2区

Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Tail Risk and Robust Portfolio Decisions

Xing Jin, Dan Luo, Xudong Zeng*

informs
http://pubsonline.informs.org/journal/mnsc
MANAGEMENT SCIENCE
Articles in Advance, pp. 1–22
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Tail Risk and Robust Portfolio Decisions

Xing Jin,[a] Dan Luo,[b,c] Xudong Zeng[b,d,*]

[a] Department of Mathematical Finance, Tianjin University of Finance and Economics, Tianjin 300222, China; [b] School of Finance, Shanghai University of Finance and Economics, Shanghai 200433, China; [c] Shanghai Key Laboratory of Financial Information Technology, Shanghai 200433, China; [d] Institute of Scientific Computation and Financial Data Analysis, Shanghai University of Finance and Economics, Shanghai 200433, China

*Corresponding author

Contact: xingjin2019@outlook.com, https://orcid.org/0000-0001-8651-0930 (XJ); luo.dan@mail.shufe.edu.cn, https://orcid.org/0000-0001-9009-5215 (DL); zeng.xudong@mail.shufe.edu.cn, https://orcid.org/0000-0002-5079-8378 (XZ)

Received: April 18, 2017
Revised: January 14, 2019; November 30, 2019
Accepted: February 8, 2020
Published Online in Articles in Advance: June 9, 2020

https://doi.org/10.1287/mnsc.2020.3615

Abstract. This paper formulates a portfolio choice problem in a multiasset incomplete market characterized by ambiguous jumps and arbitrary tail assumptions. We derive the optimal portfolio in closed form through a decomposition approach. We show that, due to fear of tail incidents, an investor diminishes portfolio diversification, and even more so under heavy-tailed jumps that intensify misspecification concerns. We then implement our model in international equity markets to quantify the impact of tail risk on portfolio selection, through comparisons between a normal and a slowly decaying jump size distribution. We find that, without jump ambiguity, constant relative risk aversion (CRRA) investors increase their jump exposures merely slightly and suffer negligible wealth losses from underestimating tail risk, given that the first two moments of the jump size distributions are preserved regardless of the tail properties. In stark contrast, sizable portfolio rebalancing and subsequent wealth losses are encountered in the presence of jump ambiguity.

History: Accepted by David Simchi-Levi, finance.
Funding: D. Luo is supported in part by the Natural Science Foundation of China [Grant 71972123/G0205]. X. Zeng gratefully acknowledges the support of the Natural Science Foundation of China [Grant 71771142].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2020.3615.

Keywords: tail risk • jump ambiguity • robust decisions • portfolio selection

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] (Liu et al. 2005, Jin and Zhang 2012, Branger and Larsen 2013, Drechsler 2013, Aït-Sahalia and Matthys 2019). Recently, Barro and Jin (2011) and Bollerslev and Todorov (2011a) documented that consumption/return jump sizes follow slowly decaying distributions with heavy tails leading to severe tail events, which are formally outside of the traditional framework of normally distributed jumps with merely 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.

The goal of this paper is to investigate the effects of jump tail behavior on optimal portfolio formation when the investor takes caution with misspecifications in the jump distribution. This consideration of jump ambiguity is especially relevant because heavy-tailed jump models are most challenging to pin down and thus naturally raise the investor's misspecification concerns. The investor implements robust decision making (Hansen and Sargent 2008) to guard against the worst-case alternative model. We show that the heavy tail of a slowly decaying jump distribution is largely exacerbated under the worst-case scenario, and in reaction, investors diminish their jump exposure and diversification of the optimal portfolio. To provide quantitative evaluation of tail risk in portfolio management, we apply our model to international equity indices and industry portfolios. The results confirm the aforementioned theoretical findings and indicate sizeable welfare losses of an ambiguity-averse investor from ignoring extreme tail events.

We study a portfolio choice problem in a multiasset jump-diffusion market that is incomplete. The investor is averse to ambiguity in the jump size distribution (and jump frequency). Following the robust control framework of Anderson et al. (2003),[2] we solve a Hamilton-Jacobi-Bellman (HJB) 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. We solve for the worst-case probability explicitly through our 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.

Although not centering on tail risk, several studies have considered ambiguity in jump size distributions. For instance, Liu et al. (2005), Drechsler (2013), and Aït-Sahalia and Matthys (2019) pursue parametric alternative models. Their analyses are subject to two limitations. First, the parametric approach requires the alternative models to be in the same distribution family as the reference model. Second, fully analytic solutions are worked out in special cases while approximations are generally implemented. To our knowledge, the present paper is the first to exploit a nonparametric method, which relaxes both restrictions, to systematically examine tail risk. Our "worst case" of the entire neighborhood of nearby models is worse than that from a subset of parametric alternative models. Furthermore, our nonparametric method readily applies to any jump distribution with equal efforts. This flexibility greatly facilitates studies of jump distributions with wide-range tail properties, from slowly decaying power law to normal.[3] Our method also enables us to track the investor's aversion to potential model misspecifications when the reference models are sufficiently distinct.

Exploiting the closed-form solutions, we show that, due to the fear of severe tail events in the worst-case scenario, investors reduce their jump exposure when they become more ambiguity averse, and ambiguity aversion toward systemic jumps diminishes portfolio diversification. An extremely ambiguity-averse investor may choose zero exposure to jump risk, hence not participating in the financial market. This result provides an alternative explanation for the nonparticipation puzzle.[4] Furthermore, we confirm that ambiguity-averse investors reduce more of their jump exposure if the jump distribution exhibits a fatter left tail that further diminishes portfolio diversification.[5]

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 implement our model in 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 later, 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 30.2% under relatively low risk aversion. For a robustness check, we apply our model to an alternative data set of five industry portfolios. We observe welfare losses of similar magnitude for the ambiguity-averse investor.

Expecting the distinct statistic properties of the two jump size distributions, we determine the investor's degree of ambiguity aversion toward different jump models using the model detection error probability (DEP) approach of Anderson et al. (2003). The models are compared under the same DEP. We affirm that, given 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 both distributions, we find that wealth equivalent losses are largely reduced. Hence, the aforementioned 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 local mean-variance property of the CRRA utility that fails to capture higher moments (Hong et al. 2007, Cvitanić et al. 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.[6] 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.[7]

Our paper is closely related to the recent studies on the measurement of tail risk, including Barro and Jin (2011), Bollerslev and Todorov (2011a,b), Kelly and Jiang (2014), Bollerslev et al. (2015), Agarwal et al. (2017), and others. These authors also examine the pricing of tail risk in stocks and hedge funds. Building on these researches, we focus on the portfolio implications of tail risk, especially when the investor acknowledges the difficulties in estimating heavy-tailed jump size distributions.

One major contribution of this paper is that we solve the optimal portfolio choice problem in closed form for an arbitrary jump distribution in the multiasset reference model. Jin and Zhang (2012) also study the optimal portfolio problem in a multiasset jump-diffusion model but only consider ambiguity in jump intensity. They use "fictitious" stocks to address the case in which the total number of Brownian motions and jumps exceeds the number of risky assets and do not provide the optimal portfolio explicitly, even without jump ambiguity. Compared with the parametric approach of Liu et al. (2005), our nonparametric approach enables us to easily analyze different types of jump distributions possibly with heavy tails, other than the prevalent normal distribution. They consider only one stock. Drechsler (2013) constructs an equilibrium model and shows that fundamentals and model uncertainty may explain a wide range of asset prices. The author adopts parametric alternative models and relies on an analytical approximation to the solution. Branger and Larsen (2013) show pronounced differences between ambiguity aversion with respect to diffusion risk and jump risk. They assume only jump intensity ambiguity and consider a single-stock model. Aït-Sahalia and Matthys (2019) derive robust consumption and portfolio policies of an investor with recursive preferences, in a model with one risky asset following a Lévy jump-diffusion process. None of the prior studies analyze the role of jump tail behavior in optimal portfolio selection.

The optimal portfolio is obtained through decomposition into exposures to diffusion risk and jump risk. The two components may be solved independently with a remarkable reduction of computational complexity, especially for cases with a large number of risky assets. This decomposition method is based on orthogonal vector decompositions and generalizes the approach originally proposed by Aït-Sahalia et al. (2009), where their decomposition relies on special structures of the model parameters. Our approach is more flexible, free of such restrictions. It can be easily applied to other relevant multiasset jump-diffusion models (e.g., those studied by Jin and Zhang 2012 and Jin et al. 2018).

The concurrent work of Jin et al. (2018) develops a pathwise optimization procedure to solve for the optimal portfolio and the worst-case probability. They demonstrate the notable difference between the parametric and nonparametric approaches for jump ambiguity under a normal reference model. The present paper significantly differs from theirs in terms of methodology and economic scope. First, their duality-based pathwise method cannot be trivially applied to our model because the prices of diffusion and jump risks cannot be separated here. Specifically, their model consists of the same number of risky assets as the total number of risk resources; in contrast, the number of risk resources can exceed the number of stocks here. Our model is more widely used in the literature. Second and more importantly, apart from their focus on methodology, the theme of this paper is to examine the implications of jump tail behavior on portfolio selection, both theoretically and empirically. The explicitly solved jump exposures and portfolio weights through decomposition are vital to our theoretic analyses on tail risk, whereas their "fictitious" completion method cannot produce such analytic results.

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 explains diversification in risky assets as an implication of our solution. Section 5 gauges the economic significance of tail risk for an ambiguity-averse investor. Section 6 concludes. All proofs are collected in the appendices.

2. Merton's Portfolio Problem with Jump Ambiguity

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 $(\Omega, \mathscr{F}, P)$, where Ω is the set of states of nature with generic element ω, $\mathscr{F}$ is the σ-algebra of observable events, and P is a probability measure on $(\Omega, \mathscr{F})$. 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 modeled by the linear stochastic differential equation:

$$\frac{dS_{i,t}}{S_{i,t-}} = \mu_i dt + \sum_{j=1}^{m} \sigma_{i,j} dB_{j,t} + \sum_{k=1}^{n} J_{i,k} Y_{k,t} dN_{k,t}, \quad i = 1, 2, \ldots, m, \tag{1}$$

where μ_i is the (constant) growth rate, $\mathbf{B}_t = (B_{1,t}, \ldots, B_{m,t})'$ is an m-dimensional standard Brownian motion, and $\mathbf{N}_t = (N_{1,t}, \ldots, N_{n,t})'$ is a standard 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 λ_k. The amplitude $Y_{k,t}$ of the type k jump has the probability density $\Phi_k(t, dy)$, independent of $\mathbf{B}_t$ and $\mathbf{N}_t$. 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 $\Sigma = (\sigma_{i,j})_{1\le i,j\le m}$ and the jump coefficient matrix $\mathbf{J} = (J_{i,k})_{1\le i\le m, 1\le k\le n}$. The parameter $J_{i,k} \in [0,1]$ is a jump scaling coefficient for each i, k. We focus on the realistic case of $n \le m$ (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 $\text{rank}(\mathbf{J}) = n$.

