
An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks
Author(s): Jun Liu, Jun Pan, Tan Wang
Source: *The Review of Financial Studies*, Vol. 18, No. 1 (Spring, 2005), pp. 131-164
Published by: Oxford University Press. Sponsor: The Society for Financial Studies.
Stable URL: http://www.jstor.org/stable/3598069
Accessed: 11/07/2009 13:52

An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks

Jun Liu
Anderson School at UCLA

Jun Pan
MIT Sloan School of Management, CCFR and NBER

Tan Wang
Sauder School of Business at UBC and CCFR

This article studies the asset pricing implication of imprecise knowledge about rare events. Modeling rare events as jumps in the aggregate endowment, we explicitly solve the equilibrium asset prices in a pure-exchange economy with a representative agent who is averse not only to risk but also to model uncertainty with respect to rare events. The equilibrium equity premium has three components: the diffusive- and jump-risk premiums, both driven by risk aversion; and the "rare-event premium," driven exclusively by uncertainty aversion. To disentangle the rare-event premiums from the standard risk-based premiums, we examine the equilibrium prices of options across moneyness or, equivalently, across varying sensitivities to rare events. We find that uncertainty aversion toward rare events plays an important role in explaining the pricing differentials among options across moneyness, particularly the prevalent "smirk" patterns documented in the index options market.

Sometimes, the strangest things happen and the least expected occurs. In financial markets, the mere possibility of extreme events, no matter how unlikely, could have a profound impact. One such example is the so-called "peso problem," often attributed to Milton Friedman for his comments about the Mexican peso market of the early 1970s.[1] Existing literature acknowledges the importance of rare events by adding a new type of risk

We thank Torben Andersen, David Bates, John Cox, Larry Epstein, Lars Hansen, John Heaton, Michael Johannes, Monika Piazzesi, Bryan Routledge, Jacob Sagi, Raman Uppal, Pietro Veronesi, Jiang Wang, an anonymous referee, and seminar participants at CMU, Texas Austin, MD, the 2002 NBER summer institute, the 2003 AFA meetings, the Cleveland Fed Workshop on Robustness, and UIUC for helpful comments. We are especially grateful to detailed and insightful comments from Ken Singleton (the editor). Tan Wang acknowledges financial support from the Social Sciences and Humanities Research Council of Canada. Jun Pan thanks the research support from the MIT Laboratory for Financial Engineering.
Address correspondence to: Jun Pan, MIT Sloan School of Management, Cambridge, MA 02142, or e-mail: junpan@mit.edu.

[1] Since 1954, the exchange rate between the U.S. dollar and the Mexican peso has been fixed. At the same time, the interest rate on Mexican bank deposits exceeded that on comparable U.S. bank deposits. In the presence of the fixed exchange rate, this interest rate differential might seem to be an anomaly to most people, but it was fully justified when in August 1976 the peso was allowed to float against the dollar and its value fell by 46%. See, for example, Sill (2000) for a more detailed description.

The Review of Financial Studies Vol. 18, No. 1
doi:10.1093/rfs/hhi011 Advance Access publication November 3 2004

(event risk) to traditional models, while keeping the investor's preference intact.[2] Implicitly, it is assumed that the existence of rare events affects the investor's portfolio of risks, but not their decision-making process.

This article begins with a simple yet important question: Could it be that investors treat rare events somewhat differently from common, more frequent events? Models with the added feature of rare events are easy to build but much harder to estimate with adequate precision. After all, rare events are infrequent by definition. How could we then ask our investors to have full faith in the rare-event model we build for them?

Indeed, some decisions we make just once or twice in a lifetime—leaving little room to learn from experiences, while some we make everyday. Naturally, we treat the two differently. Likewise, in financial markets we see daily fluctuations and rare events of extreme magnitudes. In dealing with the first type of risks, one might have reasonable faith in the model built by financial economists. For the second type of risks, however, one cannot help but feel a tremendous amount of uncertainty about the model. And if market participants are uncertainty averse in the sense of Knight (1921) and Ellsberg (1961), then the uncertainty about rare events will eventually find its way into financial prices in the form of a premium.

To formally investigate this possibility of "rare-event premium," we adopt an equilibrium setting with one representative agent and one perishable good. The stock in this economy is a claim to the aggregate endowment, which is affected by two types of random shocks. One is a standard diffusive component, and the other is pure jump, capturing rare events with low frequency and sudden occurrence. While the probability laws of both types of shocks can be estimated using existing data, the precision for rare events is much lower than that for normal shocks. As a result, in addition to balancing between risk and return according to the estimated probability law, the investor factors into his decision the possibility that the estimated law for the rare event may not be correct. As a result, his asset demand depends not only on the trade-off between risk and return, but also on the trade-off between uncertainty and return.

In equilibrium, which is solved in closed form, these effects show up in the total equity premium as three components: the usual risk premiums for diffusive and jump risks, and the uncertainty premium for rare events. While the first two components are generated by the investor's risk

[2] For example, in an effort to explain the equity-premium puzzle, Rietz (1988) introduces a low probability crash state to the two-state Markov-chain model used by Mehra and Prescott (1985). Naik and Lee (1990) add a jump component to the aggregate endowment in a pure-exchange economy and investigate the equilibrium property. More recently, the effect of event risk on investor's portfolio allocation with or without derivatives are examined by Liu and Pan (2003), Liu, Longstaff, and Pan (2003) and Das and Uppal (2001). Dufresne and Hugonnier (2001) study the impact of event risk on pricing and hedging of contingent claims.

aversion, the last one is linked exclusively to his uncertainty aversion toward rare events. To test these predictions of our model, however, data on equity returns alone are not sufficient. Either aversion coefficient can be adjusted to match an observed total equity premium, making it impossible to differentiate the effect of uncertainty aversion from that of risk aversion.

Our model becomes empirically more relevant as options are included in our analysis. Unlike equity, options are sensitive to rare and normal events in markedly different ways. For example, deep-out-of-the-money put options are extremely sensitive to market crashes. Options with varying degrees of moneyness therefore provide a wealth of information for us to examine the importance of uncertainty aversion to rare events. For options on the aggregate market (e.g., the S&P 500 index), two empirical facts are well documented: (1) options, including at-the-money (ATM) options, are typically priced with a premium [Jackwerth and Rubinstein (1996)]; (2) this premium is more pronounced for out-of-the-money (OTM) puts than for ATM options, generating a "smirk" pattern in the cross-sectional plot of option-implied volatility against the option's strike price [Rubinstein (1994)].

As a benchmark, we first examine the standard model without uncertainty aversion. Calibrating the model to the equity return data, we examine its prediction on options.[3] We find that this model cannot produce the level of premium that has been documented for at-the-money options. Moreover, in contrast to the pronounced "smirk" pattern documented in the empirical literature, this model generates an almost flat pattern. In other words, with risk aversion as the only source of risk premium, this model cannot reconcile the premium observed in the equity market with that in ATM options, nor can it reconcile the premium implicit in ATM options with that in OTM put options.

Here, the key observation is that moving from equity to ATM options, and then to deep-OTM put options, these securities become increasingly more sensitive to rare events. Excluding the investor's uncertainty aversion to this specific component, and relying entirely on risk aversion, one cannot simultaneously explain the market-observed premiums implicit in these securities: fitting it to one security, the model misses out on the others. Conversely, if risk aversion were the only source for the premiums implicit in options, then one had to use a risk-aversion coefficient

[3] It should be noted that our model cannot resolve the issue of "excess volatility." That is, the observed volatility of the aggregate equity market is significantly higher than that of the aggregate consumption, while in our model they are the same. In calibrating the model with or without uncertainty aversion, we face the problem of which volatility to calibrate. Since the main objective of this calibration exercise is to explore the link between the equity market and the options market, we choose to calibrate the model using information from the equity market. That is, we examine the model's implication on the options market after fitting it to the equity market.

for the rare events and another for the diffusive risk to reconcile the premiums implicit in these securities simultaneously.[4]

In comparison, the model incorporating uncertainty aversion toward rare events does a much better job in reconciling the premiums implicit in all these securities with varying degree of sensitivity to rare events. In particular, the models with uncertainty aversion can generate significant premiums for ATM options as well as pronounced "smirk" patterns for options with different degrees of moneyness.[5]

Our approach to model uncertainty falls under the general literature that accounts for imprecise knowledge about the probability distribution with respect to the fundamental risks in the economy. Among others, recent studies include Gilboa and Schmeidler (1989), Epstein and Wang (1994), Anderson, Hansen, and Sargent (2000), Chen and Epstein (2002), Hansen and Sargent (2001), Epstein and Miao (2003), Routledge and Zin (2002), Maenhout (2001), and Uppal and Wang (2003). The literature on learning provides an alternative framework to examine the effect of imprecise knowledge about the fundamentals.[6] Given that rare events are infrequent by nature, learning seems to be a less important issue in our setting. Furthermore, given that rare events are typically of high impact, thinking through worst-case scenarios seems to be a more natural reaction to uncertainty about rare events.

The robust control framework adopted in this article closely follows that of Anderson, Hansen, and Sargent (2000). In this framework, the agent deals with model uncertainty as follows. First, to protect himself against the unreliable aspects of the reference model estimated using existing data, the agent evaluates the future prospects under alternative models. Second, acknowledging the fact that the reference model is indeed the best statistical characterization of the data, he penalizes the choice of the alternative model by how far it deviates from the reference model. Our approach, however, differs from that of Anderson, Hansen, and Sargent (2000) in one important dimension.[7] Specifically, our investor is worried

[4] By introducing a crash aversion component to the standard power-utility framework, Bates (2001) recently proposes a model that can effectively provide a separate risk-aversion coefficient for jump risk, disentangling the market price of jump risk from that of diffusive risk. The economic source of such a crash aversion, however, remains to be explored.

[5] It is true that in such a model one can fit to one security using a particular risk-aversion coefficient and still have one more degree of freedom from the uncertainty-aversion coefficient to fit the other security. The empirical implication of our model, however, is not only about two securities. Instead, it applies to options across all degrees of moneyness.

[6] Among others, David and Veronesi (2000) and Yan (2000) study the impact of learning on option prices, and Comon (2000) studies learning about rare events. For learning under model uncertainty, see Epstein and Schneider (2002) and Knox (2002).

[7] Another important difference is that we provide a more general version of the distance measure between the alternative and reference models. The "relative entropy" measure adopted by Anderson, Hansen, and Sargent (2000) is a special case of our proposed measure. This extended form of distance measure is important in handling uncertainty aversion toward the jump component. Specifically, under the "relative

about model misspecifications with respect to rare events, while feeling reasonably comfortable with the diffusive component of the model. This differential treatment with respect to the nature of the risk sets our approach apart from that of Anderson, Hansen, and Sargent (2000) in terms of methodology as well as empirical implications.

Recently, there have been observations on the equivalence between a number of robust-control preferences and recursive utility [Maenhout (2001) and Skiadas (2003)]. A related issue is the economic implication of the normalization factor introduced to the robust-control framework by Maenhout (2001), which we adopt in this article. Although by introducing rare events and focusing on uncertainty aversion only to rare events, our article is no longer under the framework considered in these articles, it is nevertheless important for us to understand the real economic driving force behind our result. Relating to the equivalence result involving recursive utility, we consider an economy that is identical to ours except that, instead of uncertainty aversion, the representative agent has a continuous-time Epstein and Zin (1989) recursive utility. We derive the equilibrium pricing kernel explicitly, and show that it prices the diffusive and jump shocks in the same way as the standard power utility. In particular, the rare-event premium component, which is linked directly to rare-event uncertainty in our setting, cannot be generated by the recursive utility.[8] Relating to the economic implication of the normalization factor, we consider an example involving a general form of normalization. We show that although the specific form of normalization affects the specific solution of the problem, the fact that our main result builds on uncertainty aversion toward rare events is not affected in any qualitative fashion by the choice of normalization.

