Tail Risk, Robust Portfolio Choice, and Asset Prices*

Xing Jin†, Dan Luo‡, and Xudong Zeng§

February 16, 2017

Abstract

Equity jumps and consumption disasters exhibit slowly-decaying tail behavior admitting severe downside risk; moreover, heavy-tailed distributions governing these rare events are most challenging to estimate. This paper formulates and solves in closed form a portfolio choice problem in a multi-asset incomplete market characterized by ambiguous jumps. We find that, due to fear of tail incidents, an investor diminishes portfolio diversification, and more so under heavy-tailed jumps which intensify misspecification concerns. In the presence of jump ambiguity, calibration exercises show sizable wealth losses from underestimating tail risk and that heavy-tailed consumption disasters effectively help explain the variance premium and option prices.

JEL classification: G01, G11, G12
Keywords: Tail risk, Jump ambiguity, Portfolio choice, Nonparticipation, Asset pricing

---

*We thank Andrew Carverhill, Hua He, Qingfu Liu, and seminar participants of the 2016 China Finance Review International conference, the 2016 China Finance Annual Meeting, and the 2nd Shanghai Risk Forum for helpful comments.

†Warwick Business School, Coventry, CV4 7AL, UK; E-mail: Xing.Jin@wbs.ac.uk.
‡School of Finance, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, 200433, China; E-mail: huo.dan@mail.shufe.edu.cn.
§School of Finance, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, 200433, China; E-mail: zeng.xudong@mail.shufe.edu.cn.
1 Introduction

Since the seminal work of Merton (1971), numerous studies have demonstrated the abrupt and substantial impact of jump risk on optimal portfolio selection. Aware of misspecifications in an imprecisely-estimated rare jump model, an investor would make conservative investment decisions to ensure reasonable portfolio performance across the reference model and nearby models\(^1\) (see, e.g., Liu, Pan, and Wang (2005), Jin and Zhang (2012), and Drechsler (2013)). Recently, Barro and Jin (2011) and Bollerslev and Todorov (2011a) document that consumption/return jump sizes follow slowly-decaying distributions with heavy tails admitting severe tail events, which are formally outside of the traditional framework of normally distributed jumps with only light tails. However, the effect of jump tail behavior on optimal portfolio selection, which has been greatly emphasized since the financial crisis, still lacks close examination. Meanwhile, Barro (2006) and Gabaix (2012) show that a rare disaster model calibrated to international consumption data can explain a wide range of asset pricing puzzles,\(^2\) but we still lack a proper understanding of the effect of tail properties of consumption disasters on asset prices.

The goal of this paper is to investigate the effects of jump tail behavior on optimal portfolio formation and asset pricing, when an investor explicitly acknowledges model misspecifications regarding the jump distribution. This consideration of jump ambiguity is especially relevant here because heavy-tailed jump models are most challenging to pin down. The investor implements robust decision-making (e.g., Hansen and Sargent (2008)) to guard against the worst case alternative model. We show that the heavy tail of a slowly-decaying jump distribution gets largely exacerbated under the worst case

---

\(^1\)The reference model is the “best” estimate from data and nearby/alternative models are those that are statistically difficult to separate from the reference model.

scenario and, in reaction, the investor diminishes the jump exposure and diversification of her optimal portfolio. When calibrated to international equity indices, our model confirms the above theoretical findings and indicates sizeable welfare losses of an ambiguity averse investor from ignoring extreme tail events of the systemic jump. And in market equilibrium, ambiguity aversion towards heavy-tailed consumption disasters effectively induces the otherwise puzzlingly high variance and option smirk premia.

We first study a portfolio choice problem in a multi-asset jump-diffusion market which is incomplete. An investor is averse to ambiguity in the jump size distribution (and jump frequency). Following the robust control framework of Anderson, Hansen, and Sargent (2003), we solve a Hamilton-Jacobi-Bellman equation altered with a penalty term and with a minimization over all alternative models to obtain the optimal portfolio under ambiguity aversion. As any nearby model could potentially be the “true” model, we consider the entire “neighborhood” of alternative models using a nonparametric method. Our “worst case” is thus worse than those in the literature which confine to certain parametric alternative models. Furthermore, our nonparametric method readily applies to any jump distributions, either slowly-decaying power laws or normal, with equal efforts. We solve for the worst case probability and the optimal portfolio both in closed forms through a novel decomposition approach. The great transparency of our method allows us to easily track the investor’s aversion to potential model misspecifications when the reference models are sufficiently distinct.

Exploiting the closed-form solutions, we find: (1) Due to the fear of severe tail events in the worst case scenario, an investor reduces her jump exposure when she becomes more ambiguity averse. (2) Ambiguity aversion towards systemic jumps diminishes portfolio diversification. (3) An extremely ambiguity averse investor may choose zero exposure to jump risk hence not to participate in the financial market. This result provides an alternative explanation for the nonparticipation puzzle (e.g., Campbell (2006)).\(^3\) (4) An

\(^3\)Campbell (2006) finds that a significant portion of households do not hold any risky assets; however, modern asset allocation theories suggest that they invest in all available risky assets. Cao, Wang,
ambiguity averse investor reduces more of her jump exposure if the jump distribution exhibits a fatter left tail, which further diminishes portfolio diversification. ⁴

Due to the high impact of tail events, it is of particular importance to quantify the effects of tail risk on the investor’s portfolio holdings and economic welfare. For this purpose, we calibrate our model to an economy consisting of seven international equity indices. We arm the investor with a constant relative risk aversion (CRRA) utility function and solve the optimal portfolio choice problem with two different jump size distributions: Merton’s normal distribution which is unable to capture heavy jump tails and the power law distribution proposed by Barro and Jin (2011) which exhibits a heavier tail than the normal. To gain a clear understanding, we require the latter heavy-tailed distribution to have the same first two moments as the normal distribution. As we explain below, this arrangement allows any discrepancies in the portfolio outcomes to be primarily ascribed to the jump tail behavior, which is the focus of this study.

The optimal portfolio strategies corresponding to the two jump size distributions are notably separable when the investor is averse to jump ambiguity. For example, at the 20-year investment horizon, the investor’s total jump exposure under moderate risk aversion is 26.7% with uncertain normal jumps and 17.6% with uncertain heavy-tailed jumps. In economic terms, failing to accommodate tail fatness leads to a 9.3% loss in the investor’s certainty equivalent wealth under moderate risk aversion. This loss reaches as high as 35.5% under relatively low risk aversion.

Expecting the distinct statistic properties of the two jump size distributions, we determine the investor’s degree of ambiguity aversion towards different jump models using the model detection error probability (DEP) approach of Anderson, Hansen, and Sargent (2003). The models are compared under the same DEP. We affirm that, given

---


and Zhang (2005) and Easley and O’Hara (2009) show that nonparticipation may arise from the rational decisions of traders with ambiguity aversion to uncertain expected returns or variance. We propose a different explanation through aversion to jump ambiguity.
the same DEP, the investor exhibits higher ambiguity aversion under the heavy-tailed distribution than under the normal distribution. Restricting the level of ambiguity aversion to be identical for the both distributions, we find that wealth equivalent losses are largely reduced. Hence, the above improvements in economic welfare may be primarily traced back to the inability of the investor to separate the alternative models from the reference model with heavy jump tails.

Surprisingly, the difference between the optimal portfolio strategies under the two jump size distributions is negligible in an expected utility framework with an ambiguity neutral investor. A possible explanation for this result lies in the locally mean-variance property of the CRRA utility which fails to capture higher moments (e.g., Hong, Tu, and Zhou (2007) and Cvitanić, Polimenis, and Zapatero (2008)). Because the first two moments of both jump size distributions are perfectly matched, the negligible difference in portfolio holdings may be anticipated when there is no jump ambiguity.\(^5\) On the contrary, under the worst case models with jump ambiguity, these moments clearly deviate under different tail assumptions leading to divergent portfolio holdings. Our results emphasize the essential importance of recognizing tail thickness for an investor in the presence of jump ambiguity.\(^6\)

We now turn to study the asset pricing implications of consumption jump distributions in a general equilibrium model. The aggregate consumption follows a jump-diffusion process. We adopt the single power law distribution for consumption disasters estimated in Barro and Jin (2011). Their calibrated consumption process can resolve the equity premium puzzle with a simple CRRA utility function and modest risk aversion, while they do not examine the variance premium and option prices. For comparison, we consider a normal distribution for consumption jumps (e.g., Liu, Pan, and Wang

\(^5\)Particularly, Cvitanić, Polimenis, and Zapatero (2008) demonstrate that, for a CRRA investor, a jump model will perform closely to a diffusion model with the same return to variance ratio, if volatility is not very high. Under our calibration, the two jump models share the same first two moments thus naturally generate similar portfolio outcomes to that of the same diffusion model.

\(^6\)Of course, a naïve investor without ambiguity aversion shall encounter much larger economic losses when the worst case scenario occurs.
that preserves the first two moments of the heavy-tailed power law. The results indicate that the variance premium is strengthened in the worst case of the heavy-tailed jump distribution over the normal distribution. Moreover, the heavy-tailed jump distribution also enhances the ability for the ambiguity model to generate steep option implied volatility smirk curves. Further collaborating the results in Liu, Pan, and Wang (2005), we find that, with ambiguity aversion, the standard deviation of the normal distribution has to be lifted up to 50% to generate the same premia under the power-law. Thus the slowly-decaying disaster distribution of Barro and Jin (2011), combined with the investor’s robust consumption decisions, provides an effective way to reconcile the high variance and smirk premia in the market data.

Our paper is closely related to recent studies that focus on estimating tail risk of asset returns (e.g., Barro and Jin (2011) and Bollerslev and Todorov (2011a)). Bollerslev and Todorov (2011b) find large time-varying compensation for fears of disasters which accounts for a large percentage of the average equity and variance risk premia. Kelly and Jiang (2014) and Bollerslev, Todorov, and Xu (2015) show that time variations in tail risk may explain cross-sectional expected stock returns and predictability in the aggregate market. Agarwal, Ruenzi, and Weigert (forthcoming) construct a new systematic tail risk measure for equity-oriented hedge funds and find that tail risk affects cross-sectional fund returns. Building on these works, we demonstrate in the present paper that jump tails are amplified in the worst-case scenario of the ambiguity averse investor and bears important portfolio choice and asset pricing consequences. The proposed nonparametric approach enables us to examine jump distributions with diverse tail assumptions.

For portfolio choice under jump ambiguity, Jin and Zhang (2012) use “fictitious” stocks to address a case such that the total number of Brownian motions and jumps is greater than the number of risky assets. Therefore, the optimal portfolio is not explicitly provided for this case in their study. We solve this case directly and explicitly using
a decomposition technique. Moreover, they do not provide a method to determine the worst case probability. Compared to the parametric approach of Liu, Pan, and Wang (2005), our nonparametric approach enables us to easily analyze different types of jump distributions possibly with heavy tails, besides the popular normal distribution. Furthermore, we consider a multi-asset model while they consider only one stock. Drechsler (2013) constructs an equilibrium model and shows that fundamentals and model uncertainty may explain option price and the variance premium. He adopts parametric alternative models and relies on an analytical approximation to the solution, instead of obtaining a precise solution by using a nonparametric approach as in the present paper. Branger and Larsen (2013) show pronounced differences between ambiguity aversion with respect to diffusion risk and jump risk. They do not assume ambiguity in jump size and consider a single stock model. All the prior studies do not analyze the role of jump tail behavior in the presence of ambiguity aversion.

We solve for the worst case probability explicitly through the nonparametric method, and then for the optimal portfolio in closed form through a novel decomposition into two components representing exposures to diffusion risk and jump risk, respectively. These two components may be solved independently with remarkable reduction of computational complexity, especially for cases with a large number of risky assets. For the modeling of the market and methodology to solve for the optimal portfolio, our paper is related to previous studies regarding portfolio choice in a multi-asset jump-diffusion market. Das and Uppal (2004) show that systemic jumps reduce the gains from international diversification. Their model is absent of model uncertainty. To obtain explicit portfolio policies, we develop a novel decomposition approach that does not rely on a special structure of the model parameters, which is critical for the approach proposed by Aït-Sahalia, Cacho-Diaz, and Hurd (2009).

The remainder of the paper is organized as follows. Section 2 introduces Merton’s dynamic portfolio choice problem extended with ambiguity aversion. Section 3 derives
the worst case probability and the optimal portfolio in closed forms. Section 4 studies the effects of tail risk on portfolio choice and explain nonparticipation as implications of our solution. Section 5 is devoted to a calibration exercise to evaluate the investor’s fear of uncertain jump tails. Section 6 examines the asset pricing implications of uncertain consumption disasters. Section 7 concludes. All proofs are collected in the appendices.

2 Merton’s problem and ambiguity aversion

This section formulates a portfolio choice problem with ambiguity and ambiguity aversion in a continuous-time incomplete financial market. We fix a complete probability space \((\Omega, \mathcal{F}, P)\), where \(\Omega\) is the set of states of nature with generic element \(\omega\), \(\mathcal{F}\) is the \(\sigma\)-algebra of observable events, and \(P\) is a probability measure on \((\Omega, \mathcal{F})\). The market includes \(m + 1\) assets traded continuously on the time horizon \([0, T]\). One risk-free asset, called a bond, pays a risk-free interest rate \(r\). The remaining \(m\) assets, called stocks, are risky; their prices are modelled by the linear stochastic differential equation:

\[
\frac{dS_{i,t}}{S_{i,t}} = \mu_i dt + \sum_{j=1}^{m} \sigma_{i,j} dB_{j,t} + \sum_{k=1}^{n} J_{i,k} Y_{k,t} dN_{k,t}, \quad i = 1, 2, ..., m,
\]

where \(B_t = (B_{1,t}, ..., B_{m,t})'\) is an \(m\)-dimensional standard Brownian motion, and \(N_t = (N_{1,t}, ..., N_{n,t})'\) is an \(n\)-dimensional multivariate Poisson process. The symbol ‘\(\)' denotes transposition of a vector. \(N_{k,t}\) counts the number of type \(k\) jumps in the stock price up to time \(t\) with a constant intensity \(\lambda_k\). The amplitude \(Y_{k,t}\) of the type \(k\) jump has the probability density \(\Phi_k(t, dy)\). We examine mixed jumps in this study hence \(Y_{k,t}\) takes value from \((-1, \infty)\).

The diffusion coefficient \(\sigma_{i,j}\) and jump scale \(J_{i,k}\) are all constants. Define the diffusion coefficient matrix \(\Sigma = (\sigma_{i,j})_{1 \leq i,j \leq m}\) and the jump coefficient matrix \(J = (J_{i,k})_{1 \leq i \leq m, 1 \leq k \leq n}\). \(J_{i,k} \in [0, 1]\) is a jump scaling coefficient for each \(i, k\). We focus on the realistic case of \(n \leq m\), i.e., the number of jumps is not greater than the number of risky assets. Without
loss of generality, we assume $\text{rank}(\Sigma) = m$ and $\text{rank}(J) = n$.

The Brownian motions represent frequent small movements in the stock prices, while the jump processes capture occasional large shocks to the market. Both $B_t$ and $N_t$ are defined on the probability space $(\Omega, \mathcal{F}, P)$. The flow of information in the economy is given by the natural filtration, i.e., the right-continuous and augmented filtration $\{\mathcal{F}_t\}_{t \in [0,T]} = \{\mathcal{F}^B_t \vee \mathcal{F}^N_t, t \in [0,T]\}$, where $\mathcal{F}^B_t = \sigma(B_s; 0 \leq s \leq t)$, and $\mathcal{F}^N_t = \sigma(N_s; 0 \leq s \leq t)$. Observable events are eventually known, i.e., $\mathcal{F} = \mathcal{F}_T$.

