Dynamic Asset Allocation with Uncertain Jump Risks: A Pathwise Optimization Approach*

Xing Jin

Warwick Business School, University of Warwick, Coventry, CV4 7AL, UK

telephone: +44 2476575698; Email: Xing.Jin@wbs.ac.uk.

Dan Luo

School of Finance, Shanghai University of Finance and Economics,

777 Guoding Road, Shanghai, 200433, China; Email: luo.dan@mail.shufe.edu.cn.

Xudong Zeng

School of Finance, Shanghai University of Finance and Economics,

777 Guoding Road, Shanghai, 200433, China; Email: zeng.xudong@mail.shufe.edu.cn.

December 22, 2016

Abstract

This paper studies the dynamic portfolio choice problem with ambiguous jump risks in a multi-dimensional jump-diffusion framework. We formulate a continuous-time model of incomplete market with uncertain jumps. We develop an efficient pathwise optimization procedure based on the martingale methods and minimax results to obtain closed-form solutions for the indirect utility function and the probability of the worst scenario. We then introduce an orthogonal decomposition method for the multi-dimensional problem to derive the optimal portfolio strategy explicitly under ambiguity aversion to jump risks. Finally, we calibrate our model to real market data drawn from ten international indices and illustrate our results by numerical examples. The certainty equivalent losses affirm the importance of jump uncertainty in optimal portfolio choice.

MSC2000 subject classification: Primary: 91B28, 93E20; secondary: 91A15
OR/MS subject classification: Primary: finance, portfolio; secondary: games/group decisions, stochastic
Keywords: Optimal investment, Ambiguity aversion, Duality method, Jump diffusion, Worst-case scenario

*We are grateful to an Associate Editor and two anonymous referees for their insightful and detailed comments that substantially improved this article. We thank Jaksa Cvitanic, Hening Liu, Jun Liu, and seminar participants at Manchester Business School, Shanghai University of Finance and Economics, Xi'an Jiaotong-Liverpool University, and 2013 China International Conference of Finance for helpful comments.

1 Introduction

A number of empirical and theoretical studies have demonstrated that jump risks have a substantial impact on optimal portfolio formation. For example, in a single-stock double-jump model, Liu et al. [24] find that an investor is less willing to take leveraged or short positions than in a standard diffusion model, due to the investor's inability to hedge jump risks through continuous rebalancing. In an international market setting, Das and Uppal [10] find that systemic jumps reduce the gain from international diversification and that leveraged portfolios may incur large losses upon the strike of jumps. Meanwhile, estimation of jump models usually lacks precision because jumps, especially those associated with disaster events, are inherently rare hence difficult to identify. The reference model associated with the point estimate is highly likely to deviate from the "true" data-generating one. Accordingly, aversion to model uncertainty, or ambiguity aversion,[1] is incorporated into dynamic asset allocation problems wherein an investor encounters jump risks (see, e.g., Liu et al. [25], Jin and Zhang [19], Branger and Larsen [7], and Drechsler [12]).

In this paper, we propose an efficient pathwise optimization approach to solve portfolio choice problems in multi-asset and multi-state-variable jump-diffusion models. Under these models, an investor, facing both jump and diffusion risks, is averse not only to the risk of loss but also to the uncertainty regarding the imprecise estimation of the jump processes. For analytic tractability, our robust control framework closely resembles that of Liu et al. [25]. Our portfolio method addresses uncertainty regarding the jump size distribution without assuming a parametric form for alternative jump size distributions; this enhanced generality distinguishes our work from previous studies, e.g., Liu et al. [25], Jin and Zhang [19], Branger and Larsen [7], and Drechsler [12].

As is well understood, it is extremely difficult to find the solution to an optimal

[1] Following the literature, we use the terms *uncertain*(*uncertainty*) and *ambiguous*(*ambiguity*) synonymously.

portfolio selection problem in an incomplete market in which there are a large number of assets and state variables, especially when model uncertainty is present. One usually uses either the HJB equation approach or the duality-martingale methods to tackle the problem. Application of the HJB equation to a high-dimensional problem is confined by the curse of dimensionality, and the martingale methods are not readily extended to an incomplete market because there are infinitely many martingale measures. In this paper, we develop a new approach based on the martingale methods and minimax results to deriving closed form solutions up to solving a set of pathwise optimization problems for the probability of the worst case scenario and the indirect value function. We then solve the corresponding optimal portfolio by an orthogonal decomposition.

Equipped with the theoretical results, we conduct a calibration exercise to apply our approach to gauge the effects of uncertain jump risks. In an economy consisting of ten international indices, we consider a constant relative risk aversion (CRRA) utility function and solve the optimal portfolio choice problem with normally distributed return jump size. We find that the total risky allocations are reduced due to the uncertain jump risks relative to the optimal portfolio weights without jump ambiguity and ambiguity aversion. In economic terms, failing to accommodate uncertain jump risks leads to as high as a 95% loss in the investor's certainty equivalent wealth for a 20-year investment horizon in the worst case, under a moderate magnitude of ambiguity aversion. This result confirms the importance of jump uncertainty in portfolio choice. Similarly, constraining to parametric alternative jump size distributions instead of more general nonparametric alternatives registers a notable 33% loss in certainty equivalent wealth for a 20-year investment when the investor is less risk averse while relatively highly uncertainty averse.

Our approach to solving the optimal portfolio choice problem is closely related to the work of Jin and Zhang [19] who use a decomposition approach based on an HJB equation. However, they focus on uncertain jump frequency while do not touch uncertain jump size distribution. Moreover, their approach is based on the HJB equation for CRRA utility functions and is not easy to extend to more general HARA utility functions. In contrast, in the present paper we do not rely on the dynamic programming principle and instead we develop a pathwise optimization method based on a duality-martingale approach in combination with minimax results. Our approach enables us to obtain the worst case probability and in turn to study the effects of ambiguous jump size distribution on portfolio choice theoretically and empirically. Furthermore, our method is certainly desirable for tackling possibly large scale problems and rigorously studying the existence of solutions, and can be extended to study more general HARA utility functions. Besides, we develop an alternative decomposition method which can easily solve the multi-dimensional portfolio choice problem after the worst probability is already obtained by our pathwise optimization approach.

Our paper is also related to several papers in the operations research literature regarding robust portfolio choice. By using the martingale method, Seifried [32] proposes a pathwise approach to study optimal investment for worst-case scenario in a non-probabilistic jump model, which is different from the probabilistic jump model in the present paper. The martingale approach used there may not be easily extended to deal with such a case with state variables (e.g. stochastic volatilities) as in our model. Moreover, the present paper follows the line of robust control approach proposed by Hansen and Sargent [17],[18], dealing with portfolio choice under ambiguity. A special case (infinite ambiguity aversion) of our objective function corresponds to the max-min problem studied in Seifried's paper. Goldfarb and Iyengar [15] also study portfolio selection problems under uncertainty, but they consider a framework of mean-variance. Laeven and Stadje [23] investigate the problems of optimal portfolio choice and indifference valuation in a general continuous-time setting with time-consistent ambiguity-averse preferences and a general and possibly infinite activity jump part in the asset price processes. The solutions are characterized as solutions to backward stochastic differential equations (BSDEs). The present paper is different from these mentioned studies either in mathematical models or in methodologies. For other related literature, we refer to Pennanen [29] regarding duality approach, Zhao and Ziemba [35] regarding asset allocation with transaction costs; etc.

The rest of the paper is organized as follows. In the next section, we present the framework for Merton's dynamic portfolio selection problem and demonstrate how it can be extended to incorporate ambiguity aversion. In Section 3, we develop a pathwise optimization approach using the martingale methods and minimax results. The worst case probability of jumps is obtained. The proof of the main result is divided into three subsections. We then find the optimal portfolio choice under the worst case probability in Section 4. Section 5 is an extension to HARA utility functions. Section 6 is devoted to a calibration exercise for a model consisting of ten international indices to evaluate an investor's fear of uncertain jump risks. Section 7 concludes. The proof of Proposition 1 is collected in Appendix A.

2 Merton's problem and ambiguity aversion

In this section we formulate a model of incomplete financial market in continuous time. Asset prices follow a multi-dimensional jump-diffusion process on the fixed time horizon $[0,T]$, $0<T<\infty$. We consider a complete probability space $(\Omega,\mathcal{F},P)$, where Ω is the set of states of nature with generic element ω, $\mathcal{F}$ is the σ-algebra of observable events and P is a probability measure on $(\Omega,\mathcal{F})$.