The Brownian motions represent frequent small movements in the stock prices, whereas the jump processes capture occasional large shocks to the market. Both $\mathbf{B}_t$ and $\mathbf{N}_t$ 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 $\{\mathcal{F}_t\}_{t\in[0,T]} = \{\mathcal{F}_t^B \vee \mathcal{F}_t^N,\ t \in [0, T]\}$, where $\mathcal{F}_t^B = \sigma(\mathbf{B}_s; 0 \le s \le t)$, and $\mathcal{F}_t^N = \sigma(\mathbf{N}_s; 0 \le s \le t)$). 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 aforementioned $m+1$ assets. Let $\pi_t = (\pi_{1,t}, \ldots, \pi_{m,t})'$ denote a portfolio, where $\pi_{i,t}$ $(1 \le i \le m)$ is the proportion of total wealth invested in the ith risky asset at time t and is $\mathcal{F}_t$-predictable. Any portfolio policy π_t has an associated wealth process W_t that evolves as follows:

$$W_t = W_0 + \int_0^t rW_s ds + \int_0^t W_s\pi_s'(\mu - r\mathbf{1}_m)ds + \int_0^t W_s\pi_s'\Sigma d\mathbf{B}_s + \int_0^t W_{s-}\pi_{s-}'\mathbf{J}\mathbf{Y}_s d\mathbf{N}_s, \tag{2}$$

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

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

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

$$u(w_0) = \max_{\pi\in\mathcal{A}(w_0)} \mathbf{E}[U(W_T)], \tag{3}$$

where $U(W) = W^{1-\gamma}/(1-\gamma)$, with constant relative risk aversion coefficient $\gamma > 0$ and $\gamma \ne 1$;[8] $\mathbf{E}[\cdot]$ 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 that 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, whereas 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.[9] 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 also explicitly recognizing 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 his or her portfolio would perform reasonably well even if the worst-case scenario occurs. However, choosing any model other than the reference model results in a penalty 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 on $\mathcal{F}_t$:

$$\left.\frac{dP(\zeta)}{dP}\right|_{\mathcal{F}_t} = \zeta_t = \prod_{k=1}^n \zeta_{k,t}, \tag{4}$$

where $\zeta_{k,t}$ is modeled by the stochastic differential equation

$$d\zeta_{k,t} = d\zeta_{k,t}(\vartheta_k, \varphi_k) = \int_A (\vartheta_k(t)\varphi_k(t,y) - 1)\zeta_{k,t-}q_k(dt, dy), \quad t \in [0, T], \tag{5}$$

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

$$\int_A \varphi_k(t,y)\Phi_k(t,dy) = 1, \quad k = 1, \ldots, n, \tag{6}$$

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

Under the probability measure $P(\zeta)$, the kth jump intensity λ_k and the density function $\Phi_k(t, dy)$ are changed into $\vartheta_k\lambda_k$ and $\varphi_k(t, y)\Phi_k(t, dy)$ 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, $\mathbf{E}^{\zeta}[\ln(\frac{dP(\zeta)}{dP}|_{\mathcal{F}_T})]$, of Anderson et al. (2003) as follows:

$$\mathbf{E}^{\zeta}\left[\ln\left(\frac{dP(\zeta)}{dP}\bigg|_{\mathcal{F}_T}\right)\right] = \mathbf{E}^{\zeta}[\ln\zeta_T] = \mathbf{E}[\zeta_T\ln(\zeta_T)] = \int_0^T H(\zeta_s)ds, \tag{7}$$

where

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

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

Ambiguity-averse investors search for an optimal portfolio to maximize their CRRA utility function against the worst case of the jump-diffusion model. Two specifications of robustness are frequently explored in the literature. For the first specification of ambiguity aversion, known as multiplier robustness, following the robust decision framework developed by Anderson et al. (2003) and Maenhout (2004),[10] we solve a robust version of the Hamilton-Jacobi-Bellman equation for the investor's value function $V(W,t)$:

$$\begin{aligned}0 = \max_{\pi}\inf_{P(\zeta)\in\mathcal{P}}\bigg\{&\frac{\partial V(W,t)}{\partial t} + W\big(\pi'\big(\mu - r\mathbf{1}_m\big)+r\big)\\ &\times\frac{\partial V(W,t)}{\partial W} + \frac{1}{2}W^2\pi'\Sigma\Sigma'\pi\frac{\partial^2 V(W,t)}{\partial W^2}\\ &+\sum_{k=1}^{n}\lambda_k\mathbf{E}^{\zeta_k}\big[V\big(W + W\pi'\mathbf{J}_k Y_{k,t},t\big)\\ &- V(W,t)\big] + \frac{1}{\theta_t}H(\zeta_t)\bigg\},\end{aligned} \tag{8}$$

with $V(W,T) = \frac{W^{1-\gamma}}{1-\gamma}$. The expectation operator $\mathbf{E}^{\zeta_k}[\cdot]$ is defined for any functional f of the random jump size $Y_{k,t}$; thus, $\mathbf{E}^{\zeta_k}[f(Y_{k,t})] = \mathbf{E}[f(Y_{k,t})\vartheta_k(t)\varphi_k(t,Y_{k,t})]$. We later specify the model parameter θ_t measuring the preference for robustness.

The minimization in (8) clarifies the investor's pessimistic reaction to uncertain jumps. He or she makes decisions by first minimizing the expected utility among all alternative models, every one of which is statistically indistinguishable from the reference model and thus is possibly the true model. One distinctive feature of our setup is that we investigate model misspecifications in the entire neighborhood of the reference model, whereas Liu et al. (2005) and Drechsler (2013) consider only a subset of the neighborhood. In particular, these authors use a parametric approach to choose the worst jump size and jump intensity, whereas 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, whereas the worst-case jump size distribution does not necessarily fall in the same distribution family of the reference model in our study. Our setup is distinguished from those of Jin and Zhang (2012) and Branger and Larsen (2013). The jump size distributions in the latter two papers are assumed to be known, and are restricted to being identical across the reference and alternative models.

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, the investor 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 θ_t dictates the investor's attitude toward alternative models. We assume θ_t takes the following form (Maenhout 2004, 2006):

$$\theta_t = \frac{\phi}{\big(1-\gamma\big)V(W,t)}, \tag{9}$$

where ϕ will be referred to as the ambiguity aversion coefficient, with a larger ϕ indicating a lesser penalty and a higher level of ambiguity aversion.

At the extreme, when $\phi\to\infty$, the investor makes decisions based on the worst case among all possibilities, treating the reference model as equal to any alternative model; when $\phi\to 0$, the investor considers only the reference model without any fear of jump ambiguity. In our calibration exercises, the levels of ambiguity aversion ϕ (or the levels of penalties) are determined by the DEP approach of Anderson et al. (2003). Models with different jump distributions are compared under the same DEP, implying the same level of jump ambiguity. For different jump size distributions, ϕ may differ under the same DEP, although we cannot separate ambiguity from ambiguity aversion, both better understood as a unified preference for robustness,[11] under a given jump size distribution. More detailed discussions and examples are provided in Section 5.2.

An alternative specification of ambiguity aversion called locally constrained robustness is also used in the literature (Trojani and Vanini 2002, Aït-Sahalia and Matthys 2019).[12] Specifically, for a CRRA investor with $\gamma\neq 1$,[13] we solve a constrained version of the Hamilton-Jacobi-Bellman equation for the investor's value function $V(W,t)$ as follows:

$$\begin{aligned}0 = \max_{\pi}\inf_{P(\zeta)\in\mathcal{P}}\bigg\{&\frac{\partial V(W,t)}{\partial t} + W\big(\pi'\big(\mu - r\mathbf{1}_m\big)+r\big)\\ &\times\frac{\partial V(W,t)}{\partial W} + \frac{1}{2}W^2\pi'\Sigma\Sigma'\pi\frac{\partial^2 V(W,t)}{\partial W^2}\\ &+\sum_{k=1}^{n}\lambda_k\mathbf{E}^{\zeta_k}\big[V\big(W + W\pi'\mathbf{J}_k Y_{k,t},t\big)\\ &- V(W,t)\big]\bigg\},\end{aligned}$$

subject to $H(\zeta_t)\le\eta$, for a fixed $\eta>0$ and $V(W,T) = \frac{W^{1-\gamma}}{1-\gamma}$.

The locally constrained approach enjoys several advantages over the multiplier approach. For instance, the locally constrained approach admits a recursive multiple prior utility representation, and unlike the multiplier specification, the set of alternative beliefs under ambiguity is prespecified by the constraint $H(\zeta_t) \leq \eta$. For more detailed discussions, see Trojani and Vanini (2002) and Aït-Sahalia and Matthys (2019). In particular, Trojani and Vanini (2004) analyze different functional form solutions and various specifications of robustness. Interestingly, because there are no state variables in our model, it turns out that, for such a case, the locally constrained approach results in the same solution as that from the multiplier approach when we consider interior solutions. We provide mathematical derivations of the equivalency in Section OA.1 of the online appendix.

We are now ready to solve the portfolio choice problems with our multiplier specification defined earlier in this section. For notational convenience, we will suppress the dependence of $Y_{k,t}$, $\vartheta_k(t)$, $\varphi_k(t, y)$, $\Phi_k(t, dy)$, and π_t on t hereafter.

3. Optimal Portfolio: 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 π. The results are summarized in Proposition 1. Second, we propose a decomposition approach to determine the optimal portfolio weights under ambiguity aversion. The results are summarized in Proposition 2.

Proposition 1. *For any admissible portfolio π, the solution to the inner minimization of (8) for the multiplier specification of ambiguity aversion is given by*

$$\varphi_k^*(y_k) = \varphi_k^*(y_k, \pi) = \frac{1}{\vartheta_k^*} \exp\left(\frac{\phi}{\gamma - 1}\left(\left(1 + \pi'\mathbf{J}_k y_k\right)^{1-\gamma} - 1\right)\right), \tag{10}$$

$$\vartheta_k^* = \vartheta_k^*(\pi) = \mathbf{E}\left[\exp\left(\frac{\phi}{\gamma - 1}\left((1 + \pi'\mathbf{J}_k Y_k)^{1-\gamma} - 1\right)\right)\right], \tag{11}$$

for $k = 1, \ldots, n$.

Suppose the optimal portfolio π^* has been obtained; then, Proposition 1 provides the worst probability by φ_k^* and ϑ_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 $(\tilde{\pi}^*)' = (\pi^*)'\mathbf{J} = (\tilde{\pi}_1^*, \ldots, \tilde{\pi}_n^*)$. We refer to $\tilde{\pi}_k^*$ as the exposure to the kth jump risk. Then, the jump intensity in the worst probability is given by

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

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 when 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 ϕ indeed reinforces this tendency. 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 kth jump size of the worst case is given by

$$\Phi_k^*(dy_k) = \varphi_k^*(y_k)\Phi_k(dy_k), \tag{13}$$

where φ_k^* can be regarded as a weighting function. For the typical case of $\tilde{\pi}_k^* > 0$, we can verify that φ_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. It is worth mentioning that the worst-case density $\Phi_k^*(dy_k)$ is not in the same parametric family of the reference density $\Phi_k(dy_k)$ in general, given the expression (10) for $\varphi_k^*(y_k)$.