We consider an investor armed with the utility function $U(x)$ and endowed with initial wealth $w_0$; this wealth will be invested in the above-mentioned $m+1$ assets. Let $\pi_t = (\pi_{1,t}, ..., \pi_{m,t})'$ denote a portfolio, where $\pi_{i,t}$ ($1 \leq i \leq m$) is the proportion of total wealth invested in the $i$-th risky asset at time $t$ and is $\mathcal{F}_t$-predictable. Any portfolio policy $\pi_t$ has an associated wealth process $W_t$ that evolves as:

$$W_t = W_0 + \int_0^t rW_s ds + \int_0^t W_s \pi'_s (\mu - r1_m) ds$$
$$+ \int_0^t W_s \pi'_s \Sigma dB_s + \int_0^t W_s - \pi'_s J Y_t dN_s,$$

where $1_m$ denotes the $m$-dimensional column vector of ones, and $Y_t$ is a diagonal matrix with diagonal entries $Y_{1,t}, ..., Y_{n,t}$.

A portfolio rule $\pi_t$ is said to be admissible if the corresponding wealth process satisfies $W_t \geq 0$ almost surely. We use $\mathcal{A}(w_0)$ to denote the set of all admissible portfolios, given initial wealth $W_0 = w_0$. Since we consider mixed jumps, that is, $Y_{k,t}$ may take any value from $(-1, \infty)$, an admissible portfolio $\pi$ must satisfy $\pi' J_k \in [0, 1]$ for each $k$, where $J_k$ is the $k$-th column of $J$.

In the traditional Merton’s portfolio choice problem without ambiguity aversion, the investor attempts to maximize expected utility from terminal wealth:

$$u(w_0) = \max_{\pi \in \mathcal{A}(w_0)} E[U(W_T)],$$
where $U(x)$ is nondecreasing and concave on $\mathcal{R} = (-\infty, \infty)$, $\mathbb{E}[\cdot]$ denotes expectation under the reference probability measure $P$, and $u(\cdot)$ is the value function.

Now we extend Merton’s problem to incorporate ambiguity aversion. Suppose the investor fears possible model misspecifications and makes investment decisions to guard against the worst case scenario. Specifically, in our model, the rare disasters are typically high impact events, while the parameters of the underlying jump processes are difficult to estimate with adequate accuracy due to scarcity of data. We therefore focus on the investor’s ambiguity aversion with regard to uncertain jump parameters.\(^7\) In other words, the investor’s problem stems from a class of prior models generated by imprecise estimates of the jump parameters governing the jump size distribution and jump intensity. The investor considers the point estimate and the corresponding reference model to be the most reliable, while she also explicitly recognizes that the competing models are difficult to distinguish statistically from the reference model. As a result, the investor makes a precautionary portfolio choice to guard against the competing alternatives such that her portfolio would perform reasonably well even if the worst case scenario occurs. However, choosing any model other than the reference model is penalized because the choice represents a deviation from the most likely model.

We introduce a set of probability measures, denoted by $\mathcal{P}$, that specify alternative models of concern. Let $P$ be the probability measure associated with the reference model. Each probability measure $P(\zeta) \in \mathcal{P}$ has a Radon-Nikodym derivative with respect to $P$:

$$
\frac{dP(\zeta)}{dP} = \zeta_T = \prod_{k=1}^{n} \zeta_{k,T},
$$

\(^7\)As in Liu, Pan, and Wang (2005), the diffusion parameters are less of our concern since the corresponding decision-making can be founded on abundant daily fluctuations of the asset prices.
where $\zeta_{k,T}$ is modeled by the stochastic differential equation:

$$
\mathrm{d}\zeta_{k,t} = \mathrm{d}\zeta_{k,t} (\vartheta_k, \phi_k) = \int_A (\vartheta_k(t) \phi_k(t,y) - 1) \zeta_{k,t} \Phi_k(t,dy) \mathrm{d}t, \quad t \in [0,T],
$$

with $\zeta_{k,0} = 1$ and $q(dt,dy) = (q_1(dt,dy), ..., q_n(dt,dy))$ where $q_k(dt,dy) = \mathrm{d}N_k(t) - \lambda_k \Phi_k(t,dy) dt$, $k = 1, ..., n$. $\vartheta_k(t)$ and $\phi_k(t,y)$ are positive stochastic processes, and $\phi_k(t,y)$ satisfies the following relationship:

$$
\int_A \phi_k(t,y) \Phi_k(t,dy) = 1, \quad k = 1, ..., n,
$$

where $A = (-1, \infty)$ is the support of the $k$-th jump size $Y_{k,t}$.

Under the probability measure $P(\zeta)$, the $k$-th jump intensity $\lambda_k$ and the density function $\Phi_k(t,dy)$ are changed into $\vartheta_k \lambda_k$ and $\phi_k(t,y) \Phi_k(t,dy)$ in the alternative model, respectively. (See, e.g., Theorem T10 of Bremaud (1981)). To measure the distance between the probability measure $P$ and $P(\zeta)$, we use the relative entropy, $E^\zeta \left[ \ln \left( \frac{dP(\zeta)}{dP} \right) \right]$, of Anderson, Hansen, and Sargent (2003) as follows.

$$
E^\zeta \left[ \ln \left( \frac{dP(\zeta)}{dP} \right) \right] = E^\zeta [\ln \zeta_T] = E[\zeta_T \ln(\zeta_T)] = \int_0^T H(\zeta_s) ds,
$$

where

$$
H(\zeta_s) = \sum_{k=1}^n \int_A \lambda_k \left[ \vartheta_k(s) \phi_k(s,y) \log(\vartheta_k(s) \phi_k(s,y)) + 1 - \vartheta_k(s) \phi_k(s,y) \right] \Phi_k(s,dy)
$$

is obtained from (4) and (5) by applying Ito’s lemma for jump processes to $\zeta_t \ln(\zeta_t)$.

The ambiguity averse investor searches for an optimal portfolio to maximize her CRRA utility function against the worst-case of the jump-diffusion model. Following the robust decision framework developed by Anderson, Hansen, and Sargent (2003), we solve a robust version of the Hamilton-Jacobi-Bellman equation for the investor’s value

---

function $V(W,t)$ as follows.

$$0 = \max_{\pi} \inf_{\pi(\zeta) \in \mathcal{P}} \left\{ \frac{\partial V(W,t)}{\partial \pi} + W(\pi'(\mu - r 1_m) + r) \frac{\partial V(W,t)}{\partial W} + \frac{1}{2} W^2 \pi' \Sigma \Sigma' \pi \frac{\partial^2 V(W,t)}{\partial W^2} 
+ \sum_{k=1}^n \lambda_k \mathbb{E}^{\pi_k}[V(W + W \pi' J_k Y_{k,t}, t) - V] + \frac{1}{\theta} H(\zeta_t) \right\}, \quad (8)$$

with $V(W,T) = \frac{W^{1-\gamma}}{1-\gamma}$. The expectation operator $\mathbb{E}^{\pi_k}[\cdot]$ is defined for any functional $f$ of the random jump size $Y_{k,t}$ thus $\mathbb{E}^{\pi_k}[f(Y_{k,t})] = \mathbb{E}[f(Y_{k,t}) \varphi_k(t,Y_{k,t})]$. The minimization in (8) clarifies the investor’s pessimistic reaction to uncertain jumps. She makes decisions by first minimizing the expected utility among all alternative models, every model of which is statistically indistinguishable from the reference model thus is possibly true model. One distinctive feature of our setup is that we investigate model misspecifications in the entire neighborhood of the reference model, while Liu, Pan, and Wang (2005) and Drechsler (2013) consider only a subset of the neighborhood. Particularly, these authors use a parametric approach to choose the worst jump size and jump intensity while we apply a nonparametric method to choose the worst case. Hence, the worst case log-price jump size distributions remain normal or Gamma-distributed in their models, while the worst case jump size distribution does not necessarily fall in the same distribution family of the reference model in our study.

Meanwhile, the investor penalizes any deviation from the reference model by the relative entropy, shown in the last term of (8), which measures the distance between the alternative probability and the reference probability. Hence she weighs the reference model more heavily than the alternative models because the former is the individual’s “best guess” of the probability law. The model preference variable $\theta_t$ dictates the investor’s attitude towards alternative models. We assume $\theta_t$ takes the following form (e.g., Maenhout (2004, 2006)):

$$\theta_t = \frac{\phi}{(1 - \gamma)V(W,t)}, \quad (9)$$
where $\phi$ will be referred to as the ambiguity aversion coefficient, with a larger $\phi$ indicating a lesser penalty and a higher level of ambiguity aversion.

At the extreme, when $\phi = \infty$, the investor makes decisions based on the worst case among all possibilities, treating the reference model equal to any alternative model; when $\phi = 0$, the investor considers only the reference model without any fear of jump ambiguity. In our calibration exercises, the levels of ambiguity aversion $\phi$ (or the levels of penalties) are determined by the DEP approach of Anderson, Hansen, and Sargent (2003). Models with different jump distributions are compared under the same DEP implying the same level of jump ambiguity. More detailed discussions are provided later in Section 5.2.

We are now ready to solve the portfolio choice problem defined above. For notational convenience, we will suppress the dependence of $Y_{k,t}$, $\vartheta(t)$, $\varphi(t,y)$, $\Phi_k(t,dy)$, and $\pi_t$ on $t$ hereafter.

3 A closed-form solution

In this section, we solve the HJB equation (8) explicitly in two steps. First, we solve the inner minimization problem and obtain the worst probability for any admissible portfolio $\pi$. The results are summarized in Proposition 1 below. Second, we propose a decomposition approach to determine the optimal portfolio weights under ambiguity aversion. The results are summarized in Proposition 2 in the sequel.

**PROPOSITION 1.** For any admissible portfolio $\pi$, the solution to the inner minimization of (8) is given by

$$\varphi_k^*(y) = \varphi_k(y, \pi) = \frac{1}{\vartheta_k^*} \exp \left( \frac{\phi}{\gamma - 1} \left( (1 + \pi' J_k y_k)^{1-\gamma} - 1 \right) \right),$$

$$\vartheta_k^* = \vartheta_k^*(\pi) = \mathbb{E} \left[ \exp \left( \frac{\phi}{\gamma - 1} \left( (1 + \pi' J_k Y_k)^{1-\gamma} - 1 \right) \right) \right].$$
for \( k = 1, \ldots, n \).

Suppose the optimal portfolio \( \pi^* \) has been obtained, then Proposition 1 provides the worst probability by \( \varphi_k^* \) and \( \vartheta_k^* \) as we substitute \( \pi = \pi^* \) into (10) and (11). We see that the worst probability under our nonparametric approach has an identical functional form regardless of the jump distribution of \( Y \). On the contrary, an alternative parametric approach has to solve for the worst case separately for each reference jump distribution. Our method thus greatly facilitates the examination of jump tail risk and allows us to easily track the investor’s aversion to potential model misspecifications when the reference models embody distinct tail properties, e.g., either slowly-decaying distributions (following power laws) with heavy tails or normal distributions with light tails.

Let \((\tilde{\pi}_k^*)' = (\pi_k^*)' J = (\tilde{\pi}_1^*, \ldots, \tilde{\pi}_n^*)\). We refer to \( \tilde{\pi}_k^* \) as the exposure to the \( k \)-th jump risk. Then the jump intensity in the worst probability is given by

\[
\tilde{\lambda}_k^* = \lambda_k \mathbb{E} \left[ \exp \left( \frac{\phi}{\gamma - 1} \left( (\tilde{\pi}_k^* Y_k + 1)^{1-\gamma} - 1 \right) \right) \right] = \lambda_k \vartheta_k^*(\pi^*). \tag{12}
\]

Recall that for each \( k \) we consider a mixed jump size random variable \( Y_k \in (-1, \infty) \), hence \( \tilde{\pi}_k^* \) must take value from \([0, 1] \). Therefore, it is not always true that \( \tilde{\lambda}_k^* \geq \lambda_k \). In the particular case that the jump size is a negative constant, it is clear that \( \tilde{\lambda}_k^* > \lambda_k \) if \( \tilde{\pi}_k^* > 0 \). That is, the investor with positive exposure to the downward jump fears more frequent jumps. Moreover, the ambiguity aversion coefficient \( \phi \) reinforces this tendency indeed. In general, whether the worst intensity is strengthened relative to the reference intensity depends on the expectation in (12) or really on the distribution of the jump size. Our empirical studies in Section 5 show that, for the representative case of seven international indices, the worst intensity tends to be higher while the jump exposure tends to be positive for a jump with a negative expected size.

As for the jump size distribution under the worst probability, by Proposition 1, the
density of the $k$-th jump size of the worst case is given by

$$\Phi^*_k(dy_k) = \varphi^*_k(y_k)\Phi_k(dy_k), \quad (13)$$

where $\varphi^*_k$ can be regarded as a weighting function. For the typical case of $\hat{\pi}^*_k > 0$, we can verify that $\varphi^*_k$ is a decreasing function of the jump size $y_k$. Therefore, the ambiguity averse investor pessimistically attaches more weight to more negative jumps and less weight to more positive jumps. As a result, the weighting function leads to a lower expected jump size and a more negatively skewed and less positively skewed jump size distribution in the worst case model relative to those in the reference model.

Finally, note that, for the expectation term in (8), $\vartheta^*_k$ and $\varphi^*_k(y_k)$ always come together as the product $\vartheta^*_k\varphi^*_k(y_k)$, which decreases in $y_k$. Hence the worst probability twists this expectation towards the negative side of the jump size distribution and further impacts the optimal portfolio choice as we discuss in Section 4.

We now turn to finding the optimal portfolio policies under ambiguity aversion. Our solution method is based on a decomposition transformation as follows. Let

$$\hat{J} = \Sigma^{-1}J, \quad \hat{\mu} = \Sigma^{-1}(\mu - r\mathbf{1}_m). \quad (14)$$

Note that $\hat{J}$ is an $m \times n$ matrix with rank $n$. We treat each column of $\hat{J}$ as a vector of $\mathbb{R}^{m \times 1}$. For $m > n$, we can find unit vectors $\alpha_1, ..., \alpha_{m-n} \in \mathbb{R}^{m \times 1}$ such that $\alpha_k, k = 1, ..., m-n$, is orthogonal to each column of $\hat{J}$. Denote the matrix with columns $\alpha_1, ..., \alpha_{m-n}$ by $\hat{J}_\perp$. Then each $m$-vector can be decomposed on $\hat{J}$ and $\hat{J}_\perp$. In particular, the decomposition of $\hat{\mu}$ on the space $\hat{J}$ and its orthogonal space $\hat{J}_\perp$ can be written as

$$\hat{\mu} = \tilde{\mu} + \mu_\perp, \quad (15)$$
where \( \bar{\mu} \) is an \( m \)-vector in the space generated by the columns of \( \hat{J} \) and can be expressed as \( \bar{\mu} = \hat{J}\mu^0 \) with \( \mu^0 \in \mathbb{R}^m \); \( \mu_\perp \) is an \( m \)-vector in the space generated by the columns of \( \hat{J}_\perp \) and can be represented by \( \mu_\perp = \hat{J}_\perp\mu^0_\perp \) with \( \mu^0_\perp \in \mathbb{R}^{m-n} \). Note that the decomposition is unique. By multiplying (15) from the left by \( \hat{J} (\hat{J}'\hat{J})^{-1}\hat{J}' \) and noticing that \( \hat{J}'\hat{J}_\perp = 0_{m \times (m-n)} \), we can find \( \bar{\mu} = \hat{J}(\hat{J}'\hat{J})^{-1}\hat{J}'\mu \). Similarly \( \mu_\perp = \hat{J}_\perp(\hat{J}'\hat{J}_\perp)^{-1}\hat{J}'\mu \). We achieve the optimal portfolio with ambiguity aversion in a closed-form decomposition as follows.