The market considered in this paper includes $m+1$ assets traded continuously on the time horizon $[0,T]$. One of these assets, which is risk-free, has a price $S_{0,t}$ evolving according to the differential equation

$$dS_{0,t} = S_{0,t}r(X_t)dt, \qquad S_{0,0}=1. \tag{1}$$

The process $X_t=(X_{1,t},...,X_{l,t})^\top$ is an l-dimensional vector representing the state variables of the economy, where $^\top$ denotes transpose of the row vector. X_t may include stochastic volatilities and stochastic interest rate as its components. For analytical tractability, we assume that X_t follows a pure diffusion process:

$$dX_t = b^X(X_t)dt + \sigma^X(X_t)dB_t, \tag{2}$$

where $b^X(X_t)$ is an l-dimensional vector function and $\sigma^X(X_t)$ is an $l \times d$ matrix function of X_t, respectively. $\sigma^X(X_t)$ has diffusion coefficient row vectors $\sigma_i^X(X_t), i = 1, ..., l$. $B_t = (B_{1,t}, ..., B_{d,t})^\top$ is a d-dimensional standard Brownian motion.

The remaining m assets, called stocks, are risky. The price vector $S_t = (S_{1,t}, ..., S_{m,t})^\top$ is modelled by the linear stochastic differential equation

$$dS_t = diag(S_{t-})\left(b(X_t)dt + \Sigma_b(X_t)dB_t + \Sigma_q(X_t)YdN_t\right),$$

where $b(X_t)$ is an m-dimensional vector function; $\Sigma_b(X_t)$ is an $m \times d$ matrix with diffusion coefficient row vectors $\sigma_i^b(X_t), i = 1, ..., m$; $\Sigma_q(X_t)$ is an $m \times (n-d)$ matrix, with jump coefficient row vectors $\sigma_i^q, i = 1, ..., m$; Let $\Sigma = [\Sigma_b, \Sigma_q]$.[2] Y is a diagonal matrix with diagonal entries $Y_1, ..., Y_{n-d}$. Y_k representing the amplitude of type k jump has probability density $\Phi_k(t, dz)$. $N_t = (N_{1,t}, ..., N_{n-d,t})^\top$ is an $(n-d)$-dimensional multivariate Poisson process. $N_{k,t}$ admits stochastic intensity $\lambda_k(X_t)$. Our results can be extended to infinite activity jump processes. We assume that N_t is independent of B_t.[3]

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

[2] Without loss of generality, we assume that $rank(\Sigma) = m$ if $m \leq n$; $rank(\Sigma) = n$ if $m > n$ to avoid redundant stocks in the model. See Section 3 for more discussion on the number of stocks and the number of risk sources.

[3] The state variable process and the stock price vector in our model are governed by the same Brownian motion $B(t)$. Note that when $d \geq 2$, the instantaneous correlation between the diffusions of X_t ($\sigma_i^X dB_t$) and stock return ($\sigma_j^b dB_t$) may range from -1 to 1 for each $i = 1, ..., l$ and $j = 1, ..., m$. Hence in general the state variables are not perfectly correlated with the continuous part of stock prices even if they are driven by the same multi-dimensional Brownian motion.

$\mathcal{F}_t^B = \sigma(B_s; 0 \le s \le t)$, and $\mathcal{F}_t^N = \sigma(N_s; 0 \le s \le t)$. Observable events are eventually known, i.e., $\mathcal{F} = \mathcal{F}_T$.

We consider that an investor is endowed with initial wealth W_0; this wealth is invested in the above-mentioned $m+1$ assets. Let $\pi_t = (\pi_{1,t}, ..., \pi_{m,t})^\top$ denote a portfolio, where $\pi_{k,t}$ is the proportion of total wealth invested in the k-th stock at time t and is $\mathcal{F}_t$-predictable. Any portfolio policy π_t has an associated wealth process W_t that evolves as

$$\begin{aligned} W_t &= W_0 + \int_0^t r(X_s) W_s ds + \int_0^t W_s \pi_s^\top (b(X_s) - r(X_s)\mathbf{1}_m) ds \\ &\quad + \int_0^t W_s \pi_s^\top \Sigma_b(X_s) dB_s + \int_0^t W_{s-} \pi_{s-}^\top \Sigma_q(X_s) Y dN_s, \end{aligned} \tag{3}$$

where we use $\mathbf{1}_m$ to denote the m-dimensional column vector of ones. A portfolio policy 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 trading strategies, given initial wealth $W_0 = w_0$, and we denote by $\mathcal{W}(w_0)$ the family of all wealth processes generated by admissible trading strategies in $\mathcal{A}(w_0)$.

Given a portfolio π in equation (3), the vectors

$$\widetilde{\pi}_b = (\widetilde{\pi}_{b1}, ..., \widetilde{\pi}_{bd}) = \pi_t^\top \Sigma_b(X_t) \quad \text{and} \quad \widetilde{\pi}_q = (\widetilde{\pi}_{q1}, ..., \widetilde{\pi}_{q(n-d)}) = \pi_t^\top \Sigma_q(X_t)$$

measure the exposures or sensitivities to diffusion and jump risks, respectively. In particular, $\widetilde{\pi}_{qk}$ reduces to the portfolio weight of stock in a single-stock jump-diffusion model studied, e.g., by Liu et al. [24], while in the multi-stock jump-diffusion models in the present paper, the investor reacts to the k-th jump risk by choosing $\widetilde{\pi}_{qk}$ appropriately.

The traditional Merton's problem without ambiguity aversion is that the investor attempts to maximize the following quantity

$$J(w_0) = \max_{\pi \in \mathcal{A}(w_0)} \mathbf{E}\left[u(W_T)\right],$$

where the utility function $u(x)$ is non-decreasing and concave on $\mathcal{R} = (-\infty, \infty)$, and $\mathbf{E}[\cdot]$ denotes the expectation under the natural probability measure P.

Our next step is to incorporate ambiguity aversion into Merton's problem. Suppose that an investor fears possible model misspecifications and makes investment decisions to guard against the worst case scenario. Rare disasters in our model are typically high impact events, while the parameters of the underlying jump processes are difficult to estimate with adequate accuracy. We therefore focus on the investor's ambiguity aversion with regard to uncertain jump parameters to address the issues raised in the introduction. In other words, the investor's problem stems from a class of prior models generated by imprecise estimates of the jump parameters governing, e.g., the jump intensity and jump size distribution. The investor considers the point estimates and the corresponding model (called the reference model) to be the most reliable, while she also explicitly recognizes that the competing models are difficult to distinguish statistically from the reference model. As a result, the investor makes a precautionary portfolio choice to guard against the competing alternatives such that her portfolio performs reasonably well even if the worst case scenario occurs. However, choosing any model other than the reference model is penalized because the selection is a deviation from the most likely model.

Before defining the utility function that incorporates ambiguity aversion and deviation penalty, we introduce a set of probability measures, denoted by $\mathcal{P}$, that specify alternative models of concern. To this end, we define the martingale differential as

$$q(dt, dz) = (q_1(dt, dz), ..., q_{n-d}(dt, dz)),$$

where

$$q_k(dt, dz) = dN_k(t) - \lambda_k(X_t)\Phi_k(t, dz)dt, \quad k = 1, ..., n-d.$$

Note that P is the probability measure associated with the reference model. Each probability measure $P(\zeta) \in \mathcal{P}$ has a Radon-Nikodym derivative, $\frac{dP(\zeta)}{dP} = \zeta_T = \prod_{k=1}^{n-d} \zeta_T^{(k)}$, with respect to P, where the process $\zeta_T^{(k)}$ is modelled by the stochastic differential equation

$$\zeta_T^{(k)} = \zeta_0^{(k)} + \int_0^T \int_{A_k} (\vartheta_k(s)\psi_k(s, z) - 1)\zeta_{s-}^{(k)} q_k(ds, dz), \tag{4}$$

with $\zeta_0^{(k)} = 1$. Note that $\vartheta_k(s)$ and $\psi_k(s, z)$ are positive stochastic processes, and $\psi_k(s, z)$ satisfies the following relationship

$$\int_{A_k} \psi_k(t, z)\Phi_k(t, dz) = 1, \quad k = 1, ..., n-d, \tag{5}$$

where A_k is the support of the size of the k-th jump. In particular, we set $A_k = (0, \infty)$ for a positive jump, $A_k = (-1, 0)$ for a negative jump, and $A_k = (-1, \infty)$ for a mixed jump.