The rest of the article is organized as follows. Section 1 sets up the framework of robust control for rare events. Section 2 solves the optimal portfolio and consumption problem for an investor who exhibits aversions to both risk and uncertainty. Section 3 provides the equilibrium results. Section 4 examines the implication of rare-event uncertainty on option pricing. Section 5 concludes the article. Technical details, including proofs of all three propositions, are collected in the appendices.

---

entropy" measure, the robust control problem is not well defined for the jump case. For pure-diffusion models, however, our extended distance measure is equivalent to the "relative entropy" measure.

[8] This result also serves to strengthen our calibration exercises involving options. The recursive utility considered in our example has two free parameters: one for risk aversion and the other for elasticity of intertemporal substitution. Similarly, in our framework, the utility function also has two parameters: one for risk aversion and the other for uncertainty aversion. In this respect, we are comparing two utility functions on equal footing, although the economic motivations for the two utility functions are distinctly different. We show that the recursive utility cannot resolve the smile puzzle. The intuition is as follows. Although it has two free parameters, the standard recursive utility has one risk-aversion coefficient to price both the diffusive and rare-event risks, while the additional parameter associated with the intertemporal substitution affects the risk-free rate. In effect, it does not have the additional coefficient to control the market price of rare events separately from the market price of diffusive shocks.

1. Robust Control for Rare Events

Our setting is that of a pure exchange economy with one representative agent and one perishable consumption good [Lucas (1978)]. As usual, the economy is endowed with a stochastic flow of the consumption good. For the purpose of modeling rare events, we adopt a jump-diffusion model for the rate of endowment flow $\{Y_t, 0 \leq t \leq T\}$. Specifically, we fix a probability space $(\Omega, \mathcal{F}, P)$ and information filtration $(\mathcal{F}_t)$, and assume that Y is a Markov process in $\mathbb{R}$ solving the stochastic differential equation

$$dY_t = \mu Y_t dt + \sigma Y_t dB_t + (e^{Z_t} - 1) Y_{t-} dN_t, \tag{1}$$

where $Y_0 > 0$, B is a standard Brownian motion and N is a Poisson process. In the absence of the jump component, this endowment flow model is the standard geometric Brownian motion with constant mean growth rate $\mu \geq 0$ and constant volatility $\sigma > 0$. Jump arrivals are dictated by the Poisson process N with intensity $\lambda > 0$. Given jump arrival at time t, the jump amplitude is controlled by Z_t, which is normally distributed with mean μ_J and standard deviation σ_J. Consequently, the mean percentage jump in the endowment flow is $k = \exp(\mu_J + \sigma_J^2/2) - 1$, given jump arrival. In the spirit of robust control over worse-case scenarios, we focus our attention on undesirable event risk. Specifically, we assume $k \leq 0$. At different jump times $t \neq s$, Z_t and Z_s are independent, and all three types of random shocks B, N, and Z are assumed to be independent. This specification of aggregate endowment follows from Naik and Lee (1990). It provides the most parsimonious framework for us to incorporate both normal and rare events.[9]

We deviate from the standard approach by considering a representative agent who, in addition to being risk averse, exhibits uncertainty aversion in the sense of Knight (1921) and Ellsberg (1961). The infrequent nature of the rare events in our setting provides a reasonable motivation for such a deviation. Given his limited ability to assess the likelihood or magnitude of such events, the representative agent considers alternative models to protect himself against possible model misspecifications.

To focus on the effect of jump uncertainty, we restrict the representative agent to a prespecified set of alternative models that differ only in terms of the jump component. Letting P be the probability measure associated with the reference model [Equation (1)], the alternative model is defined by its probability measure $P(\xi)$, where $\xi_T = dP(\xi)/dP$ is its

[9] One feature not incorporated in this model is stochastic volatility. Given that our objective is to evaluate the effect of imprecise information about rare events and contrast it with normal events, adding stochastic volatility is not expected to bring in any new insight.

Radon–Nikodym derivative with respect to P,

$$d\xi_t = \left(e^{a+bZ_t - b\mu_J - \frac{1}{2}b^2\sigma_J^2} - 1\right)\xi_{t-}dN_t - (e^a - 1)\lambda\xi_t dt, \tag{2}$$

where a and b are predictable processes,[10] and where $\xi_0 = 1$. By construction, the process $\{\xi_t, 0 \leq t \leq T\}$ is a martingale of mean 1. The measure $P(\xi)$ thus defined is indeed a probability measure.

Effectively, ξ changes the agent's probability assessment with respect to the jump component without altering his view about the diffusive component.[11] More specifically, under the alternative measure $P(\xi)$ defined by ξ, the jump arrival intensity λ^ξ and the mean jump size k^ξ change from their counterparts λ and k in the reference measure P to

$$\lambda^\xi = \lambda e^a, \quad 1 + k^\xi = (1+k)e^{b\sigma_J^2}. \tag{3}$$

A detailed derivation of Equation (3) can be found in Appendix A.

The agent operates under the reference model by choosing $a = 0$ and $b = 0$, and ventures into other models by choosing some other a and b. Let $\mathcal{P}$ be the entire collection of such models defined by a and b. We are now ready to define our agent's utility when robust control over the set $\mathcal{P}$ is his concern. For ease of exposition, we start our specification in a discrete-time setting, leaving its continuous-time limit to the end of this section. Fixing the time period at Δ, we define his time-t utility recursively by

$$U_t = \frac{c_t^{1-\gamma}}{1-\gamma}\Delta + e^{-\rho\Delta} \inf_{P(\xi)\in\mathcal{P}} \left\{ \frac{1}{\phi}\psi(E_t^\xi(U_{t+\Delta}))E_t^\xi\left[h\left(\ln\frac{\xi_{t+\Delta}}{\xi_t}\right)\right] + E_t^\xi(U_{t+\Delta}) \right\}$$
$$\text{and} \quad U_T = 0, \tag{4}$$

where c_t is his time-t consumption, $\rho > 0$ is a constant discount rate, and $\psi(E_t^\xi(U_{t+\Delta}))$ is a normalization factor introduced for analytical tractability [Maenhout (2001)]. To keep the penalty term positive, we let $\psi(x) = (1-\gamma)x$ for the case of $\gamma \neq 1$ and $\psi(x) = 1$ for the log-utility case.

The specification in Equation (4) implies that any chosen alternative model $P(\xi) \in \mathcal{P}$ can affect the representative agent in two different ways. On the one hand, in an effort to protect himself against model uncertainty associated with the jump component, the agent evaluates his future prospect $E_t^\xi(U_{t+1})$ under alternative measures $P(\xi) \in \mathcal{P}$. Naturally, he focuses

[10] That is to say, a_t and b_t are fixed just before time t. See, for example, Andersen, Borgan, Gill and Keiding (1992).

[11] It is also important to notice that while the agent is free to deviate his probability assessment about the jump component, he cannot change the state of nature. That is, an event with probability 0 in P remains so in $P(\xi)$. In other words, our construction of ξ in Equation (2) ensures P and $P(\xi)$ to be equivalent measures.

on other jump models that provide prospects worse than the reference models P, hence the infimum over $P(\xi) \in \mathcal{P}$ in Equation (4). On the other hand, he knows that statistically P is the best representation of the existing data. With this in mind, he penalizes his choice of $P(\xi)$ according to how much it deviates from the reference P. This discrepancy or distance measure is captured in this article by $E_t^{\xi}[h(\ln(\xi_{t+1}/\xi_t))]$, where for some $\beta > 0$ and any $x \in \mathbb{R}$,

$$h(x) = x + \beta(e^x - 1). \tag{5}$$

Intuitively, the further away the alternative model is from the reference model P, the larger the distance measure. Conversely, when the alternative model is the reference model, we have $\xi \equiv 1$ with a distance measure of 0. Finally, to control this trade-off between "impact on future prospects" and "distance from the reference model," we introduce a constant parameter $\phi > 0$ in Equation (4). With a higher ϕ, the agent puts less weight on how far away the alternative model is from the reference model and, effectively, more weight on how it would worsen his future prospect. In other words, an agent with higher ϕ exhibits higher aversion to model uncertainty.

The agent's utility function in Equation (4) is similar to that in Anderson, Hansen, and Sargent (2000). Our approach, however, differs from theirs in two ways. First, we restrict the agent to a prespecified set $\mathcal{P}$ of alternative models that differ from the reference model only in their jump components. As a result, the uncertainty aversion exhibited by the agent only applies to the jump component of the model. This distinction becomes important as we later take the model to option pricing because options are sensitive to diffusive shocks and jumps in different ways.

In fact, we can further apply this idea and modify the set $\mathcal{P}$ so the agent can express his uncertainty toward one specific part of the jump component. For example, by restricting $b = 0$ in the definition of ξ in Equation (2), we build a subset $\mathcal{P}^a \subset \mathcal{P}$ of alternative models that is different from the reference model only in terms of the likelihood of jump arrival. Applying this subset to the utility definition of Equation (4), we effectively assume that the agent has doubt about the jump-timing aspect of the model, while he is comfortable with the jump-magnitude part of the model. Similarly, by letting $a = 0$ in Equation (2), we build a class $\mathcal{P}^b$ of alternative models that is different from the reference model only in terms of jump size. An agent who searches over $\mathcal{P}^b$ instead of $\mathcal{P}$ finds the jump-magnitude aspect of the model unreliable, while having full faith in the jump-timing aspect of the model. Finally, by letting $a = 0$ and $b = 0$, we reduce the set $\mathcal{P}^0$ to a singleton that contains only the reference model. Effectively, this is the standard case of a risk-averse investor.

Second, we extend the discrepancy (or distance) measure of Anderson, Hansen, and Sargent (2000) to a more general form. Specifically, our "extended entropy" measure is reduced to their "relative entropy" when β approaches to zero. Given that $h(x)$ is convex and $h(0)=0$, the result of Wang (2003) can be used to provide an axiomatic foundation for our specification (his Theorem 5.1, part a). As it will become clear later, this extended form of distance measure is important in handling uncertainty aversion toward the jump component. In particular, the minimization problem specified in Equation (4) does not have an interior global minimum for the "relative entropy" case.[12] For pure diffusion models, however, it is easy to show that our extended distance measure is equivalent to the "relative entropy" case.

Our utility specification also differs from Anderson, Hansen, and Sargent (2000) in the normalization factor ψ, which we adopt from Maenhout (2001) for analytical tractability. A couple of issues have been raised in the literature regarding this normalization factor. One relates to its effect on the equivalence between a number of robust-control preferences and recursive utility [see Maenhout (2001) and Skiadas (2003)]; the other relates to its effect on the link between the robust-control framework and that of Gilboa and Schmeidler (1989) [see Pathak (2000)]. In this respect, the utility function adopted in this article is not a multiperiod extension of Gilboa and Schmeidler (1989). It is, however, a utility function motivated by uncertaintyaversion toward rare events.[13] Applying this utility to the asset-pricing framework of this article, the most important issue for us to resolve is that the asset-pricing implication involving rare-event premiums is indeed driven by uncertainty aversion toward rare events and not by recursive utility or a particular form of the normalization factor. We clarify these issues by showing that (1) our main result regarding rare-event premiums cannot be generated by a continuous-time Epstein and Zin (1989) recursive utility (Appendix D); (2) the choice of normalization factor does not affect, in any qualitative fashion, the fact that our main result involving rare-event premiums builds on uncertainty aversion toward rare events (Appendix E).

