returns. In this appendix, we first introduce the standardized skewed-t distribution (see, for example, Lambert et al. (2012) and Bauwens and Laurent (2002)). Then, we report the Maximum Likelihood estimates for the key parameters of these distribution using our sample of S&P 500 returns.

The return on the S&P 500, r_t, follows the process

$$r_t = \mu + \varepsilon_t \tag{A.1}$$

$$\varepsilon_t = \sigma \varsigma_t, \tag{A.2}$$

where the random variable ς_t is $SKST\left(0,1,\xi,v\right)$ distributed; that is, it follows a standardized skewed-t distribution with parameters $v > 2$ (the number of degrees of freedom) and $\xi > 0$ (a parameter related to skewness). The density of this function is given by

$$f\left(\varsigma_t|\xi,v\right) = \begin{cases} \frac{2}{\left(\xi+\frac{1}{\xi}\right)} sg\left[\xi\left(s\varsigma_t+m\right)|v\right] & if \quad \varsigma_t < -m/s \\ \frac{2}{\left(\xi+\frac{1}{\xi}\right)} sg\left[\xi\left(s\varsigma_t+m\right)/\xi|v\right] & if \quad \varsigma_t \geqslant -m/s \end{cases}, \tag{A.3}$$

where $g\left(.|v\right)$ is a symmetric (zero mean and unit variance) Student-t density with v degrees of freedom, denoted $x \sim ST\left(0,1,v\right)$, defined by

$$g\left(x|v\right) = \frac{\Gamma\left(\frac{v-1}{2}\right)}{\sqrt{\pi\left(v-2\right)}\Gamma\left(\frac{v}{2}\right)}\left[1+\frac{x^2}{v-2}\right]^{-(v+1)/2} \tag{A.4}$$

where $\Gamma\left(.\right)$ is an Euler's gamma function.

The constants $m = m\left(\xi,v\right)$ and $s = \sqrt{s^2\left(\xi,v\right)}$ are, respectively, the mean and standard deviation of the non-standardized skewed-t density, $SKST\left(m, s^2, \xi, v\right)$, and are defined as follows:

$$m\left(\xi, v\right) = \frac{\Gamma\left(\frac{v-1}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\Gamma\left(\frac{v}{2}\right)}\left(\xi - \frac{1}{\xi}\right) \tag{A.5}$$

$$s^2\left(\xi, v\right) = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2. \tag{A.6}$$

The parameters ξ and v are related to the distribution's skewness and kurtosis. Because ξ^2 can be shown to be equal to the ratio of the probability masses above and below the mode, the distribution has zero skewness when $\xi = 1$, negative skewness when $\xi < 1$, and positive skewness when $\xi > 1$. The fatness of tails (kurtosis) decreases with the degrees of freedom parameter, v, but converges to a skewed normal as $v \to \infty$. When $v \to \infty$ and $\xi = 1$, this distribution collapses to a standard normal distribution.

Table C.1 reports estimates and standard errors for $\theta = (\mu, \sigma, \xi, v)$, using monthly log returns on the S&P 500 over the period running from 1960 to 2011. All parameters are statistically significant at the 1% level. ξ is significantly below 1, implying negative skewness. The ratio of the probability masses above and below the mode is equal to $\xi^2 = 0.71$, implying that the degree of negative skewness is also important in economic terms. The estimated degrees of freedom parameter, v, equals 6.33, indicating that the return distribution is not only left skewed but also fat tailed.

Table C.1: **Parameters of the Skewed-t Distribution**

This table reports the estimated parameters for the Skewed-t distribution described in equations (B.1) to (B.6). The standard errors are reported in parenthesis.

	μ	σ	ξ	v
Estimate	0.0077	0.0437	0.8448	6.3301
(s.e.)	(0.0017)	(0.0019)	(0.0487)	(1.5862)

Figure D.1: **Sensitivity of CPT Parameter Estimates to Sample Moments**

This figure shows the standardized sensitivities of the CPT parameter estimates to the following sample moments: the equity premium, the variance premium, and the put and call option returns. To facilitate the interpretation of sensitivities, option returns are orthogonalized with respect to the equity and the variance premium. Panel A shows the sensitivities for the setting in which the scale is also estimated (specification #6 in Table 3), while panel B shows the sensitivities for the benchmark CPT setting (specification (1) in Table 3). Standardized sensitivities are calculated following Gentzkow and Shapiro (2015) as the sensitivity of the expected value of each parameter to a one-standard-deviation change in the realization of each moment condition.