By Ito's lemma for jump processes, the Radon-Nikodym derivative ζ_t can be represented as:

$$\zeta_t = \left(\prod_{k=1}^{n-d} \prod_{i=1}^{N_k(t)} (\vartheta_k(t_i^k)\psi(t_i^k, z_i^k)) \right) \exp\left(\int_0^t \int_{A_k} (1 - \vartheta_k(s)\psi_k(s, z))\lambda_k(X_s)\Phi_k(s, dz)ds \right). \tag{6}$$

where t_i^k is the ith jump time of the kth type of jump up to t and z_i^k is the corresponding jump size. From now on, we suppress the dependence of $\lambda_k(X_t)$, $\vartheta_k(t)$, $\Phi_k(t, dz)$, and $\psi_k(t, z)$ on t and X_t for notational convenience on occasions of no confusion.

By Theorem T10 of Bremaud (1981), under the probability measure $P(\zeta)$, the intensity λ_k and the density function $\Phi_k(dz)$ are changed into $\vartheta_k\lambda_k$ and $\psi_k(z)\Phi_k(dz)$ in the alternative model for each $k = 1, ..., n-d$.

In the remainder of this paper, we use Θ_k to denote the set of all possible values of $\vartheta_k(t)$. For the k-th jump size, we use Ψ_k to denote the set of all possible nonnegative functions of $\psi_k(t,z)$ given by (5). In general, we let $\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_{n-d}$ and $\Psi = \Psi_1 \times \Psi_2 \times \cdots \times \Psi_{n-d}$ and we let $\mathcal{P}$ denote the set of all alternative probabilities determined by Θ and Ψ.

It is worth mentioning that the set of all alternative densities defined by (5) differs from that defined by equation (2) in Liu et al. [25]: we investigate model misspecification in the entire neighborhood of the reference model, while Liu et al. [25] consider only a subset of the neighborhood. In fact, every model in the neighborhood is statistically indistinguishable from the reference model and thus is possibly a true model. In particular, Liu et al. [25] use a parametric approach to choose the worst jump size and jump intensity while we apply a non-parametric method to choose the worst case. Hence, the worst case jump size distribution remains lognormal in the model of Liu et al. [25], while the worst case jump size distribution is not necessarily lognormal in our model.

We now define the utility function with ambiguity aversion. Following Liu et al. [25], we make some modifications to Merton's problem described above. We begin by formulating a utility function in a discrete-time setting and then, by taking the limit, arrive at the utility function for our continuous-time models. Specifically, for a fixed time period Δt, the time-t utility in discrete time is given in a recursive manner by

$$U_t = \inf_{P(\zeta)\in\mathcal{P}} \left\{ \Lambda\left(\mathbf{E}_t^{\zeta}(U_{t+\Delta t})\right) \sum_{k=1}^{n-d} \frac{1}{\phi_k} \mathbf{E}_t^{\zeta}\left[\ln\left(\frac{\zeta_{t+\Delta t}^{(k)}}{\zeta_t^{(k)}}\right)\right] + \mathbf{E}_t^{\zeta}(U_{t+\Delta t}) \right\}, \tag{7}$$

with $U_T = u(W_T)$, and $\mathbf{E}_t^{\zeta}$ denoting the conditional expectation under the probability $P(\zeta)$.[4] As in Liu et al. [25], $\mathbf{E}_t^{\zeta}\left[\ln\left(\frac{\zeta_{t+\Delta t}}{\zeta_t}\right)\right]$ measures the discrepancy between probabil-

[4]In (7), the utility is in sense of almost surely as usual, and the infimum refers to the essential infimum. On the other hand, by an abuse of notation, $P(\zeta) \in \mathcal{P}$ is a set of the time-t conditional probabilities determined by $\frac{\zeta_T}{\zeta_t}$. The setting of utility function (7) may be traced back to Anderson et al. [2]. For general risk functions defined from an axiomatic basis, a conditional formulation of dynamic

ities $P(\zeta)$ and P, which is the standard measure of entropy. The coefficient ϕ_k represents the magnitude of ambiguity aversion to the kth jump. A larger value of ϕ_k indicates a higher ambiguity aversion preference of the investor. The minimization problem reflects aversion to ambiguity of the investor who worries about the imprecise estimation of model parameters. Therefore, the investor makes decisions to guard against the worst scenario. $\Lambda(x)$ is a normalization factor and, for tractability, we assume $\Lambda(x) = (1-\gamma)x$ with $\gamma > 1$ following Maenhout [27]. As is well understood, the preference defined in (7) is dynamically consistent because it is defined recursively (see Epstein and Schneider [13] and Wang [33]). Then Merton's problem under ambiguity aversion is given by the following max-min problem.

$$
\begin{aligned}
J(t, W_t, X_t) &= \sup_{\pi}\{U_t\} \\
&= \sup_{\pi}\inf_{\zeta}\left\{\Lambda\left(\mathbf{E}_t^{\zeta}(U_{t+\Delta t})\right)\sum_{k=1}^{n-d}\frac{1}{\phi_k}\mathbf{E}_t^{\zeta}\left[\ln\left(\frac{\zeta_{t+\Delta t}^{(k)}}{\zeta_t^{(k)}}\right)\right] + \mathbf{E}_t^{\zeta}(U_{t+\Delta t})\right\}, \qquad (8)
\end{aligned}
$$

with $U_T = u(W_T)$.

Remark 1: Liu et al. [25] consider a general measure $\mathbf{E}_t^{\zeta}\left[h(\ln\frac{\zeta_{t+\Delta t}^{(k)}}{\zeta_t^{(k)}})\right]$, where $h(x) = x + \beta(e^x - 1)$ with $\beta > 0$. On contrary, we follow the "relative entropy" of Anderson, et al. [2] and Maehout [27] corresponding to the case of $h(x) = x$ in this paper. Liu et al. [25] introduce the "extended entropy" ($\beta \neq 0$) because they find that "the minimization problem ... does not have an interior global minimum for the relative entropy case." However, given $\gamma > 1$, we do find that an interior global minimum for the portfolio choice problem in their parameterized model in the relative entropy case (i.e. $\beta = 0$). In fact, our approach indicates that an interior minimum exists for our portfolio choice problem in the non-parameterized model, hence implying that an interior global minimum exists for the parameterized model as well, since our minimum is a lower bound of theirs.

programming equations with minimax problem can be found in Ruszczynski and Shapiro [30] where, unlike our model here, there is no penalty function. We thank an anonymous referee for pointing those out to us.

Remark 2: For $h(x) = x + \beta(e^x - 1)$ with $\beta > 0$, we will have an extra term in the integrand in Lemma A1 of Appendix A. The new one is:

$$\mathbf{E}_t^\zeta[\vartheta_k(s)\psi_k(s,z)\ln(\vartheta_k\psi_k(s,z)) + 1 - \vartheta_k(s)\psi_k(s,z) + \beta(1 - \vartheta_k(s)\psi_k(s,z))^2].$$

As a result, we are unable to get a closed-form solution for $\hat{x}_2$ given by (32) in Lemma 3 and $\hat{x}_2$ can be solved numerically. In the other words, our approach still works, however, unfortunately, we can only obtain the worst case density and intensity in an inexplicit form. To focus on our major purpose of illustrating our approach and applications by a closed-form solution, we shall not consider this case in this paper.

For analytic tractability, we first consider a constant relative risk aversion (CRRA) function of $u(x)$ as follows

$$u(x) = \begin{cases} \frac{x^{1-\gamma}}{1-\gamma}, & \forall x > 0, \\ -\infty, & \forall x \leq 0. \end{cases} \tag{9}$$

and extend to the more general HARA utility function in Section 5.

For practical relevance, we assume that the relative risk aversion coefficient γ is greater than one. Our approach is extended to the logarithm utility function in Appendix B.[5]

In the following Proposition 1, by letting Δt tend toward zero, we obtain the continuous-time version of the utility function with ambiguity aversion defined in (7), and the corresponding Merton's problem under ambiguity.

Proposition 1 *Under Assumption A in Appendix A, the continuous-time version of the utility with ambiguity aversion in equation (7) is given by*

$$U_t = \inf_\zeta \mathbf{E}_t^\zeta \left[e^{\int_t^T H_s ds} u(W_T) \right], \tag{10}$$

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

where

$$H_t = H(\zeta_t) = (1-\gamma)\sum_{k=1}^{n-d}\frac{\lambda_k}{\phi_k}\int_{A_k}[\vartheta_k(t)\psi_k(t,z)\ln(\vartheta_k(t)\psi_k(t,z))+1-\vartheta_k(t)\psi_k(t,z)]\Phi_k(dz),$$

with $H(\zeta_t) \leq 0$.