**PROPOSITION 2.** The optimal portfolio with ambiguity aversion is given by

\[
\pi^* = (\Sigma')^{-1}(\bar{\pi}^* + \pi^*_\perp),
\]

with

\[
\pi^*_\perp = \frac{1}{\gamma}\mu_\perp,
\]

\[
\bar{\pi}^* = \arg\max_{\pi} -\frac{\gamma}{2}\bar{\pi}'\bar{\pi} + \bar{\mu}'\bar{\pi} + \frac{1}{1-\gamma}\sum_{k=1}^{n}\lambda_k E[\zeta^*_k((1 + \bar{\pi}'\hat{J}_kY_k)^{1-\gamma} - 1)],
\]

where \( \zeta^*_k = \zeta_k(\vartheta^*_k, \varphi^*_k) \); \( \vartheta^*_k \) and \( \varphi^*_k \) are given by (10) and (11), respectively.

Since \( \pi'J_k = \bar{\pi}'\hat{J}_k \) from the proof of Proposition 2 in Appendix A, \( \vartheta^*_k \) and \( \varphi^*_k \) are both functions of \( \bar{\pi}^* \) only by Proposition 1. Hence we can solve (17) for \( \bar{\pi}^* \) independent of \( \pi^*_\perp \). To simplify the analysis, unless otherwise stated, we assume in this section that the maximization in (17) is achieved at an interior point. The first order condition of (17), along with the optimality of \( \zeta^*_k \), yields the following equation for \( \bar{\pi}^* \):

\[
0 = -\gamma\bar{\pi} + \bar{\mu} + \sum_{k=1}^{n}\lambda_k E[(1 + \bar{\pi}'\hat{J}_kY_k)^{-\gamma}Y_k\vartheta^*_k(\varphi^*_k(Y_k))]\hat{J}_k.
\]  

Multiplying \( \hat{J}_k' \) from the left hand side on (18) and denoting \( \bar{\pi}_k = \hat{J}_k'\bar{\pi} \), we obtain the
equations for $\tilde{\pi}_k$:

$$-\gamma \tilde{\pi}_k + \hat{J}_k \mu \bigg[ (1 + \tilde{\pi}_l Y_l)^{-\gamma} Y_l \bigg](\hat{J}_k^t \hat{J}_l) = 0, \quad k = 1, \ldots, n. \quad (19)$$

It is straightforward to show the existence and uniqueness of the solution to the above equation system. Furthermore, due to the uniqueness of the orthogonal decomposition, we have $\bar{\pi}' = \hat{J}(\hat{J}' \hat{J})^{-1} \tilde{\pi}$. The existence and uniqueness of the solution to the equation system (18) are therefore also guaranteed.

Our decomposition approach may largely reduce the computational complexity of a model with a large number of risky assets. Solving the optimal portfolio directly from the first order condition of the HJB equation (8) with respect to $\pi$ involves solving $m$ (the number of risky assets) non-linear equations, while, with our decomposition approach, it suffices to solve the $n$ (the number of jump types) non-linear equations of $\tilde{\pi}$. Since the number of jump types is usually much less than the number of risky assets (e.g. Das and Uppal (2004) study one systemic jump with six indices and Jin and Zhang (2012) consider two types of jumps with four risky assets.), our approach is essentially useful for studying multi-asset jump-diffusion models.

To illustrate the optimal portfolio obtained from Proposition 2, we discuss two specific examples. Before going into details, we mention that we can derive from Proposition 2 the optimal portfolio without ambiguity aversion by letting $\zeta_k^* = 1$ or $\theta_k^* = \varphi_k^* = 1$ in (18). This corresponds to the case $\phi = 0$ such that any deviation from the reference model is penalized infinitely.

The first example focuses on $J = 0$, thus the model comprises no jumps. Under this simplified version, $\hat{J} = 0$ and $\hat{J}_\perp$ is an invertible $m \times m$ matrix. There is no decomposition component on $\hat{J}$ for any vector. Hence the optimal portfolio is

$$\pi^* = (\Sigma')^{-1}\pi_\perp^* = (\Sigma')^{-1}\frac{1}{\gamma} \hat{J}_\perp (\hat{J}_\perp^t \hat{J}_\perp)^{-1} J'_\perp \Sigma^{-1}(\mu - r 1_m) = \frac{1}{\gamma} (\Sigma \Sigma')^{-1}(\mu - r 1_m), \quad (20)$$
which is the familiar optimal portfolio in a multi-asset diffusion market.

The second example, complementary to the first one, considers the case of \( m = n \). The number of jump types now matches the number of risky assets. We have \( \hat{J}_\perp = 0 \) hence there is no decomposition component on \( \hat{J}_\perp \). We have the optimal portfolio

\[
\pi^* = (\Sigma')^{-1}\bar{\pi}^*.
\]

When we further restrict \( m = n = 1 \), we get the solution of the optimal portfolio for the case of one stock considered in Liu, Longstaff, and Pan (2003).

Taken together, the above two examples suggest that \( \bar{\pi}^* \) can be regarded as closely related to the exposure to jump risk while \( \pi^*_\perp \) closely related to the exposure to diffusion risk. Furthermore, by (14) and the decomposition of \( \pi \) in Proposition 2, the relation between \( \bar{\pi}^* \) and the \( k \)-th jump exposure \( \tilde{\pi}^*_k \) can be clearly seen from the following result:

\[
\tilde{\pi}^*_k = J'_{k} \pi^* = \hat{J}_k^* \bar{\pi}^*.
\]

In the following, we refer to \( \bar{\pi}^* \) as the exposure to jump risk and in the cases of no confusion, we also refer to \( \tilde{\pi}^*_k \) as the exposure to jump risk for short.

Proposition 1 and 2 explicitly solve for the worst case probability and the optimal portfolio through a decomposition into two independent components. Other decomposition methods have also been pioneered by the previous literature. For instance, Jin and Zhang (2012) rely on a more complicated “fictitious assets” method to solve the incomplete market case we address, and they do not consider ambiguity in the jump size distribution. Additionally, our decomposition method does not rely on a special structure of the model coefficients as requested by the decomposition technique developed by Aït-Sahalia, Cacho-Diaz, and Hurd (2009).

4 Tail risk, diversification, and non-participation

This section focuses on the effects of ambiguity aversion on the jump tail behavior and the optimal portfolio. We start by disclosing that the jump tails in the worst case
scenarios are directly affected by ambiguity aversion. We then carefully examine the optimal portfolio by studying the sensitivity of the jump exposure to jump ambiguity. We finally demonstrate the crucial role played by jump tail assumptions in determining the robust optimal portfolio.

**PROPOSITION 3.** There exists \( y^* < 0 \) such that for any \( y < y^* \), \( Pr^*(Y_k < y|\phi) > Pr(Y_k < y) \) where \( Pr^*(\cdot|\phi) \) is the worst probability measure given the ambiguity aversion \( \phi \). Moreover, for any \( \phi_1 > \phi_2 \), there exists \( \hat{y} \) such that for any \( y < \hat{y} \), \( Pr^*(Y_k < y|\phi_1) > Pr^*(Y_k < y|\phi_2) \).

This proposition affirms that there is a larger tail in the worst case probability than in the reference probability, and the tail is larger if an investor is more ambiguity averse. Thus tail risk in the worst case scenario is greater for an investor with a higher level of ambiguity aversion. Figure 1 illustrates this result for a normal jump size distribution in the reference model. It is clear that the worst case density has a larger left tail as \( \phi \) becomes larger.

Turning to the optimal portfolio, we again assume interior solution \( \tilde{\pi}^* \). It follows from (19) that

\[
\tilde{\pi}^* = \frac{1}{\gamma} \hat{J}' \hat{\mu} + \frac{1}{\gamma} \sum_{k=1}^{n} \lambda_k E^\phi[(1 + \tilde{\pi}_k^* Y_k)^{-\gamma} Y_k](\hat{J}' \hat{J}_k).
\]

(21)

We see that only the second term of (21) involves the jump exposure. The first term is independent of the ambiguity aversion coefficient \( \phi \). Thus we can study the sensitivity of the jump exposure to the investor’s attitude toward model uncertainty by investigating the derivative of \( \tilde{\pi}^* \) with respect to \( \phi \). A proposition regarding the sensitivity of the jump exposure with respect to jump parameters then follows.

To gain clear intuition, we focus on the simple case \( n = 1 \), that is, only one type of jumps is admitted in the model. This allows us to omit the subscript \( k \) in the rest of this section.
PROPOSITION 4. Suppose \( n = 1 \). Then

(i) \( \frac{d\tilde{\pi}^*}{d\sigma} < 0 \),

(ii) \( \frac{d\tilde{\pi}^*}{d(J^\dagger J)} > 0 \) if and only if

\[
\frac{d(J^\dagger \hat{\mu})}{d(J^\dagger J)} + \lambda \mathbf{E}^{\pi^*}[(1 + \tilde{\pi}^*Y)^{-\gamma}Y] > 0. \tag{22}
\]

The result (i) provides a link between the jump exposure and ambiguity aversion. This result shows that as the level of ambiguity aversion becomes higher, the investor will be more averse to uncertain jumps and reduces her jump exposure accordingly.

To illustrate the result (ii) and its condition (22), further simplifications seem warranted. Here we consider the simplest case of identical risky assets with no correlation beyond the systemic jump. Specifically, we assume \( \Sigma \) is a diagonal matrix with identical diagonal entries of \( \sigma_1 \) and \( \mu \) has identical entries of \( \mu_1 \). We further assume \( J = 1_m \). Then \( J^\dagger J = \sigma_1^{-2}m \) and \( J^\dagger \hat{\mu} = J^\dagger \mu = (\mu_1 - r)\sigma_1^{-2}m \) by (14). Condition (22) becomes:

\[
\mu_1 - r + \lambda \mathbf{E}^{\pi^*}[(1 + \tilde{\pi}^*Y)^{-\gamma}Y] > 0. \tag{23}
\]

Two observations follow immediately from (23). On the one hand, we conclude that (23) holds by (21) since \( \tilde{\pi}^* \in (0, 1) \) and \( \tilde{\pi}^* = \frac{m}{\sigma_1^2}(\mu_1 - r + \lambda \mathbf{E}^{\pi^*}[(1 + \tilde{\pi}^*Y)^{-\gamma}Y]) \). Therefore, in this case, the jump exposure \( \tilde{\pi}^* \) increases with \( \sigma_1^{-2}m \). Holding constant the variance \( \sigma_1^2 \) of each risky asset, we find that the jump exposure increases with \( m \), the total number of risky assets. Note that \( \tilde{\pi}^* = J^\dagger \pi^* = \sum_{i=1}^{m} \pi_i^* \), hence the total risky investment also increases with the number of risky assets.

On the other hand, by (21),

\[
\tilde{\pi}^* = \pi^* \sigma_1/m = \frac{1}{\gamma \sigma_1} \left( (\mu_1 - r) + \lambda \mathbf{E}^{\pi^*}[(1 + \tilde{\pi}^*Y)^{-\gamma}Y] \right) 1_m.
\]
Each component of $\bar{\pi}^\ast$ is decreasing in $m$ as one can check that

$$\partial E^\zeta[(1 + \tilde{\pi}^\ast Y)^{-\gamma} Y] / \partial m = -\gamma \partial \bar{\pi}^\ast / \partial m) E^\zeta[(1 + \tilde{\pi}^\ast Y)^{-\gamma - 1} Y^2] < 0.$$ 

Hence the jump exposure of the optimal portfolio on each risky asset, referred to as individual jump exposure, decreases as the investor takes more risky assets. Put differently, portfolio diversification is diminished due to the systemic jump risk.

We plot the effect of jump ambiguity on portfolio diversification in Figure 2. Under the above simplifying assumptions for asset returns, the upper-left panel shows that the individual jump exposure $\bar{\pi}_i$ decreases towards zero as the number of risky assets increases. This is to be expected since the total jump exposure cannot exceed one for the mix jumps considered here, as we see in the lower-left panel. The optimal holding for each individual asset also approaches zeros when the number of asset increases, as plotted in the upper-right panel. This limiting behavior indicates a unique feature of diversification under jump risk and stands in stark contrast to that under i.i.d. risky assets in a pure diffusion model. In the latter model, exposure to diffusion risk of each risky asset remains constant to the number of assets as indicated by (20).

Meanwhile, we see that a more ambiguity averse investor holds even less risky positions. When accessing a larger number of risky assets, her positions in each individual asset draw to zero more closely. Thus ambiguity aversion further suppresses portfolio diversification when the investor lacks confidence in the systemic jump risk model.

One possible explanation is that a larger total jump exposure tends to raise the investor’s fear of uncertain jump severity. Given a sufficiently small negative (large positive) jump size $y_k$, the weighting function $\varphi_k^\ast$ in (13) is increasing (decreasing) in the jump exposure $\tilde{\pi}_k$.\(^9\) The investor with greater exposure to jumps will weight the left tail of the jump size distribution even more heavily and the right tail more lightly. The expected mean jump size in the worst case also becomes more negative. Conse-

\(^9\)This can be proved by taking derivative of $\varphi_k^\ast$ with respect to $\tilde{\pi}_k$. 

20
quently, expanding the investment opportunity set with more risky assets could result in higher total jump exposure and meanwhile marginally bring down each individual jump exposure.

By Proposition 3 and 4, when an investor is extremely ambiguity averse with \( \phi = \infty \), the worst case density will have the largest tail and the jump exposure will be minimal. Since the optimal jump exposure \( \tilde{\pi}^* \) must lie in \([0, 1]\) for mixed jumps, the jump exposure may drop to zero leading to an optimal portfolio of zero positions in all risky assets. This result is closely related to the nonparticipation puzzle: Modern asset allocation theories suggest that investors allocate certain non-zero percentages of their wealth to all available risky assets. However, empirical studies document that many households do not invest in risky assets at all. For example, Campbell (2006) provides evidence that at the eightieth percentile of wealth, almost 20% of households do not possess any public equity. Mankiw and Zeldes (1991) report similar findings. In the following, we show that aversion to jump ambiguity in our model may indeed generate nonparticipation.

An extremely ambiguity averse investor with \( \phi = \infty \) disregards the reference model and treats it equal to all alternative models. In such a case, any non-zero position in the risky assets results in infinity loss when the worst case scenario occurs. To demonstrate this, we note that the second component in the HJB equation becomes

\[
\inf_{P(\cdot)} \frac{1}{1 - \gamma} \sum_{k=1}^{n} \lambda_k \mathbb{E}_k \left[ (1 + \pi' J_k Y_k)^{1-\gamma} - 1 \right].
\]  

(24)

We consider mix jumps and \( Y_k \) can take negative and positive values. If \( \pi' J_k > 0 \), we can find \( \psi(\cdot) \), such that \( \int_A (1 + \pi' J_k y)^{1-\gamma} \psi(y) \Phi_k(dy) - 1 > 0 \).\(^{10}\) Then we can find a

\(^{10}\)We can do this because \( \psi \) “shifts” the major distribution of \( Y_k \) to the negative side. Then for most values of \( Y_k \), \( (1 + \pi' J_k Y_k) < 1 \) and \( \mathbb{E}[(1 + \pi' J_k Y_k)^{1-\gamma} \psi(Y_k)] > 1 \).
sequence \( \zeta_i = \zeta(\vartheta^{(i)}, \psi) \) where \( \vartheta^{(i)} \to \infty \) as \( i \to \infty \), such that

\[
\inf_{P(\zeta)} \frac{1}{1-\gamma} \lambda_k \mathbb{E} \left[ (1 + \pi' J_k Y_k)^{1-\gamma} - 1 \right] \\
\leq \frac{1}{1-\gamma} \lambda_k \vartheta^{(i)} \left( \int_A (1 + \pi' J_k y)^{1-\gamma} \psi(y) \Phi(dy) - 1 \right) \to -\infty,
\]

as \( i \to \infty \). Similarly, for a case such that \( \pi' J_k < 0 \), we can show that the quantity (24) is \( -\infty \) as well. Intuitively, an investor with extreme ambiguity aversion assumes that jumps against her positions will occur at an infinite frequency and cause infinite losses. Hence the optimal portfolio \((\pi^*)' J_k \) must be zero at which the quantity (24) is zero. Moreover, when \( m = n \) and \( \text{rank}(J) = n \), that is, the number of independent jumps is identical to the number of risky assets, it follows from \( (\pi^*)' J = 0 \) that \( \pi^* = 0 \) or nonparticipation arises.