Finally, note that for the expectation term in (8), ϑ_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 toward the negative side of the jump size distribution and further impacts the optimal portfolio choice, as we will 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

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

Note that $\hat{J}$ is an $m \times n$ matrix with rank n. We treat each column of $\hat{J}$ as a vector of $\mathcal{R}^{m\times 1}$. For $m > n$, we can find unit vectors $\alpha_1, \ldots, \alpha_{m-n} \in \mathcal{R}^{m\times 1}$ such that $\alpha_k, k = 1, \ldots, m-n$ is orthogonal to each column of $\hat{J}$. Denote the matrix with columns $\alpha_1, \ldots, \alpha_{m-n}$ 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

$$\hat{\mu} = \bar{\mu} + \mu_\perp, \tag{15}$$

where $\bar{\mu}$ is an m-vector in the space generated by the columns of $\hat{J}$ and can be expressed as $\bar{\mu} = \hat{J}\bar{\mu}^0$ with $\mu^0 \in R^n$; and $\mu_\perp$ is an m-vector in the space generated by the columns of $\hat{J}_\perp$ and can be represented by $\mu_\perp = \hat{J}_\perp \mu_\perp^0$ with $\mu_\perp^0 \in \mathcal{R}^{(m-n)\times 1}$. 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 $\hat{J}'\hat{J}_\perp = 0_{n\times(m-n)}$, we can find $\bar{\mu} = \hat{J}(\hat{J}'\hat{J})^{-1}\hat{J}'\hat{\mu}$. Similarly, we can find $\mu_\perp = \hat{J}_\perp(\hat{J}_\perp'\hat{J}_\perp)^{-1}\hat{J}_\perp'\hat{\mu}$.

Let $\hat{\pi} = \Sigma'\pi$. The decomposition of $\hat{\pi}$ onto $\hat{J}$ and $\hat{J}_\perp$ is given by

$$\hat{\pi} = \bar{\pi} + \pi_\perp.$$

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

$$\pi_\perp^* = \frac{1}{\gamma}\mu_\perp, \tag{16}$$

$$\begin{aligned}\bar{\pi}^* = \arg\max_{\bar{\pi}} &-\frac{\gamma}{2}\bar{\pi}'\bar{\pi} + \bar{\mu}'\bar{\pi}\\ &+ \frac{1}{1-\gamma}\sum_{k=1}^{n}\lambda_k\left(\mathbf{E}^{\zeta_k^*}\left[\left(1+\bar{\pi}'\hat{J}_kY_k\right)^{1-\gamma} - 1\right]\right) + \frac{1}{\phi}H(\zeta_k^*),\end{aligned} \tag{17}$$

where $\zeta_k^* = \zeta_k(\vartheta_k^*, \varphi_k^*)$; ϑ_k^* and φ_k^* are given by (10) and (11), respectively.

Because $\pi'\mathbf{J}_k = \bar{\pi}'\hat{J}_k$ from the proof of Proposition 2 in Appendix A, ϑ_k^* and φ_k^* are both functions of $\bar{\pi}$ only by Proposition 1. Hence, we can solve (17) for $\bar{\pi}^*$ independent of $\bar{\pi}_\perp^*$. To simplify the analysis, unless otherwise stated, in this section we assume that the maximization in (17) is achieved at an interior point.

The first-order condition of (17), along with the first-order condition for the optimality of ζ_k^* given in Appendix A, yields the following equation for $\bar{\pi}^*$:

$$0 = -\gamma\bar{\pi} + \bar{\mu} + \sum_{k=1}^{n}\lambda_k\mathbf{E}\left[\left(1+\bar{\pi}'\hat{J}_kY_k\right)^{-\gamma}Y_k\vartheta_k^*\varphi_k^*(Y_k)\right]\hat{J}_k. \tag{18}$$

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

$$-\gamma\tilde{\pi}_k + \hat{J}_k'\bar{\mu} + \sum_{l=1}^{n}\lambda_l\mathbf{E}^{\zeta_l^*}\left[(1+\tilde{\pi}_lY_l)^{-\gamma}Y_l\right]\left(\hat{J}_k'\hat{J}_l\right) = 0, \quad k = 1, \ldots, n. \tag{19}$$

It is straightforward to show the existence and uniqueness of the solution to the aforementioned algebraic equation system (e.g., by the fixed-point theorem or the contraction mapping principle). Furthermore, because $\bar{\pi}$ is the orthogonal projection of $\hat{\pi} = \Sigma'\pi$ in the subspace $\hat{J}$ of $\mathcal{R}^{m\times 1}$ (see (A.7) in Appendix A), due to the uniqueness of the orthogonal decomposition, we have $\bar{\pi} = \hat{J}(\hat{J}'\hat{J})^{-1}\hat{J}'\hat{\pi} = \hat{J}(\hat{J}'\hat{J})^{-1}\hat{J}'\bar{\pi} = \hat{J}(\hat{J}'\hat{J})^{-1}\tilde{\pi}$. The existence and uniqueness of the solution to the equation system (18) are therefore also guaranteed.

One advantage of our decomposition approach is that it may significantly reduce the computational complexity when there is a large number of risky assets and a relative small number of jump types (e.g., the case of six international indices with one systemic jump studied by Das and Uppal 2004, and the case of four risky assets with two types of jumps studied by Jin and Zhang 2012). Solving the optimal portfolio directly from the first-order condition of the HJB Equation (8) with respect to π involves m (the number of risky assets) nonlinear equations, which bring in more computational burden than the n (the number of jump types) nonlinear equations of $\tilde{\pi}$ obtained by our decomposition approach, given that n is much smaller than m. More specifically, in the model by Das and Uppal (2004), they need to solve m nonlinear equations that are computationally challenging when m is large. In stark contrast, we only need to solve one nonlinear equation regardless of m. In sum, our approach is useful for studying multiasset jump-diffusion models, especially in the case of a large number of risky assets.

Our decomposition method is based on orthogonal vector decompositions and generalizes that of Aït-Sahalia et al. (2009). These authors rely on specific assumptions of model coefficients (i.e., identical jump size scales and a block-structure variance-covariance to make an orthogonal decomposition). We extend their approach by removing such restrictions on model parameters. We provide a relatively easy decomposition method that has broad applications. For example, Jin and

Zhang (2012) and Jin et al. (2018) consider a multiasset jump-diffusion model with n risky assets, d Brownian motions, and $n-d$ types of jumps. Our model differs from theirs in that we do not require the number of risky assets to equal the total number of risk sources. However, our method of orthogonal vector decompositions can be applied to their model and obtain the same results. Undoubtedly, our approach can also be applied to the model by Aït-Sahalia et al. (2009) and obtain identical results.[14]

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 $\vartheta_k^* = \varphi_k^* = 1$ in (18). This corresponds to the case $\phi \to 0$ such that any deviation from the reference model is penalized infinitely.

The first example focuses on $\mathbf{J} = 0$; 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

$$\begin{aligned}\pi^* &= (\Sigma')^{-1}\pi_\perp^* = (\Sigma')^{-1}\frac{1}{\gamma}\hat{J}_\perp\left(\hat{J}_\perp'\hat{J}_\perp\right)^{-1}J_\perp'\Sigma^{-1}(\mu - r\mathbf{1}_m)\\ &= \frac{1}{\gamma}(\Sigma\Sigma')^{-1}(\mu - r\mathbf{1}_m),\end{aligned} \tag{20}$$

where $\mathbf{1}_m$ denotes the m-dimensional column vector of ones. This is the familiar optimal portfolio in a multiasset 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 by Liu et al. (2003).

Taken together, the two aforementioned examples suggest that $\bar{\pi}^*$ can be regarded as closely related to the exposure to jump risk, whereas $\pi_\perp^*$ is closely related to the exposure to diffusion risk. Furthermore, by (14) and the decomposition of π^* in Proposition 2, the relation between $\bar{\pi}^*$ and the kth jump exposure $\tilde{\pi}_k^*$ can be clearly seen from the following result: $\tilde{\pi}_k^* = \mathbf{J}_k'\pi^* = \hat{J}_k'\bar{\pi}^*$. 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.

4. Tail Risk, Diversification, and Nonparticipation

This section studies the effects of ambiguity aversion on the jump tail behavior and the optimal portfolio. We start by showing 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 robust portfolio selection.

Proposition 3. *There exits $y^* < 0$ 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 ϕ. Moreover, for any $\phi_1 > \phi_2$, there exists $\hat{y}$ such that for any $y < \hat{y}$, $\Pr^*(Y_k < y|\phi_1) > \Pr^*(Y_k < y|\phi_2)$.*

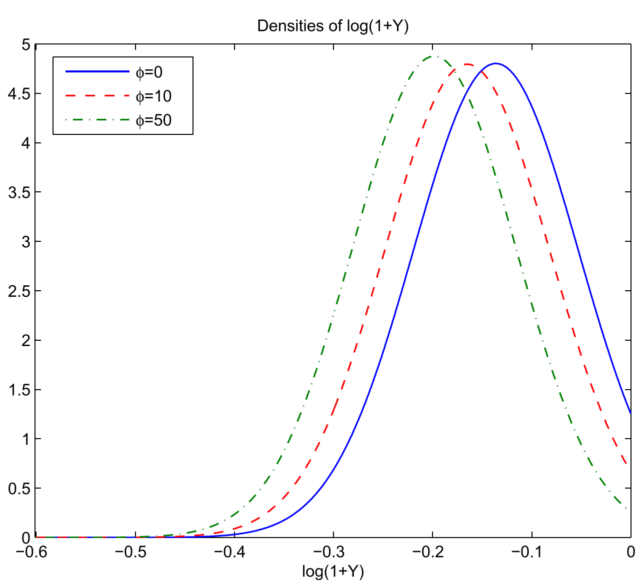
This proposition affirms that there is a larger left tail in the worst-case probability than in the reference probability, and the tail is larger if an investor is more ambiguity averse. Thus, tail risk in the worst-case scenario is greater for an investor with a higher level of ambiguity aversion. Figure 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 ϕ becomes larger.

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

$$\tilde{\pi}^* = \frac{1}{\gamma}\hat{J}'\bar{\mu} + \frac{1}{\gamma}\sum_{k=1}^{n}\lambda_k \mathbf{E}^{\zeta_k^*}\left[\left(1 + \tilde{\pi}_k^* Y_k\right)^{-\gamma} Y_k\right]\left(\hat{J}'\hat{J}_k\right). \tag{21}$$

We see that only the second term of (21) involves the jump exposure. The first term is independent of the ambiguity aversion coefficient ϕ. Thus, we can study

Figure 1. (Color online) The Reference Density and the Worst-Case Densities at Different Ambiguity Aversion Levels



Notes. 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.

the sensitivity of the jump exposure to the investor's attitude toward model uncertainty by investigating the derivative of $\tilde{\pi}^*$ with respect to ϕ.

To gain clear intuition, we focus on the simplest case of $n = 1$—that is, only one type of jump is admitted in the model. This allows us to omit the subscript k in the rest of this section unless otherwise indicated. We have the following results (with proofs in Appendix A):

$$\text{i. } \frac{d\tilde{\pi}^*}{d\phi} < 0, \tag{22}$$

$$\text{ii. } \frac{d\tilde{\pi}^*}{d\left(\hat{J}'\hat{J}\right)} > 0 \text{ if and only if } \frac{d\left(\hat{J}'\bar{\mu}\right)}{d\left(\hat{J}'\hat{J}\right)} + \lambda \mathbf{E}^{\zeta^*}\left[\left(1+\tilde{\pi}^*Y\right)^{-\gamma}Y\right] > 0. \tag{23}$$

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 the jump exposure accordingly.