Furthermore, the corresponding Merton's portfolio choice problem under ambiguity and ambiguity aversion in continuous time is given by

$$J(t, W_t, X_t) = \sup_{\pi}\{U_t\} = \inf_{\zeta}\sup_{\pi}\mathbf{E}_t^{\zeta}\left[e^{\int_t^T H_s ds}u(W_T)\right]. \tag{11}$$

Proof. See Appendix A. ■

The form of the indirect value function $J(t, W_t, X_t)$ in (11) is a new and key result with an attractive feature in the present paper, though the result (10) is the same as (21) in Jin and Zhang [19]. The maximization problem in the "inf sup" problem, which is given by the second equality of (11), is an investment optimization problem under a new probability determined by the Radon-Nikodym derivative ζ_t, and it becomes much more tractable. Thus, the new expression makes it possible to use the duality method to evaluate the optimal expected utility function given by (10). In general, it is much more difficult to solve the original "sup inf" problem defined by the first equality of (11). As opposed to the ambiguity-neutral case where $H_t = 0$, the expected utility for an ambiguity-averse investor is damaged by the discount factor $\exp(\int_t^T H_s ds)$ since $H_t \leq 0$.

3 Main Results

In this section, we present our main result which provides a closed-form solution to the dynamic portfolio choice problem under ambiguous jumps. The proof is left in the next subsections.

As shown in Bardhan and Chao [4], once unpredictable jumps are included in the model, the market is inherently incomplete, regardless of whether $m \geq n$ or $m < n$, where m is the number of risky assets, and n is the total number of Brownian motions and jumps. In contrast, in a pure-diffusion economy, increasing the number of traded assets can always complete the market. In Theorem 1 below, we consider the case $m = n$, in which the number of risky assets is equal to the total number of diffusions and jumps. Our approach to solving the portfolio choice problem is especially powerful in this case.

For the case $m < n$, Jin and Zhang [19] adopt the "fictitious completing" approach developed by Cvitaniĉ and Karatzas [9] to show that solving the portfolio selection problem in the original market can be converted into solving one in a set of fictitious markets. In particular, the number of risky assets is equal to the sum of the diffusions and jumps, that is, $m = n$ in each fictitious market, and hence, the results developed in the present paper can be used to solve the optimal portfolio selection problem in each fictitious market. We follow this exact completion method in our calibration exercise in Section 6.

In a market with asset returns following the jump-diffusion processes characterized in the last section, Bardhan and Chao [4] point out that if $m > n$ and there are no arbitrage opportunities, $m-n$ assets in the market are redundant and can be removed accordingly. This case is similar to that of a pure-diffusion economy in terms of spanning of risks although our market remains incomplete. We can simply focus on n non-redundant assets for the portfolio choice problem.

For illustrative purposes only, we focus on the most widely used case in the literature: mixed jump size $A_k = (-1, \infty)$. Given any $k \in \{1, ..., n-d\}$, we define the set

$$\widetilde{A}_k = \left\{ c_k : 0 \leq c_k < 1 - \frac{1}{\gamma} \right\},$$

which is associated with the set of feasible k-th jump exposures and alternative k-th jump size distributions. We let $\mathbb{Q}_\zeta$ denote the set of martingale measures under the probability $P(\zeta)$. We will specify this set with more details in the next subsection. The following theorem is our main result which gives a closed-form solution to the indirect value function and the worst case probability.

Theorem 1 *Suppose $m = n$, that is, the number of risky assets is equal to the total number of diffusions and jumps. For the portfolio choice problem (11) in Proposition 1, we have the following duality result:*

$$J(t, W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma} \left(\sup_{\zeta} \sup_{\xi \in \mathbb{Q}_\zeta} \mathbf{E}_t^{\zeta} \left[e^{\frac{1}{\gamma} \int_t^T (H_s + (1-\gamma) r) ds} \xi_\delta(t,T)^{1-\frac{1}{\gamma}} \right] \right)^{\gamma}. \tag{12}$$

Moreover,

$$\begin{aligned} & \sup_{\zeta} \sup_{\xi \in \mathbb{Q}_\zeta} \mathbf{E}_t^{\zeta} \left[e^{\frac{1}{\gamma} \int_t^T (H_s + (1-\gamma) r) ds} \xi_\delta(t,T)^{1-\frac{1}{\gamma}} \right] \\ = \; & \mathbf{E}_t \left[\left(e^{-\int_t^T r ds} \xi_\delta^b(t,T) \right)^{1-\frac{1}{\gamma}} \exp\left(\int_t^T \sum_{k=1}^{n-d} \inf_{c_k \in \widetilde{A}_k} G_k(s, c_k) ds \right) \right] \\ \equiv \; & f(t, X_t), \end{aligned} \tag{13}$$

where ξ_δ and ξ_δ^b are defined in the next section, and

$$G_k(s, c_k) = -c_k \theta_k^q - \frac{\lambda_k(1-\gamma)}{\gamma \phi_k} \int_{A_k} [e^{\chi_k(z, c_k)} - 1] \Phi_k(s, dz), \tag{14}$$

with

$$\chi_k(z, c_k) = \frac{\phi_k}{(1-\gamma)} \left[1 - \left(1 - \frac{\gamma c_k}{1-\gamma} z \right)^{1-\gamma} \right].$$

In particular, the intensity of the k-th jump of the worst case is given by

$$\lambda_k^* = \lambda_k \int_{A_k} e^{\chi_k(z, c_k^*)} \Phi(t, dz)$$

and the density of the k-th jump size of the worst case is given by

$$\Phi_k^*(t, dz) = \frac{e^{\chi_k(z, c_k^*)} \Phi_k(t, dz)}{\int_{A_k} e^{\chi_k(z, c_k^*)} \Phi(t, dz)}, \tag{15}$$

where c_k^* is the optimal solution to the minimization problem in (13).

In (13), we have translated the original optimization problem over the stochastic processes ζ and ξ into a pathwise minimization problem. The former, as is well understood, is notoriously difficult to solve due to the infinitely many Radon-Nikodym derivatives ζ_t and martingale measures ξ_t and due to the lack of a closed-form solution for the expectation $\mathbf{E}_t^\zeta[\cdot]$. The latter is $n-d$ minimization problems over a subset in the one-dimensional real space $\mathcal{R}$ and is straightforward to solve. Meanwhile, the pathwise minimization problem is free of the curse of dimensionality caused by $n-d$, the number of jumps, and thus, it can lead to a significant reduction in the computation burden when $n-d$ is large. In short, $f(t, X_t)$ can be evaluated by the standard Monte Carlo method in combination with the pathwise minimization problem. As a result, the optimal portfolio can be derived through an HJB equation satisfied by the indirect value function $J(t, W_t, X_t)$ in (12), which in turn will be obtained in Section 4.

In particular, by letting $\phi_k \to 0$ in the function $G_k(t, c_k)$, the indirect value function $J(t, W_t, X_t)$ for the case without ambiguity can be obtained as

$$J(t, W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma} \left(\mathbf{E}_t \left[\left(e^{-\int_t^T r ds} \xi_\delta^b(t, T) \right)^{1-\frac{1}{\gamma}} \exp \left(\sum_{k=1}^{n-d} \int_t^T \int_{A_k} \inf_{c_k} g_k(z, c_k) \Phi_k(s, dz) ds \right) \right] \right)^\gamma,$$

where

$$g_k(z, c_k) = -c_k \theta_k^q + \frac{\lambda_k}{\gamma} \left[\left(1 - \frac{\gamma c_k}{1-\gamma} z \right)^{1-\gamma} - 1 \right].$$

We now turn to the interpretation of the worst case density $\Phi_k^*(t, dz)$. We can consider the function $\psi_k^* = \exp\{\chi_k(z, c_k^*)\}$ as a weighting function. Since $c_k^* \in \widetilde{A}_k$, that is, $c_k^* \geq 0$, we can show that ψ_k^* is a non-increasing function of jump size z. This result means that the ambiguity-averse investor pessimistically attaches more weight to more negative jumps and less weight to more positive jumps, implying a smaller expected jump size, more negatively skewed and less positively skewed jump size distribution in the worst case model relative to that in the reference model.

3.1 Proof of Theorem 1

To prove Theorem 1, we apply the duality method provided in Kramkov and Schachermayer ([21],[22]) and Schied and Wu [31] together with the minimax theorem in Borwein and Zhuang [6] and Proposition 1 given above.

In order to calculate $J(t, W_t, X_t) = \sup_\pi \{U_t\}$, we now lay out the necessary notations. As in Section 2, we use $P(\zeta)$ to denote the probability defined by the Radon-Nikodym derivative ζ given by (4) with $\vartheta_1(t), ..., \vartheta_{n-d}(t)$ and $\psi_1(t, z), ..., \psi_{n-d}(t, z)$. We use $\mathbf{E}^\zeta(\cdot)$ to denote the expectation under $P(\zeta)$. According to the discussion in the previous section, the jump intensities and the jump size distributions under $P(\zeta)$ are given by

$$\lambda_k^\zeta = \vartheta_k(t)\lambda_k,$$
$$\Phi_k^\zeta(t, dz) = \psi_k(t, z)\Phi_k(t, dz),$$

for $k = 1, ..., n-d$. We let $\mathbb{Q}_\zeta$ be the family of all densities of equivalent local martingale measures with respect to the probability $P(\zeta)$.