Consider another case in which the market consists of one systemic jump (\( n = 1 \)) and each component of the vector \( J \) is positive. We further assume that the investor is not allowed to short sell stocks, that is, \( \pi_k \geq 0, k = 1, ..., m \). If the investor is extremely ambiguity averse to the systemic jump, then \( \pi_k^* = 0, k = 1, ..., m \), because \( (\pi^*)' J = 0 \). Nonparticipation occurs. When short selling is prohibited, diversification works only if correlations among the stocks are low; however, correlations caused by the systemic jump may be very high in the worst case scenario such that diversification fails to work, resulting in nonparticipation due to the fear of the jump.

We have shown that ambiguity aversion enlarges the worst case jump tail risk and in turn impacts portfolio diversification. The above analyses generally apply to any jump distribution. The following proposition directly examines the importance of tail assumptions in the reference models. Precisely, we say one jump size \( Y^p \) has a larger left tail than another jump size \( Y^n \) if there exists \( y_0 < 0 \) such that \( Pr(Y^p < y) > Pr(Y^n < y) \) for any \( y < y_0 \).
PROPOSITION 5. Suppose one jump size $Y^n$ has a larger left tail than another jump size $Y^p$. Let $\tilde{\pi}^{p,*}$ be the jump exposure with respect to $Y^p$ and let $\tilde{\pi}^{n,*}$ be the jump exposure with respect to $Y^n$. Then under mild conditions we have $\tilde{\pi}^{p,*} < \tilde{\pi}^{n,*}$.

Hence acknowledging a heavy-tailed jump distribution, an ambiguity averse investor tends to reduce her jump exposure and further diminishes the diversification of the optimal portfolio. Given the crucial role played by jump tail behavior, we quantify the impact of tail risk on optimal portfolios through a calibration exercise in the next section, where we deliberately consider jump distributions common in the literature but with distinct tail properties.

5 Calibrating the effects of uncertain jump tails

The recent financial crises has fueled a renewed interest in modeling, estimating, and deriving the implications of extreme tail events. Bollerslev and Todorov (2011a) empirically illustrate that the traditional normally distributed jump size proposed by Merton (1976) severely underestimates the likelihood of “large” jumps. Furthermore, they show that typically larger jumps are formally outside of this traditional framework because the tail of a normal distribution decays too quickly. Guided by the extreme value theory, these scholars show that the distribution for extreme events can be well approximated by a power law that captures the slow tail decay for financial returns typically reported in the literature. Kelly and Jiang (2014) and Bollerslev, Todorov, and Xu (2015) construct tail risk measures using stocks and options, respectively, and both studies indicate that tail risk has strong predictive power for aggregate asset returns.\textsuperscript{11} As Merton’s normal jump size model is intensively used in prior studies, a natural question to ask is the following: What is the economic loss of ignoring the heavy jump tails from an asset allocation perspective?

\textsuperscript{11}Wang (2015) provides evidence that tail risk is a common priced factor in the international equity markets.
In this section, we solve the optimal portfolio choice problem empirically using two models: One has a normally distributed jump size that fails to capture fat tails and the other, aligning with Barro and Jin (2011), adopts a power law distribution for extreme events. In addition, we incorporate ambiguity into the jump size distributions in the models because the heavy-tailed distribution may be even harder to determine. We conduct a calibration exercise with a large number of stocks to affirm the economic relevance of tail risk on portfolio selection.

5.1 Model calibration

For the purposes mentioned above, we first utilize a jump-diffusion model with a normally distributed jump size as summarized in the following equation.

$$\frac{dS_{i,t}}{S_{i,t}} = \mu_i dt + \sum_{j=1}^{m} \sigma_{i,j} dB_j(t) + J_i Y dN_t, \quad i = 1, 2, ..., m, \tag{25}$$

where $Y = \exp(\mu J + \sigma J \varepsilon) - 1$ and $\varepsilon$ is a standard normal random variable; $E(dN_i) = \lambda dt$; $B_1(t)$ to $B_m(t)$ are standard independent Brownian motions and independent of $Y$; $m$ is the total number of stocks. Only one type of jumps is considered in this model, i.e. $n = 1$; jump scale $J_i \in [0, 1]$. We denote $J = (J_1, ..., J_m)'.$

We calibrate the model to the monthly continuously compounded returns on the equity indices of seven developed countries. The developed countries include the United States (U.S.), the United Kingdom (U.K.), Germany (GE), France (FR), Canada (CA), Sweden (SD), and Japan (JP). We collect the beginning-of-month equity index levels from finance.yahoo.com. Due to data availability, our sample period is January 1993 to December 2015.

Table 1 reports the descriptive statistics of the monthly return series. All seven indices exhibit negative skewness and high excess kurtosis. Our sample comprises the Asian crisis of 1997, the hedge fund crisis of late 1998, the financial crisis of 2008 and
the European sovereign-debt crisis of 2010 and 2011. Large return shocks during those turbulent periods contribute to the high kurtosis of the returns. Occasional large market crashes result in a negative skewness of the returns. Pairwise correlations among the equity index returns are unanimously higher than 56%. This result indicates a close linkage of the international equity markets.

We estimate the jump-diffusion model using the method of moments approach provided by Das and Uppal (2004) and Jin and Zhang (2012). The first four unconditional moments of the multivariate return series are considered. Following Das and Uppal (2004), we derive in closed form the characteristic function of the continuously compounded stock returns. We then differentiate the characteristic function to obtain the moments. Let $\bar{Y}_i = \ln(J_i Y + 1)$. For $i, j = 1, 2, ..., m$ ($m = 7$),

\[
\begin{align*}
\text{mean} &= t(\mu_i - 0.5 \sum_{k=1}^m \sigma_{ik}^2 + \lambda \mathbb{E}[\bar{Y}_i]), \\
\text{covariance} &= t(\sum_{k=1}^m \sigma_{ik} \sigma_{jk} + \lambda \mathbb{E}[\bar{Y}_i \bar{Y}_j]), \\
\text{coskewness} &= \frac{t \lambda \mathbb{E}[(\bar{Y}_i)^u (\bar{Y}_j)^v]}{\text{variance}_{ij}^{u+v}}, \\
\text{excess kurtosis} &= \frac{t \lambda \mathbb{E}[(\bar{Y}_i)^4]}{\text{variance}_{ij}^2},
\end{align*}
\]

where $\mathbb{E}[(\bar{Y}_i)^u (\bar{Y}_j)^v] = \int_{-\infty}^{+\infty} (\bar{Y}_i)^u (\bar{Y}_j)^v f(\varepsilon) d\varepsilon$ with $u = 1, 2; v = 0, 1; ...$ and $f(\cdot)$ is the standard normal density. This integral may be easily evaluated using the numeric quadrature method. We first use the $7 \times 7$ co-skewness conditions and $7 \times 1$ kurtosis conditions to estimate the 10 jump parameters $(J_i, \mu_J, \sigma_J, \lambda)$ by minimizing the sum of squared deviations of the model moments from those in the data. We then derive $\mu_i$ and $\sigma_{ij}$ by exactly matching the $7 \times 1$ mean conditions and the $7 \times 7$ covariance conditions, respectively.

Table 2 presents the parameter estimates on a monthly basis. Panel A indicates that the average jump size is -13.6% for the developed countries evaluated. This result is consistent with the negative skewness of the return series. The standard deviation of jump size is 8.3%. Therefore, a 95% confidence interval for the jump size is (-30.2%,
3\%). As shown in the moment condition in equation (26), large-sized jumps are crucial to match the high excess kurtosis of the data. The jump intensity is estimated to be 0.075. Simultaneous jumps among the seven markets are expected to occur about once every 13 months, or once every 1.1 years. This is broadly consistent with the literature finding that equity indices jump approximately once or twice a year.

To address the extreme tail risk of stock returns documented in the literature, we introduce an alternative tail distribution of jump size aside from the normal distribution above. We adopt the single power law distribution of Barro and Jin (2011). These scholars collect a panel of international consumption disasters and show that the empirical distribution of properly transformed large consumption drops may be reasonably approximated by a power law.\textsuperscript{12} It is well-known that, under certain conditions (e.g., Naik and Lee (1990)), the stock price is a linear function of consumption in market equilibrium; therefore, both share the same jump size distribution. Let $\eta = \frac{1}{1+Y}$ and $\eta$ follows a single power law distribution with its density as follows:

$$\nu(\eta) = \alpha\eta_0^\alpha \eta^{-(\alpha+1)}, \quad \eta \geq \eta_0 > 1.$$  \hfill (27)

$\eta_0$ and $\alpha$ are fixed by matching the first two moments of this distribution to those obtained previously under the normally distributed jumps in log-price. We obtain $\eta_0 = 1.0547$ and $\alpha = 12.0362$. Hereafter, for simplicity, we will refer to the power law jump size distribution as the tail distribution, and the associated jump model as the tail jump model. Similarly for the normal jump size distribution.

The reference (original) densities and the corresponding worst-case densities are provided in Figure 3. We see that the tail distribution has less weight on mid-sized jumps and more weight on large-sized jumps than the normal. This difference is even

\textsuperscript{12}In an unreported exercise, we also explore the tail distribution proposed by Bollerslev and Todorov (2011a). Let $Z = e^{[X]} - 1$, where $X$ denotes jumps in log-price. Bollerslev and Todorov (2011a) take $Z$ to follow a power law. Note that $Z$ may be interpreted as “discrete” price jumps for $X > 0$; however, there is no such intuitive interpretation for $X < 0$. We obtain qualitatively similar results (available upon request) regarding optimal portfolio weights and investor’s welfare under this power law.
more significant in the worst case scenarios. Note that the density in the worst case scenario shifts toward larger-sized jumps in the left for either jump distribution.

To gauge the performance of our models at fitting the index return data, we report in Table 3 the theoretic moments reconstructed using the model parameter estimates and the moments computed directly from data. We see that the fitting of the third and fourth moments is reasonably good for the normal jump model. The skewness of the tail jump model is more negative than that of the normal jump model, reflecting the large left tail risk of the tail distribution.\footnote{Comparison of other coskewness conditions reaches essentially the same conclusions. The detailed results are available upon request.} Tail risk is made apparent by the much higher excess kurtosis in the tail jump model relative to the normal jump model. In the following section, we discuss portfolio choice and the worst case probabilities implied by the two models.

5.2 International asset allocation with ambiguity aversion

Since the two jump size distributions pursued here display distinct tail behaviors, an investor may hold different degrees of ambiguity aversion towards the two models. To make the results more comparable, we identify the appropriate ambiguity aversion coefficients $\phi$ based on the same DEP (derived in Appendix B) under the two distributions; a DEP is the probability that an ambiguity averse investor incorrectly rejects the worst case model in favor of the reference model in a likelihood ratio test. In particular, greater ambiguity aversion (higher $\phi$) implies lower DEP because it becomes increasingly easier to distinguish the worst case model from the reference model. Following the literature (e.g., Anderson, Hansen, and Sargent (2003), Maenhout (2004, 2006), and Drechsler (2013)), we require a DEP of no less than 10\% to determine a reasonable ambiguity aversion coefficient.

Table 4 reports the optimal portfolio holdings in the aggregate stocks of the seven countries. For exposition, we list the results regarding three DEPs equal to 10\%, 15\%,
and 20%. The corresponding \( \phi \) values are also listed in the table. The annual interest rate is set at 5% and the risk aversion coefficient \( \gamma \) is 3, 2, or 1.1. We consider only jump size ambiguity in this exercise to focus on effects of ambiguity aversion on jump size distribution. Similar results are obtained when jump intensity ambiguity is also encountered.

### 5.2.1 Aversion to jump ambiguity

We present portfolio weights \((\pi_k, k = 1, \ldots, 7)\) together with the exposure to jump risk \(\tilde{\pi}\) for the two jump models in Table 4. We see significant differences in these weights across models. For example, when the DEP is 10% and \(\gamma\) is 3, the allocation to the aggregate stock of the U.S. is 93.6% under the normal jump model but increases to 95.1% under the tail jump model, and the allocation to the aggregate stock of JP is -89.5% under the normal jump model but reduces to -81.3% under the tail jump model. The exposure to jump risk is 26.7% under normal jumps; in contrast, the same exposure reduces to 17.6% under the tail distribution. Given the same DEP, the jump exposure under the tail distribution is generally lower than that under the normal. This significantly reduced jump exposure demonstrates an investor’s fear of uncertain extreme tail events.

The jump exposures in both models shift down as the DEP becomes smaller or the investor becomes more ambiguity averse, implying that a more ambiguity averse investor will be more fearful of jump uncertainty. In particular, the corresponding results without ambiguity are listed in the last two columns of Table 4. Clearly, the exposures to jump risk are reduced in both models when ambiguity aversion is incorporated. Note that for different levels of risk aversion \(\gamma\), the ambiguity aversion coefficient \(\phi\) must change to reach a given DEP. For example, for the tail distribution with DEP=15%, the ambiguity aversion is 20.4 when \(\gamma = 3\) but is 8 when \(\gamma = 1.1\). This occurs because the worst density depends on \(\gamma\) and the jump exposure \(\tilde{\pi}\) decreases with \(\gamma\). Thus a less risk averse investor is less fearful of jump risk. Apparently, from (10) in Proposition 1, the worst density
\( \varphi_k^* \) increases with the jump exposure \( \tilde{\pi} = \pi'J_k \) for \( y_k < 0 \) and decreases with the jump exposure \( \tilde{\pi} \) for \( y_k > 0 \). As a result, under the same level of ambiguity aversion, the worst probability shifts further away from the reference probability when risk aversion is smaller, leading to a lower DEP.

Interestingly, the differences in optimal portfolio weights and the optimal jump exposures between the two models without ambiguity aversion are negligible. A possible explanation is that the first two moments in the two models without ambiguity aversion are perfectly matched but the moments in the worst cases of the two models may deviate. In addition, the first two moments dominate the higher moments for a CRRA investor when solving the optimal asset allocation problem. Put differently, the CRRA utility does not sufficiently capture the investor’s concern regarding the extreme downside risk modeled by the tail distribution (27). This result is consistent with Hong, Tu, and Zhou (2007) and Cvitanić, Polimenis, and Zapatero (2008): these scholars note that the CRRA utility function represents a locally mean-variance preference that does not capture higher moments.