Finally, the continuous-time limit of our utility specification [Equation (4)] can be derived as

$$U_t = \inf_{\{a,b\}} \left\{ E_t^{\xi} \left[\int_t^T e^{-\rho\{s-t\}} \left\{ \frac{1}{\phi} \psi(U_s) H(a_s, b_s) + \frac{c_s^{1-\gamma}}{1-\gamma} \right\} ds \right] \right\}, \tag{6}$$

[12] Roughly speaking, the penalty function in Anderson, Hansen, and Sargent (2000) is not strong enough to counterbalance the "loss in future prospect" for an agent with risk-aversion coefficient $\gamma > 1$. As a result, the investor's concern about a misspecification in the jump magnitude makes him go overboard to the case of total ruin.

[13] See Wang (2003) for an axiomatic foundation in a static setting.

where H is the component associated with the distance measure and can be calculated explicitly as[14]

$$H(a,b) = \lambda\left[1 + \left(a + \frac{1}{2}b^2\sigma_J^2 - 1\right)e^a + \beta(1 + (e^{a+b^2\sigma_J^2} - 2)e^a)\right]. \quad (7)$$

Given this, the investor's objective is to optimize his time-0 utility function U_0.

2. The Optimal Consumption and Portfolio Choice

As in the standard setting, there exists a market where shares of the aggregate endowment are traded as stocks. At any time t, the dividend payout rate of the stock is Y_t, and the ex-dividend price of the stock is denoted by S_t. In addition, there is a risk-free bond market with instantaneous interest rate r_t. The investor starts with a positive initial wealth W_0, trades competitively in the securities market, and consumes the proceeds. At any time t, he invests a fraction θ_t of his wealth in the stock market, $1-\theta_t$ in the risk-free bond, and consumes c_t, satisfying the usual budget constraint.

Having the equilibrium solution in mind, we consider stock prices of the form $S_t = A(t)Y_t$ and constant risk-free rate r, where $A(t)$ is a deterministic function of t with $A(T)=0$. Under the reference measure P, the stock price follows,

$$dS_t = \left(\mu + \frac{A'(t)}{A(t)}\right)S_t dt + \sigma S_t dB_t + (e_t^Z - 1)S_{t-}dN_t. \quad (8)$$

And the budget constraint of the investor becomes

$$\begin{aligned} dW_t = &\left[r + \theta_t\left(\mu - r + \frac{1+A'(t)}{A(t)}\right)\right]W_t dt + \theta_t W_t \sigma dB_t \\ &+ \theta_{t-}W_{t-}(e_t^Z - 1)dN_t - c_t dt. \end{aligned} \quad (9)$$

Given this budget constraint, our investor's problem is to choose hisconsumption and investment plans $\{c, \theta\}$ so as to optimize his utility. Let J_t be the indirect utility function of the investor,

$$J(t, W) = \sup_{\{c,\theta\}} U_t, \quad (10)$$

where U_t is the continuous-time limit of the utility function defined by Equation (4). The following proposition provides the Hamilton–Jacobi–Bellman (HJB) equation for J.

[14] See the proof of Proposition 1 in Appendix for the derivation.

Proposition 1. *The investor's indirect utility J, defined by Equation (10), has the terminal condition $J(T, W) = 0$ and satisfies the following HJB equation,*

$$\sup_{c,\theta}\left\{u(c) - \rho J(t, W) + \mathcal{A}J(t, W) + \inf_{a,b}\left\{\lambda e^a(E^{Z(b)}[J(t, W(1+(e^Z-1)\theta))]\right.\right.$$
$$-J(t, W)) + \frac{1}{\phi}\psi(J)\lambda\left[1 + \left(a + \frac{1}{2}b^2\sigma_J^2 - 1\right)e^a\right.$$
$$\left.\left.\left.+\beta(1+(e^{a+b^2\sigma_J^2}-2)e^a)\right]\right\}\right\} = 0, \tag{11}$$

where $E^{Z(b)}(.)$ denotes the expectation with respect to Z under the alternative measure associated with b. That is, for any function f,

$$E^{Z(b)}(f(Z)) = E(e^{bZ - b\mu_J - \frac{1}{2}b^2\sigma_J^2} f(Z)). \tag{12}$$

The term $\mathcal{A}\ J(t,W)$ in the HJB equation [Equation (11)] is the usual infinitesimal generator for the diffusion component of the wealth dynamics,

$$\mathcal{A}J = J_t + \left[r + \theta\left(\mu - r + \frac{A'(t)+1}{A(t)}\right)\right]WJ_W - cJ_W + \frac{\sigma^2}{2}\theta^2 W^2 J_{WW}, \tag{13}$$

where J_t is the derivative of the indirect utility J with respect to t, and J_W and J_{WW} are its first and second derivatives with respective to W.

The intuition behind the HJB equation [Equation (9)] exactly parallels that of its discrete time counterpart, Equation (4). Specifically, compared with the standard HJB equation for jump diffusions, the HJB equation in Equation (11) has two important modifications. First, the risk associated with the jump component is evaluated at all possible alternative models indexed by (a, b), reflecting the investor's precaution against model uncertainty with respect to the jump component. Second, it incorporates an additional term in the last two lines of Equation (11), penalizing the choice of the alternative model by its distance from the reference model. The following proposition provides the solution to the HJB equation.

Proposition 2. *The solution to the HJB equation is given by*

$$J(t, W)\frac{W^{1-\gamma}}{1-\gamma}f(t)^\gamma, \tag{14}$$

where $f(t)$ is a time-dependent coefficient satisfying the ordinary differential Equation (B.4) in Appendix B with the terminal condition $f(T) = 0$. The optimal consumption plan is given by $c_t^* = W_t^*/f(t)$, where W^* is the optimal *wealth process. Finally, the optimal solutions θ^*, a^*, and b^* satisfy*

$$\left(\mu - r + \frac{1 + A'(t)}{A(t)}\right) - \gamma\theta\sigma^2 + \lambda e^a E^{Z(b)}[(1 + (e^Z - 1)\theta)^{-\gamma}(e^Z - 1)] = 0, \tag{15}$$

$$\frac{1-\gamma}{\phi}\left(a + \frac{1}{2}b^2\sigma_J^2 + 2\beta(e^{a+b^2\sigma_J^2} - 1)\right) + E^{Z(b)}[(1 + (e^Z - 1)\theta)^{1-\gamma}] - 1 = 0, \tag{16}$$

$$\frac{1-\gamma}{\phi} b\sigma_J^2(1 + 2\beta e^{a+b^2\sigma_J^2}) + \frac{\partial}{\partial\beta} E^{Z(b)}[(1 + (e^Z - 1)\theta)^{1-\gamma}) = 0, \tag{17}$$

where $E^{Z(b)}(.)$ defined in Equation (12) is the expectation with respect to Z under the alternative measure associated with b.

3. Market Equilibrium

In equilibrium, the representative agent invests all his wealth in the stock market $\theta_t = 1$ and consumes the aggregate endowment $c_t = Y_t$ at any time $t \leq T$. The solution to market equilibrium and the pricing kernel are summarized by the following proposition.

Proposition 3. *In equilibrium, the total (cum-dividend) equity premium is*

$$\text{Total equity premium} = \gamma\sigma^2 + \lambda k - \lambda^Q k^Q, \tag{18}$$

where $k = \exp(\mu_J + \sigma_J^2/2) - 1$ is the mean percentage jump size of the aggregate endowment, and λ^Q and k^Q are defined by[15]

$$\begin{aligned} \lambda^Q &= \lambda \exp\left(-\gamma\mu_J + \frac{1}{2}\gamma^2\sigma_J^2 + a^* - b^*\gamma\sigma_J^2\right), \\ k^Q &= (1 + k)\exp((b^* - \gamma)\sigma_J^2) - 1, \end{aligned} \tag{19}$$

and a^* and b^* are the solution of the following nonlinear equations:

$$a + \frac{1}{2}b^2\sigma_J^2 + 2\beta(e^{a+b^2\sigma_J^2} - 1) + \frac{\phi}{1-\gamma}([(1 + k)e^{(b-\frac{1}{2}\gamma)\sigma_J^2}]^{1-\gamma} - 1) = 0 \tag{20}$$

$$b(1 + 2\beta^{a+b^2\sigma_J^2}) + \phi[(1 + k)e^{(b-\frac{1}{2}\gamma)\sigma_J^2}]^{1-\gamma} = 0. \tag{21}$$

[15] As will become clear in the next section, λ^Q and k^Q are the risk-neutral counterparts of λ and k.

The equilibrium riskfree rate r is

$$r = \rho + \gamma\mu - \frac{1}{2}\gamma(\gamma+1)\sigma^2 + \lambda^*(1 - (1+k^*)^{-\gamma} e^{\frac{1}{2}\gamma(1+\gamma)\sigma_J^2}) + \chi^*, \quad (22)$$

where $\lambda^* = \lambda \exp(a^*)$ *and* $k^* = (1+k)\exp(b^*\sigma_J^2) - 1$, *and where*

$$\chi^* = -\frac{1-\gamma}{\phi}\lambda\left[1 + \left(a^* + \frac{1}{2}(b^*)^2\sigma_J^2 - 1\right)e^{a^*} + \beta(1 + (e^{a^*+(b^*)^2\sigma_J^2} - 2)e^{a^*})\right]. \quad (23)$$

Finally, the equilibrium pricing kernel is given by

$$d\pi_t = -r\pi_t dt - \gamma\sigma\pi_t dB_t + (e^{a^*+(b^*-\gamma)Z - b^*\mu_J - \frac{1}{2}(b^*)^2\sigma_J^2} - 1)\pi_{t-} dN_t - \lambda(e^{a^*-\gamma(\mu_J+b^*\sigma_J^2)+\frac{1}{2}\gamma^2\sigma_J^2} - 1)\pi_t dt. \quad (24)$$

To understand how the investor's uncertainty aversion affects the equilibrium asset prices, let us first take away the feature of uncertainty aversion by setting $a \equiv 0$ and $b \equiv 0$, or $\phi \to 0$. Our results in Equations (18) and (22) are then reduced to those of Naik and Lee (1990)—the standard case of a risk-averse investor with no uncertainty aversion. In this case, the total equity premium is attributed exclusively to risk aversion:

$$\begin{aligned} &\text{Diffusive risk premium} = \gamma\sigma^2, \\ &\text{Jump-risk premium} = \lambda k - \bar{\lambda}\bar{k}, \end{aligned} \quad (25)$$

where $\bar{\lambda}$ and $\bar{k}$ are the counterparts of λ^Q and k^Q when the uncertainty aversion ϕ is set to zero:

$$\bar{\lambda} = \lambda \exp\left(-\gamma\mu_J + \frac{1}{2}\gamma^2\sigma_J^2\right), \quad \bar{k} = (1+k)\exp(-\gamma\sigma_J^2) - 1. \quad (26)$$

Quite intuitively, both types of risk premiums approach zero when the risk-aversion coefficient γ approaches zero and are positive for any risk-averse investors ($\gamma > 0$).

When the investor exhibits uncertainty aversion ($\phi > 0$), there is one additional component in the equity premium:

$$\text{Rare-event premium} = \bar{\lambda}\bar{k} - \lambda^Q k^Q. \quad (27)$$

It is important to emphasize that while the magnitude of this part of the equity premium depends on the risk-aversion parameter of the investor, it is the uncertainty aversion of the investor that gives rise to this premium. Specifically, the rare-event premium remains positive even when we take the limit $\gamma \to 0$, while it becomes zero when the investor's model uncertainty aversion ϕ approaches zero. The following two examples highlight this feature of the rare-event premium by considering the extreme case where the investor is risk neutral ($\gamma = 0$).