To illustrate result (ii) and the sufficient and necessary condition in (23), further simplifications seem warranted. Here, we consider the simplest case of identical risky assets with no correlation beyond the systemic jump. Specifically, we assume Σ is a diagonal matrix with identical diagonal entries of σ_1, and μ has identical entries of μ_1. We further assume $\mathbf{J} = \mathbf{1}_m$. Then, $\hat{J} = \sigma_1^{-1}\mathbf{1}_m$, $\hat{J}'\hat{J} = \sigma_1^{-2}m$, and $\hat{J}'\bar{\mu} = \hat{J}'\hat{\mu} = (\mu_1 - r)\sigma_1^{-2}m$ by (14). The condition in (23) becomes

$$\mu_1 - r + \lambda \mathbf{E}^{\zeta^*}\left[\left(1+\tilde{\pi}^*Y\right)^{-\gamma}Y\right] > 0. \tag{24}$$

Two observations follow immediately from (24). On the one hand, we conclude that (24) holds by (21) because $\tilde{\pi}^* \in (0,1)$ and $\tilde{\pi}^* = m(\mu_1 - r + \lambda \mathbf{E}^{\zeta^*}[(1+\tilde{\pi}^*Y)^{-\gamma}Y])/\gamma\sigma_1^2$. Therefore, in this case, the jump exposure $\tilde{\pi}^*$ increases with $\sigma_1^{-2}m$. Holding constant the variance σ_1^2 of each risky asset, we find that the jump exposure increases with m, the total number of risky assets. Note that $\tilde{\pi}^* = \mathbf{J}'\pi^* = \sum_{i=1}^m \pi_i^*$; hence, the total risky investment also increases with the number of risky assets.

On the other hand, it follows from $\bar{\pi} = \hat{J}(\hat{J}'\hat{J})^{-1}\tilde{\pi}$ that

$$\bar{\pi}^* = \tilde{\pi}^*\sigma_1/m\mathbf{1}_m = \frac{1}{\gamma\sigma_1}\left((\mu_1 - r) + \lambda \mathbf{E}^{\zeta^*}\left[\left(1+\tilde{\pi}^*Y\right)^{-\gamma}Y\right]\right)\mathbf{1}_m.$$

Each component of $\bar{\pi}^*$ is decreasing in m as one can check that[15]

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

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 systemic jump risk.

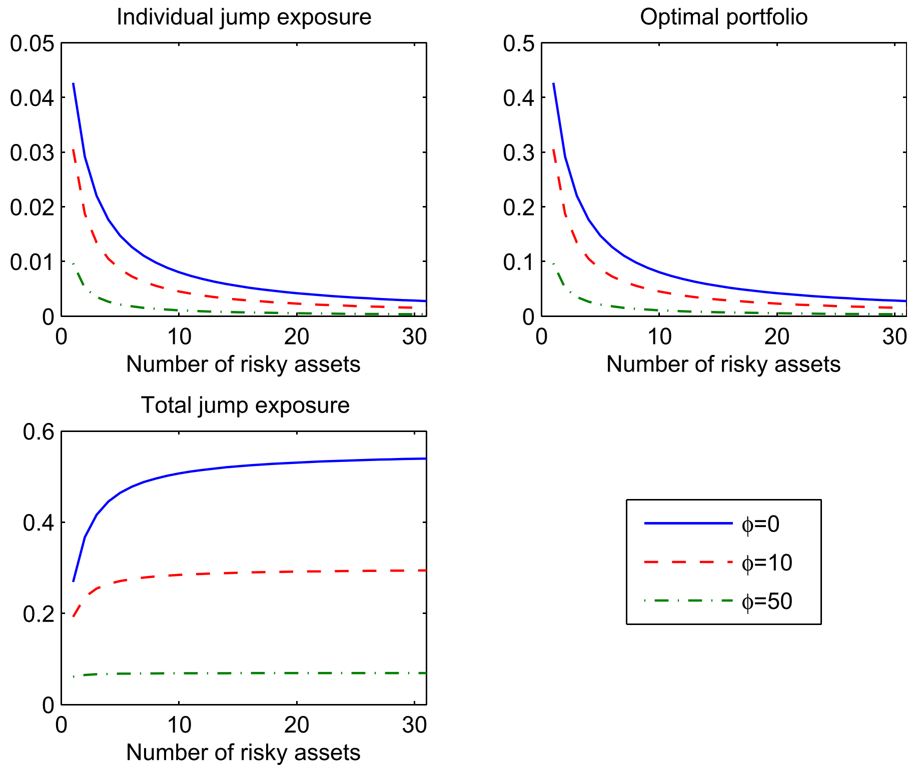
We plot the effect of jump ambiguity on portfolio diversification in Figure 2. Under the aforementioned simplifying assumptions for asset returns, the upper-left panel shows that the individual jump exposure $\bar{\pi}_i$ decreases toward zero as the number of risky assets increases. This is to be expected because 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 zero when the number of assets 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 independent identically distributed (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 positions that are even less risky. When accessing a larger number of risky assets, this investor's positions in each individual asset draw more closely to zero. Thus, ambiguity aversion further suppresses portfolio diversification when the investor lacks confidence in the systemic jump 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 φ_k^* in (13) is increasing (decreasing) in the jump exposure $\tilde{\pi}_k^*$.[16] 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. Consequently, expanding the investment opportunity set with more risky assets, for the benefit of diversifying diffusive risk, could result in higher total jump exposure and, meanwhile, marginally bring down each individual jump exposure.

By Proposition 3 and the aforementioned result i, when an investor is extremely ambiguity averse with $\phi \to \infty$, the worst-case density will have the largest tail and the jump exposure will be minimal. In this case, the investor disregards the reference model and treats it equal to all alternative models. This result is closely related to the *nonparticipation puzzle*: modern asset allocation theories suggest that investors allocate certain nonzero 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 80th percentile of

Figure 2. (Color online) Ambiguity Aversion and Portfolio Diversification



Notes. We show the individual and total jump exposures and the optimal portfolio weights of the investor when investing into more risky assets. The model parameters are taken from Table 2, and we use the average values for $\mathbf{J}$, σ_n, and μ. We consider three levels of ambiguity aversion in the plotting.

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 induce nonparticipation.

We can show that in our model, the total jump exposure for the kth jump, $\tilde{\pi}_k^* = (\pi^*)'\mathbf{J}_k$, for the mixed jumps Y_k with support $(-1, \infty)$, must be zero as $\phi \to \infty$.[17] Intuitively, an investor with extreme ambiguity aversion makes decisions as if jumps against his or her positions would occur at an infinite frequency and cause infinite losses. The investor ensures that his or her optimal portfolio in total has no exposure to the systemic jump risk, eliminating jump ambiguity at the same time ($\vartheta_k^* = \varphi_k^*(y_k) = 1$ when $\tilde{\pi}_k^* = 0$). Recall that we have assumed $\mathbf{J}$, without loss of generality, to have full column rank. Then, for the case in which $m = n$, we have $\pi_i^* = 0$, $i = 1, \ldots, m$, because $(\pi^*)'\mathbf{J} = 0$. Nonparticipation occurs. For the case in which $m > n$, if the investor is further not allowed to short sell stocks—that is, $\pi_i \geq 0$—then $\pi_i^* = 0, i = 1, \ldots, m$, because $(\pi^*)'\mathbf{J} = 0$. Nonparticipation again occurs. When short selling is prohibited, diversification of jump uncertainty fails to work because long positions are necessarily accompanied by short positions to neutralize the systemic jumps. Without the short selling constraint, the investor may well (selectively) participate the market to harvest diffusive risk premiums.[18]

In the context of ambiguity, Uppal and Wang (2003) show that an ambiguous return distribution results in an underdiversified portfolio relative to the standard mean-variance portfolio. In static mean-variance frameworks, Boyle et al. (2012) and Liu and Zeng (2017) show that expected return ambiguity and return correlation ambiguity may cause nonparticipation. Cao et al. (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. In our paper, exploiting the tractability of our nonparametric approach, we explicitly show that the reduction of diversification is attributed to suppressed jump exposure under jump ambiguity. Intuitively, extreme ambiguity aversion may well induce nonparticipation of the market. However, to the best of our knowledge, little research explicitly demonstrates this result, especially in a dynamic setting. Our paper indicates that nonparticipation may occur at the presence of aversion to jump ambiguity in the dynamic model and thus enriches the literature by providing a new channel to explain this phenomenon. Interestingly, we also show in the last paragraph that even an extremely ambiguity-averse investor may (selectively) participate in the market, because he or she can cross hedge the systemic jump risk.

We have shown that ambiguity aversion enlarges the worst-case jump tail risk and in turn impacts portfolio diversification. The previous analyses generally apply to any jump distribution. The following proposition directly examines the importance of tail assumptions in the reference models. To be precise, we define one jump size Y^p to have 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 4. *Suppose one jump size Y^p has a larger left tail than another jump size Y^n. 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 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 have 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 et al. (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.[19] Because Merton's normal jump size model has been 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?

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 earlier, we first utilize a jump-diffusion model with a normally distributed jump size as summarized in the following equation:

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

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

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, the United Kingdom, Germany (GE), France (FR), Canada (CA), Sweden (SD), and Japan (JP). We collect the beginning-of-month equity index levels from finance.yahoo.com.[20] 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. The skewness is highly statistically significant for all seven indices, and the kurtosis is also highly statistically significant for all indices except JP. 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

Table 1. Summary Statistics for Equity Returns

	U.S.	U.K.	GE	FR	CA	SD	JP
Panel A: Central moments							
Mean	0.0056	0.0029	0.0070	0.0052	0.0050	0.0090	0.0035
Standard deviation	0.0426	0.0401	0.0626	0.0455	0.0433	0.0528	0.0549
Skew	−0.8552	−0.6984	−0.8666	−0.8549	−1.2256	−0.6742	−0.5143
Significance level	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005
ExKurt	1.7615	0.7352	2.6257	2.1204	4.0930	1.7548	0.4362
Significance level	0.0000	0.0120	0.0000	0.0000	0.0000	0.0000	0.1362
Panel B: Correlation							
U.S.							
U.K.	0.8141						
GE	0.7634	0.7799					
FR	0.6968	0.7270	0.7149				
CA	0.7747	0.6940	0.6346	0.5674			
SD	0.6254	0.6555	0.7031	0.6278	0.6020		
JP	0.7542	0.8210	0.8780	0.7471	0.6465	0.6959	

Notes. 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.

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 $i, j = 1, 2, \ldots, m$ $(m = 7)$,

$$\text{mean} = t\left(\mu_i - 0.5\sum_{k=1}^{m}\sigma_{ik}^2 + \lambda \mathrm{E}\left[\bar{Y}_i\right]\right),$$

$$\text{covariance} = t\left(\sum_{k=1}^{m}\sigma_{ik}\sigma_{jk} + \lambda \mathrm{E}\left[\bar{Y}_i\bar{Y}_j\right]\right),$$

$$\text{coskewness} = \frac{t\lambda \mathrm{E}\left[\left(\bar{Y}_i\right)^2\bar{Y}_j\right]}{\text{variance}_i\left(\text{variance}_j\right)^{0.5}},$$

$$\text{excess kurtosis} = \frac{t\lambda \mathrm{E}\left[\left(\bar{Y}_i\right)^4\right]}{\left(\text{variance}_i\right)^2}, \tag{26}$$

where $\mathbf{E}[(\bar{Y}_i)^u(\bar{Y}_j)^v] = \int_{-\infty}^{+\infty}(\bar{Y}_i)^u(\bar{Y}_j)^v f(\varepsilon)d\varepsilon$ with $u=1,2,\ldots;$ $v=0,1,\ldots;$ and $f(\cdot)$ is the standard normal density. This integral may be easily evaluated using the numeric quadrature method. We first use the 7×7 coskewness conditions and 7×1 kurtosis conditions to estimate the 10 jump parameters (J_i, μ_J, σ_J, λ) by minimizing the sum of squared deviations of the model moments from those in the data. We then derive μ_i and σ_{ij} by exactly matching the 7×1 mean conditions and the 7×7 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 a year (Eraker et al. 2003, Das and Uppal 2004).