To investigate the impact of higher moments, Hong, Tu, and Zhou (2007) use the Disappointment Aversion (DA) preference of Ang, Bekaert, and Liu (2005). It is well known that the DA preference is particularly useful for analyzing the extreme downside risk because by selecting a reference point, the preference weights the outcomes below the reference point more heavily than those above it. It appears that we may use the DA preference of Ang, Bekaert, and Liu (2005) to investigate an investor’s concern about extreme downside risk.\(^{14}\) More specifically, we want to know if an investor with DA preference behaves significantly differently in the two jump models considered here. This question is beyond the scope of the present paper and left for future research.

\(^{14}\)An interesting alternative preference for this purpose is the downside loss-averse utility considered in Jarrow and Zhao (2006).
5.2.2 Economic welfare

We now gauge the economic significance of the differences in the optimal portfolio weights between the two models with and without ambiguity aversion. Specifically, we assume that the jump size follows the tail distribution given by (27) in the true model, and we calculate the certainty equivalent loss (CEL) from adopting the suboptimal portfolio in the normal jump model. Let $\pi^{(1)}$ and $\pi^{(2)}$ denote the optimal portfolios in the true model and in the normal jump model, respectively. Then the CEL is defined as the percentage of initial wealth an investor is willing to sacrifice to switch from $\pi^{(2)}$ to $\pi^{(1)}$. Equivalently, the CEL solves the following equation:

$$V_{\pi^{(2)}}(W,t) = V(W(1 - CEL), t),$$

where $V(W,t)$ and $V_{\pi^{(2)}}(W,t)$ are the value functions obtained by implementing the portfolios $\pi^{(1)}$ and $\pi^{(2)}$ in the true model, respectively. The value function $V(W,t)$ is calculated in the previous section. The value function corresponding to the portfolio $\pi^{(2)}$ is given by

$$V_{\pi^{(2)}}(W,t) = \inf_\zeta E_t^\zeta \left[ e^{\int_t^T \frac{1}{2} \gamma H(\zeta_s) ds} \frac{W_t^{1-\gamma}}{1 - \gamma} \right].$$

Unlike the calculation of $V(W,t)$, the worst case Radon-Nikodym derivative $\zeta$ for the value function $V_{\pi^{(2)}}(W,t)$ is endogenously determined for the fixed strategy $\pi^{(2)}$ (Flor and Larsen (2014)).

The wealth losses are reported in Table 5. We see that the CELs from adopting suboptimal portfolios are significant for a wide range of risk aversion. For instance, picking the wrong model causes CELs of 9.3% at $\gamma = 3$ and 35.5% at $\gamma = 1.1$ at a 20-year investment horizon. Generally, when it is relatively easy to identify the alternative model from the reference model (with a low DEP), or the alternative model is relatively far away from the reference model, the assumptions regarding the jump tail behaviors are important to an investor with low risk aversion because, as indicated in Table 4, the
investor takes relatively large jump exposures.

By further restricting $\phi$ to be identical for both distributions, we confirm that the CEL under ambiguity aversion is largely reduced. For example, when $\gamma = 1.1$, given $\phi = 10.8$ for both distributions, the CEL is 8.3% for a 20-year investment, much less than the CEL of 35.5% obtained when the degrees of ambiguity aversion for the two distributions are determined by matching DEPs. Table 4 indicates that, when matching a DEP=10%, $\phi$ is 15.6 for the tail distribution but 10.8 for the normal distribution. Thus, the investor is much more concerned about model uncertainty with the tail distribution as the true jump distribution. Accordingly, the welfare effect of uncertain jump tails may be primarily attributed to the investor’s high ambiguity aversion under the tail distribution which is even harder to separate from its alternatives.

Consistent with the negligible differences in portfolio outcomes provided in Table 4, the two reference models with different jump distributions result in CELs as small as $10^{-6}$ for 20-year investments. This result confirms that if the first two moments of the jump size distributions can be matched accurately, it may be safe for a CRRA investor to opt for an alternative distribution (e.g., normal) that shares the same first two moments. A minor economic loss will be caused by this misspecification of the jump size distribution when there is no jump ambiguity. However, in the presence of jump ambiguity, an ambiguity neutral investor who takes the reference model as true shall encounter much larger losses than an ambiguity averse investor in case that the worst case scenario occurs.

6 Tail risk and asset pricing

We have shown that tail risk significantly impacts optimal portfolio selection. We naturally speculate that it may also impact asset prices. The idea that aggregate consumption (dividends) is subject to rare disasters has a long tradition in finance (e.g., Rietz
(1988) and Barro (2006)). Recently, Barro and Jin (2011) calibrate the consumption process to international macroeconomic data and demonstrate that consumption jumps closely follow a slowly-decaying power law. They show that the calibrated consumption disaster model could resolve the equity premium puzzle with a low risk aversion and a CRRA utility function without ambiguity aversion. However, they do not discuss their model implications on derivatives pricing and the variance premium. Because fat tails of the jump size distributions are further exacerbated in the worst case scenarios, we expect the heavy-tailed consumption disasters to contribute even more to the “option smirk” pattern and variance premium under ambiguity aversion. We discuss this issue in this section.

In a pure-exchange economy, considering only one risky asset, a stock, \( (m = 1) \) and only one type of jump in our model \( (n = 1) \), we focus on the impact of the jump size distribution on asset prices. The dividend payout rate \( D_t \) (equivalent to the consumption rate in equilibrium) follows a jump-diffusion process

\[
dD_t = \mu D_t dt + \sigma D_t + Y_t D_t dN_t, \tag{30}
\]

with \( D_0 > 0 \). The pricing kernel in the equilibrium model becomes:

\[
d\kappa_t / \kappa_t = -r dt - \gamma \sigma dB_t + ((1 + Y_t)^{-\gamma} - 1)dN_t - \vartheta^* \lambda E^\zeta^* [(1 + Y_t)^{-\gamma} - 1] dt, \tag{31}
\]

where under the worst-case probability measure \( \zeta^* \), \( Y_t \) is the random jump size with density \( \varphi^*(y) \Phi(dy) \); \( N_t \) is a Poisson process with intensity \( \lambda^* = \vartheta^* \lambda \); \( \vartheta^* \) and \( \varphi^*(y) \) are defined by (10).

The stock price \( S(t) = A(t)D_t \) where \( A(t) \) is a deterministic function of \( t \) satisfying

\[15\] The insightful framework of Barro and Jin (2011) stimulates many follow-up studies (e.g., Wachter (2013), Wachter and Tsai (2016)).

\[16\] Detailed derivations are collected in Appendix C.
\[ A(T) = 0. \] We have

\[
dS_t = (\mu + A'(t)/A(t))S_t dt + \sigma S_t dB_t + Y_t S_t - dN_t, \tag{32}
\]

where \(Y_t\) follows the exact power law distribution calibrated by Barro and Jin (2011).\(^{17}\) For comparison, we calibrate an alternative normal distribution for \(Y_t\) to have the same mean and variance as the tail distribution. We set \(\sigma = 0.15\) for the entire study in this section. Since our focus is not to address the excess volatility of stock returns, we calibrate the volatility of dividend/consumption growth rate to match the volatility of stock returns as in Liu, Pan, and Wang (2005).\(^{18}\) The stock price under the risk-neutral probability \(Q\) is given by

\[
dS_t = (r - q)S_t dt + \sigma S_t d\mathcal{B}_t^Q + Y_t S_t - d\mathcal{N}_t^Q - \mathbb{E}^Q[Y_t]\lambda^Q S_t dt, \tag{33}
\]

where the risk-free interest rate \(r = 5\%\); the payout rate \(q\) is assumed to be a constant of \(3\%\) for simplicity as in Liu, Pan, and Wang (2005). Under \(Q\), \(d\mathcal{B}_t^Q\) is a standard Brownian motion; \(d\mathcal{N}_t^Q\) is a Poisson process with intensity \(\lambda^Q = \mathbb{E}^\varepsilon[(1 + Y_t)^{-\gamma}]\theta^* \lambda\); and \(Y_t\) has a density of \((1 + y)^{-\gamma} \varphi^*(y) \Phi(\Phi(dy)) / \mathbb{E}^\varepsilon[(1 + Y_t)^{-\gamma}]\).

The stock price dynamics under both the pricing and real measures enable us to characterize derivatives prices and the variance premium. We start with the variance

\(^{17}\)For simplicity, we opt for the single power law of Barro and Jin (2011). Let \(\log(1 + Y) = \log(1/\eta).\) Then \(\eta \sim f(x) = \alpha (1.105)^{\alpha x^{-(\alpha + 1)}}\) for \(x \geq 1/(1 - 0.095)\), where \(\alpha = 6.27\). We have \(\mathbb{E}[\log(1 + Y)] = -25.6\%\) and \(\text{Std}[\log(1 + Y)] = 15.0\%\). The jump intensity is \(\lambda = 0.0380\). The double power law of Barro and Jin (2011) has a lower upper-tail exponent of \(\alpha = 4.16\) thus the effects of heavy jump tails will be even more pronounced under this calibration.

\(^{18}\)Gabaix (2012) and Wachter (2013) propose a method to resolve the excess stock volatility puzzle by introducing time-varying disaster risk.
premium, \( v_{p_t} \), which is defined as follows.

\[
v_{p_t} = \int_t^{t+\Delta t} \mathbb{E}^Q[(d \log(S_s))^2] - \int_t^{t+\Delta t} \mathbb{E}[(d \log(S_s))^2] = \int_t^{t+\Delta t} \mathbb{E}^Q[(\log(1+Y_s))^2] \lambda ds - \int_t^{t+\Delta t} \mathbb{E}[(\log(1+Y_s))^2] \lambda ds = \int_t^{t+\Delta t} \left( \mathbb{E}[Z_s^2 e^{\gamma/\tau(e^{Z_s(1-\gamma)}-1)-\gamma Z_s}] - \mathbb{E}[Z_s^2] \right) \lambda ds, \tag{34}
\]

where \( Z_s = \log(1+Y_s) \) and \( \Delta t \) is the time period for variance measurement (equal to one month here). The variance premium can be further decomposed into two terms:

\[
v_{p_t} = \int_t^{t+\Delta t} \left( \mathbb{E}[Z_s^2 e^{-\gamma Z_s}] - \mathbb{E}[Z_s^2] \right) \lambda ds + \int_t^{t+\Delta t} \left( \mathbb{E}[Z_s^2 e^{\gamma/\tau(e^{Z_s(1-\gamma)}-1)-\gamma Z_s}] - \mathbb{E}[Z_s^2 e^{-\gamma Z_s}] \right) \lambda ds,
\]

where \( v_{p_t}^0 \) denotes the variance premium due to risk aversion but without ambiguity aversion and \( v_{p_t}^* \) denotes the extra variance premium corresponding to aversion to jump ambiguity.

The following proposition provides a sufficient condition under which the variance premium is positive.

**PROPOSITION 6.** \( v_{p_t}^0 > 0 \) if \( \mathbb{E}[Z_t^3] < 0 \); \( v_{p_t}^* > 0 \) if \( \mathbb{E}[Z_t^2 Y_t] < 0 \).

Particularly, for the normal jump size, \( Z_t \sim \text{normal} (\mu_J, \sigma_J) \), the condition \( \mathbb{E}[Z_t^2 Y_t] < 0 \) is equivalent to \( \mu_J + \sigma_J^2/2 < 0 \), which is generally satisfied in our calibration exercises. So is the condition \( \mathbb{E}[Z_t^3] < 0 \). The worst case density shifting to the left side has a more negative mean jump size, resulting in a larger variance premium. The tail distribution in the reference model further contributes to this premium since its fat tails get further amplified in the worst case.

Table 6 presents the variance premium over future one month in basis points (e.g., in terms of VIX^2/12). For \( \gamma = 2 \) and a modest DEP=36%, the variance premia are 6.15 and 10.54 under the normal and tail distributions, respectively. The data counterpart
reported by Drechsler (2013) is 10.55 for the sample period 1990 to 2009. Thus, the variance premium is significantly larger under the tail distribution with slowly-decaying tail behavior. Under the tail distribution, the variance premium without model uncertainty, \( v_p^0 \), is only 4.37. Thus jump ambiguity contributes a major portion of the total variance premium. The same effect is limited under the normal distribution, with variance premium increases mildly from 3.30 to 6.15 when ambiguity aversion is incorporated.

We now study the smirk premium as another example of the asset pricing implications of tail risk. The smirk premium\(^{19}\) with ambiguity aversion is 7.14% under the tail distribution versus 6.63% under the normal. In contrast, the smirk premia for the tail and normal distributions without jump ambiguity are much lower, computed at 5.61% and 5.60%, respectively. The average option smirk premium is 8.40% for the S&P500 index options from 1996 to 2013 (based on OptionMetrics data). Hence the representative investor’s aversion to uncertain jump risk largely improves the model’s ability to generate the high smirk premium observed in the market data.\(^{20}\) Moreover, heavy-tailed consumption jumps bring the smirk premium closer to data only when model mis-specification is incorporated, as witnessed by the negligible difference of the smirk premia from the both models without jump ambiguity. In sum, the investor pays large premia when purchasing index options to hedge uncertain consumption disaster risk, especially when a heavy tail characterizes the strike of rare disasters.

Before leaving this section, we further substantiate our results by calibrating the normal consumption jump model to match the risk premia generated under the advocated heavy-tailed jump model. We find that, without consideration of jump ambiguity, we need to double the standard deviation of the normal distribution to generate the same variance premium as that under the tail distribution. Even when jump ambiguity

\(^{19}\)We define the smirk premium as the difference between option implied volatility at \( K/S = 0.9 \) and \( K/S = 1.0 \), where \( K \) is the strike price. Option maturity is one month.

\(^{20}\)We focus on the effect of tail risk thus do not introduce any state variable into our model. We could generate an even higher option smirk premium to match the data by, e.g., incorporating time-varying disaster risk. It leads to stochastic volatility which is negatively correlated with stock returns. This leverage effect will make the implied volatility curve even steeper. See Seo and Wachter (2016).
is admitted, we still have to increase the standard deviation of the normal distribution by 40% to achieve the same variance premium, and by 30% to achieve the same smirk premium, as those under the tail distribution. Hence, the slowly-decaying consumption disaster distribution estimated by Barro and Jin (2011), combined with the investor’s robust decision-making regarding potential model mis-specifications, effectively induces the high variance and option smirk premia observed in the market.

7 Conclusion

This paper studies the effects of jump tail behavior on optimal portfolio choice and asset prices in the presence of jump ambiguity. We solve the portfolio choice problem in a multi-asset incomplete market using a novel decomposition technique. Both the optimal portfolio and the worst-case probability are obtained in closed forms. We then construct an equilibrium model and examine the variance premium and option prices under rare consumption disasters. To quantify the effects of the heavy jump tails, we calibrate the model to international equity indices for the portfolio model and borrow the slowly-decaying consumption disaster distribution estimated by Barro and Jin (2011).