In the first case, the investor is worried about model misspecification with respect to the jump arrival intensity, that is, how frequently the jumps occur. He performs robust control by searching over the subset $\mathcal{P}^a$ defined by $a \in R$ and $b \equiv 0$. Setting $b = 0$ and $\gamma = 0$, Equation (20) reduces to

$$a + 2\beta(e^a - 1) + \phi k = 0. \tag{28}$$

For the case of adverse event risk ($k < 0$), we can see from Equation (28) that $a^* > 0$ if and only if the investor exhibits uncertainty aversion ($\phi > 0$). The rare-event premium in this case is

$$\bar{\lambda}\bar{k} - \lambda^Q k^Q = \lambda k(1 - e^{a^*}),$$

which is positive if and only if $\phi > 0$.

In the second case, the investor is worried about model misspecification with respect to the jump size. This time, he performs robust control by searching over the subset $\mathcal{P}^b$ defined by $b \in R$ and $a \equiv 0$. Setting $a = 0$ and $\gamma = 0$, Equation (21) reduces to

$$b = -\phi \frac{(1+k)e^{b\sigma_J^2}}{1 + 2\beta e^{b^2\sigma_J^2}}, \tag{29}$$

which indicates that $b^* < 0$ when there is uncertainty aversion ($\phi > 0$). The rare-event premium in this case is

$$\bar{\lambda}\bar{k} - \lambda^Q k^Q = \lambda k(1 + k)e^{b^*\sigma_J^2},$$

which is again positive if and only if $\phi > 0$.

These two cases are the simplest examples of our more general results. In addition to providing some important intuition behind our results, they also deliver a quite important point. That is, the aversion toward model uncertainty is independent of that toward risk, and the effect of uncertainty aversion becomes most prominent with respect to rare events. Indeed, the fact that our model allows such separation of total equity premium into risk and rare-event components is crucial for our analysis. As emphasized in the introduction, our contention is that investors treat rare events differently from more common events and such differential treatment will be reflected in asset prices. The decomposition of the equity premium characterized in Proposition 3 allows us to study the effect on prices and can potentially lead to empirically testable implications with respect to the different components of the equity premium.

To elaborate on the last point and set the stage for the next section, we note that if there is no model uncertainty, or if the investor is uncertainty neutral ($\phi = 0$), then according to Equations (25) and (26), both diffusive and jump-risk premiums are linked by just one risk-aversion coefficient γ. This constraint can, in fact, be tested using equity and equity options, which have different sensitivities to the diffusive and jump risks. In such an equilibrium, the pricing kernel that links the equity to the equity options is controlled by just one risk-aversion coefficient γ. On the other hand, empirical studies [e.g., Pan (2002) and Jackwerth (2000)] using time-series data from both markets (the S&P 500 index and option) indicate that the pricing kernel linking the two markets cannot be supported by such an equilibrium.[16] In particular, the "data-implied γ" for the jump risk is considerably larger than that for the diffusive risk.

We close this section by discussing the asset-pricing implication of the normalization factor ψ in more detail. For this, we focus on the equilibrium pricing kernel derived in Equation (24), which can be rewritten as

$$\pi_t = e^{-\rho t} e^{-\chi^* t} \xi_t^* Y_t^{-\gamma}, \tag{30}$$

where χ^* is a constant defined in Equation (3) and where ξ^* is the Randon–Nikodym derivative that defines the optimal alternative measure $P(\xi^*)$. The shocks to the pricing kernel consist of two parts: $Y_t^{-\gamma}$ generates the diffusive- and jump-risk premiums; and ξ_t^* generates the rare-event premium. It is easy to see that the presence of a nontrivial ξ^* in the pricing kernel derives from the investor's consideration over alternative measures regarding rare events. In other words, in our specific setting, rare-event premiums can be traced to the investor's uncertainty aversion toward rare events.

To understand the extent to which different normalization factors affect this link, in Appendix E we consider an example with a more general form of the normalization factor. We show that the particular form of normalization affects (1) the risk-free rate through its direct impact on intertemporal substitution; (2) the optimal solution of ξ^*. For the more general cases, the optimal ξ^* cannot be solved in closed form, although the uncertainty aversion aspect of the utility will lead ξ^* toward measures giving worse prospects than the reference measure.

More importantly, we show that, regardless of the specific choice of normalization, the shocks to the pricing kernel still consist of $Y_t^{-\gamma}$ and ξ_t^* as in Equatin (30). Similar to our earlier discussion, the presence of a nontrivial ξ^* in the pricing kernel can be traced back to the investor's consideration over alternative measures regarding rare events. Thus, while

[16] This relies on our specification of the aggregate consumption process. If one is willing to relax this specification, then one can always find an equilibrium to support any given pricing kernel, including the empirical pricing kernel that links the equity and equity options markets. For example, for a power utility with risk-aversion coefficient γ, one can back out a consumption process by equating marginal utility to the empirical pricing kernel.

a more general normalization factor might provide a more complicated ξ^*, the important link between rare-event premia and uncertainty aversion toward rare events still survives.[17]

4. The Rare-Event Premiums in Options

To further disentangle the rare-event premiums from the standard risk premiums, we turn our attention to the options market. Using the equilibrium pricing kernel π (Proposition 3), we can readily price any derivative securities in this economy. Specifically, let Q be the risk-neutral measure defined by the equilibrium pricing kernel π such that $e^{rT}\pi_T/\pi_0 = dQ/dP$. It can be shown that the risk-neutral dynamics of the ex-dividend stock price follows:

$$dS_t = (r-q)S_t dt + \sigma S_t dB_t^Q + (e^{Z_t}-1)S_{t-}dN_t - \lambda^Q k^Q dt, \qquad (31)$$

where r is the risk-free rate and q is the dividend payout rate,[18] and where under Q, B^Q is a standard Brownian motion and N_t is a Poisson process with intensity λ^Q. Given jump arrival at time t, the percentage jump amplitude is lognormally distributed with the risk-neutral mean k^Q. Both risk-neutral parameters λ^Q and k^Q are defined earlier in Equation (19). European-style option pricing for this model is a modification of the Black and Scholes (1973) formula, and has been established in Merton (1976). For completeness of the article, the pricing formula is provided in Appendix C.

What makes the option market valuable for our analysis is that, unlike equity, options have different sensitivities to diffusions and jumps. For example, a deep OTM put option is extremely sensitive to negative price jumps but exhibits little sensitivity to diffusive price movements. This nonlinear feature inherent in the option market enables us to disentangle the three components of the total equity premium (Proposition 3) that are otherwise impossible to separate using equity returns alone. This "observational equivalence" with respect to equity returns is further illustrated in Table 1.

[17] Unless, of course, one considers a normalization factor that effectively prevents the investor from choosing alternative measures, resulting in a trivial optimal solution of $\xi^* \equiv 1$.

[18] For the rest of our analysis, we will set the risk-free rate at $r = 5\%$ and the dividend yield at $q = 3\%$. In other words, we are not using the equilibrium interest rate and the dividend yield. This is without much loss of generality. Specifically, the parameter ρ can be used to match the desired level of r. The dividend payout ratio q is slightly more complicated, since it is in fact time varying in our setting. For an equilibrium horizon T that is sufficiently large compared with the maturity of the options to be considered, we can use the result for the infinite horizon case, and take $q = 1/\alpha$, where α, given by Equation (B.6), can be calibrated by the free parameter μ. Finally, as our analysis focuses on comparing the prices of options with different moneyness, the effect of r and q will be minor as long as the same r and q are used to price all options.

Table 1.
The three components of the equity premium, jump case 1

	Aversion		Premium (%)			
Jump parameters	ϕ	γ	Diffusive risk	Jump risk	Rare event	Total premium
$\lambda = 1/3$	0	3.47	7.80	0.20	0	
	10	3.15	7.09	0.19	0.72	8%
$\mu_J = -1\%$	20	2.62	5.91	0.15	1.94	

Table 1 details a simple calibration exercise with parameters for the reference model P set as follows. For the diffusive component, the volatility is set at $\sigma = 15\%$. For the jump component,[19] the arrival intensity is $\lambda = 1/3$, and the random jump amplitude is normal with mean $\mu_J = -1\%$ and standard deviation $\sigma_J = 4\%$. It should be noted that our model cannot resolve the issue of "excess volatility."

As a result, we face the problem of which set of data the model should be calibrated to: the aggregate equity market or the aggregate consumption. For example, if we were to fit the model directly to the data on aggregate consumption, the equity volatility would be around 2%, and the equity options would be severely underpriced simply because of this low volatility level. Given that the main objective of this calibration exercise is to explore the link between the equity market and the options market, calibrating the model to the aggregate equity market seems to be a more reasonable choice. For this reason, the set of model parameters are chosen to fit the data on the S&P 500 index market.

Given this reference model, three different scenarios are considered for the representative agent's risk aversion γ and uncertainty aversion ϕ. As shown in Table 1, each scenario corresponds to an economy with a distinct level of uncertainty aversion ϕ and yields a distinct composition of the diffusive-risk premium, the jump-risk premium, and the rare-event premium. For example, the rare-event premium is zero when the representative agent exhibits no aversion to model uncertainty, and increases to 1.94% per year when the uncertainty aversion coefficient becomes $\phi = 20$. These predictions of our model, however, cannot be tested if we focus only on the equity return data. As shown in Table 1, for a fixed level of uncertainty aversion ϕ, one can always adjust the level of risk aversion γ so that the total equity premium is fixed at 8% per year, although the economic sources of the respective equity premiums differ significantly from one scenario to another. To be able to decompose the total equity

[19] The jump parameters are close to those reported by Pan (2000) for the S&P 500 index. Alternative jump parameters will be considered in later examples.

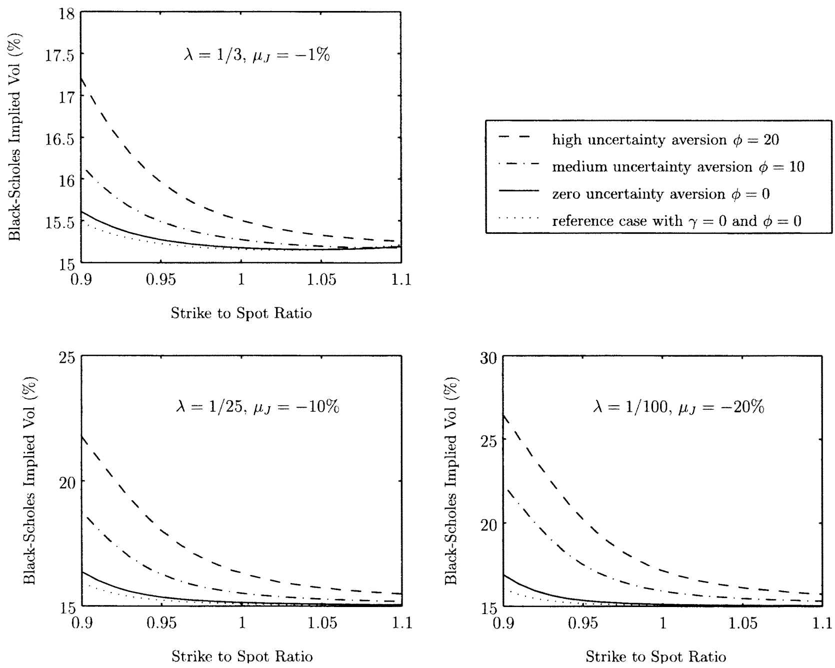

Figure 1
The equilibrium "smile" curves

premium into its three components, we need to take our model one step further to the options data.