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 from earlier. 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.[21] 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 η follows a single power law distribution with its density as follows:

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

η_0 and α 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, we call the model with normal jump size distribution the jump size model.

Table 2. Parameter Estimates

	U.S.	U.K.	GE	FR	CA	SD	JP
				Panel A: Jump parameters			
μ_J	−0.1363						
σ_J	0.0831						
λ	0.0751						
J	0.5381	0.4401	0.8208	0.5856	0.6503	0.6464	0.5447
				Panel B: Other parameters			
σ_{n1}	0.0362	0.0271	0.0344	0.0221	0.0225	0.0220	0.0346
σ_{n2}	0	0.0232	0.0165	0.0121	0.0042	0.0128	0.0195
σ_{n3}	0	0	0.0349	0.0061	−0.0025	0.0114	0.0201
σ_{n4}	0	0	0	0.0282	−0.0030	0.0047	0.0066
σ_{n5}	0	0	0	0	0.0242	0.0052	0.0051
σ_{n6}	0	0	0	0	0	0.0350	0.0029
σ_{n7}	0	0	0	0	0	0	0.0208
μ	0.0115	0.0078	0.0166	0.0117	0.0120	0.0164	0.0101

Notes. 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 the model from those in the data. All parameter estimates and moments are on a monthly basis. The sample period is January 1993 to December 2015.

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 midsized jumps and more weight on large-sized jumps than the normal. This difference is even 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, in Table 3 we report 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.[22] 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

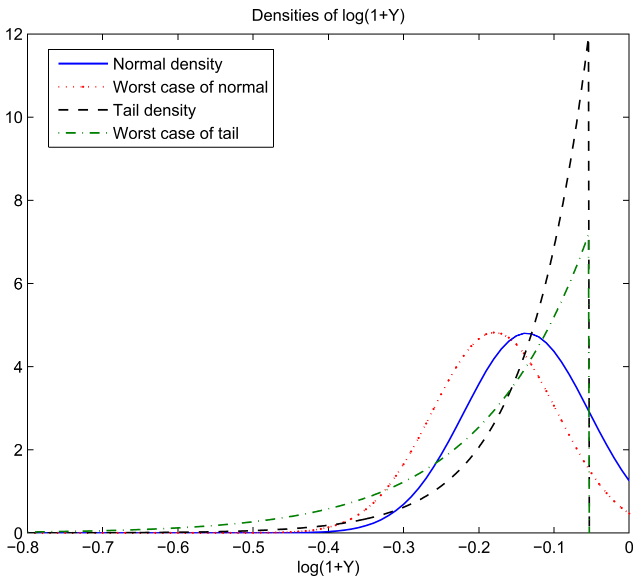
Because the two jump size distributions pursued here display distinct tail behaviors, an investor may hold different degrees of ambiguity aversion toward the two models. To make the results more comparable, we identify the appropriate ambiguity aversion coefficients ϕ based on the same DEP (derived in 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 ϕ) implies lower DEP because it becomes increasingly easier to distinguish the worst-case model from the reference model. Following the literature (Anderson et al. 2003; Maenhout 2004, 2006; Drechsler 2013), we require a

Table 3. Moment Comparison

	U.S.	U.K.	GE	FR	CA	SD	JP
Skew: data	−0.8552	−0.6984	−0.8666	−0.8549	−1.2256	−0.6742	−0.5143
Skew: normal	−0.6873	−0.4366	−0.8493	−0.7377	−1.1945	−0.6473	−0.3337
Skew: tail	−0.7797	−0.4901	−1.0006	−0.8417	−1.3739	−0.7441	−0.3789
ExKurt: data	1.7615	0.7352	2.6257	2.1204	4.093	1.7548	0.4362
ExKurt: normal	1.8947	1.0314	2.5389	2.0858	3.9752	1.7558	0.7234
ExKurt: tail	2.8153	1.4975	4.0933	3.1367	6.0829	2.6839	1.0766

Notes. This table reports return moments reconstructed using the theoretic moment conditions in Equation (26) and the model parameter estimates in Table 2 (labeled "normal"). For comparison, we also report the empirical moments computed from the return data (labeled "data") and the moments computed under the calibrated tail distribution following a power law (labeled "tail"). ExKurt, excess kurtosis.

Figure 3. (Color online) The Reference Densities and the Worst-Case Densities Under Different Jump Distributions



Notes. We plot the reference densities as the normal and tail distributions that 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.

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 ϕ values are also listed in the table. The annual interest rate is set at 5%, and the risk aversion coefficient γ is 5, 3, or 1. The last case corresponds to the log utility function. See Appendix C for relevant results. 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 γ 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 γ, the ambiguity aversion coefficient ϕ 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 7.3 when $\gamma = 1$. This occurs because the worst density depends on γ, and the jump exposure $\tilde{\pi}$ decreases with γ. Thus, a less risk-averse investor is less fearful of jump risk. Apparently, from (10) in Proposition 1, the worst density φ_k^* increases with the jump exposure $\tilde{\pi} = \pi'\mathbf{J}_k$ 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 et al. (2007) and Cvitanić et al. (2008): These scholars note that the CRRA utility function represents a local mean-variance preference that does not capture higher moments.

To investigate the impact of higher moments, Hong et al. (2007) use the disappointment aversion (DA) preference of Ang et al. (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 et al. (2005) to investigate an investor's concern about extreme downside risk.[23] 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

Table 4. Optimal Portfolios at Different DEPs

	DEP = 10%		DEP = 15%		DEP = 20%		No ambiguity	
	Normal	Tail	Normal	Tail	Normal	Tail	Normal	Tail
Panel A: $\gamma = 5$								
ϕ	47.6	60.9	29.6	33.6	20.2	22.3	0	0
U.S.	0.5617	0.5705	0.5554	0.5629	0.5503	0.5571	0.5242	0.5245
U.K.	−1.0955	−1.0815	−1.1054	−1.0935	−1.1135	−1.1027	−1.1547	−1.1544
GE	0.4970	0.4568	0.5254	0.4913	0.5486	0.5177	0.6673	0.6662
FR	0.0053	−0.0274	0.0284	0.0006	0.0473	0.0221	0.1440	0.1431
CA	−0.2977	−0.3496	−0.2609	−0.3050	−0.2309	−0.2709	−0.0774	−0.0788
SD	0.6443	0.6342	0.6515	0.6429	0.6573	0.6495	0.6872	0.6870
JP	−0.5378	−0.4886	−0.5728	−0.5309	−0.6012	−0.5632	−0.7466	−0.7453
$\tilde{\pi}$	0.1610	0.1063	0.1998	0.1533	0.2314	0.1892	0.3930	0.3915
Panel B: $\gamma = 3$								
ϕ	28.6	36.8	18.1	20.4	12.3	13.6	0	0
U.S.	0.9364	0.9511	0.9264	0.9387	0.9179	0.9293	0.8755	0.8763
U.K.	−1.8255	−1.8022	−1.8413	−1.8218	−1.8546	−1.8367	−1.9217	−1.9204
GE	0.8274	0.7605	0.8726	0.8167	0.9111	0.8596	1.1040	1.1001
FR	0.0080	−0.0465	0.0449	−0.0007	0.0763	0.0343	0.2334	0.2302
CA	−0.4973	−0.5839	−0.4388	−0.5111	−0.3890	−0.4556	−0.1395	−0.1445
SD	1.0736	1.0567	1.0850	1.0709	1.0947	1.0817	1.1434	1.1424
JP	−0.8953	−0.8133	−0.9508	−0.8822	−0.9980	−0.9349	−1.2345	−1.2297
$\tilde{\pi}$	0.2671	0.1759	0.3288	0.2526	0.3812	0.3111	0.6440	0.6387
Panel C: $\gamma = 1$ (log utility)								
ϕ	9.9	13.2	6.3	7.3	4.7	4.7	0	0
U.S.	2.8145	2.8588	2.7877	2.8246	2.7770	2.7978	2.7770	2.7770
U.K.	−5.4680	−5.3979	−5.5103	−5.4520	−5.5272	−5.4943	−5.5272	−5.5272
GE	2.4576	2.2562	2.5793	2.4116	2.6279	2.5334	2.6279	2.6279
FR	0.0041	−0.1600	0.1032	−0.0334	0.1428	0.0658	0.1428	0.1428
CA	−1.5237	−1.7843	−1.3663	−1.5832	−1.3034	−1.4257	−1.3034	−1.3034
SD	3.2147	3.1639	3.2453	3.2031	3.2576	3.2337	3.2576	3.2576
JP	−2.6558	−2.4088	−2.8050	−2.5994	−2.8646	−2.7487	−2.8646	−2.8646
$\tilde{\pi}$	0.7679	0.4935	0.9337	0.7053	0.9999	0.8711	0.9999	0.9999

Notes. This table reports optimal portfolio weights at three DEP levels. The ambiguity aversion coefficients ϕ 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.

beyond the scope of the present paper and left for future research.

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:

$$V^{\pi^{(2)}}(W,t) = V(W(1-\mathrm{CEL}),t), \tag{28}$$

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

$$V^{\pi^{(2)}}(W,t) = \inf_{\zeta} \mathbf{E}_t^{\zeta}\left[e^{\int_t^T \frac{1-\gamma}{\phi} H(\zeta_s)ds} \frac{W_T^{1-\gamma}}{1-\gamma} \right]. \tag{29}$$

Unlike the calculation of $V(W,t)$, the worst-case Radon-Nikodym derivative ζ 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 30.2% at $\gamma = 1$ at a 20-year

Table 5. CELs: Tail (True) vs. Normal

	Investment horizon (in years)		
DEP	1	10	20
	Panel A: $\gamma = 5$		
10%	0.0028	0.0279	0.0550
15%	0.0011	0.0118	0.0236
20%	0.0007	0.0071	0.0143
	Panel B: $\gamma = 3$		
10%	0.0048	0.0474	0.0925
15%	0.0020	0.0196	0.0388
20%	0.0012	0.0124	0.0247
	Panel C: $\gamma = 1$ (log utility)		
10%	0.0177	0.1643	0.3016
15%	0.0074	0.0716	0.1381
20%	0.0016	0.0161	0.0320

Notes. 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.

investment horizon. As the investor becomes less risk averse, the differences in jump exposures and stock holdings widen as the investor's risk positions enlarge under both tail distributions. As a result, the CELs from adopting the suboptimal strategies also rise. 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 ϕ to be identical for both distributions, we confirm that the CEL under ambiguity aversion is largely reduced. For example, when $\gamma = 1$, given $\phi = 9.9$ for both distributions, the CEL is 9.2% for a 20-year investment, much less than the CEL of 30.2% 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%, ϕ is 13.2 for the tail distribution but 9.9 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 that 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 the worst-case scenario occurs.