We find that, due to the fear of severe tail incidents in the worst case scenario, an ambiguity averse investor diminishes portfolio diversification and may not participate in market if she is extremely ambiguity averse. Moreover, a jump distribution with a fatter left tail diminishes diversification even further by lowering the optimal jump exposure. Pursuing jump size distributions that differ largely in their tail behavior, we show that the economic losses from ignoring heavy jump tails are negligible in an expected utility model without jump ambiguity, as long as the first two moments of jump size distributions are put on an equal footing. In stark contrast, underestimating tail risk may result in sizeable wealth losses in the presence of jump ambiguity. Furthermore, ambiguity aversion towards heavy-tailed consumption disasters effectively enhances the
Appendix A. Proofs of propositions

Proof of Proposition 1 and Proposition 2

Proof. We conjecture $V(W, t) = U(W)h(t)$. Substituting the conjecture into (8), we obtain an equation for $h(t)$ as follows.

$$0 = \max_\pi \left\{ \frac{1}{1 - \gamma} \frac{dh(t)}{dt} - \frac{\gamma}{2} \pi'\Sigma'\pi + [\pi'(\mu - r_1)1_m + r] \right\}
+ \inf_{P(\xi) \in P} \left\{ \frac{1}{1 - \gamma} \sum_{k=1}^n \lambda_k E^\xi [(1 + \pi'J_kY_k)^{1-\gamma} - 1] + \frac{1}{\phi} H(\xi) \right\}
= \max_\pi \left\{ \frac{1}{1 - \gamma} \frac{dh(t)}{dt} - \frac{\gamma}{2} \pi'\Sigma'\pi + [\pi'(\mu - r_1)1_m + r] \right\}
+ \inf_{\xi} \frac{1}{1 - \gamma} \sum_{k=1}^n \lambda_k \int_A ((1 + \pi'J_ky_k)^{1-\gamma} - 1) \vartheta_k \varphi_k(y_k) \Phi_k(dy_k)
+ \frac{\lambda_k}{\phi} \int_A (\vartheta_k \varphi_k(y_k) \log(\vartheta_k \varphi_k(y_k))) + 1 - \vartheta_k \varphi_k(y_k)) \Phi_k(dy_k) \right\}. \tag{A.1}

Applying calculus of variations (e.g. Weinstock (1974))\textsuperscript{21}, we find the minimizer $\vartheta_k \varphi_k(\cdot)$ for the inner minimization problem by solving

$$\frac{1}{1 - \gamma} ((1 + \pi'J_ky_k)^{1-\gamma} - 1) + \frac{1}{\phi} \ln(\vartheta_k \varphi_k(y_k)) = 0. \tag{A.2}
$$

Then

$$\varphi_k^*(y_k) \vartheta_k^* = \exp\left( \frac{\phi}{\gamma - 1} ((1 + \pi'J_ky_k)^{1-\gamma} - 1) \right). \tag{A.3}
$$

Noting that $E[\varphi_k^*(Y_k)] = 1$, by taking expectation on both sides of (A.3), it follows that

$$\varphi_k^*(y_k) = \frac{1}{\vartheta_k^*} \exp\left( \frac{\phi}{\gamma - 1} ((1 + \pi'J_ky_k)^{1-\gamma} - 1) \right), \tag{A.4}
$$

$$\vartheta_k^* = E \left[ \exp\left( \frac{\phi}{\gamma - 1} ((1 + \pi'J_ky_k)^{1-\gamma} - 1) \right) \right]. \tag{A.5}
$$

Having found the worst probability $\zeta^*$ for any $\pi$, we next use the decomposition technique to find the optimal portfolio under ambiguity aversion. Let

$$\hat{\pi} = \Sigma'\pi. \tag{A.6}
$$

\textsuperscript{21}The idea of calculus of variations is to find the “first order condition” with respect to a function, in our case, to $\vartheta \varphi$. The optimal solution is obtained from the equation for the first order condition.
We decompose \( \hat{\pi} \) onto the space \( \hat{J} \) and its orthogonal space \( \hat{J}_\perp \):

\[
\hat{\pi} = \bar{\pi} + \pi_\perp,
\]

(A.7)

with \( \bar{\pi} \) in \( \hat{J} \) and \( \pi_\perp \) in \( \hat{J}_\perp \). Then equation (A.1) can be written as follows.

\[
0 = \frac{1}{1 - \gamma} \frac{dh_t}{dt} + r + \max_{\pi_\perp} \left( -\frac{\gamma}{2} \pi'_\perp \pi_\perp + \pi'_\perp \mu_\perp \right) \]

(A.8)

\[
+ \max_{\pi} -\frac{\gamma}{2} \bar{\pi}' \bar{\pi} + \bar{\pi}' \bar{\mu} + \frac{1}{1 - \gamma} \sum_{k=1}^{n} \lambda_k \mathbb{E}^{\xi_k}[(1 + \bar{\pi}' \hat{J}_k Y_k)^{1-\gamma} - 1] + \frac{1}{\phi} H(\zeta^*_t).
\]

Hence,

\[
\pi_\perp^* = \frac{1}{\gamma} \mu_\perp,
\]

(A.9)

and

\[
\pi^* = \arg \max_{\pi} -\frac{\gamma}{2} \bar{\pi}' \bar{\pi} + \bar{\pi}' \bar{\mu} + \frac{1}{1 - \gamma} \sum_{k=1}^{n} \lambda_k \left( \mathbb{E}^{\xi_k}[(1 + \bar{\pi}' \hat{J}_k Y_k)^{1-\gamma} - 1] \right) + \frac{1}{\phi} H(\zeta^*_t).
\]

(A.10)

The first order condition with respect to \( \bar{\pi} \) gives

\[
-\gamma \bar{\pi} + \bar{\mu} + \sum_{k=1}^{n} \lambda_k \mathbb{E}^{\xi_k}[(1 + \bar{\pi}' \hat{J}_k Y_k)^{-\gamma} Y_k] \hat{J}_k = 0.
\]

(A.11)

Note that we use the optimality of \( \zeta^*_k \) in the above derivation.

The optimal portfolio is given by

\[
\pi^* = (\Sigma')^{-1} \hat{\pi}^* = (\Sigma')^{-1} (\bar{\pi}^* + \pi^*_\perp).
\]

Finally, the worst case probability \( \zeta^* \) is obtained by substituting \( \pi = \pi^* \) into (A.4) and (A.5). The value function \( V(W,t) = U(W)h^*(t) \), where \( h^*(t) \) satisfies (A.1) with the optimizers given by \( \pi^* \) and \( \zeta^* \), respectively.

\[\square\]

**Proof of Proposition 3**

*Proof.* By Proposition 1,

\[
\varphi^*(y) = \frac{\exp \left( \frac{\phi}{\gamma - 1} ((1 + \pi^* y)^{1-\gamma} - 1) \right)}{\int_A \exp \left( \frac{\phi}{\gamma - 1} ((1 + \pi^* x)^{1-\gamma} - 1) \right) \Phi(dx)}.
\]
Let \( f(y) = \exp \left( \frac{\phi}{\gamma - 1}((1 + \hat{\pi}y)^{1-\gamma} - 1) \right) \). Then \( \varphi^*(y) = f(y) / \int_A f(x) \Phi(dx) \). Note that \( f(y) \) is monotonically decreasing in \( y \). Hence \( f(-1) > f(y) > f(\infty) \). Then

\[
f(-1) > \int_A f(y) \Phi(dy) > f(\infty).
\]

Since \( f(y) \) is a continuous function in \( y \), there exists \( y^* \) such that \( f(y^*) = \int_A f(y) \Phi(dy) \). Then for any \( y < y^* \), \( \varphi^*(y) > \varphi^*(y^*) = 1 \). Thus for any \( y \leq y^* \),

\[
Pr^*(Y_k < y|\phi) = \int_{-1}^{y} \varphi^*(y) \Phi(dy) > \int_{-1}^{y} \Phi(dy) = Pr(Y_k < y). \quad (A.12)
\]

To prove the second part of the proposition, we write \( \varphi(y; \phi) = \varphi^*(y) \) and denote \( g(y) = \frac{1}{\gamma - 1}((1 + \hat{\pi}y)^{1-\gamma} - 1) \). Then, for \( a \in (-1, \infty) \),

\[
\frac{d\varphi(a; \phi)}{d\phi} = f(a) \frac{g(a) \int_A f(y) \Phi(dy) - \int_A f(y) g(y) \Phi(dy)}{(\int_A f(y) \Phi(dy))^2}
\]

\[
= \frac{\int_{-1}^{a} (g(a) - g(y)) f(y) \Phi(dy) + \int_{-1}^{\infty} (g(a) - g(y)) f(y) \Phi(dy)}{f^{-1}(a)(\int_A f(y) \Phi(dy))^2}.
\]

The first term in the numerator is negative and approaches zero as \( a \) goes to \(-1\), while the second term is positive. Hence there exists a \( y^{**} \), such that for any \( y < y^{**} \), \( \varphi(y; \phi) \) is an increasing function in \( \phi \). And for any \( \phi_1 > \phi_2 \), there exists \( \hat{y} \), such that for any \( y < \hat{y} \)

\[
Pr^*(Y_k < y|\phi_1) = \int_{-1}^{y} \varphi(y; \phi_1) \Phi(dy) > \int_{-1}^{y} \varphi(y; \phi_2) \Phi(dy) = Pr^*(Y_k < y|\phi_2).
\]

\[
\square
\]

**Proof of Proposition 4**

We prove (2) as follows. A proof for (1) is similar to the proof for (2). We omit it.

**Proof.** We assume \( n = 1 \).\(^{22}\) Multiplying \( \hat{J}' \) from the left hand side on both sides of (18), and taking derivative of \( \hat{\pi}^* = \hat{J}' \hat{\pi}^* \) with respect to \( \hat{J}' \hat{J} \), we obtain

\[
\frac{1}{\gamma} \frac{\partial \hat{\pi}^*}{\partial (\hat{J}' \hat{J})} = \frac{\partial \hat{J}' \mu}{\partial (\hat{J}' \hat{J})} + \lambda \mathbb{E} \left[ (1 + \hat{\pi}^* Y)^{-\gamma} Y e^\frac{\phi}{\gamma - 1}(1 + \hat{\pi}^* Y)^{1-\gamma-1} \right]
\]

\[
+ \lambda (-\gamma) \mathbb{E} \left[ (1 + \hat{\pi}^* Y)^{-\gamma-1} Y^2 e^\frac{\phi}{\gamma - 1}(1 + \hat{\pi}^* Y)^{1-\gamma-1} \right] (\hat{J}' \hat{J}) \frac{\partial \hat{\pi}^*}{\partial (\hat{J}' \hat{J})}
\]

\[
+ \lambda \mathbb{E} \left[ (1 + \hat{\pi}^* Y)^{-\gamma} Y e^\frac{\phi}{\gamma - 1}(1 + \hat{\pi}^* Y)^{1-\gamma-1} \phi(-1)(1 + \hat{\pi}^* Y)^{-\gamma} \right] (\hat{J}' \hat{J}) \frac{\partial \hat{\pi}^*}{\partial (\hat{J}' \hat{J})}.
\]

\(^{22}\)For more types of jumps, the results may be mixed due to interactions among jumps, or between jumps and diffusions.
We can solve $\frac{\partial \pi^*}{\partial (\bar{J}^T \bar{J})}$ from the above equality, and it turns out that $\frac{\partial \pi^*}{\partial (\bar{J}^T \bar{J})} > 0$ if and only if the sum of the first two term of the right hand side is positive, or the condition (22) is satisfied.

Proof of Proposition 5

Proof. We introduce the following mild conditions on the density function $f^p(y)$ of $Y^p$ and $f^n(y)$ of $Y^n$ to prove the proposition. It worths mentioning that these conditions may be relaxed.

(1). Larger left tail: There exists $y_0 < 0$, such that $f^p(y) > f^n(y)$ for $y < y_0$.

(2). There exists $y_1 \in (y_0, 0)$ such that $f^p(y) > f^n(y)$, for $y \in (y_1, 0)$ and $f^p(y) < f^n(y)$ for $y \in (y_0, y_1)$.

(3). Smaller right tail: $f^p(y) < f^n(y)$ for $y > 0$ and $\int_{y_1}^{\infty} f^p(y)dy < \int_{y_1}^{\infty} f^n(y)dy$.

Condition (1) is a definition for the “slowly-decaying” or “fat” tail. Condition (2) states that there are two intersection points of the two density functions in the negative area. Condition (3) naturally follows because if the left tail is large then the right tail should be small. All three conditions are satisfied by the normal density and the power-law density studied in the calibration exercise of this paper.

Consider an interior solution $\pi^* \in (0, 1)$. It follows from (19) that

$$\pi^* = \frac{1}{\gamma} \bar{J}^T \bar{\mu} + \frac{1}{\gamma} \lambda \mathcal{E}^{\mathcal{C}^*}_{\gamma}[(1 + \pi^* Y)^{-\gamma} Y](\bar{J}^T \bar{J}).$$

(A.13)

Then $\pi^{*,n}$ and $\pi^{*,p}$ are the solutions to (A.13) regarding $Y^n$ and $Y^p$, respectively.

Assume $\pi^{*,p} > \pi^{*,n}$. Using monotonic properties of the function $\varphi^*(y, \pi)$ regarding $y$ and $\pi$ and under conditions (1)-(3) above, we show that

$$\lambda \mathcal{E}^{\mathcal{C}^*}_{\gamma}[(1 + \pi^{*,p} Y^p)^{-\gamma} Y^p](\bar{J}^T \bar{J}) < \lambda \mathcal{E}^{\mathcal{C}^*}_{\gamma}[(1 + \pi^{*,n} Y^n)^{-\gamma} Y^n](\bar{J}^T \bar{J}),$$

(A.14)

where $\zeta^{*,p}$ and $\zeta^{*,n}$ denote the worst probabilities regarding $Y^p$ and $Y^n$, respectively. The reason is that

$$\mathcal{E}^{\mathcal{C}^*}_{\gamma}[(1 + \pi^{*,p} Y^p)^{-\gamma} Y^p] - \mathcal{E}^{\mathcal{C}^*}_{\gamma}[(1 + \pi^{*,n} Y^n)^{-\gamma} Y^n]$$

$$= \int_{-\infty}^{\infty} G^p(y)f^p(y)dy - \int_{-\infty}^{\infty} G^n(y)f^n(y)dy$$

$$= \int_{-\infty}^{\infty} (G^p(y) - G^n(y))f^p(y)dy + \int_{-\infty}^{\infty} G^n(y)(f^p(y) - f^n(y))dy,$$

(A.15)

where $G^p(y) = (1+\pi^{*,p} y)^{-\gamma} y e^{-\frac{1}{\gamma^e}((1+\pi^{*,p} y)^{1-\gamma}-1)}$ and $G^n(y) = (1+\pi^{*,n} y)^{-\gamma} y e^{-\frac{1}{\gamma^e}((1+\pi^{*,n} y)^{1-\gamma}-1)}$.