We now introduce more notations. Since the matrix $\Sigma = [\Sigma_b, \Sigma_q]$ is assumed to be invertible, we define

$$\begin{pmatrix} \theta^b \\ \theta^q \end{pmatrix} = \Sigma^{-1}(b - r\mathbf{1}_n), \tag{16}$$

where $\theta^b = (\theta_1^b, ..., \theta_d^b)^\top$ and $\theta^q = (\theta_1^q, ..., \theta_{n-d}^q)^\top$. We now introduce a characterization result of $\mathbb{Q}_\zeta$ developed in Bardhan and Chao [4]. Let Γ^{loc} denote the family of triples

$\delta = (v, \theta, \varphi)$, such that

$$\begin{aligned} v(t) &= (v_1(t), ..., v_d(t))^\top, \\ \theta(t) &= (\theta_1(t), ..., \theta_{n-d}(t))^\top, \\ \varphi(t) &= (\varphi_1(t,z), ..., \varphi_{n-d}(t,z))^\top, \end{aligned}$$

are predictable processes; θ and φ are strictly positive; φ satisfies

$$\int_{A_k} \varphi_k(t,z)\Phi_k^\zeta(dz) = 1, \tag{17}$$

for $t \in [0,T]$ and $k = 1, ..., n-d$, and the following equation holds:

$$\Sigma_b v(t) - \Sigma_q(\lambda^\zeta \bullet \theta(t) \bullet \widetilde{\alpha}) = b - r\mathbf{1}_m.$$

Or equivalently, by (16),

$$v(t) = \theta^b, \quad and \quad \lambda^\zeta \bullet \theta(t) \bullet \widetilde{\alpha} = -\theta^q, \tag{18}$$

where $\lambda^\zeta \bullet \theta(t) \bullet \widetilde{\alpha} := (\lambda_1^\zeta \theta_1(t)\widetilde{\alpha}_1, ..., \lambda_{n-d}^\zeta \theta_{n-d}(t)\widetilde{\alpha}_{n-d})^\top$ and

$$\widetilde{\alpha} = (\widetilde{\alpha}_1, ..., \widetilde{\alpha}_{n-d}), \quad \widetilde{\alpha}_k = \int_{A_k} z\varphi_k(t,z)\Phi_k^\zeta(t,dz),$$

for $t \geq 0$ and $k = 1, ..., n-d$. For each $\delta \in \Gamma^{loc}$, define the local martingale,

$$\xi_\delta(t) = \xi_\delta^b(t)\xi_\delta^q(t), \tag{19}$$

where

$$\xi_\delta^b(t) = \exp\left(- \int_0^t v^T(s)dB(s) - \frac{1}{2}\int_0^t ||v(s)||^2 ds \right),$$

$$\xi_\delta^q(t) = \left(\prod_{k=1}^{n-d}\prod_{i=1}^{N_k(t)}(\theta_k(t_i^k)\varphi_k(t_i^k, z_i^k))\right)\exp\left(\int_0^t\int_{A_k}(1-\theta_k(s)\varphi_k(s,z))\lambda_k^\zeta\Phi_k^\zeta(s,dz)ds\right).(20)$$

In particular, $\xi_\delta(t)$ is a supermartingale under $P(\zeta)$ for each $\delta \in \Gamma^{loc}$ since it is non-negative. We use Γ to denote the subset of Γ^{loc} for which $\xi_\delta(t)$ is a martingale.

The following lemma is one of the main results in Bardhan and Chao [4] and plays a key role in our paper.

Lemma 1 *A measure $Q \in \mathbb{Q}_\zeta$ if and only if there exists a triple $\delta \in \Gamma$, such that the Radon-Nikodym derivative $\frac{dQ}{dP} = \xi_\delta(t)$.*

Proof. See Bardhan and Chao [4]. ■

As characterized above, each probability in $\mathbb{Q}_\zeta$ can be represented by a Radon-Nikodym derivative $\xi_\delta(t) = \xi_\delta^b(t)\xi_\delta^q(t)$. Set $\xi_\delta(t,T) = \xi_\delta(T)/\xi_\delta(t)$ and $\xi_\delta^b(t,T) = \xi_\delta^b(T)/\xi_\delta^b(t)$.

Equipped with Proposition 1 and Lemma 1, we are able to solve the portfolio problem by using the duality method developed in Kramkov and Schachermayer ([21], [22]) and Schied and Wu [31]. First we define the convex conjugate of $u(x)$:

$$V(y) = (-u)^*(-y) = \sup_{x>0}(u(x) - xy) = \frac{\gamma}{1-\gamma}y^{1-\frac{1}{\gamma}}, \tag{21}$$

where $(-u)^*(y)$ is the convex conjugate of $-u(x)$. Note that

$$\mathbf{E}_t^\zeta\left[\exp\left(\int_t^T H_s ds\right)\frac{W_T^{1-\gamma}}{1-\gamma}\right] = D_t\mathbf{E}_t^\zeta\left[\eta(t,T)\frac{W_T^{1-\gamma}}{1-\gamma}\right],$$

where $D_t = \mathbf{E}_t^\zeta\left[\exp\left(\int_t^T H_s ds\right)\right]$ and $\eta(t,T) = \frac{\exp\left(\int_t^T H_s ds\right)}{D_t}$. According to Schied and Wu [31],

$$J(t, W_t, X_t) = \inf_\zeta D_t \inf_{y>0}(v(y) + W_t y), \tag{22}$$

where

$$v(y) = \inf_{\xi\in\mathbb{Q}_\zeta}\mathbf{E}_t^\zeta\left[\eta(t,T)V\left(y\frac{\xi_\delta(t,T)\exp\left(-\int_t^T rds\right)}{\eta(t,T)}\right)\right].$$

Thus, by using (21) and by noticing $\gamma > 1$,

$$v(y) = \frac{\gamma}{1-\gamma} y^{1-\frac{1}{\gamma}} \sup_{\xi \in \mathbb{Q}_\zeta} \mathbf{E}_t^\zeta \left[\exp\left(-\left(1 - \frac{1}{\gamma}\right) \int_t^T r ds \right) \eta(t,T)^{\frac{1}{\gamma}} \xi_\delta(t,T)^{1-\frac{1}{\gamma}} \right],$$

and consequently, by solving the minimization problem $\inf_{y>0}$ in (22) using the definition of $\eta(t,T)$,

$$J(t, W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma} \left(\sup_\zeta \sup_{\xi \in \mathbb{Q}_\zeta} \mathbf{E}_t^\zeta \left[\exp\left(\frac{1}{\gamma} \int_t^T (H_s + (1-\gamma) r) ds \right) \xi_\delta(t,T)^{1-\frac{1}{\gamma}} \right] \right)^\gamma. \tag{23}$$

This result completes the proof of (12).

Given the result (23), we now turn to the proof of (13) of Theorem 1. The proof is broken into several lemmas that are organized into two subsections. The key step is to show that the maximization problem "$\sup_\zeta \sup_{\xi \in \mathbb{Q}_\zeta}$" and the expectation "$\mathbf{E}$" in (23) can be exchangeable, leading to the pathwise optimization problem in (13). The proof of the exchangeability is presented in Section 3.3. The Fenchel Duality Theorem plays an important role in the proofs below. For more details about this theorem and relevant notation, see Chapter 7 of Luenberger [26].