5.3. Further Results and Discussions

As another application of our model, we investigate optimal portfolio choice for the data set of five industry portfolios in the U.S. market.[24] We obtain the monthly data from Kenneth French's online data library for the sample period January 1993 to December 2017. Panel A of Table 6 reports summary statistics, and Panel B reports parameter estimates for the industry portfolio returns. We see that the skewness is highly statistically significant for all industries, and that the kurtosis is also highly statistically significant for all industries except the health sector. The jump size has a mean of −15.29% and a standard deviation of 9.47%. The jumps occur about once every 2.5 years on average. The corresponding power law has the lower bound $\eta_0 = 1.0600$ and exponent $\alpha = 10.5602$. Panels C and D report the investor's optimal portfolio weights and economic welfare. We see noticeable differences in jump exposures and portfolio holdings for this new set of assets. Under DEP = 10% and $\gamma = 3$, the investor incurs a CEL of 19.71% in 20 years.[25] These results confirm the importance of uncertain tail events for portfolio selection in different market settings.

Our analysis so far has been partial in nature, from the perspective of a small marginal investor (Merton 1971). The optimal holdings, which in some cases exceed 1, may not be easily implemented in practice (Black and Litterman 1992). To alleviate this concern and gauge the significance of tail risk from the perspective of a representative investor, we conduct an asset pricing study under ambiguous tail risk, where the equilibrium demand for the risky stock is unit.[26] Aversion to extreme tail events in this setup, which suppresses risk-bearing ability of the investor, translates into higher risk compensations to reach market equilibrium. We find that tail risk effectively enhances the smirk premium and variance risk premium for the ambiguity-averse investor, whereas the effect is largely attenuated for an investor without ambiguity aversion. We relegate the model details to Section OA.2 of the online appendix.

Table 6. Investment Strategies with Industry Portfolios

	Consumer	Manufacture	HiTech	Health	Other
Panel A: Summary statistics					
Mean	0.0087	0.0095	0.0103	0.0097	0.0086
Standard deviation	0.0371	0.0414	0.0614	0.0419	0.0502
Skew	−0.5724	−0.6493	−0.4787	−0.3762	−0.7488
Significance level	0.0000	0.0000	0.0007	0.0075	0.0000
ExKurt	1.6689	1.8220	1.5065	0.3050	2.6987
Significance level	0.0000	0.0000	0.0000	0.2769	0.0000
Panel B: Model estimates					
μ_J	−0.1529				
σ_J	0.0947				
λ	0.0347				
J	0.4967	0.5602	0.7606	0.4155	0.7298
Skew: model	−0.5192	−0.5521	−0.4565	−0.2054	−0.7301
ExKurt: model	1.6879	1.8364	1.4370	0.4886	2.6837

Panel C: Optimal portfolio weights (γ = 3)

	DEP = 10%		DEP = 15%		DEP = 20%		No ambiguity	
	Normal	Tail	Normal	Tail	Normal	Tail	Normal	Tail
ϕ	24.5	26.1	16.7	17.1	12.4	11.9	0	0
Consumer	2.5437	2.5337	2.5478	2.5393	2.5509	2.5441	2.5670	2.5667
Manufacture	0.5766	0.5272	0.5974	0.5547	0.6130	0.5790	0.6932	0.6917
HiTech	−0.0615	−0.0927	−0.0483	−0.0753	−0.0385	−0.0600	0.0122	0.0112
Health	0.6640	0.6560	0.6674	0.6605	0.6700	0.6644	0.6831	0.6828
Other	−1.9495	−2.0254	−1.9177	−1.9832	−1.8939	−1.9459	−1.7709	−1.7731
$\tilde{\pi}$	0.3929	0.2777	0.4412	0.3418	0.4773	0.3983	0.6640	0.6606

Panel D: CELs (γ = 3)

	DEP = 10%	DEP = 15%	DEP = 20%	No ambiguity
CELs (1 year)	0.0109	0.0052	0.0022	0.0000
CELs (10 years)	0.1040	0.0507	0.0216	0.0000
CELs (20 years)	0.1971	0.0988	0.0429	0.0000

Notes. This table reports summary statistics, parameter estimates, model moments, and portfolio decisions and outcomes for five industry portfolios in the U.S. market. The sample period is January 1993 to December 2017. ExKurt, excess kurtosis.

6. 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 multiasset incomplete market using a new decomposition approach. Both the optimal portfolio and the worst-case probability are obtained in closed forms. To quantify the effects of the thickness of jump tails, we calibrate the model to international equity indices and compare the optimal portfolios under both normally distributed jump sizes and heavy-tailed jump sizes obeying a power law.

We demonstrate 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 the market if he or 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 find that the economic losses from ignoring heavy jump tails are negligible in an expected utility model without jump ambiguity, as long as the jump size distributions are put on an equal footing in terms of the first two moments. In stark contrast, underestimating tail risk may result in sizeable wealth losses in the presence of jump ambiguity. Given the level of jump ambiguity (DEP) in our model, the investor exhibits higher ambiguity aversion toward jump distributions with heavier tails.

Our study emphasizes the importance of jump tail behavior in optimal portfolio formation. It also works as a springboard for further investigating the asset pricing implications of consumption disaster risk (Barro and Jin 2011). The current analysis can be extended to consider even broader asset classes and market settings (e.g., currency and currency options). The highly tractable approach proposed here provides a convenience for future research along this dimension.

Acknowledgments

The authors are grateful to the Department Editor Tyler Shumway, an associate editor, and two anonymous referees for their insightful and detailed comments that substantially improved the article. The authors also thank Andrew Carverhill; Jerome Detemple; Vadim Elenev; Hua He; Qingfu Liu; Raman Uppal; Harold Zhang; and seminar participants of the 2017 European Finance Association Annual Meeting, the 2017 China International Conference in Finance, the 2016 China Finance Review International Conference, the 2016 China Finance Annual Meeting, and the 2nd Shanghai Risk Forum for helpful comments.

Appendix A. Proofs of Propositions and Results

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:

$$\begin{aligned}
0 &= \max_{\pi}\left\{\frac{1}{1-\gamma}\frac{1}{h(t)}\frac{dh(t)}{dt} - \frac{\gamma}{2}\pi'\Sigma\Sigma'\pi + \left[\pi'\left(\mu - r\mathbf{1}_m\right) + r\right]\right.\\
&\quad \left. + \inf_{P(\zeta)\in\mathcal{P}}\frac{1}{1-\gamma}\sum_{k=1}^{n}\lambda_k \mathbf{E}^{\zeta_k}\left[(1+\pi'\mathbf{J}_k Y_k)^{1-\gamma} - 1\right] + \frac{1}{\phi}H(\zeta_t)\right\}\\
&= \max_{\pi}\left\{\frac{1}{1-\gamma}\frac{1}{h(t)}\frac{dh(t)}{dt} - \frac{\gamma}{2}\pi'\Sigma\Sigma'\pi + \left[\pi'\left(\mu - r\mathbf{1}_m\right) + r\right]\right.\\
&\quad + \inf_{\zeta}\frac{1}{1-\gamma}\sum_{k=1}^{n}\lambda_k \int_A \left((1+\pi'\mathbf{J}_k y_k)^{1-\gamma} - 1\right)\vartheta_k\varphi_k(y_k)\Phi_k(dy_k).\\
&\quad \left. + \frac{\lambda_k}{\phi}\int_A \left(\vartheta_k\varphi_k(y_k)\log(\vartheta_k\varphi_k(y_k)) + 1 - \vartheta_k\varphi_k(y_k)\right)\Phi_k(dy_k)\right\}.
\end{aligned} \tag{A.1}$$

Applying calculus of variations (Weinstock 1974),[27] we find the minimizer $\vartheta_k\varphi_k(\cdot)$ for the inner minimization problem by solving

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

Then,

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

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

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

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

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

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

We decompose $\hat{\pi}$ onto the space $\hat{J}$ and its orthogonal space $\hat{J}_\perp$:

$$\hat{\pi} = \bar{\pi} + \pi_\perp, \tag{A.7}$$

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

$$\begin{aligned}
0 &= \frac{1}{1-\gamma}\frac{1}{h(t)}\frac{dh_t}{dt} + r + \max_{\pi_\perp}\left(-\frac{\gamma}{2}\pi_\perp'\pi_\perp + \pi_\perp'\mu_\perp\right)\\
&\quad + \max_{\bar{\pi}} -\frac{\gamma}{2}\bar{\pi}'\bar{\pi} + \bar{\pi}'\bar{\mu} + \frac{1}{1-\gamma}\sum_{k=1}^{n}\lambda_k \mathbf{E}^{\zeta_k^*}\\
&\quad \times\left[\left(1+\bar{\pi}'\hat{J}_k Y_k\right)^{1-\gamma} - 1\right] + \frac{1}{\phi}H(\zeta_t^*).
\end{aligned} \tag{A.8}$$

Hence,

$$\pi_\perp^* = \frac{1}{\gamma}\mu_\perp, \tag{A.9}$$

and

$$\begin{aligned}
\bar{\pi}^* &= \arg\max_{\bar{\pi}} -\frac{\gamma}{2}\bar{\pi}'\bar{\pi} + \bar{\mu}'\bar{\pi} + \frac{1}{1-\gamma}\sum_{k=1}^{n}\lambda_k\\
&\quad \times\left(\mathbf{E}^{\zeta_k^*}\left[\left(1+\bar{\pi}'\hat{J}_k Y_k\right)^{1-\gamma} - 1\right]\right) + \frac{1}{\phi}H(\zeta_t^*).
\end{aligned} \tag{A.10}$$

The first-order condition with respect to $\bar{\pi}$ gives

$$-\gamma\bar{\pi} + \bar{\mu} + \sum_{k=1}^{n}\lambda_k \mathbf{E}^{\zeta_k^*}\left[\left(1+\bar{\pi}'\hat{J}_k Y_k\right)^{-\gamma} Y_k\right]\hat{J}_k = 0. \tag{A.11}$$

Note that we have used the first-order condition for the optimality of ζ_k^* (i.e. (A.2) in the aforementioned derivation).