Given $\pi^{*,p} > \pi^{*,n}$, we have $G^p(y) < G^n(y)$. Hence the first term in (A.15) is negative. Using the mild condition (3), we find $\int_{0}^{\infty} G^n(y)(f^p(y) - f^n(y))dy < 0$. In addition, since

---

23For cases with one or more than two intersection points, the proposition can be proved similarly.
$G^n(y)$ increases in $y$ for $y < 0$, using the mild condition (1) and (3), we obtain
\[
\int_{-\infty}^{0} G^n(y)(f^p(y) - f^n(y))dy \\
= \int_{-\infty}^{y_0} G^n(y)(f^p(y) - f^n(y))dy + \int_{y_0}^{y_1} G^n(y)(f^p(y) - f^n(y))dy + \int_{y_1}^{0} G^n(y)(f^p(y) - f^n(y))dy \\
< G^n(y_0) \int_{-\infty}^{y_0} (f^p(y) - f^n(y))dy + G^n(y_0) \int_{y_0}^{y_1} (f^p(y) - f^n(y))dy \\
= G^n(y_0) \int_{-\infty}^{y_1} (f^p(y) - f^n(y))dy < 0,
\]
where the last inequality is due to the condition (3). Hence (A.14) is proved. Then by (A.13), $\tilde{\pi}^{*,p} < \tilde{\pi}^{*,n}$, a contradiction with the assumption. Hence we must have $\tilde{\pi}^{*,p} < \tilde{\pi}^{*,n}$, i.e. the jump exposure under the slowly-decaying jump density $f^p$ is smaller than that under $f^n$.

\[\square\]

**Proof of Proposition 6**

**Proof.** Let $F(\pi) = \mathbb{E}[Z^2e^{\frac{\phi}{1+y}}(1+\pi Y)^{1-\gamma-1}e^{-\gamma Y}]$, $\pi \in [0, 1]$. Then
\[
F'(\pi) = \mathbb{E}[Z^2e^{\frac{\phi}{1+y}}(1+\pi Y)^{1-\gamma-1}(-\phi)(1+\pi Y)^{-\gamma}Ye^{-\gamma Y}] \\
= \mathbb{E}[1_{Y>0}Z^2e^{\frac{\phi}{1+y}}(1+\pi Y)^{1-\gamma-1}(-\phi)(1+\pi Y)^{-\gamma}Ye^{-\gamma Y}] \\
+ \mathbb{E}[1_{Y\leq 0}Z^2e^{\frac{\phi}{1+y}}(1+\pi Y)^{1-\gamma-1}(-\phi)(1+\pi Y)^{-\gamma}Ye^{-\gamma Y}] \\
\geq \phi \mathbb{E}[1_{Y>0}Z^2(-Y)] + \phi \mathbb{E}[1_{Y\leq 0}Z^2(-Y)] = \phi \mathbb{E}[Z^2(-Y)].
\]
Thus if $\mathbb{E}[Z^2(-Y)] > 0$, then $F'(\pi) > 0$ and $vp_t = F(1) - F(0) > 0$. Similarly we can prove the second part of the proposition. \[\square\]

**Appendix B. Detection-error probabilities**

Let $P$ be the probability measure associated with the reference model. The worst case probability measure $P(\zeta^*) \in \mathcal{P}$ has a Radon-Nikodym derivative, $\frac{dP(\zeta^*)}{dP} = \zeta^*_t = \prod_{k=1}^{m} \zeta^*_{t,k}$, with respect to $P$, where $\zeta^*_{t,k}$ is modelled by the stochastic differential equation
\[
\zeta^*_{t,k} = \zeta^*_{0,k} + \int_{0}^{t} \int_{A} (\varphi^*_{k}(s,y) - 1)\zeta^*_{s-k}g_k(ds,dy),
\]

41
where $\zeta_{0,k}^* = 1$. We assume only ambiguity of jump distribution. Then, by Ito’s lemma,

$$\ln(\zeta_{T,k}^*) = \lambda_k \int_0^T \int_A [\ln(\varphi_k^*(s,y)) + 1 - \varphi_k^*(s,y)] \Phi_k(s,dy)ds + \int_0^T \int_A \ln(\varphi_k^*(s,y))q_k(ds,dy)$$

$$= \lambda_k \int_0^T \int_A \ln(\varphi_k^*(s,y))\Phi_k(s,dy)ds + \int_0^T \int_A \ln(\varphi_k^*(s,y))q_k(ds,dy)$$

$$= \int_0^T \int_A \ln(\varphi_k^*(s,y))dN_{k,s}ds$$

with the second equality following from

$$\int_A \varphi_k^*(s,y)\Phi_k(s,dy) = \int_A \Phi_k(s,dy) = 1.$$ 

If the reference model with probability $P$ is the true model and $\zeta_{T,k}^* > 1$ or $\ln(\zeta_{T,k}^*) > 0$, the investor will reject $P$ for $P(\zeta^*)$ mistakenly. The corresponding probability of this event is

$$Pr(\ln(\zeta_{T,k}^*) > 0) = Pr \left( \int_0^T \int_A \ln(\varphi_k^*(s,y))dN_{k,s}ds > 0 \right). \quad (B.1)$$

The density of the jump size is $\Phi_k(t,dy)$.

Likewise, if the worst case model with probability $P(\zeta^*)$ is the true model and $\zeta_{T,k}^* < 1$ (or $\ln(\zeta_{T,k}^*) < 0$), then the investor will mistakenly reject $P^* = P(\zeta^*)$ for $P$. The corresponding probability of this event is

$$Pr^*(\ln(\zeta_{T,k}^*) < 0) = Pr^* \left( \int_0^T \int_A \ln(\varphi_k^*(s,y))dN_{k,s}ds < 0 \right). \quad (B.2)$$

The density of the jump size is $\varphi_k^*(t,y)\Phi_k(t,dy)$.

The detection-error probability $\varepsilon_T(\phi)$ is given by

$$\varepsilon_T(\phi) = \frac{1}{2}Pr(\ln(\zeta_{T,k}^*) > 0) + \frac{1}{2}Pr^*(\ln(\zeta_{T,k}^*) < 0).$$

We can employ a Monte Carlo approach to determine $\varepsilon_T(\phi)$ for each $\phi$.

**Appendix C. An equilibrium model with uncertain jump risk**

We consider an equilibrium model with a similar setup to Liu, Pan, and Wang (2005). These authors consider parametric jump sizes in both the reference and worst case models. We employ the nonparametric method detailed in Section 2. Hence we allow for reference jump size distributions with diverse tail properties and do not restrict the alternative models to certain distribution families.
The dividend payout rate $D_t$ satisfies

$$dD_t = \mu D_t dt + \sigma D_t dB_t + Y_t D_t - dN_t$$

with $D_0 > 0$. We consider the (equilibrium) stock price $S_t = A_t D_t$ where $A_t$ is a deterministic function of $t$ satisfying $A_T = 0$. Then $S_t$ under the physical measure follows

$$dS_t = (\mu + A_t' / A_t) S_t dt + \sigma S_t dB_t + Y_t S_t - dN_t,$$

And the budget constraint of the investor becomes

$$dW_t = (r + \pi_t (\mu - r + (A_t' + 1) / A_t)) W_t dt + \pi_t W_t \sigma dB_t + \pi_t - W_t - Y_t dN_t - c_t dt,$$  \hspace{1cm} (C.1)

where $\pi_t$ is the fraction of wealth invested in the stock, and $c_t$ is the consumption rate.

Analogous to the framework of Liu, Pan, and Wang (2005) (or Jin and Zhang (2012)), the investor’s utility in the worst case under ambiguity aversion (with probability measure $\zeta^*$) can be derived as follows.

$$U_t = \mathbb{E}_t^* \left[ \int_t^T e^{\int_s^t \left( \frac{1}{\phi} H(\zeta^*_s) - \rho \right) ds} \frac{c_t^{1-\gamma}}{1-\gamma} ds \right],$$

where $\rho$ is a constant discount rate. Given the budget constraint, the investor needs to choose consumption and investment plans $(c, \pi)$ to maximize utility. Let $\bar{V}$ be the indirect utility function of the investor. Then

$$\bar{V}(W, t) = \sup_{c, \pi} \mathbb{E}_t^*,$$

and $\bar{V}$ satisfies the following HJB equation:

$$\sup_{c, \pi} \frac{c_t^{1-\gamma}}{1-\gamma} - \rho \bar{V} + \mathcal{A} \bar{V}(W, t) + \lambda \mathbb{E}_t^* \left[ \bar{V}(W(1 + \pi Y_t), t) - \bar{V} \right] + \frac{1 - \gamma}{\phi} \bar{V} H(\zeta^*_t) = 0,$$  \hspace{1cm} (C.2)

where $\mathcal{A} \bar{V} = \frac{\partial \bar{V}}{\partial t} + \left( r + \pi \left( \mu - r + \frac{A_t' + 1}{A_t} \right) \right) W \frac{\partial \bar{V}}{\partial W} - c \frac{\partial^2 \bar{V}}{\partial W^2} + \frac{1}{2} \sigma^2 \pi^2 W^2 \frac{\partial^2 \bar{V}}{\partial W^2}$. Conjecture $\bar{V}(W, t) = \frac{W^{1-\gamma}}{1-\gamma} f(t)^{\gamma}$ and substitute it into the above HJB equation. The first order condition with respect to $c$ gives $c^* = W f^{-1}$. Then the HJB equation becomes

$$\sup_{\pi} \frac{\gamma}{1-\gamma} \frac{1 + f'}{f} - \rho \frac{1}{1-\gamma} + r + \pi \left( \mu - r + \frac{A_t' + 1}{A_t} \right) - \frac{1}{2} \gamma \sigma^2 \pi^2$$

$$+ \frac{1}{1-\gamma} \lambda \mathbb{E}_t^* [(1 + \pi Y_t)^{1-\gamma} - 1] + \frac{1}{\phi} H(\zeta^*_t) = 0.$$  \hspace{1cm} (C.3)

The first order condition with respect to $\pi$ gives

$$\mu - r + \frac{A_t' + 1}{A_t} - \gamma \pi \sigma^2 + \lambda \mathbb{E}_t^* [(1 + \pi Y_t)^{-\gamma} Y_t] = 0.$$
In equilibrium, $\pi = 1$ and $A = f$, (C.3) becomes

$$\frac{1}{1-\gamma} \frac{1 + f'}{f} - \rho \frac{1}{1-\gamma} + r + (\mu - r) - \frac{1}{2} \gamma \sigma^2 + \frac{1}{1-\gamma} \lambda \mathbb{E}^{\mathcal{C}}[(1 + Y_t)^{1-\gamma} - 1] + \frac{1}{\phi} H(\zeta^*_t) = 0.$$  

(C.4)

Hence $A'(t) = f'(t) = A(t) = \alpha(1 - e^{-\frac{T-t}{\alpha}})$, where

$$1/\alpha = \rho - (1-\gamma)\mu + \frac{1}{2} \gamma (1-\gamma) \sigma^2 - \lambda \mathbb{E}^{\mathcal{C}}[(1 + Y_t)^{1-\gamma} - 1] - \frac{1-\gamma}{\phi} H(\zeta^*_t).$$

Meanwhile, it follows the first order condition of $\pi$ that

$$\mu + \frac{1}{\alpha} - r = \gamma \sigma^2 - \lambda \mathbb{E}^{\mathcal{C}}[(1 + Y_t)^{-\gamma} Y_t].$$

Note that this equation gives the cum-dividend equity premium. Plugging $\alpha$ in the above, we have the interest rate:

$$r = \rho + \gamma \mu - \frac{1}{2} (1 + \gamma) \gamma \sigma^2 - \lambda \mathbb{E}^{\mathcal{C}}[(1 + Y_t)^{1-\gamma} - 1] + \lambda \mathbb{E}^{\mathcal{C}}[(1 + Y_t)^{-\gamma} Y_t] - \frac{1-\gamma}{\phi} H(\zeta^*_t).$$  

(C.5)

Next, we derive the pricing kernel. By Duffie and Skiadas (1994), the pricing kernel is given by

$$\kappa_t = e^{\int_0^t (-\rho + \frac{1-\gamma}{\alpha} H(\zeta^*_s)) ds} e^{-\gamma \tau_d} = e^{\int_0^t (-\rho + \frac{1-\gamma}{\alpha} H(\zeta^*_s)) ds} D(t)^{-\gamma}.$$  

(C.6)

Then by Ito’s lemma, we have

$$d\kappa_t/\kappa_{t-} = -rdt - \gamma\sigma dB_t + ((1 + Y_t)^{-\gamma} - 1)dN_t - \varphi^* \lambda \mathbb{E}^{\mathcal{C}}[((1 + Y_t)^{-\gamma} - 1)] dt;$$  

(C.7)

where we exploit (C.5) to simplify the $dt$ term. Thus, under the risk-neutral probability measure $(Q)$,

$$dS_t/S_{t-} = (r - q) dt + \sigma dB_t^Q + Y_t dN_t^Q - \mathbb{E}^{Q}[Y_t] \lambda^Q dt$$  

(C.8)

where $q$ is the time-varying payout rate and may be approximated by a constant for simplicity, as noted in Liu, Pan, and Wang (2005); $N_t^Q$ is a Poisson process with intensity $\lambda^Q = \mathbb{E}^{\mathcal{C}}[(1 + Y_t)^{-\gamma} \varphi^* \lambda]$ and $Y_t$ has a density $(1 + y)^{-\gamma} \varphi^*(y) \Phi(dy) / \mathbb{E}^{\mathcal{C}}[(1 + Y_t)^{-\gamma}].$

Now we may price financial derivatives, e.g., European call options, as follows.

$$C_t = e^{-r(T_0 - t)} \mathbb{E}^{Q}[(S_{T_0} - K)^+],$$  

(C.9)

where $C_t$ is the option price, $T_0$ is the option maturity date, and $K$ is the strike price. Put values can be obtained from put-call parity. We choose $r = 5\%$, $q = 3\%$ and use Monte Carlo method to directly evaluate the above expectation.

Finally, we describe our method to select the ambiguity aversion coefficient $\phi$ in the
equilibrium analysis. In principle, we have to select a $\phi$ such that the DEP is at least 10%. In equilibrium, the density for log stock price jump $Z_t = \log(1 + Y_t)$ is:

$$
\Phi^*(z) = \frac{\exp \left( \frac{\phi}{\gamma - 1} \left( e^{z(1-\gamma)} - 1 \right) \right)}{\mathbb{E} \left[ \exp \left( \frac{\phi}{\gamma - 1} \left( e^{Z_t(1-\gamma)} - 1 \right) \right) \right]} \Phi(dz).
$$

Since the numerator increases with $-z$ at an extremely high order, the equilibrium density shifts to negative infinity for most reference density functions $\Phi(dz)$, including the normal and power-law distribution. In this case, we cannot find a non-zero $\phi$ such that the DEP is above 10%.\(^{24}\) Economically, this issue is related to the antipuzzle of equity premium, e.g., Weitzman (2007) shows that the equity premium is infinite for lognormal stock prices with unknown variance. Barro and Jin (2011) point out that when the risk aversion coefficient is higher than the tail exponent of the power law governing consumption disasters, their model also encounters infinite equity premium.

To address this issue, Liu, Pan, and Wang (2005) consider an enhanced penalty function preventing an infinite worst case. Barro and Jin (2011) suggest instead that an infinite equity premium will be avoided if an upper bound is put on the disaster size, i.e., by requiring $-Y_t < 1$. For tractability, we opt for the simple relative entropy measure which allows us to solve the model explicitly (Maenhout (2004)). We follow Barro and Jin (2011) and consider a truncated version of the jump size density function in this paper. Specifically, we cut off the density at $Z_t = -0.91$ (or $Y_t = -60\%$). The observed disasters for OECD countries seldom exceed this upper bound in reality (Barro and Ursúa (2008)). We get qualitatively similar results when we select different cutoffs for $Z_t$. Therefore, our main results on the effect of jump tails on asset prices under ambiguity aversion are robust to this choice.

References


\(^{24}\)When the optimal portfolio $\pi^* \in [0, 1)$, the exponential function $\exp \left( \frac{\phi}{\gamma - 1} \left( (1 + \pi^*(e^{z} - 1))^{1-\gamma} - 1 \right) \right)$ in the worst case density converges to $\exp \left( \frac{\phi}{\gamma - 1} \left( (1 - \pi^*)^{1-\gamma} - 1 \right) \right)$ as $z$ goes to $-\infty$. Therefore, we do not encounter this infinite worst case problem in our portfolio choice study.