In the following subsection, we provide several auxiliary lemmas for proving (13). The key result is Lemma 4 which is used directly in the subsection 3.2. Readers may skip Section 3.2 and read Section 3.3 first. To help readers better understand the main idea of the proof, we present the following result proved in Section 3.3 to change the objective function of the optimization problem in (23). The result is:

$$\begin{aligned} & \mathbf{E}^\zeta \left[\exp\left(\frac{1}{\gamma} \int_0^T (H_s + (1-\gamma) r) ds \right) \xi_\delta(0,T)^{1-\frac{1}{\gamma}} \right] \qquad (24) \\ = \; & \mathbf{E} \left[e^{-\left(1-\frac{1}{\gamma}\right) \int_0^T r ds} \left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}} \exp\left(\sum_{k=1}^{n-d} \lambda_k \int_0^T \int_{A_k} f_k(z, \theta_k(t), \varphi_k(z), \vartheta_k(t), \psi_k(z)) \Phi_k(dz) dt \right) \right] \end{aligned}$$

where, by letting $x_1(z) = \theta_k(t)\varphi_k(z)\vartheta_k(t)\psi_k(z)$ and $x_2(z) = \vartheta_k(t)\psi_k(z)$,

$$
\begin{aligned}
& f_k(z, \theta_k(t), \varphi_k(z), \vartheta_k(t), \psi_k(z)) \qquad (25) \\
= \; & (x_1(z))^{1-\frac{1}{\gamma}} x_2(z)^{\frac{1}{\gamma}} - \left(1 - \frac{1}{\gamma}\right) x_1(z) - \frac{1}{\gamma} x_2(z) + \frac{1}{\gamma} h(x_2(z))
\end{aligned}
$$

subject to

$$
\begin{aligned}
\int_{A_k} \psi_k(z)\Phi_k(dz) &= 1, \\
\int_{A_k} \varphi_k(z)\psi_k(z)\Phi_k(dz) &= 1, \\
\int_{A_k} \theta_k(t)\varphi_k(z)\vartheta_k(t)\psi_k(z) z \Phi_k(dz) &= -\frac{\theta_k^q}{\lambda_k},
\end{aligned}
$$

for $k \in \{1, ..., n-d\}$. The function h is given by (26) at the beginning of next section. The proof of Theorem 1 will be based on an optimization problem with the objective function f_k. As a result, the optimization problem in (23) with respect to two stochastic processes ζ_t and ξ_t is converted into a set of pathwise optimization problems, which significantly relieves the computation burden for solving the indirect value function $J(t, W_t, X_t)$.

3.2 Auxiliary results for the proof of (13)

We fix $k \in \{1, ..., n-d\}$ and define

$$
h(x) = \frac{(1-\gamma)}{\phi_k}[x\ln(x) + 1 - x], \; x > 0, \qquad (26)
$$

which is the integrand in the function H_t given in Proposition 1. Now we apply the Fenchel Duality Theorem to solve the following optimization problem:

$$
\sup_{\substack{x=(x_1,x_2)\in\mathcal{X} \\ x\geq 0}} \int_{A_k} \left[x_1^{1-\frac{1}{\gamma}}(z) x_2^{\frac{1}{\gamma}}(z) - \left(1-\frac{1}{\gamma}\right) x_1(z) - \frac{1}{\gamma} x_2(z) + \frac{1}{\gamma} h(x_2(z)) \right] \Phi_k(dz), \qquad (27)
$$

subject to the constraint

$$\int_{A_k} x_1(z) z \Phi_k(dz) = -\frac{\theta_k^q}{\lambda_k}. \tag{28}$$

As will be clear in next section, the problem (13) reduces to the problem (27). The constraint in (28) is obtained from the second equation of (18). It will be illustrated in next section that the optimization problem (27) with the constraint (28) corresponds to the optimization problem with the objective function (25) given in the last section. Define

$$\overline{\Phi}_k(z) = \int_{-\infty}^{z} |s| \Phi_k(ds).$$

Define a linear normal space $\mathcal{X}$ of functions as follows:

$$\mathcal{X} = \left\{ x(z) = (x_1(z), x_2(z)) : \int_{A_k} |x_1(z)| \overline{\Phi}_k(dz) + \int_{A_k} |x_2(z)| \Phi_k(dz) < \infty \right\},$$

with norm

$$||x|| = \int_{A_k} |x_1(z)| \overline{\Phi}_k(dz) + \int_{A_k} |x_2(z)| \Phi_k(dz).$$

Then, the dual space $\mathcal{X}^*$ of $\mathcal{X}$ is

$$\mathcal{X}^* = \{x^*(z) = (x_1^*(z), x_2^*(z)) : x_1^*(z) \in L^\infty(\overline{\Phi}_k), x_2^*(z) \in L^\infty(\Phi_k)\}.$$

Define a concave function: For $x = (x_1, x_2)$, let

$$g_0(x) = \begin{cases} x_1^{1-\frac{1}{\gamma}} x_2^{\frac{1}{\gamma}} - \left(1 - \frac{1}{\gamma}\right) x_1 - \frac{1}{\gamma} x_2 + \frac{1}{\gamma} h(x_2), & \forall x_1, x_2 \geq 0, \\ -\infty, & otherwise. \end{cases} \tag{29}$$

Then (27) is equivalent to the following problem:

$$\sup_{x \in \mathcal{X}} \int_{A_k} g_0(x(z)) \Phi_k(dz),$$

subject to

$$\int_{A_k} x_1(z) Sgn(z) \overline{\Phi}_k(dz) = -\frac{\theta_k^q}{\lambda_k},$$

where

$$Sgn(z) = \begin{cases} -1, & \forall z < 0, \\ 0, & \forall z = 0, \\ 1, & \forall z > 0. \end{cases}$$

To employ the Fenchel Duality Theorem to solve the above problem, we lay out relevant notations below. Set

$$\begin{aligned} C &= \left\{ x \in \mathcal{X} : \int_{A_k} x_1(z) Sgn(z) \overline{\Phi}_k(dz) = -\frac{\theta_k^q}{\lambda_k} \right\}, D = \mathcal{X}, \\ f(x) &= \begin{cases} 0, \text{ if } x \in C \\ \infty, \text{ else} \end{cases}, \quad g(x) = \int_{A_k} g_0(x(z)) \Phi_k(dz). \end{aligned} \tag{30}$$

We first calculate the functional f^* conjugate to f, given by

$$f^*(x^*) = \sup_{x \in C} [\langle x, x^* \rangle - f(x)] = \sup_{x \in C} \left(\int_{A_k} x_1(z) x_1^*(z) \overline{\Phi}_k(dz) + \int_{A_k} x_2(z) x_2^*(z) \Phi_k(dz) \right),$$

where

$$\langle x, x^* \rangle = \int_{A_k} x_1(z) x_1^*(z) \overline{\Phi}_k(dz) + \int_{A_k} x_2(z) x_2^*(z) \Phi_k(dz), \quad x \in \mathcal{X} \;\; and \; x^* \in \mathcal{X}^*.$$

Lemma 2 *The conjugate space C^* of $f^*(x^*)$ is given by*

$$C^* = \{ x^* : f^*(x^*) < \infty \} = \{ (cSgn(z), 0) : z \in A_k, c \in \mathcal{R} \},$$

and

$$f^*(x^*) = \sup_{x \in C}[\langle x, x^* \rangle] = -\frac{c\theta_k^q}{\lambda_k}, \quad for \quad x^* = (cSgn(z), 0) \in C^*.$$

Proof. Define a linear functional on $\mathcal{X}$ as for any $x \in \mathcal{X}$,

$$f_1(x) = \int_{A_k} x_1(z) Sgn(z) \overline{\Phi}_k(dz),$$

and its zero space is given by

$$Ker(f_1) = \{x \in \mathcal{X} : f_1(x) = 0\}.$$

Note that for any $x^{(1)} \in Ker(f_1)$, $x^{(2)} \in C$ and integer N, $Nx^{(1)} + x^{(2)} \in C$. Thus, we must have $\langle x^{(1)}, x^* \rangle = 0$ in order that

$$f^*(x^*) = \sup_{x \in C}[\langle x, x^* \rangle] < \infty.$$

By Lemma 1 on Page 188 in Luenberger [26], there exists a constant c, such that $\langle x, x^* \rangle = cf_1(x)$ for any $x \in C$. That is,

$$\int_{A_k} x_1(z) x_1^*(z) \overline{\Phi}_k(dz) + \int_{A_k} x_2(z) x_2^*(z) \Phi_k(dz) = \int_{A_k} cx_1(z) Sgn(z) \overline{\Phi}_k(dz),$$

implying $x_1^*(z) = cSgn(z)$, $x_2^*(z) = 0$ and $C^* = \{(cSgn(z), 0) : c \in \mathcal{R}\}$. Moreover, by the definition of set C,

$$f^*(x^*) = \sup_{x \in C}[\langle x, x^* \rangle] = \sup_{x \in C} \left(\int_{A_k} cx_1(z) Sgn(z) \overline{\Phi}_k(dz) \right) = -\frac{c\theta_k^q}{\lambda_k},$$

completing the proof. ■

We now turn to the calculation of the concave conjugate functional g^* of g. According to the definition, for $x^* \in \mathcal{X}^*$,

$$\begin{aligned} g^*(x^*) &= \inf_{x \in D} [\langle x, x^* \rangle - g(x)] \\ &= \inf_{x \in D} \int_{A_k} [x_1(z)x_1^*(z)|z| + x_2(z)x_2^*(z) - g_0(x(z))]\Phi_k(dz). \end{aligned}$$

The *conjugate space* of $g^*(x^*)$ is $D^* = \{x^* : g^*(x^*) > -\infty\}$. When using the Fenchel Duality Theorem, we only need to calculate $g^*(x^*)$ for $x^* \in C^*$ since, as shown in Lemma 4 below, the infimum problem in the Fenchel Duality Theorem is taken over the set $C^* \cap D^*$ with objective function $f^*(x^*) - g^*(x^*)$. To this end, we have the following result.