To examine the option pricing implication of our model, we start with the same reference model and the same set of scenarios of uncertainty aversion as those considered in Table 1. For each scenario, we use our equilibrium model to price one-month European-style options, both calls and puts, with the ratio of strike to spot prices varying from 0.9 to 1.1. As it is standard in the literature, we quote the option prices in terms of the Black-Scholes implied volatility (BS-vol) and plot them against the respective ratios of strike to spot prices. The first panel of Figure 1 reports the "smile" curves generated by the three equilibrium models with varying degrees of uncertainty aversion. We can see that although all three scenarios are observationally equivalent with respect to the equity market, their implications on the options market are notably different.

4.1 The case of only risk aversion

Let us first consider the case of zero uncertainty aversion, where risk aversion is the only source of premiums in both equity and options. Calibrating the risk-aversion coefficient γ to match the equity premium, let us first examine the model's implication for the ATM option (puts and calls with a strike-to-spot ratio of 1). From the first panel of Figure 1, we see that the model prices such options at a BS-vol of 15.2%, which is very close in magnitude to the total market volatility $\sqrt{\sigma^2 + \lambda(\mu_J^2 + \sigma_J^2)} = 15.2\%$. The market-observed BS-vols for such ATM options, however, are known to be higher than the volatility of the underlying index returns. In other words, there is a premium implicit in such ATM options that is not captured by this model with only risk aversion.

Next, we examine this model's implication for options across moneyness. Moving the strike-to-spot ratio from 1 to 0.9, we arrive at a 10% OTM put option, which is priced by the model at 15.6% BS-vol. That is, moving 10% out of the money, the BS-vol increases from 15.2 to 15.6%. The market-observed "smile" curves, however, are much steeper than what is captured by this model.

In other words, the market views the OTM put options to be more valuable than what this model predicts. There is an additional component implicit in such OTM put options that is not captured by this model with only risk aversion.

Moving from equity to ATM options and to OTM put options, we are looking at a sequence of securities that are increasingly sensitive to rare events. At the same time, the model with only risk aversion misprices this sequence of securities with increasing proportion. As we can see from our next example, one plausible explanation is that the rare-event component is not priced properly in this model with only risk aversion.

4.2 The case of uncertainty aversion toward rare events

Let us now consider the two cases that incorporate the representative agent's uncertainty aversion. As shown in Table 1, in both cases the total equity premium has three components, two of which are driven by the representative agent's risk aversion γ and one driven by his uncertainty aversion ϕ. Comparing the case of $\phi = 20$ with the previously discussed case of $\phi = 0$, our first observation is that, even for ATM options, the two models generate different equilibrium prices. Specifically, for the case of zero uncertainty aversion, the BS-vol implied by an ATM option is 15.2%, but for the case of uncertainty aversion $\phi = 20$, the BS-vol implied by an ATM option is 15.5%. This implies that, while both cases are observationally equivalent when viewed using equity prices, the model incorporating uncertainty aversion ($\phi = 20$) predicts a premium of about 2% for one-month ATM options. This result is indeed consistent with the empirical fact that options, even those that are at the money, are priced with a premium.[20]

[20] See, for example, Jackwerth and Rubinstein (1996) and Pan (2002).

This additional premium, which is linked exclusively to the investor's uncertainty aversion toward rare events, becomes even more pronounced as we move to OTM puts, which, compared with ATM options, have more sensitivity to adverse rare events. The first panel in Figure 1 shows that a 10% OTM put option is priced at 17.2% BS-vol, compared with 15.6% BS-vol in the case of $\phi = 0$. That is, for every dollar invested in a one-month 10% OTM put option, typically used as a protection against rare events, the investor is willing to pay 10 cents more because of his uncertainty aversion toward the adverse rare events.

As shown in Pan (2002), both empirical facts — ATM options priced with a premium and OTM put options priced with an even higher premium, resulting in a pronounced "smirk" pattern — are indeed closely connected. If only risk aversion is used to explain these empirical facts, one direct implication is that the "data-implied γ" for the jump risk has to be considerably larger than that for the diffusive risk. By incorporating uncertainty aversion in this article, however, we are able to explain these empirical facts without having to incorporate an exaggerated risk-aversion coefficient for the jump risk. By doing so, we offer a simple explanation for the significant premium implicit in options, especially those put options that are deep out of the money. That is, when it comes to rare events, the investors simply do not have a reliable model. They react by assigning rare-event premiums to each financial security that is sensitive to rare events. Options with varying moneyness are sensitive to the rare events in a variety of ways, bearing different levels of rare-event premiums. Our analysis shows that a significant portion of the pronounced "smirk" pattern can be attributed to this varying degree of rare-event premiums implicit in options.

Finally, to show the robustness of our results, we modify the two key jump parameters, λ and μ_J, in the reference model considered in Table 1. In Table 2, we consider jumps that happen once every 25 years, with a mean magnitude of -10%, capturing the magnitude of major market corrections. In Table 3, jumps happen once every 100 years with a magnitude of -20%, capturing the magnitude of an event as rare as the 1987 crash. The option pricing implications of these models are reported in the

Table 2.
The three components of the equity premium, jump case 2

	Aversion		Premiums (%)			
Jump parameters	ϕ	γ	Diffusive risk	Jump risk	Rare event	Total premium
$\lambda = 1/25$	0	3.47	7.81	0.19	0	
$\mu_J = -10\%$	10	2.88	6.47	0.15	1.38	8%
	20	1.61	3.62	0.08	4.30	

Table 3.
The three components of the equity premium, jump case 3

	Aversion		Premiums (%)			
Jump parameters	ϕ	γ	Diffusive risk	Jump risk	Rare event	Total premium
$\lambda = 1/100$	0	3.47	7.81	0.19	0	
	10	2.36	5.31	0.12	2.58	8%
$\mu_J = -20\%$	20	0.68	1.54	0.03	6.43	

lower two panels in Figure 1. As we can see, although all three reference models incorporate rare events that are very different in intensity and magnitude, the impact of uncertainty aversion remains qualitatively similar.

4.3 Implications of alternative utility specifications

4.3.1 The case of recursive utility. Our specification involves two free parameters (in addition to the time discount coefficient ρ): the risk-aversion coefficient γ and the uncertainty-aversion coefficient ϕ. Compared with the standard power utility, we have one more free parameter. One may argue that with one more free parameter, it is no surprise that the "smirk" patterns can be generated. To compare our model against alternative utility functions at equal footing, we consider the case of continuous-time Epstein and Zin (1989) recursive utility, which also has two free parameters: the risk-aversion coefficient γ and the coefficient for the intertemporal substitution δ. This comparison is also of interest because of the equivalence result documented in the literature for diffusion models [Maenhout (2001) and Skiadas (2003)].

In Appendix D we show that the recursive utility results in a more complex risk-free rate, but for the purpose of pricing risks it has the same implication as a standard power utility. This result is quite intuitive given that the recursive utility is designed to separate intertemporal substitution from risk aversion. If our interest lies in how the diffusive risk is priced relative to the rare events, we need look no further than the special case of power utility, which indeed captures the risk aversion component of the recursive utility. Other than their differential implications for risk-free rates, the option-pricing implication of a recursive utility is very much the same as that of a power utility. In Appendix D, the risk-neutral jump parameters λ^Q and k^Q, which are important for option pricing, are derived explicitly and are shown to be identical to those of a power utility case.

In addition to serving as a robust check against alternative utility functions, this example also helps clarify, for our setting, the issue of equivalence between the robust-control framework and recursive utility. Specifically, we show with an explicit example, that the robust-control framework in our setting is not equivalent to the continuous-time Epstein and Zin (1989) recursive utility. This, however, does not contradict the equivalence results established by Maenhout (2001) and Skiadas (2003), since we add a new dimension to the problem: rare events and uncertainty aversion only toward rare events.

4.3.2 The case of habit formation. An alternative preference of interest is the external habit formation model of Campbell and Cochrane (1999), which is shown to generate rich dynamics for asset prices from consumption data. This utility specification is of particular interest because it is capable of resolving the "excess volatility" and equity-premium puzzle, which our model does not explain. It is therefore important for us to understand if such habit-formation models can explain the option-smirk puzzle. To some extent, this analysis also serves to clarify the key difference between the equity-premium puzzle and the option-smirk puzzle.

At the heart of the option-smirk puzzle is the differential pricing of options with varying sensitivities to rare-event risk. For a preference to generate the observed level of option smirk, the associated equilibrium pricing kernel should have the ability to price rare-event risk separately from the diffusive risk.[21] Standard formations of the habit model such as that in Campbell and Cochrane, in contrast, assume that the shock to habit is perfectly correlated with the shock to consumption (the endowment). As such, the habit-model-implied pricing kernel, though following a richer dynamic process than in the standard CRRA model, effectively does not price the diffusive and jump components of the endowment process differently.[22] We therefore conjecture that, as formulated and calibrated in recent studies, the habit model will not generate the observed smile in option prices. Indeed, as preliminary evidence in support of our conjecture, we took the model-implied option prices computed by Bansal, Gallant and Tauchen (2002) from their calibrated habit model and

[21] This is best illustrated by comparing our model against a model with constant relative risk aversion (CRRA) preference with no uncertainty aversion toward rare events. As shown in Equation (30), the equilibrium pricing kernel of our model is proportional to $\xi_t^* Y_t^{-\gamma}$, while that of the CRRA preference is proportional to $Y_t^{-\gamma}$. The additional term ξ_t^* in our model is the key to our model's ability to generate option smirks. Economically, it adds a layer to the market price of rare-event risk that is above and beyond that associated with risk aversion, and this extra degree of freedom arises from uncertainty aversion toward rare-event risk.

[22] Specifically, the pricing kernel generated by the habit formation preference of Campbell and Cochrane can be shown to be proportional to $S_t^{-\gamma} Y_t^{-\gamma}$, where S_t is the surplus consumption ratio and Y_t is the aggregate consumption (which equals aggregate endowment in equilibrium). In their external habit specification, the dynamics of $s = \log(S)$ follows

$$s_t = (1-\phi)\bar{s} + \phi s_{t-1} + \lambda(s_{t-1})(y_t - y_{t-1} - g),$$

where ϕ, $\bar{s}$, and g are parameters, $y = \log Y$, and $\lambda(s_{t-1})$ is the sensitivity function. Effectively, by introducing an external habit through the surplus consumption ratio S, the habit formation preference of Campbell and Cochrane generates an equilibrium pricing kernel proportional to $Y_t^{-\gamma'_{t-1}}$, where $\gamma'_{t-1} = \gamma(1 + \lambda(s_{t-1}))$ is the implied state-dependent risk-aversion coefficient.

converted the prices to BS-vols using a constant risk-free rate of 5% and dividend payout rate of 2%. These calculations generate inverted options smirks contrary both to the data and to the implications of our model with rare-event premiums.[23]

Moving beyond the standard formation of habit, one could add an exogenous shock to the habit so that it is not perfectly correlated with the consumption shock. For example, one could allow the jump component of the endowment to affect the habit more severely than the diffusive component. These models would do better in explaining the option smirks than the standard habit models. It would be important, however, to develop an economic explanation for why the habit shock has the requisite correlation patterns with the diffusive and jump components of endowments to generate the option smirks. In contrast, option smirks arise naturally in our model because of uncertainty aversion toward rare events.

4.4 Features of the underlying shocks vs. the pricing kernel

The various utility specifications examined in our calibration exercises effectively lead us to various forms of pricing kernels, which in turn play an important role in pricing options and shaping the smile curves. Given that option prices also depend on the underlying stock dynamics, it is therefore natural to question the role played by the underlying stock dynamics in generating smile curves.