Figure 1. The reference density and the worst-case densities at different ambiguity aversion levels. We plot the reference density of the normal distribution. For comparison, we show the worst case densities for the normal under $\phi = 10$ and 50, respectively. We set $\gamma = 3$. The model parameters are taken from Table 2.
Figure 2. Ambiguity aversion and portfolio diversification. We show the individual and total jump exposures and the optimal portfolio weights of the investor when she invests into more risky assets. The model parameters are taken from Table 2 and we use the average values for $J$, $\sigma_n$, and $\mu$. We consider three levels of ambiguity aversion in the plotting.
Figure 3. The reference densities and the worst-case densities under different jump distributions. We plot the reference densities as the normal and tail distributions which share the same first two moments. The worst case densities for the two distributions are also shown for comparison under $\phi = 20$ and $\gamma = 3$. “Worst case of normal” denotes the density of the jump size in the worst case scenario when the jump size in log-price is normally distributed in the reference model.
Table 1 Summary Statistics for Equity Returns

This table reports summary statistics for the monthly continuously compounded returns of seven equity indices. Skew and ExKurt denote return skewness and excess kurtosis, respectively. The sample period is January 1993 to December 2015.

<table>
<thead>
<tr>
<th></th>
<th>U.S.</th>
<th>U.K.</th>
<th>GE</th>
<th>FR</th>
<th>CA</th>
<th>SD</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.0056</td>
<td>0.0029</td>
<td>0.0070</td>
<td>0.0052</td>
<td>0.0050</td>
<td>0.0090</td>
<td>0.0035</td>
</tr>
<tr>
<td><strong>S.T. Dev.</strong></td>
<td>0.0426</td>
<td>0.0401</td>
<td>0.0626</td>
<td>0.0455</td>
<td>0.0433</td>
<td>0.0528</td>
<td>0.0549</td>
</tr>
<tr>
<td><strong>Skew</strong></td>
<td>-0.8552</td>
<td>-0.6984</td>
<td>-0.8666</td>
<td>-0.8549</td>
<td>-1.2256</td>
<td>-0.6742</td>
<td>-0.5143</td>
</tr>
<tr>
<td><strong>ExKurt</strong></td>
<td>1.7615</td>
<td>0.7352</td>
<td>2.6257</td>
<td>2.1204</td>
<td>4.0930</td>
<td>1.7548</td>
<td>0.4362</td>
</tr>
</tbody>
</table>

Panel B: Correlation

<table>
<thead>
<tr>
<th></th>
<th>U.S.</th>
<th>U.K.</th>
<th>GE</th>
<th>FR</th>
<th>CA</th>
<th>SD</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>U.S.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.K.</td>
<td>0.8141</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>0.7634</td>
<td>0.7799</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>0.6968</td>
<td>0.7270</td>
<td>0.7149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>0.7747</td>
<td>0.6940</td>
<td>0.6346</td>
<td>0.5674</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>0.6254</td>
<td>0.6555</td>
<td>0.7031</td>
<td>0.6278</td>
<td>0.6020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JP</td>
<td>0.7542</td>
<td>0.8210</td>
<td>0.8780</td>
<td>0.7471</td>
<td>0.6465</td>
<td>0.6959</td>
<td></td>
</tr>
</tbody>
</table>
This table reports parameter estimates of the multivariate jump-diffusion model of stock index returns. We estimate the parameters by minimizing the sum of squared deviations of the return moments implied by model from those in data. All the parameter estimates and moments are on the monthly basis. The sample period is January 1993 to December 2015.

<table>
<thead>
<tr>
<th>U.S.</th>
<th>U.K.</th>
<th>GE</th>
<th>FR</th>
<th>CA</th>
<th>SD</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Jump parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_J$</td>
<td>-0.1363</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_J$</td>
<td>0.0831</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0751</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$J$</td>
<td>0.5381</td>
<td>0.4401</td>
<td>0.8208</td>
<td>0.5856</td>
<td>0.6503</td>
<td>0.6464</td>
</tr>
<tr>
<td>Panel B: Other parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{n1}$</td>
<td>0.0362</td>
<td>0.0271</td>
<td>0.0344</td>
<td>0.0221</td>
<td>0.0225</td>
<td>0.0220</td>
</tr>
<tr>
<td>$\sigma_{n2}$</td>
<td>0</td>
<td>0.0232</td>
<td>0.0165</td>
<td>0.0121</td>
<td>0.0042</td>
<td>0.0128</td>
</tr>
<tr>
<td>$\sigma_{n3}$</td>
<td>0</td>
<td>0</td>
<td>0.0349</td>
<td>0.0061</td>
<td>-0.0025</td>
<td>0.0114</td>
</tr>
<tr>
<td>$\sigma_{n4}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0282</td>
<td>-0.0030</td>
<td>0.0047</td>
</tr>
<tr>
<td>$\sigma_{n5}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0242</td>
<td>0.0052</td>
</tr>
<tr>
<td>$\sigma_{n6}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0350</td>
</tr>
<tr>
<td>$\sigma_{n7}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.0115</td>
<td>0.0078</td>
<td>0.0166</td>
<td>0.0117</td>
<td>0.0120</td>
<td>0.0164</td>
</tr>
</tbody>
</table>
Table 3  Moment Comparison

This table reports return moments reconstructed using the theoretic moment conditions in equations (26) and the model parameter estimates in Table 2 (labelled “Normal”). For comparison, we also report the empirical moments computed from the return data (labelled “Data”) and the moments computed under the calibrated tail distribution following a power law (labelled “Tail”).

<table>
<thead>
<tr>
<th></th>
<th>U.S.</th>
<th>U.K.</th>
<th>GE</th>
<th>FR</th>
<th>CA</th>
<th>SD</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skew: Data</td>
<td>-0.8552</td>
<td>-0.6984</td>
<td>-0.8666</td>
<td>-0.8549</td>
<td>-1.2256</td>
<td>-0.6742</td>
<td>-0.5143</td>
</tr>
<tr>
<td>Skew: Normal</td>
<td>-0.6873</td>
<td>-0.4366</td>
<td>-0.8493</td>
<td>-0.7377</td>
<td>-1.1945</td>
<td>-0.6473</td>
<td>-0.3337</td>
</tr>
<tr>
<td>Skew: Tail</td>
<td>-0.7797</td>
<td>-0.4901</td>
<td>-1.0006</td>
<td>-0.8417</td>
<td>-1.3739</td>
<td>-0.7441</td>
<td>-0.3789</td>
</tr>
<tr>
<td>ExKurt: Data</td>
<td>1.7615</td>
<td>0.7352</td>
<td>2.6257</td>
<td>2.1204</td>
<td>4.093</td>
<td>1.7548</td>
<td>0.4362</td>
</tr>
<tr>
<td>ExKurt: Normal</td>
<td>1.8947</td>
<td>1.0314</td>
<td>2.5389</td>
<td>2.0858</td>
<td>3.9752</td>
<td>1.7558</td>
<td>0.7234</td>
</tr>
<tr>
<td>ExKurt: Tail</td>
<td>2.8153</td>
<td>1.4975</td>
<td>4.0933</td>
<td>3.1367</td>
<td>6.0829</td>
<td>2.6839</td>
<td>1.0766</td>
</tr>
</tbody>
</table>
Table 4 Optimal Portfolios at Different Detection-Error Probabilities (DEPs)

This table reports optimal portfolio weights at three DEP levels. The ambiguity aversion coefficients \( \phi \) are determined by the corresponding DEPs. \( \tilde{\pi} \) denotes exposure to jump risk. Results under normal and tail jumps are presented together for comparison. The last two columns correspond to the optimal portfolios and the exposures to jump risk without ambiguity.

<table>
<thead>
<tr>
<th>Panel A: ( \gamma = 3 )</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>No Ambiguity</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi )</td>
<td>28.6</td>
<td>36.8</td>
<td>18.1</td>
<td>20.4</td>
<td>12.3</td>
<td>13.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>U.S.</td>
<td>0.9364</td>
<td>0.9511</td>
<td>0.9264</td>
<td>0.9387</td>
<td>0.9179</td>
<td>0.9293</td>
<td>0.8755</td>
<td>0.8763</td>
<td></td>
</tr>
<tr>
<td>U.K.</td>
<td>-1.8255</td>
<td>-1.8022</td>
<td>-1.8413</td>
<td>-1.8218</td>
<td>-1.8546</td>
<td>-1.8367</td>
<td>-1.9217</td>
<td>-1.9204</td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>0.8274</td>
<td>0.7605</td>
<td>0.8726</td>
<td>0.8167</td>
<td>0.9111</td>
<td>0.8596</td>
<td>1.1040</td>
<td>1.1001</td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>0.0080</td>
<td>-0.0465</td>
<td>0.0449</td>
<td>-0.0007</td>
<td>0.0763</td>
<td>0.0343</td>
<td>0.2334</td>
<td>0.2302</td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>-0.4973</td>
<td>-0.5839</td>
<td>-0.4388</td>
<td>-0.5111</td>
<td>-0.3890</td>
<td>-0.4556</td>
<td>-0.1395</td>
<td>-0.1445</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.0736</td>
<td>1.0567</td>
<td>1.0850</td>
<td>1.0709</td>
<td>1.0947</td>
<td>1.0817</td>
<td>1.1434</td>
<td>1.1424</td>
<td></td>
</tr>
<tr>
<td>JP</td>
<td>-0.8953</td>
<td>-0.8133</td>
<td>-0.9508</td>
<td>-0.8822</td>
<td>-0.9980</td>
<td>-0.9349</td>
<td>-1.2345</td>
<td>-1.2297</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\pi} )</td>
<td>0.2671</td>
<td>0.1759</td>
<td>0.3288</td>
<td>0.2526</td>
<td>0.3812</td>
<td>0.3111</td>
<td>0.6440</td>
<td>0.6387</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: ( \gamma = 2 )</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>No Ambiguity</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi )</td>
<td>19.6</td>
<td>25.7</td>
<td>12.5</td>
<td>14.4</td>
<td>8.1</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>U.S.</td>
<td>1.4058</td>
<td>1.4281</td>
<td>1.3913</td>
<td>1.4104</td>
<td>1.3774</td>
<td>1.3944</td>
<td>1.3166</td>
<td>1.3188</td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>1.2355</td>
<td>1.1340</td>
<td>1.3010</td>
<td>1.2145</td>
<td>1.3645</td>
<td>1.2871</td>
<td>1.6406</td>
<td>1.6304</td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>0.0075</td>
<td>-0.0752</td>
<td>0.0609</td>
<td>-0.0097</td>
<td>0.1126</td>
<td>0.0496</td>
<td>0.3375</td>
<td>0.3292</td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>-0.7532</td>
<td>-0.8845</td>
<td>-0.6684</td>
<td>-0.7804</td>
<td>-0.5864</td>
<td>-0.6864</td>
<td>-0.2292</td>
<td>-0.2424</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.6090</td>
<td>1.5834</td>
<td>1.6255</td>
<td>1.6037</td>
<td>1.6415</td>
<td>1.6220</td>
<td>1.7111</td>
<td>1.7086</td>
<td></td>
</tr>
<tr>
<td>JP</td>
<td>-1.3361</td>
<td>-1.2117</td>
<td>-1.4165</td>
<td>-1.3103</td>
<td>-1.4943</td>
<td>-1.3994</td>
<td>-1.8328</td>
<td>-1.8203</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\pi} )</td>
<td>0.3931</td>
<td>0.2548</td>
<td>0.4824</td>
<td>0.3644</td>
<td>0.5688</td>
<td>0.4634</td>
<td>0.945</td>
<td>0.9311</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: ( \gamma = 1.1 )</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>Normal</th>
<th>Tail</th>
<th>No Ambiguity</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi )</td>
<td>10.8</td>
<td>15.6</td>
<td>6.8</td>
<td>8</td>
<td>4.8</td>
<td>5.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>U.S.</td>
<td>2.5576</td>
<td>2.6020</td>
<td>2.5318</td>
<td>2.5669</td>
<td>2.5125</td>
<td>2.5419</td>
<td>2.5099</td>
<td>2.5099</td>
<td></td>
</tr>
<tr>
<td>U.K.</td>
<td>-4.9725</td>
<td>-4.9023</td>
<td>-5.0133</td>
<td>-4.9579</td>
<td>-5.0438</td>
<td>-4.9974</td>
<td>-5.0480</td>
<td>-5.0480</td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>2.2387</td>
<td>2.0371</td>
<td>2.3561</td>
<td>2.1968</td>
<td>2.4439</td>
<td>2.3103</td>
<td>2.4557</td>
<td>2.4557</td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>0.0074</td>
<td>-0.1569</td>
<td>0.1030</td>
<td>-0.0268</td>
<td>0.1745</td>
<td>0.0657</td>
<td>0.1842</td>
<td>0.1842</td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>-1.3793</td>
<td>-1.6402</td>
<td>-1.2275</td>
<td>-1.4335</td>
<td>-1.1140</td>
<td>-1.2867</td>
<td>-1.0986</td>
<td>-1.0986</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>2.9235</td>
<td>2.8727</td>
<td>2.9531</td>
<td>2.9130</td>
<td>2.9753</td>
<td>2.9416</td>
<td>2.9783</td>
<td>2.9783</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\pi} )</td>
<td>0.7043</td>
<td>0.4296</td>
<td>0.8642</td>
<td>0.6472</td>
<td>0.9838</td>
<td>0.8018</td>
<td>0.9999</td>
<td>0.9999</td>
<td></td>
</tr>
</tbody>
</table>
Table 5  Certainty Equivalent Losses (CELs): Tail (True) versus Normal

This table reports the CELs when the investor fails to properly account for heavy jump tails. The jump size distribution in log-price under the true model is the tail distribution obeying a power law. The CEL is incurred when the investor switches to the suboptimal portfolio under the alternative model with light-tailed normal jump sizes.

<table>
<thead>
<tr>
<th>DEP</th>
<th>Investment horizon (in years)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>10</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Panel A: $\gamma = 3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>0.0048</td>
<td>0.0474</td>
<td>0.0925</td>
<td></td>
</tr>
<tr>
<td>15%</td>
<td>0.0020</td>
<td>0.0196</td>
<td>0.0388</td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td>0.0012</td>
<td>0.0124</td>
<td>0.0247</td>
<td></td>
</tr>
<tr>
<td>Panel B: $\gamma = 2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>0.0080</td>
<td>0.0770</td>
<td>0.1481</td>
<td></td>
</tr>
<tr>
<td>15%</td>
<td>0.0034</td>
<td>0.0338</td>
<td>0.0665</td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td>0.0019</td>
<td>0.0193</td>
<td>0.0382</td>
<td></td>
</tr>
<tr>
<td>Panel C: $\gamma = 1.1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>0.0217</td>
<td>0.1969</td>
<td>0.3551</td>
<td></td>
</tr>
<tr>
<td>15%</td>
<td>0.0072</td>
<td>0.0698</td>
<td>0.1348</td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td>0.0037</td>
<td>0.0368</td>
<td>0.0723</td>
<td></td>
</tr>
</tbody>
</table>
Table 6 Variance and Option Smirk Premia

This table reports variance premium (per month in basis points) and option implied volatility smirk premium (in percentages) for both the normal and tail jump distributions. The smirk premium is defined as the difference between implied volatilities at $K/S = 0.9$ and $K/S = 1.0$. We set $\gamma = 2$ and the ambiguity aversion coefficients are determined at DEP=36%.

<table>
<thead>
<tr>
<th></th>
<th>Variance premium</th>
<th>Smirk premium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Real data</td>
<td>10.55</td>
<td>8.40</td>
</tr>
<tr>
<td>Panel B: W/O jump ambiguity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Jump</td>
<td>3.30</td>
<td>5.60</td>
</tr>
<tr>
<td>Tail jump</td>
<td>4.37</td>
<td>5.61</td>
</tr>
<tr>
<td>Panel C: W/ jump ambiguity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Jump</td>
<td>6.15</td>
<td>6.63</td>
</tr>
<tr>
<td>Tail Jump</td>
<td>10.54</td>
<td>7.14</td>
</tr>
</tbody>
</table>