Lemma 3 *For* $x^* = (cSgn(z), 0) \in C^* \cap D^*$,

$$\begin{aligned} g^*(x^*) &= \inf_{x \in \mathcal{X}} \int_{A_k} [x_1(z)x_1^*(z)|z| + x_2(z)x_2^*(z) - g_0(x(z))]\Phi_k(dz) \\ &= \int_{A_k} \inf_{x \in \mathcal{R}^2} [x_1 cz - g_0(x)]\Phi_k(dz). \qquad (31) \\ &= \int_{A_k} \frac{1-\gamma}{\gamma\phi_k} [\exp\{\chi_k(z,c)\} - 1]\Phi_k(dz), \end{aligned}$$

where

$$\chi_k(z,c) = \frac{\phi_k}{1-\gamma}\left[1 - \left(\frac{\gamma c}{\gamma - 1} z + 1\right)^{1-\gamma}\right].$$

Furthermore,

$$C^* \cap D^* = \left\{(cSgn(z), 0) : c < 1 - \frac{1}{\gamma}\right\},$$

Proof. The inequality $\geq$ in (31) is trivial, namely,

$$\begin{aligned} &\inf_{x \in \mathcal{X}} \int_{A_k} [x_1(z)x_1^*(z)|z| + x_2(z)x_2^*(z) - g_0(x(z))]\Phi_k(dz) \\ \geq &\int_{A_k} \inf_{x=(x_1,x_2) \in R^2} [cx_1 z - g_0(x)]\Phi_k(dz). \end{aligned}$$

We now prove $\leq$. We solve the optimization problem

$$\inf_{x \in R^2} [c x_1 z - g_0(x)],$$

where g_0 is defined in (29). It is easy to obtain the optimal solution as

$$\begin{aligned} \widehat{x}_1 &= \left(\frac{\gamma c}{\gamma - 1} z + 1 \right)^{-\gamma} \widehat{x}_2, \qquad (32) \\ \widehat{x}_2 &= \exp \left\{ \frac{\phi_k}{1-\gamma} \left[1 - \left(\frac{\gamma c}{\gamma - 1} z + 1 \right)^{1-\gamma} \right] \right\} \equiv \exp \left\{ \chi_k(z, c) \right\}. \end{aligned}$$

and the corresponding optimal objective function is

$$\inf_{x \in R^2} [c x_1 z - g_0(x)] = \frac{1-\gamma}{\gamma} \frac{1}{\phi_k} [\exp \left\{ \chi_k(z, c) \right\} - 1].$$

And furthermore, on the one hand, $(\widehat{x}_1, \widehat{x}_2) \in \mathcal{X}$ for $0 \leq c < 1 - \frac{1}{\gamma}$, implying $\leq$ since $A_k = (-1, \infty)$. On the other hand,

$$\begin{aligned} C^* \cap D^* &= \left\{ (cSgn(z), 0) : \int_{A_k} \frac{1-\gamma}{\gamma} \frac{1}{\phi_k} [\exp \left\{ \chi_k(z, c) \right\} - 1] \Phi_k(dz) > -\infty \right\} \\ &= \left\{ (cSgn(z), 0) : 0 \leq c < 1 - \frac{1}{\gamma} \right\}. \end{aligned}$$

■

Without causing any confusion, we set

$$C^* \cap D^* = \left\{ (c, 0) : 0 \leq c < 1 - \frac{1}{\gamma} \right\}. \qquad (33)$$

Consequently, by the Fenchel Duality Theorem, we can establish the following result.

Lemma 4

$$
\begin{aligned}
&\sup_{\substack{x=(x_1,x_2)\in C\\ x\geq 0}} \int_{A_k} \left[x_1^{1-\frac{1}{\gamma}}(z) x_2^{\frac{1}{\gamma}}(z) - \left(1-\frac{1}{\gamma}\right) x_1(z) - \frac{1}{\gamma} x_2(z) + \frac{1}{\gamma} h(x_2(z)) \right] \Phi_k(dz) \\
= \ & \inf_{c\in C^*\cap D^*} \left[-\frac{c\theta_k^q}{\lambda_k} - \frac{1-\gamma}{\gamma}\frac{1}{\phi_k} \int_{A_k} [\exp\{\chi(z,c)\} - 1]\Phi_k(dz) \right].
\end{aligned}
$$

Proof. By using the definition of functions $f(x)$ and $g(x)$ in (30),we obtain

$$
\begin{aligned}
&\sup_{\substack{x=(x_1,x_2)\in C\\ x\geq 0}} \int_{A_k} \left[x_1^{1-\frac{1}{\gamma}}(z) x_2^{\frac{1}{\gamma}}(z) - \left(1-\frac{1}{\gamma}\right) x_1(z) - \frac{1}{\gamma} x_2(z) + \frac{1}{\gamma} h(x_2(z)) \right] \Phi_k(dz) \\
= \ & \sup_{x\in C} [g(x) - f(x)] \\
= \ & \inf_{c\in C^*\cap D^*} [f^*(x^*) - g^*(x^*)] \\
= \ & \inf_{c\in C^*\cap D^*} \left[-\frac{c\theta_k^q}{\lambda_k} - \frac{1-\gamma}{\gamma}\frac{1}{\phi_k} \int_{A_k} [\exp\{\chi_k(z,c)\} - 1]\Phi_k(dz) \right],
\end{aligned}
$$

with the second equality following the Fenchel Duality Theorem and completing the proof. ■

Remark: If $(c^*, 0)$ is the optimal solution to the right hand side of the equality in Lemma 4, the plugging the value into (32), we obtain the optimal solution to the left hand side of the equality, or the problem (22).

3.3 Proof of (13)

The aim of this section is to establish the exchangeability between the maximization problem "$\sup_\zeta \sup_{\xi\in\mathbb{Q}_\zeta}$" and the expectation "$\mathbf{E}$" in (23) and the result (24). We then apply the results in the last section to prove (13). For simplicity, we let $t = 0$ in the proof. In this case, as mentioned in Bardhan and Chao [4], the set Γ comprises the vectors $\delta = (v, \theta, \varphi, \vartheta, \psi)$, with θ, φ, ϑ, and ψ being strictly positive, satisfying (18) or equivalently,

$$
\begin{aligned}
v(t) &= \theta^b, \\
\int_{A_k} \theta_k(t)\varphi_k(t,z)\vartheta_k(t)\psi_k(t,z) z \Phi_k(dz) &= -\frac{\theta_k^q}{\lambda_k}, \\
\int_{A_k} \varphi_k(t,z)\psi_k(t,z)\Phi_k(dz) &= 1, \\
\int_{A_k} \psi_k(t,z)\Phi_k(dz) &= 1,
\end{aligned} \tag{34}
$$

for $t \geq 0$ and $k = 1, ..., n-d$. We let Γ^d denote the family of the vectors $\delta = \{(v, \theta, \varphi, \vartheta, \psi) = (v(t), \theta(t), \varphi(t,z), \vartheta_k(t), \psi(t,z))\}_{t\in[0,T]}$ satisfying four conditions in (34) and $(\theta(t), \varphi(t,z), \vartheta_k(t), \psi(t,z))$ solving the following optimization problem mentioned at the end of Section 3.1:

$$
\begin{aligned}
&\sup_{\delta\in\Gamma^d} \int_{A_k} f_k(z, \theta_k(t), \varphi_k(z), \vartheta_k(t), \psi_k(z))\Phi_k(dz) \qquad (35) \\
= &\int_{A_k} \left[(x_1(z))^{1-\frac{1}{\gamma}} x_2(z)^{\frac{1}{\gamma}} - \left(1 - \frac{1}{\gamma}\right) x_1(z) - \frac{1}{\gamma} x_2(z) + \frac{1}{\gamma} h(x_2(z)) \right] \Phi_k(dz),
\end{aligned}
$$

for $t \geq 0$ and $k \in \{1, ..., n-d\}$. From the above objective function and the second constraint in (34), the two variables $\vartheta_k(t)$ and $\psi_k(z)$ are not separable and the same holds true for the two variables $\theta_k(t)$ and $\varphi_k(z)$. For this reason, by letting $x_1(z) = \theta_k(t)\varphi_k(z)\vartheta_k(t)\psi_k(z)$ and $x_2(z) = \vartheta_k(t)\psi_k(z)$, it is straightforward to conclude that the above optimization problem with the second constraint only is equivalent to the optimization problem (27) with the constraint (28), which has been solved in the preceding subsection. After obtaining the optimal $x_1^*(z)$ and $x_2^*(z)$, we can recover the optimal $\theta^*, \varphi^*, \vartheta^*, \psi^*$ through normalization to give the solution to the above optimization problem (35). In particular, φ^* and ψ^* satisfy the third and the fourth constraints in (34). The details are presented below.