We would like to point out that to resolve the puzzle associated with smile curves, modifying the underlying stock dynamics alone is not adequate because any return process, however sophisticated, has to fit to the actual dynamics observed in the underlying stock market. Once this constraint is enforced, there is little room for different specifications of the return process to maneuver in order to generate the kind of smile curves observed in the options market. This point can be best made by examining the data from both markets nonparametrically. As reported by Jackwerth (2000), the option-implied risk-neutral return distribution is much more negatively skewed than the actual return distribution observed directly from the underlying stock market. In other words, the option-implied crash is both more frequent and more severe than that observed from the stock market.

Therefore, the pricing kernel, which links the two distributions,plays an important role in resolving this puzzle and reconciling the information from the two markets. Conversely, the empirical literature on the joint estimation of stock and option markets presents a great deal of

[23] It should be mentioned that both interest rate and dividend yield are stochastic in their models. For the purpose of understanding option smirks, however, the stochastic nature of risk-free rate or dividend yield should not play an important role. The inverted option smirk pattern implied by their equilibrium option prices stays true when different risk-free rates and dividend yields are used.

information regarding the empirical features of pricing kernels. Less, however, is known about what features of utility functions generate pricing kernels consistent with those considered in the empirical literature.

In this article, we provide such a link between utility function and pricing kernel. Specifically, we start with a utility specification motivated by uncertainty aversion toward rare events, and arrive at an equilibrium pricing kernel of the form,

$$\pi_t = e^{-\rho t} e^{-\chi^* t} \xi_t^* Y_t^{-\gamma}.$$

As can be seen from our calibration exercises, the presence of a nontrivial optimal ξ^* in the pricing kernel plays an important role in generating the "smirk" patterns in options across moneyness. At the same time, as discussed at the end of Section 3, the presence of the optimal ξ_t^* in the pricing kernel can be traced back to the utility specification that corresponds to uncertainty aversion toward rare events.

Finally, we would like to point out that there are potential alternative explanations for "smirk" patterns. For example, a nontrivial ξ_t^* could show up in the pricing kernel simply because the investor has a very pessimistic prior about the jump component. That is, he starts with the prior that the jump intensity is λ^* and the mean percentage jump is k^*. Although observationally equivalent, the economic source behind this interpretation is very different from ours. In our model, the optimal λ^* and k^* arise endogenously from robust control due to uncertainty aversion toward rare events. In the Bayesian interpretation, λ^* and k^* are a part of the investor's prior. It is important to point out that without using information from the options market, it is hard for the investor to come up with such a prior.

5. Conclusion

Motivated by the observation that models with rare events are easy to build but hard to estimate, we have developed a framework to formally investigate the asset pricing implication of imprecise knowledge about rare events. We modeled rare events by adding a jump component in aggregate endowment and modified the standard pure-exchange economy by allowing the representative agent to perform robust control [in the sense of Anderson, Hansen, and Sargent (2000)] as a precaution against possible model misspecification with respect to rare events. The equilibrium is solved explicitly.

Our results show that the total equity premium has three components: the diffusive risk premium, the jump-risk premium, and the rare-event premium. In such a framework, the standard model with only risk aversion becomes a special case with overidentifying restrictions on the three components of the total equity premium. While such restrictions do not appear if we fit the model to the equity data alone, these restrictions do become important as we apply the model to a range of securities with varying sensitivity to rare events. Our calibration exercise on equity and equity options across moneyness provides one such example. Our results suggest that uncertainty aversion toward rare events and, consequently, rare-event premiums play an important role in generating the "smirk" pattern observed for options across moneyness.

Appendix A: Changes of Probability Measures for Jumps

We first derive the arrival intensity λ^{ξ} of the Poisson process under the new probability measure $P(\xi)$. Let

$$dM = dN_t - \lambda dt$$

be the compensated Poisson process, which is a P-martingale. Applying the Girsanov theorem for point processes [see, e.g., Elliott (1982)], we have

$$dM^{P(\xi)} = dM_t - E(e^{a+bZ_t-b\mu_J-\frac{1}{2}b^2\sigma_J^2} - 1)\lambda dt = dM_t - (e^a - 1)\lambda dt = dN_t - \lambda^{\xi} dt$$

where $\lambda^{\xi} = \lambda \exp(a)$, as given in Equation (3).

Next we derive the mean percentage jump size k^{ξ} under $P(\xi)$. Let

$$dM = (e^Z - 1)S_{t-}dN_t - kS_t\lambda dt$$

be the compensated pure-jump process, which is a P-martingale. Applying the Girsonov theorem, we have

$$\begin{aligned} dM^{P(\xi)} &= dM_t - E[(e^{a+bZ_t-b\mu_J-\frac{1}{2}b^2\sigma_J^2} - 1)(e^Z - 1)]S_t\lambda dt \\ &= (e^Z - 1)S_{t-}dN_t - k^{\xi} S_t \lambda^{\xi} dt \end{aligned}$$

where $k^{\xi} = (1+k)\exp(b\sigma^2_J) - 1$, as given in Equation (3).

Appendix B: Proofs of Propositions

Proof of Proposition 1. Given zero bequest motive, it must be that $J(T, W) = 0$. The derivation of the HJB equation involves applications of Ito's lemma for jump-diffusion processes. The derivation is standard except for the penalty term. In particular, we need to calculate the continuous-time limit of the "extended entropy" measure. For this, we first let

$$\begin{aligned} E_t^{\xi}\left[h\left(\ln\frac{\xi_{t+\Delta}}{\xi_t}\right)\right] &= E_t\left[\frac{\xi_{t+\Delta}}{\xi_t}h\left(\ln\frac{\xi_{t+\Delta}}{\xi_t}\right)\right] = E_t\left(\frac{\xi_{t+\Delta}}{\xi_t}\ln\frac{\xi_{t+\Delta}}{\xi_t}\right) + \beta E_t\left(\frac{\xi_{t+\Delta}}{\xi_t}\left(\frac{\xi_{t+\Delta}}{\xi_t} - 1\right)\right) \\ &= \frac{1}{\xi}E_t(\xi_{t+\Delta}\ln\xi_{t+\Delta} - \xi_t\ln\xi_t) + \beta\frac{1}{\xi_t^2}E_t(\xi_{t+\Delta}^2 - \xi_t^2), \end{aligned} \tag{B.1}$$

where we use the martingale property $E_t(\xi_{t+\Delta}) = \xi_t$ of the Radon–Nikodym process $\{\xi\}$. Applying Ito's lemma to the processes $\{\xi\ln\xi\}$ and $\{\xi^2\}$ separately, a straightforward calculation shows that

$$\lim_{\Delta \to 0} \frac{1}{\Delta} \frac{1}{\xi_t} E_t(\xi_{t+\Delta} \ln \xi_{t+\Delta} - \xi_t \ln \xi_t) = \lambda \left(1 + \left(a + \frac{1}{2} \sigma_J^2 b^2 - 1 \right) e^a \right)$$

$$\lim_{\Delta \to 0} \frac{1}{\Delta} \frac{1}{\xi_t^2} E_t(\xi_{t+\Delta}^2 - \xi_t^2) = \lambda (1 + (e^{a+b^2 \sigma_J^2} - 2) e^a). \qquad \blacksquare$$

Proof of Proposition 2. We conjecture that the solution to the HJB equation is indeed of the form in Equation (14). The first-order condition for c becomes

$$c = f(t)^{-1} W. \tag{B.2}$$

Substituting Equations (14) and (B.2) into the HJB equation, we have

$$\begin{aligned} \sup_{c,\theta} \Bigg\{ & \frac{\gamma}{1-\gamma} \frac{1+f'(t)}{f(t)} - \frac{\rho}{1-\gamma} + r + \theta \left(\mu - r + \frac{A'(t)+1}{A(t)} \right) - \frac{1}{2} \gamma \sigma^2 \theta^2 \\ & + \inf_{a,b} \Bigg\{ \frac{1}{\phi} \lambda \left[1 + \left(a + \frac{1}{2} b^2 \sigma_J^2 - 1 \right) e^a + \beta (1 + (e^{a+b^2 \sigma_J^2} - 2) e^a) \right] \\ & + \frac{1}{1-\gamma} \lambda e^a (E^{Z(b)} [(1 + (e^Z - 1)\theta)^{1-\gamma}] - 1) \Bigg\} \Bigg\} = 0. \end{aligned}$$

The first-order conditions in θ, a, and b give Equations (15–17), respectively. Substituting the solutions θ^*, a^*, and b^* back to equation (B.3), we obtain the ordinary differential equation for $f(t)$,

$$\begin{aligned} & \frac{\gamma}{1-\gamma} \frac{1+f'(t)}{f(t)} - \frac{\rho}{1-\gamma} + r + \theta^* \left(\mu - r + \frac{A'(t)+1}{A(t)} \right) - \frac{1}{2} \gamma \sigma^2 (\theta^*)^2 \\ & + \frac{1}{\phi} \lambda \left[1 + \left(a^* + \frac{1}{2} (b^*)^2 \sigma_J^2 - 1 \right) e^{a^*} + \beta (1 + (e^{a^* + (b^*)^2 \sigma_J^2} - 2) e^{a^*}) \right] \\ & + \frac{1}{1-\gamma} \lambda e^{a^*} (E^{Z(b^*)} [(1 + (e^Z - 1)\theta^*)^{1-\gamma}] - 1) = 0 \end{aligned} \tag{B.4}$$

$\blacksquare$

Proof of Proposition 3. Applying the equilibrium condition $\theta = 1$ to the first-order conditions (16) and (17), we immediately obtain Equations (20) and (21) for the optimal a^* and b^*.

Next, the equilibrium conditions of $S_t = W_t$ and $c_t = Y_t$ imply $A(t) = f(t)$. The ordinary differential Equation (B.4) becomes

$$A'(t) = \frac{A(t)}{\alpha} - 1, \tag{B.5}$$

where the constant coefficient α is defined by

$$\begin{aligned} \frac{1}{\alpha} = & \rho - (1-\gamma)\mu + \frac{\sigma^2}{2} \gamma (1-\gamma) - \lambda e^{a^*} (e^{(1-\gamma)(\mu_J + b^* \sigma_J^2) + \frac{1}{2}(1-\gamma)^2 \sigma_J^2} - 1) \\ & - \frac{1-\gamma}{\phi} \lambda [1 + (a^* + \frac{1}{2} (b^*)^2 \sigma_J^2 - 1) e^{a^*} + \beta (1 + (e^{a^* + (b^*)^2 \sigma_J^2} - 2) e^{a^*})]. \end{aligned} \tag{B.6}$$

Under the terminal condition $A(T) = 0$, $A(t)$ can be solved uniquely,

$$A(t) = \alpha \left(1 - \exp\left(-\frac{T-t}{\alpha} \right) \right).$$

The first-order condition (15) evaluated at $\theta = 1$ gives,

$$\mu + \frac{1}{\alpha} = r + \gamma \sigma^2 - \lambda e^{a^*} \left(e^{(1-\gamma) b^* \sigma_J^2 + (1-\gamma)^2 \frac{\sigma_J^2}{2} + (1-\gamma)\mu_J} - e^{-\gamma b^* \sigma_J^2 + \gamma^2 \frac{\sigma_J^2}{2} - \gamma \mu_J} \right). \tag{B.7}$$

Using Equations (B.5) and (B.7), it is a straightforward calculation to show that the equity premium (cum-dividend) and the risk-free rate are as given in Equations (18) and (22).