The optimal portfolio is given by

$$\pi^* = (\Sigma')^{-1}\hat{\pi}^* = (\Sigma')^{-1}\left(\bar{\pi}^* + \pi_\perp^*\right).$$

Finally, the worst-case probability ζ^* is obtained by substituting $\pi = \pi^*$ into (C.1) and (C.2). The value function $V(W,t) = U(W)h^*(t)$, where $h^*(t)$ satisfies (A.1) with the optimizers given by π^* and ζ^*, respectively. □

Proof of Proposition 3

Proof. By Proposition 1,

$$\varphi^*(y) = \frac{\exp\left(\frac{\phi}{\gamma-1}\left((1+\tilde{\pi}y)^{1-\gamma} - 1\right)\right)}{\int_A \exp\left(\frac{\phi}{\gamma-1}\left((1+\tilde{\pi}x)^{1-\gamma} - 1\right)\right)\Phi(dx)}.$$

Let $f(y) = \exp(\frac{\phi}{\gamma-1}((1+\tilde{\pi}y)^{1-\gamma} - 1))$. Then, $\varphi^*(y) = f(y)/\int_A f(x)\Phi(dx)$. Note that $f(y)$ is monotonically decreasing in y because $\tilde{\pi} \in [0,1]$ for the mixed jump size (i.e., $y \in (-1,\infty)$). Hence, $f(-1) > f(y) > f(\infty)$. Then,

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

Because $f(y)$ 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 $y \le y^*$,

$$\Pr{}^*(Y_k < y|\phi) = \int_{-1}^{y}\varphi^*(x)\Phi(dx) > \int_{-1}^{y}\Phi(dx) = \Pr(Y_k < y). \tag{A.12}$$

To prove the second part of the proposition, we write $\varphi(y;\phi)=\varphi^*(y)$ and denote $g(y)=\frac{1}{\gamma-1}[(1+\tilde{\pi}y)^{1-\gamma}-1]$. Then, for $a\in(-1,\infty)$ and $\phi\in[\phi_1,\phi_2]$,

$$\frac{d\varphi(a;\phi)}{d\phi}=f(a)\frac{g(a)\int_A f(y)\Phi(dy)-\int_A f(y)g(y)\Phi(dy)}{\left(\int_A f(y)\Phi(dy)\right)^2}=\frac{\int_A(g(a)-g(y))f(y)\Phi(dy)}{f^{-1}(a)\left(\int_A f(y)\Phi(dy)\right)^2}.$$

Because $g(-1)>g(y)$ for $y\in(-1,\infty)$, to show that $\frac{d\varphi(a;\phi)}{d\phi}>0$ for $\phi\in[\phi_1,\phi_2]$, it suffices to prove that, as a goes to -1, $\int_A(g(a)-g(y))f(y)\Phi(dy)$ uniformly converges to $\int_A(g(-1)-g(y))f(y)\Phi(dy)$ with respect to $\phi\in[\phi_1,\phi_2]$. To this end, without loss of generality, we assume that $\tilde{\pi}=\tilde{\pi}_0<1$ when $\phi=0$. Then, by result i of (22), $\tilde{\pi}\le\tilde{\pi}_0<1$ for any $\phi\in[\phi_1,\phi_2]$. Thus, the function $f(y)$ is uniformly bounded with respect to $y\in(-1,\infty)$ and $\phi\in[\phi_1,\phi_2]$. Furthermore, notice that for any $y_1,y_2\in(-1,\infty)$,

$$|g(y_1)-g(y_2)|\le\frac{1}{(1-\tilde{\pi}_0)^\gamma}|y_1-y_2|.$$

Then, $g(a)$ converges uniformly to $g(-1)$ as a goes to -1 with respect to $\phi\in[\phi_1,\phi_2]$. As a result, by Lebesgue's dominated convergence theorem, as a goes to -1, $\int_A(g(a)-g(y))f(y)\Phi(dy)$ uniformly converges to $\int_A(g(-1)-g(y))f(y)\Phi(dy)$ with respect to $\phi\in[\phi_1,\phi_2]$.

Hence, there exists a y^{**}, such that for any $y<y^{**}$, $\varphi(y;\phi)$ is an increasing function in $\phi\in[\phi_1,\phi_2]$. And for any $\phi_1>\phi_2$, there exists $\hat{y}$, such that for any $y<\hat{y}$,

$$\Pr{}^*(Y_k<y|\phi_1)=\int_{-1}^{y}\varphi(x;\phi_1)\Phi(dx)>\int_{-1}^{y}\varphi(x;\phi_2)\Phi(dx)=\Pr{}^*(Y_k<y|\phi_2).\quad\square$$

Proof of (23)

We prove ii as follows. A proof for i is similar to the proof for ii and is omitted.

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

$$\begin{aligned}\frac{1}{\gamma}\frac{\partial\tilde{\pi}^*}{\partial(\hat{J}'\hat{J})}&=\frac{\partial\hat{J}'\bar{\mu}}{\partial(\hat{J}'\hat{J})}+\lambda\mathbf{E}\left[(1+\tilde{\pi}^*Y)^{-\gamma}Ye^{\frac{\phi}{\gamma-1}((1+\tilde{\pi}^*Y)^{1-\gamma}-1)}\right]\\&+\lambda(-\gamma)\mathbf{E}\left[(1+\tilde{\pi}^*Y)^{-\gamma-1}Y^2e^{\frac{\phi}{\gamma-1}((1+\tilde{\pi}^*Y)^{1-\gamma}-1)}\right](\hat{J}'\hat{J})\frac{\partial\tilde{\pi}^*}{\partial(\hat{J}'\hat{J})}\\&+\lambda\mathbf{E}\Big[(1+\tilde{\pi}^*Y)^{-\gamma}Ye^{\frac{\phi}{\gamma-1}((1+\tilde{\pi}^*Y)^{1-\gamma}-1)}\\&\times\phi(-1)(1+\tilde{\pi}^*Y)^{-\gamma}Y\Big](\hat{J}'\hat{J})\frac{\partial\tilde{\pi}^*}{\partial(\hat{J}'\hat{J})}.\end{aligned}$$

We can solve $\frac{\partial\tilde{\pi}^*}{\partial(\hat{J}'\hat{J})}$ from the previous equality, and it turns out that $\frac{\partial\tilde{\pi}^*}{\partial(\hat{J}'\hat{J})}>0$ if and only if the sum of the first two terms of the right-hand side is positive, or condition (23) is satisfied. $\square$

Proof of Proposition 4

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 is worth 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.[28] 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 $\tilde{\pi}^*\in(0,1)$. It follows from (19) that

$$\tilde{\pi}^*=\frac{1}{\gamma}\hat{J}'\bar{\mu}+\frac{1}{\gamma}\lambda\mathbf{E}^{\zeta^*}\left[(1+\tilde{\pi}^*Y)^{-\gamma}Y\right](\hat{J}'\hat{J}).\tag{A.13}$$

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

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

$$\lambda\mathbf{E}^{\zeta^{*,p}}\left[(1+\tilde{\pi}^{*,p}Y^p)^{-\gamma}Y^p\right](\hat{J}'\hat{J})<\lambda\mathbf{E}^{\zeta^{*,n}}\left[(1+\tilde{\pi}^{*,n}Y^n)^{-\gamma}Y^n\right](\hat{J}'\hat{J}),\tag{A.14}$$

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

$$\begin{aligned}&\mathbf{E}^{\zeta^{*,p}}\left[(1+\tilde{\pi}^{*,p}Y^p)^{-\gamma}Y^p\right]-\mathbf{E}^{\zeta^{*,n}}\left[(1+\tilde{\pi}^{*,n}Y^n)^{-\gamma}Y^n\right]\\&=\int_{-1}^{\infty}G^p(y)f^p(y)dy-\int_{-1}^{\infty}G^n(y)f^n(y)dy\\&=\int_{-1}^{\infty}(G^p(y)-G^n(y))f^p(y)dy+\int_{-1}^{\infty}G^n(y)(f^p(y)-f^n(y))dy,\end{aligned}\tag{A.15}$$

where $G^p(y)=(1+\tilde{\pi}^{*,p}y)^{-\gamma}ye^{\frac{\phi}{\gamma-1}((1+\tilde{\pi}^{*,p}y)^{1-\gamma}-1)}$, $G^n(y)=(1+\tilde{\pi}^{*,n}y)^{-\gamma}ye^{\frac{\phi}{\gamma-1}((1+\tilde{\pi}^{*,n}y)^{1-\gamma}-1)}$. Letting $G(\tilde{\pi},y)=(1+\tilde{\pi}y)^{-\gamma}ye^{\frac{\phi}{\gamma-1}((1+\tilde{\pi}y)^{1-\gamma}-1)}$, it is easy to check that, for a fixed y, the function $G(\tilde{\pi},y)$ is a strictly decreasing function of $\tilde{\pi}$. Given $\tilde{\pi}^{*,p}>\tilde{\pi}^{*,n}$, 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, because $G^n(y)$ increases in y for $y<0$, using the mild conditions 1 and 3, we obtain

$$\begin{aligned}\int_{-\infty}^{0}G^n(y)(f^p(y)-f^n(y))dy&=\int_{-\infty}^{y_0}G^n(y)(f^p(y)-f^n(y))dy\\&+\int_{y_0}^{y_1}G^n(y)(f^p(y)-f^n(y))dy+\int_{y_1}^{0}G^n(y)(f^p(y)-f^n(y))dy\\&<G^n(y_0)\int_{-\infty}^{y_0}(f^p(y)-f^n(y))dy+G^n(y_0)\int_{y_0}^{y_1}(f^p(y)-f^n(y))dy\\&=G^n(y_0)\int_{-\infty}^{y_1}(f^p(y)-f^n(y))dy<0,\end{aligned}$$

because $G^n(y_0)<0$ for $y_0<0$ and $\int_{-\infty}^{y_1}(f^p(y)-f^n(y))dy=\int_{y_1}^{\infty}(f^n(y)-f^p(y))dy>0$ due to 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). □

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 modeled by the stochastic differential equation:

$$\zeta_{t,k}^* = \zeta_{0,k}^* + \int_0^t \int_A (\varphi_k^*(s,y) - 1)\zeta_{s-,k}^* q_k(ds,dy),$$

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

$$\begin{aligned}\ln(\zeta_{t,k}^*) &= \lambda_k \int_0^t \int_A [\ln(\varphi_k^*(s,y)) + 1 - \varphi_k^*(s,y)]\Phi_k(s,dy)ds \\ &\quad + \int_0^t \int_A \ln(\varphi_k^*(s,y)) q_k(ds,dy) \\ &= \lambda_k \int_0^t \int_A \ln(\varphi_k^*(s,y))\Phi_k(s,dy)ds \\ &\quad + \int_0^t \int_A \ln(\varphi_k^*(s,y)) q_k(ds,dy) \\ &= \int_0^t \int_A \ln(\varphi_k^*(s,y)) dN_{k,s} ds,\end{aligned}$$

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

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

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

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

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 use a Monte Carlo approach to determine $\varepsilon_T(\phi)$ for each ϕ.

Appendix C. The Case of Log Utility Function

In the paper, we adopt a CRRA utility function $U(W) = \frac{W^{1-\gamma}}{1-\gamma}$, where $\gamma > 0$ and $\gamma \neq 1$. In this appendix, we provide the result for the case of log utility function $U(W) = \log(W)$.

Replacing (9) by $\theta_t = \phi$ and following similar steps, we obtain the following:

$$\varphi_k^*(y_k) = \frac{1}{\vartheta^*}(1 + \pi'\mathbf{J}_k y_k)^{-\phi}, \quad \text{(C.1)}$$

$$\vartheta_k^* = \mathbf{E}[(1 + \pi'\mathbf{J}_k Y_k)^{-\phi}]. \quad \text{(C.2)}$$

Because the log utility function is a limiting case of the general CRRA utility, we get closely mimicking results for the worst-case probability. In fact, (C.1) and (C.2) can be directly obtained by letting $\gamma \to 1$ in (10) and (11).

Endnotes

[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.