Finally, to see that π is indeed a pricing kernel, one can first show, via a straightforward deviation, that π produces the equilibrium risk-free rate and the total equity premium for the stock. Next, one can solve the same equilibrium problem by adding a derivative security (nonlinear in stock) with zero net supply and show that the equilibrium risk premium for the derivative security can indeed be produced by π. ■

Appendix C: The Option-Pricing Formula

The following result can be found in Merton (1976), and is included for the completeness of the article. Let C_0 denote the time-0 price of a European-style call option with exercise price K and time τ to expiration. It is a straightforward derivation to show that

$$C_0 = e^{-\lambda'\tau} \sum_{j=0}^{\infty} \frac{(\lambda'\tau)^j}{j!} BS(S_0, K, r_j, q, \sigma_j, \tau) \tag{C.8}$$

where $\lambda' = \lambda^Q (1 + k^Q)$, and for $j = 0, 1, \ldots,$

$$r_j = r - \lambda^Q k^Q + \frac{j \ln(1 + k^Q)}{\tau}, \quad \sigma_j^2 = \sigma^2 + \frac{j\sigma_J^2}{\tau},$$

and where $BS(S_0, K, r, q, \sigma, \tau)$ is the standard Black-Scholes option pricing formula with initial stock price S_0, strike price K, risk-free rate r, dividend yield q, volatility σ, and time τ to maturity. To price a put option with the same maturity and strike price, one can use the put-call parity.

Appendix D: The Case of Recursive Utility

This appendix provides the equilibrium pricing kernel and the asset-pricing implication for an agent with a continuous-time Epstein and Zin (1989) recursive utility facing the endowment process Y defined in Equation (1).

D.1. Stochastic differential utility

For a consumption process c, the representative agent's utility U is determined by

$$U = V_0,$$

with

$$V_t = E_t\left[\int_t^T f(c_s, V_s)ds\right]. \tag{D.9}$$

As pointed out by Duffie and Epstein (1992b) (p. 365), for the case of Brownian information, the above utility characterizes the continuous-time version of recursive utility; for non-Brownian information, such as the jump-diffusion case we are considering here, the above utility characterizes only a subclass of the continuous-time version of recursive utility. We will further specialize to the following case, which has the feature of separating intertemporal substitution from risk aversion [p. 420, Duffie and Epstein (1992a), see also Epstein and Zin (1989) for a discrete-time version],

$$f(c, V) = \frac{\rho}{1-\delta} \frac{c^{1-\delta} - ((1-\gamma)V)^{\frac{1-\delta}{1-\gamma}}}{((1-\gamma)V)^{\frac{1-\delta}{1-\gamma}-1}}, \tag{D.10}$$

where γ the risk-aversion coefficient and $1/\delta$ is the elasticity of intertemporal substitution. When $\delta = \gamma$, the utility function reduces to the standard time-and-state additive power utility and ρ is the constant discount rate.

D.2 The pricing kernel

Duffie and Epstein (1992b) and Duffie and Skiadas (1994) show that the pricing kernel for a single-agent economy with stochastic differential utility formulation defined by Equation (D.9) is given by

$$\pi_t = \exp\left(-\int_0^t f_V(c_s, J_s)ds\right)f_c(c_t, J_t), \tag{D.11}$$

where c is the agent's optimal consumption and J is his indirect utility function, and where f_c and f_V are first derivatives of f with respect to c and V, respectively. Applying this result to our case, and setting the optimal consumption to the endowment Y, we obtain the equilibrium pricing kernel,

$$\pi_t = \exp\left(\int_0^t \rho\frac{1-\gamma}{1-\delta}\left[\left(1-\frac{1-\delta}{1-\gamma}\right)\frac{Y_s^{1-\delta}}{((1-\gamma)J_s)^{\frac{1-\delta}{1-\gamma}}} - 1\right]ds\right)\rho Y_t^{-\delta}((1-\gamma)J_t)^{\frac{\delta-\gamma}{1-\gamma}}, \tag{D.12}$$

where the indirect utility function J is the solution to

$$J_t = E_t\left[\int_t^T f(Y_s, J_s)ds\right]. \tag{D.13}$$

To obtain the pricing kernel, we need to compute J, and we do so by conjecturing that

$$J_t = \frac{1}{1-\gamma}(l(t)Y_t)^{1-\gamma}, \tag{D.14}$$

where $l(t)$ is a deterministic function of time. Substituting Equation (D.14) into Equation (D.13), one can verify that the indirect utility is indeed of the conjectured form with $l(t)$ defined by

$$l(t) = \left(\rho\frac{e^{A(1-\delta)(T-t)} - 1}{(1-\delta)A}\right)^{\frac{1}{1-\delta}}, \tag{D.15}$$

where

$$A = -\frac{\rho}{1-\delta} + \frac{g}{1-\lambda}, \quad g = (1-\gamma)\left(\mu - \gamma\frac{\sigma^2}{2}\right) + \left(e^{(1-\gamma)\mu_J + \frac{(1-\gamma)^2}{2}\sigma_J^2} - 1\right). \tag{D.16}$$

Substituting J_t into Equation (D.12), we obtain the equilibrium pricing kernel for the economy,

$$\pi_t = \exp\left(\int_0^t \rho\frac{1-\gamma}{1-\delta}\left[\frac{\delta-\gamma}{1-\gamma}l(s)^{\delta-1} - 1\right]ds\right)\rho l(t)^{\delta-\gamma}Y_t^{-\gamma}. \tag{D.17}$$

D.3 Asset pricing

Setting $\delta = \gamma$ in Equation (D.17), we are back to the case of power utility, and the pricing kernel derived in Equation (D.17) for the recursive utility reduces to the familiar form[24]: $\pi_t = \rho\ \exp(-\rho t)\ Y_t^{-\gamma}$. Setting the power-utility case as the benchmark, we can see that the

[24] The constant coefficient ρ in π_t should not cause any confusion, since pricing kernels are determined only up to a multiplicative constant. Had we defined the normalized aggregator f in a slightly different form than that in Equation (D.10), this extra factor could have been taken care of. We chose to work with the current form of f, since it was the original form provided by Duffie and Epstein (1992a).

recursive utility results in a more complex risk-free rate, which could be time varying, through the pricing kernel's dependence on $l(t)$. Its effect on the market prices of endowment shocks, however, remains identical to the power-utility case. This can be seen from the fact that in both cases the pricing kernel depends on the endowment through $Y^{-\gamma}$.

This result is quite intuitive, given that the recursive utility introduced in Equation (D.10) is designed to separate intertemporal substitution from risk aversion. If our interest is on how the diffusive endowment shocks are priced differently from the rare events, we need look no further than the special case of power utility, which captures the risk-aversion component of the recursive utility. The intertemporal component of the recursive utility does affect preferences, but only through the risk-free rate.

From this, one can already obtain an intuitive understanding on how the recursive utility will affect option pricing. Specifically, the diffusive shocks and rare events will be priced in a very similar fashion to the benchmark case of power utility. As we discussed earlier, the power-utility case is not adequate to explain the observed pricing kernel in the empirical literature. In particular, it cannot generate the level of differential pricing of the diffusive Brownian shocks vs. the shocks due to rare events. In this respect, although the recursive utility has two preference parameters γ and δ, which bring it to equal footing with the uncertainty aversion case considered in this article, the recursive utility is not capable of generating the type of pricing kernels consistent with those reported in the empirical literature.

D.4 An explicit example

To be more concrete, we work out an explicit example by considering an economy with an infinite horizon ($T \to \infty$). For the purpose of option pricing, this is an appropriate consideration, since the maturity of an option is typically short compared with the lifespan of the economy. This specialization gives us the added convenience of a constant riskfree rate, since, when $A < 0$, we have

$$\lim_{T\to\infty} l(t) = \left(-\frac{\rho}{(1-\delta)A}\right)^{1/(1-\delta)} \tag{D.18}$$

Applying Ito's lemma to the pricing kernel (D.17) specialized for this economy, we have

$$\frac{d\pi_t}{\pi_t} = -rdt - \gamma\sigma dB_t + (e^{-\gamma Z} - 1)dN_t - \lambda(e^{-\gamma\mu_J + \frac{1}{2}\gamma^2\sigma_J^2} - 1)dt, \tag{D.19}$$

where σ, μ_J, σ_J, λ are the diffusion and jump parameters affecting the endowment process (see Section 1), and where the risk-free rate r can be derived:

$$r = \rho + \gamma\mu - \frac{\gamma(1+\gamma)}{2}\sigma^2 - \lambda(e^{-\gamma\mu_J + \frac{1}{2}\gamma^2\sigma_J^2} - 1) + \left(1 - \frac{1-\delta}{1-\gamma}\right)g, \tag{D.20}$$

where μ is the drift parameter for the endowment process (see Section 1) and g is as defined in Equation (D.16). From this, we can see that the elasticity of intertemporal substitution affects the pricing kernel only through its affect on the risk-free rate. This is consistent with our general discussion in Section D.3.

For the purpose of option pricing, let us consider a stock that has the same set of risk exposures as the endowment shock and pays out dividend at a constant rate of q:

$$\frac{dS_t}{S_t} = \mu_s dt + \sigma dB_t + (e^{Z_t} - 1)dN_t - \lambda k dt, \tag{D.21}$$

where k is the mean percentage jump size as defined in Section 1 and μ_s is the ex-dividend expected stock return. In order to determine the equilibrium expected equity return μ_s and option prices, we take advantage of risk-neutral pricing. Under the risk-neutral measure Q defined by the equilibrium pricing kernel π, it must be that

$$\frac{dS_t}{S_t} = (r-q)dt + \sigma dB_t^Q + (e^{Z_t} - 1)dN_t^Q - \lambda^Q k^Q dt, \tag{D.22}$$

where r is the risk-free rate solved in Equation (D.20), q is the dividend payout ratio, and λ^Q and k^Q are the risk-neutral jump intensity and the mean percentage jump size, respectively. Using Equations (D.19), (D.21), (D.22), and the Girsanov theorem, it is straightforward to show that

$$\lambda^Q = \lambda \exp\left(-\gamma\mu + \frac{1}{2}\gamma^2\sigma_J^2\right), \quad k^Q = (1+k)\exp(-\gamma\sigma_J^2) - 1, \tag{D.23}$$

and

$$\mu_s = r - q + \gamma\sigma^2 + \lambda k - \lambda^Q k^Q. \tag{D.24}$$

In terms of pricing risky assets (equity and options), the recursive utility provides exactly the same market prices of risks as the power utility. This can be seen by comparing the above results to those reported in Proposition 3, Equations (18) and (19) for an investor without uncertainty aversion.

Appendix E: On Normalization

In this appendix, we examine the economic impact of the normalization factor ψ of our utility specification and investigate the economic driving force behind our main result: normalization vs. uncertainty aversion.

It is clear that in a one-period model, the choice of the normalization factor amounts to constant scaling and will not affect the model-uncertainty aspect of the utility specification. In a multiperiod setting, the normalization does play a role in affecting preferences. To demonstrate that it is indeed uncertainty aversion not normalization that is driving our result, we provide the following concrete example, which extends our setting by allowing for a general form of the normalization factor.

Consider a representative agent who maximizes his utility in the following discrete-time setting,

$$U_t = u(c_t) + e^{-\rho}\min_\xi\{\Lambda(E_t^\xi[v(\xi_{t+1})], E_t^\xi[U_{t+1}]) + E_t^\xi[U_{t+1}]\}, \tag{E.25}$$

where $\lambda \geq 0$ and $\Lambda(0, \cdot) = 0$. Mapping back to our utility specification in Equation (4), we have $u(c) = c^{1-\gamma}/1-\gamma)$ and

$$\Lambda(E_t^\xi[v(\xi_{t+1})], E_t^\xi[U_{t+1}]) = \frac{1}{\phi}\psi(E_t^\xi[U_{t+1}])E_t^\xi[v(\xi_{t+1})], \tag{E.26}$$

where $\psi(E_t^\xi[U_{t+1}])$ is the normalization factor, $E_t^\xi[v(\xi_{t+1})]$ is the discrepancy or distance measure, and ϕ is the uncertainty aversion parameter.