[2] We also investigate the locally constrained specification of robustness adopted by, for example, Trojani and Vanini (2002) and Aït-Sahalia and Matthys (2019). Both approaches lead to the same results under our model setting. See Section 2 for detailed discussions. It is, however, worth mentioning that the two specifications for robustness considered in this paper do not allow the separation between ambiguity and ambiguity aversion. Klibanoff et al. (2005) propose a model where the functional form of the preference allows for the separation between ambiguity and ambiguity aversion (we thank an anonymous referee for pointing this out). Analogous to the differentiation of risk from risk aversion, the separation can allow one to perform comparative statics by varying only ambiguity aversion (or only ambiguity). It is nevertheless challenging to solve the optimal portfolio choice problem of a multiasset incomplete market in closed form under this framework.

[3] To avoid picking the functional form for the consumption disaster distribution, a popular method in asset pricing studies is to directly use the empirical distribution (Wachter 2013, Seo and Wachter 2019). In such a case, the parametric approach is undefined and only our nonparametric method can be applied.

[4] 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.

[5] Das and Uppal (2004) find that systemic jump risk may cause reduction in diversification. Focusing on model misspecification in a market with pure diffusion risk, Uppal and Wang (2003) show that an ambiguous return distribution results in an underdiversified portfolio relative to the standard mean-variance portfolio.

[6] In particular, Cvitanić et al. (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 generating portfolio outcomes similar to that of the same diffusion model.

[7] Of course, a naive investor without ambiguity aversion shall encounter much larger economic losses when the worst-case scenario happens.

[8] We examine the case of log utility ($\gamma = 1$) in Appendix C.

[9] As in Liu et al. (2005), the diffusion parameters are less of our concern because the corresponding decision making can be founded on abundant daily fluctuations of the asset prices.

[10] See Hansen and Sargent (2008) for a textbook treatment.

[11] We thank an anonymous referee for this clarification.

[12] We thank an anonymous referee for suggesting that we investigate this specification of ambiguity aversion.

[13] We discuss the case of log utility ($\gamma = 1$) in Section OA.1 of the online appendix.

[14] The derivations are available upon request.

[15] We treat m as if it could take any real number, to help infer the monotonic property for positive integer values of m.

[16] This can be proved by taking the derivative of φ_k^* with respect to $\tilde{\pi}_k$.

[17] As $\phi \to \infty$, the minimization component in the HJB Equation (A.1) becomes a term as follows:

$$\inf_{P(\zeta)} \frac{1}{1-\gamma} \sum_{k=1}^{n} \lambda_k \mathbf{E}^{\zeta_k}\big[(1+\pi'\mathbf{J}_k Y_k)^{1-\gamma} - 1\big].$$

Suppose $\gamma > 1$ (similar arguments for $0 < \gamma < 1$). If $\pi'\mathbf{J}_k > 0$, we can find $\psi(\cdot)$, such that $\int_A (1+\pi'\mathbf{J}_k y)^{1-\gamma}\psi(y)\Phi_k(dy) - 1 > 0$. ($\psi$ "shifts" the distribution of Y_k to the negative side, such that $\mathbf{E}[(1+\pi'\mathbf{J}_k Y_k)^{1-\gamma} \psi(Y_k)] > 1$.) Then, we can find a sequence $\zeta_i = \zeta(\vartheta^{(i)}, \psi)$, where $\vartheta^{(i)} \to \infty$ as $i \to \infty$, such that

$$\inf_{P(\zeta)} \frac{1}{1-\gamma} \lambda_k \mathbf{E}^{\zeta(k)}\big[(1+\pi'\mathbf{J}_k Y_k)^{1-\gamma} - 1\big] \le \frac{1}{1-\gamma}\lambda_k \vartheta^{(i)} \times \left(\int_A (1+\pi'\mathbf{J}_k y)^{1-\gamma}\psi(y)\Phi(dy) - 1\right) \to -\infty,$$

as $i \to \infty$. Similarly, for the case $\pi'\mathbf{J}_k < 0$, we can show that the term on the left of the previous inequality approaches $-\infty$ as well. As a result, the optimal portfolio $(\pi^*)'\mathbf{J_k}$ must be zero, at which the term is zero. We examine the model extensively and, in particular, find a necessary and sufficient condition for the occurrence of nonparticipation in Section OA.1.1 of the online appendix. We thank an anonymous referee for suggesting this study.

[18] See Epstein and Schneider (2010) for more discussions on ambiguity and participation in markets with pure diffusion risk.

[19] Wang (2015) provides evidence that tail risk is a common priced factor in the international equity markets.

[20] The data are frequently used by studies of international equity markets (e.g., recently by Aït-Sahalia et al. 2015).

[21] 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.

[22] Comparison of other coskewness conditions reaches essentially the same conclusions. The detailed results are available upon request.

[23] An interesting alternative preference for this purpose is the downside loss-averse utility considered by Jarrow and Zhao (2006).

[24] We thank an anonymous referee for suggesting this study.

[25] To save space, we only report the case of $\gamma = 3$. Results for $\gamma = 5$ and 1 are available upon request.

[26] We thank an anonymous referee for motivating us to pursue this study.

[27] 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.

[28] For cases with one or more than two intersection points, the proposition can be proved similarly.

References

Aït-Sahalia Y, Matthys F (2019) Robust consumption and portfolio policies when asset prices can jump. *J. Econom. Theory* 179-(January):1–56.

Aït-Sahalia Y, Cacho-Diaz J, Hurd T (2009) Portfolio choice with jumps: A closed-form solution. *Ann. Appl. Probab.* 19(2):556–584.

Aït-Sahalia Y, Cacho-Diaz J, Laeven R (2015) Modeling financial contagion using mutually exciting jump processes. *J. Financial Econom.* 117(3):585–606.

Agarwal V, Ruenzi S, Weigert F (2017) Tail risk in hedge funds: A unique view from portfolio holdings. *J. Financial Econom.* 125(3): 610–636.

Anderson E, Hansen L, Sargent T (2003) A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *J. Eur. Econom. Assoc.* 1(1):68–123.

Ang A, Bekaert G, Liu J (2005) Why stocks may disappoint? *J. Financial Econom.* 76(3):471–508.

Barro R, Jin T (2011) On the size distribution of macroeconomic disasters. *Econometrica* 79(5):1567–1589.

Black F, Litterman R (1992) Global portfolio optimization. *Financial Analysts J.* 48(5):28–43.

Bollerslev T, Todorov V (2011a) Estimation of jump tails. *Econometrica* 79(6):1727–1783.

Bollerslev T, Todorov V (2011b) Tails, fears, and risk premia. *J. Finance* 66(6):2165–2211.

Bollerslev T, Todorov V, Xu L (2015) Tail risk premia and return predictability. *J. Financial Econom.* 118(1):113–134.

Boyle P, Garlappi L, Uppal R, Wang T (2012) Keynes meets Markowitz: The trade-off between familiarity and diversification. *Management Sci.* 58(2):253–272.

Branger N, Larsen L (2013) Robust portfolio choice with uncertainty about jump and diffusion risk. *J. Banking Finance* 37(12):5036–5047.

Bremaud P (1981) *Point Processes and Queues: Martingale Dynamics* (Springer, Berlin).

Campbell J (2006) Household finance. *J. Finance* 61(4):1553–1604.

Cao H, Wang T, Zhang H (2005) Model uncertainty, limited market participation, and asset prices. *Rev. Financial Stud.* 18(4):1219–1251.

Cvitanić J, Polimenis V, Zapatero F (2008) Optimal portfolio allocation with higher moments. *Ann. Finance* 4:1–28.

Das S, Uppal R (2004) Systemic risk and international portfolio choice. *J. Finance* 59(6):2809–2834.

Drechsler I (2013) Uncertainty, time-varying fear, and asset prices. *J. Finance* 68(5):1843–1889.

Easley D, O'Hara M (2009) Ambiguity and nonparticipation: The role of regulation. *Rev. Financial. Stud.* 22(5):1817–1843.

Epstein L, Schneider M (2010) Ambiguity and asset markets. *Annual Rev. Financial Econom.* 2(1):315–346.

Eraker B, Johannes M, Polson N (2003) The impact of jumps in volatility and returns. *J. Finance* 58(3):1269–1300.

Flor C, Larsen L (2014) Robust portfolio choice with stochastic interest rates. *Ann. Finance* 10:243–265.

Hansen L, Sargent T (2008) *Robustness* (Princeton University Press, Princeton, NJ).

Hong Y, Tu J, Zhou G (2007) Asymmetries in stock returns: Statistical tests and economic evaluation. *Rev. Financial Stud.* 20(5): 1547–1581.

Jarrow R, Zhao F (2006) Downside loss aversion and portfolio management. *Management Sci.* 52(4):558–566.

Jin X, Zhang A (2012) Decomposition of optimal portfolio weight in a jump-diffusion model and its applications. *Rev. Financial Stud.* 25(9):2877–2919.

Jin X, Luo D, Zeng X (2018) Dynamic asset allocation with uncertain jump risks: A pathwise optimization approach. *Math. Oper. Res.* 43(2):347–376.

Kelly B, Jiang H (2014) Tail risk and asset prices. *Rev. Financial Stud.* 27(10):2841–2871.

Klibanoff P, Marinacci M, Mukerji S (2005) A smooth model of decision making under ambiguity. *Econometrica* 73(6):1849–1892.

Liu J, Zeng X (2017) Correlation ambiguity and under-diversification. Working paper, Rady School of Management, University of California at San Diego, San Diego.

Liu J, Longstaff F, Pan J (2003) Dynamic asset allocation with event risk. *J. Finance* 58(1):231–259.

Liu J, Pan J, Wang T (2005) An equilibrium model of rare-event premia and its implication for option smirks. *Rev. Financial Stud.* 18(1):131–164.

Maenhout P (2004) Robust portfolio rules and asset pricing. *Rev. Financial Stud.* 17(4):951–983.

Maenhout P (2006) Robust portfolio rules and detection-error probabilities for a mean-reverting risk premium. *J. Econom. Theory* 128(1):136–163.

Mankiw N, Zeldes S (1991) The consumption of stockholders and nonstockholders. *J. Financial Econom.* 29(1):97–112.

Merton R (1971) Optimum consumption and portfolio rules in a continuous-time model. *J. Econom. Theory* 3(4):373–413.

Merton R (1976) Option pricing when underlying stock returns are discontinuous. *J. Financial Econom.* 3(1–2):125–144.

Naik V, Lee M (1990) General equilibrium pricing of options on the market portfolio with discontinuous returns. *Rev. Financial Stud.* 3(4):493–521.

Seo S, Wachter J (2019) Option prices in a model with stochastic disaster risk. *Management Sci.* 65(8):3449–3947.

Trojani F, Vanini P (2002) A note on robustness in Merton's model of intertemporal consumption and portfolio choice. *J. Econom. Dynam. Control* 26(3):423–435.

Trojani F, Vanini P (2004) Robustness and ambiguity aversion in general equilibrium. *Rev. Finance* 8(2):279–324.

Uppal R, Wang T (2003) Model misspecification and under-diversification. *J. Finance* 58(6):2465–2486.

Wachter J (2013) Can time-varying risk of rare disasters explain aggregate stock market volatility? *J. Finance* 68(3):987–1035.

Wang Y (2015) An ignored risk factor in international markets: Tail risk. Working paper, Purdue University, West Lafayette, IN.

Weinstock R (1974) *Calculus of Variations: With Applications to Physics and Engineering* (Dover Publications, New York).