Assuming the existence of an optimal solution to the agent's problem, we let ξ^* be the Radon–Nikodym derivative that defines the optimal alternative measure $P(\xi^*)$, and c^* be the optimal consumption, which, in our setting, equals the representative agent's endowment in equilibrium. We will show, at the end of this appendix, that the pricing kernel is of the form

$$\pi_{t+1} = \pi_t e^{-\rho}(\Lambda_{U_t}^* + 1)\xi_{t+1}^* \frac{u'(c_{t+1}^*)}{u'(c_t^*)} \quad \text{and} \quad \pi_0 = 1, \tag{E.27}$$

where we use the notation

$$\Lambda^*_{Ut} = \Lambda_U(E_t^{\xi^*}[v(\xi^*_{t+1})], E_t^{\xi^*}[J_{t+1}]), \tag{E.28}$$

where Λ_U is the derivative of Λ with respect to the second argument of Λ and J is the indirect utility function. Note that in Anderson, Hansen, and Sargent (2000), $\Lambda^*_{Ut} = 0$, while in our case,

$$\Lambda^*_{Ut} = \frac{1-\gamma}{\phi} E_t^{\xi^*}[v(\xi^*_{t+1})].$$

The pricing kernel (E.27) can in fact provide quite a general understanding of the asset-pricing implication of our utility specification. Broadly speaking, premiums associated with risk aversion are incorporated through $u'(c^*_{t+1})/u'(c^*_t)$ and premiums associated with uncertainty aversion are incorporated through ξ^*_{t+1}. The normalization factor shows up in the pricing kernel via Λ^*_{Ut}, which is known at time t and can only affect the risk-free rate. In other words, there is no direct impact of the normalization factor on the market prices of risk or uncertainty (of course, it does have an indirect effect through the optimal ξ).

Recall that the main result of our article builds on a nontrivial solution of ξ^*_{t+1} and its presence in the pricing kernel. From the above analysis, we can see that the driving force for this result is clearly the minimization part of the utility specification, which is motivated by uncertainty aversion. Specifically, taking away the minimization part of Equation (E.25) by not allowing investors to choose alternative measures, we will have a trivial solution of $\xi^* \equiv 1$, regardless of the choice of the normalization factor. On the other hand, taking away the normalization factor, Λ no longer depends on $E_t^{\xi}[U_{t+1}]$, and we have $\Lambda_U = 0$. The exact functional form of the optimal ξ^*_{t+1} might become more complicated (to the extent that we will not be able to solve our problem in closed form), we will still have a nontrivial ξ^*_{t+1} present in the pricing kernel. In other words, the fact that our main result builds on uncertainty aversion is not affected in any qualitative fashion by the choice of normalization.

The pricing kernel in Equation (E.27) can also help us obtain some intuition regarding the equivalence result between robust control and recursive utility. For example, in Maenhout (2001), the setting is that of a pure diffusion with only one Brownian motion. We have $\xi^*_t = \exp(-a^* B_t - 1/2(a^*)^2 t)$, for some optimal value of a^*. Given that the Brownian motion is the only shock driving the aggregate endowment $Y_t = Y_0 \exp(\sigma B_t - 1/2\,\sigma^2\, t + \mu t)$, it is easy to show that the random component of ξ^*_{t+1} can be written as $Y_{t+1}^{-\theta}$ where $\theta = a^*/\sigma$. In addition, the marginal utility contributes $u'(Y_{t+1}) = Y_{t+1}^{-\gamma}$ to the pricing kernel. Combining the two, the component of the pricing kernel associated with risk aversion is of the form $Y_{t+1}^{-(\gamma+\theta)}$, while the component associated with intertemporal substitution is of the form $1 + \Lambda^*_{Ut}$. From this, we can see the possibility of, for the purpose of asset pricing, this specific robust-control problem being equivalent to a recursive utility with the risk-aversion coefficient $\gamma + \theta$. In a more general setting with multiple sources of random shocks, however, the optimal ξ^*_t might not be a function Y_t. For example, in our setting, ξ^* picks up only the Poisson component, not the Brownian component. For such cases, the equivalence to standard recursive utility does not generally apply.

Finally, we show that the pricing kernel is indeed of the form in Equation (E.27). Let J be the indirect utility function, which is a function of the state variables including the wealth process W and other state variables X affecting the endowment process $J_t = J(t, W_t, X_t)$. In the following analysis, we will suppress arguments t and X_t in function J for notational simplicity. The principle of optimality implies[25]

$$J(W_t) = \max_{c_t}\left\{u(c_t) + e^{-\rho}\min_{\xi}\{\Lambda(E_t^{\xi}[v(\xi_{t+1})], E_t^{\xi}[J(W_{t+1})]) + E_t^{\xi}[J(W_{t+1})]\}\right\}, \tag{E.29}$$

[25] We suppress the portfolio part of the optimization problem to focus on the pricing kernel. It should be clear that our results will not be affected.

where, given any security with time $t+1$ return denoted by R_{t+1}, we have,

$$W_{t+1} = (W_t - c_t)R_{t+1}. \tag{E.30}$$

The first-order condition of Equation (E.29) for c_t gives

$$u'(c_t^*) - e^{-\rho}(\Lambda_U(E_t^{\xi^*}[v(\xi_{t+1}^*)], E_t^{\xi^*}[J_{t+1}]) + 1)E_t^{\xi^*}(J_W(W_{t+1})R_{t+1}) = 0, \tag{E.31}$$

where J_W denotes the derivative of the indirect utility J with respect to wealth W. Using the fact that both $u'(c_t^*)$ and $\Lambda_{Ut}^* = \Lambda_U(E_t^{\xi^*}[v(\xi_{t+1}^*)], E_t^{\xi^*}[J_{t+1}])$ are in the time-t information set, we can rewrite the above first-order condition as

$$1 = E_t^{\xi^*}\left(e^{-\rho}\frac{\Lambda_{Ut}^* + 1}{u'(c_t^*)}J_W(W_{t+1})R_{t+1}\right) = E_t\left(e^{-\rho}\frac{\Lambda_{Ut}^* + 1}{u'(c_t^*)}\xi_{t+1}^* J_W(W_{t+1})R_{t+1}\right). \tag{E.32}$$

Using the envelope theorem, we have

$$J_W(W_{t+1}) = u'(c_{t+1}^*). \tag{E.33}$$

Using the above results, we can now verify that the π defined in Equation (E.27) is indeed a valid pricing kernel in that,

$$E_t\left(\frac{\pi_{t+1}}{\pi_t}R_{t+1}\right) = 1, \tag{E.34}$$

for any security with return R_{t+1}.

References

Andersen, P. K., O. Borgan, R. Gill, and N. Keiding, 1992, *Statistical Models Based on Counting Processes*, Springer-Verlag, New York.

Anderson, E., L. Hansen, and T. Sargent, 2000, "Robustness, Detection and the Price of Risk," working paper, University of North Carolina, University of Chicago, and Stanford University.

Bansal, R., A. R. Gallant, and G. Tauchen, 2002, "Rational Pessimism, Rational Exuberance, and Markets for Macro Risks," working paper, Duke University.

Bates, D., 2001, "The Market Price of Crash Risk," working paper, University of Iowa.

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

Campbell, J. Y., and J. H. Cochrane, 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Chen, Z., and L. Epstein, 2002, "Ambiguity, Risk and Asset Returns in Continuous Time," *Econometrica*, 70, 1403–1443.

Comon, E., 2000, "Extreme Events and the Role of Learning in Financial Markets," working paper, Harvard University.

Das, S., and R. Uppal, 2001, "Systemic Risk and International Portfolio Choice," working paper, Santa Clara University and London Business School.

David, A., and P. Veronesi, 2000, "Option Prices with Uncertain Fundamentals," working paper, Graduate School of Business, University of Chicago.

Duffie, D., and L. Epstein, 1992a, "Asset Pricing with Stochastic Differential Utility", *Review of Financial Studies*, 5, 411–436.

Duffie, D., and L. Epstein, 1992b, "Stoachstic Differential Utility", *Econometrica*, 60, 353–394.

Duffie, D., and C. Skiadas, 1994, "Continuous-Time Security Pricing: A Utility Gradient Approach", Journal of Mathematical Economics, 23, 107–132.

Dufresne, P. C., and J.-N. Hugonnier, 2001, "Event Risk, Contingent Claims and the Temporal Resolution of Uncertainty," working paper, Carnegie Mellon University

Elliott, R., 1982, *Stochastic Calculus and Applications*, Springer-Verlag, New York.

Ellsberg, D., 1961, "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

Epstein, L., and J. Miao, 2003, "A Two-Person Dynamic Equilibrium Under Ambiguity," *Journal of Economic Dynamics and Control*, 27, 1253–1288.

Epstein, L., and M. Schneider, 2002, "Learning Under Ambiguity", working paper, University of Rochester.

Epstein, L., and T. Wang, 1994, "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62, 283–322.

Epstein, L., and S. Zin, 1989, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Gilboa, I., and D. Schmeidler, 1989, "Maxmin Expected Utility Theory with Non-unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

Hansen, L., and T. Sargent, 2001, "Robust Control and Model Uncertainty," *American Economic Review*, 91, 60–66.

Jackwerth, J. C., 2000, "Recovering Risk Aversion from Options Prices and Realized Returns", *Review of Financial Studies*, 13, 433–451.

Jackwerth, J. C., and M. Rubinstein, 1996, "Recovering Probability Distributions from Option Prices", *Journal of Finance*, 51, 1611–1631.

Knight, F., 1921, "Risk, Uncertainty and Profit," Houghton, Mifflin, Boston.

Knox, T., 2002, "Learning How to Invest when Returns are Uncertain", working paper, University of Chicago.

Liu, J., F. Longstaff, and J. Pan, 2003, "Dynamic Asset allocation with Event Risk," *Journal of Finance*, 58, 231–259.

Liu, J., and J. Pan, 2003, "Dynamic Derivative Strategies," *Journal of Financial Economics*, 69, 410–430.

Lucas, R. E., 1978, "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–1445.

Maenhout, P., 2001, "Robust Portfolio Rules, Hedging and Asset Pricing," working paper, Finance Department, INSEAD.

Mehra, R., and E. Prescott, 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Merton, R., 1976, "Option Pricing When Underlying Stock Returns are Discontinuous", *Journal of Financial Economics*, 3, 125–44.

Naik, V., and M. Lee, 1990, "General Equilibrium Pricing of Options on the Market Portfolio with Discontinuous Returns," *Review of Financial Studies*, 3, 493–521.

Pan, J., 2002, "The Jump-Risk Premia Implicit in Option Prices: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3–50.

Pathak, P., 2002, "Notes on Robust Portfolio Choice", working paper, Harvard University.

Rietz, T., 1988, "The Equity Risk Premium: A Solution," *Journal of Monetary Economics*, 22, 117–131.

Routledge, B., and S. Zin, 2002, "Model Uncertainty and Liquidity," working paper, Carnegie Mellon University.

Rubinstein, M., 1994, "Implied Binomial Trees", *Journal of Finance*, 49, 771–818.

Sill, K., 2000, "Understanding Asset Values: Stock Prices, Exchange Rates, and the 'Peso Problem'," *Business Review, Federal Reserve Bank of Philadelphia*, September/October, 3–13.

Skiadas, C., 2003, "Robust Control and Recursive Utility", *Finance and Stochastics*, 7, 475–489.

Uppal, R., and T. Wang, 2003, "Model Misspecification and Under-Diversification," *Journal of Finance*, 58, 2465–2486.

Wang, T., 2003, "A Class of Multi-Prior Preferences," University of British Columbia.

Yan, H., 2000, "Uncertain Growth Prospects, Estimation Risk, and Asset Prices," working paper, University of Texas at Austin.