We prove the following lemma.

Lemma 5

$$\sup_{\zeta}\sup_{\xi\in\mathbb{Q}_\zeta}\mathbf{E}^{\zeta}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T rds}\eta^{\frac{1}{\gamma}}\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right]=\sup_{\delta\in\Gamma^d}\mathbf{E}\left[\zeta(T)e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T rds}\eta^{\frac{1}{\gamma}}\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right],$$

where $\eta=\exp\left(\int_0^T H(\zeta_s)ds\right)$. *Significantly, from (35), the right hand side is a pathwise optimization problem.*

Proof. From the definition of Γ^d and Lemma 1 in Section 3.1, we can see that given $\delta=\{(v,\theta,\varphi,\vartheta,\psi)\in\Gamma^d$, the corresponding ζ defined by (6) and ξ_δ given by (19) satisfy conditions in Section 2 and Section 3.1. Thus, by noticing that $\mathbf{E}^{\zeta}(Y)=\mathbf{E}(\zeta Y)$, it suffices to prove

$$\sup_{\zeta}\sup_{\xi\in\mathbb{Q}_\zeta}\mathbf{E}\left[\zeta(T)e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T rds}\eta^{\frac{1}{\gamma}}\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right]\le\sup_{\delta\in\Gamma^d}\mathbf{E}\left[\zeta(T)e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T rds}\eta^{\frac{1}{\gamma}}\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right].\tag{36}$$

Let $N_k(t,T)=N_k(T)-N_k(t)$ denote the number of k-th type of jump in the interval $(t,T]$. Note that for any $t\in[0,T]$,

$$\xi_\delta(T)^{1-\frac{1}{\gamma}}=\left(\xi_\delta(t)\right)^{1-\frac{1}{\gamma}}\left(\xi_\delta(t,T)\right)^{1-\frac{1}{\gamma}},$$

where

$$\begin{aligned}\xi_\delta(t,T) &= \exp\left(-\int_t^T v^T(s)dz(s)-\frac{1}{2}\int_t^T||v(s)||^2ds\right)\times\prod_{k=1}^{n-d}\prod_{i=1}^{N_k(t,T)}(\theta_k(t_i^k)\varphi_k(t_i^k,z_i^k))\\ &\quad\times\exp\left(\int_t^T\int_{A_k}(1-\theta_k(s)\varphi_k(s,z))\lambda_k\vartheta_k(t)\psi_k(t,z)\Phi_k(dz)ds\right).\end{aligned}$$

Note that $\zeta(T)$ can be decomposed in the same way. Hence the optimal $v^*(t)$, $\theta_k^*(t)$, $\varphi_k^*(s,z)$, $\vartheta_k(t)^*$ and $\psi_k^*(t,z)$ only depend on the state variables X_t. Thus, if we let Γ_X denote the family of δ with $v(t)$, $\theta_k(t)$, $\varphi_k(t,z)$, $\vartheta_k(t)$ and $\psi_k(s,z)$ only depending on the state variables X_t, then

$$\sup_{\zeta}\sup_{\xi\in\mathbb{Q}_\zeta}\mathbf{E}\left[\zeta(T)e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T rds}\eta^{\frac{1}{\gamma}}\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right]=\sup_{\delta\in\Gamma_X}\mathbf{E}\left[\zeta(T)e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T rds}\eta^{\frac{1}{\gamma}}\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right].$$

Hence, to prove (36), it suffices to show the following result:

$$\sup_{\delta\in\Gamma_X}\mathbf{E}\left[\zeta(T)e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T rds}\eta^{\frac{1}{\gamma}}\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right]\leq\sup_{\delta\in\Gamma^d}\mathbf{E}\left[\zeta(T)e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T rds}\eta^{\frac{1}{\gamma}}\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right]. \tag{37}$$

Note that, by (6), (20) and (19), we have

$$\eta^{\frac{1}{\gamma}}\zeta(T)\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}=\left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}}\widetilde{\xi}_\delta(T)\exp\left(\sum_{k=1}^{n-d}\lambda_k\int_0^T\int_{A_k}f_k(z,\theta_k(t),\varphi_k(z),\vartheta_k(t),\psi_k(z))\Phi_k(dz)dt\right),$$

where $f_k(z,\theta_k(t),\varphi_k(z),\vartheta_k(t),\psi_k(z))$ is defined in (35) and

$$\begin{aligned}\widetilde{\xi}_\delta(t) &= \prod_{k=1}^{n-d}\prod_{i=1}^{N_k(t)}(\theta_k(t_i^k)\varphi_k(t_i^k,z_i^k)\vartheta_k(t)\psi_k(t_i^k,z_i^k))^{1-\frac{1}{\gamma}}(\vartheta_k(t)\psi_k(t_i^k,z_i^k))^{\frac{1}{\gamma}}\\ &\quad\times\exp\left(\int_0^t\int_{A_k}(1-(\theta_k(s)\varphi_k(s,z)\vartheta_k(t)\psi_k(s,z))^{1-\frac{1}{\gamma}}(\vartheta_k(t)\psi_k(s,z))^{\frac{1}{\gamma}})\lambda_k\Phi_k(dz)ds\right).\end{aligned}$$

Furthermore, $\widetilde{\xi}_\delta(t)$ can be rewritten as

$$\widetilde{\xi}_\delta(t)=\left(\prod_{k=1}^{n-d}\prod_{i=1}^{N_k(t)}\widetilde{\theta}_k(t_i^k)\widetilde{\varphi}_k(t_i^k,z_i^k)\right)\exp\left(\int_0^t\int_{A_k}(1-\widetilde{\theta}_k(s)\widetilde{\varphi}_k(s,z))\lambda_k\Phi_k(dz)ds\right),$$

where

$$\begin{aligned}\widetilde{\theta}_k(s) &= (\theta_k(s))^{1-\frac{1}{\gamma}}\vartheta_k(t)\int_{E_k}(\varphi_k(s,z))^{1-\frac{1}{\gamma}}\psi_k(s,z)\Phi_k(dz),\\ \widetilde{\varphi}_k(s,z) &= \frac{(\varphi_k(s,z))^{1-\frac{1}{\gamma}}\psi_k(s,z)}{\int_{A_k}(\varphi_k(s,z))^{1-\frac{1}{\gamma}}\psi_k(s,z)\Phi_k(dz)}.\end{aligned} \tag{38}$$

Then

$$\int_{A_k} \widetilde{\varphi}_k(s,z)\Phi_k(dz) = 1.$$

And thus, for $\delta \in \Gamma_X^{loc}$, $\widetilde{\xi}_\delta(t)$ is a non-negative local martingale from C4 in Bremaud [8] and hence a supermartingale. And moreover, noticing that the state variables X_t do not include jumps, we have

$$\mathbf{E}\left[\widetilde{\xi}_\delta(T)|\mathcal{F}_T^X\right] \leq \mathbf{E}\left[\widetilde{\xi}_\delta(0)|\mathcal{F}_T^X\right] = 1, \tag{39}$$

where $\mathcal{F}_T^X$ is the σ-algebra generated by $\{X_t, 0 \leq t \leq T\}$. Hence, by (39), for $\delta \in \Gamma_X$,

$$\begin{aligned}
& \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds}\eta^{\frac{1}{\gamma}}\zeta(T)\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right] \\
= \; & \mathbf{E}\left[E\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds}\eta^{\frac{1}{\gamma}}\zeta(T)\left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}|\mathcal{F}_T^X\right]\right] \\
= \; & \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds}\left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}}\mathbf{E}\left[\widetilde{\xi}_\delta(T)|\mathcal{F}_T^X\right] e^{\left(\sum_{k=1}^{n-d}\lambda_k\int_0^T\int_{A_k} f_k(z,\theta_k(t),\varphi_k(z),\vartheta_k(t),\psi_k(z))\Phi_k(dz)dt\right)}\right] \\
\leq \; & \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds}\left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}} e^{\left(\sum_{k=1}^{n-d}\lambda_k\int_0^T\int_{A_k} f_k(z,\theta_k(t),\varphi_k(z),\vartheta_k(t),\psi_k(z))\Phi_k(dz)dt\right)}\right].
\end{aligned} \tag{40}$$

Let $\theta_k^*(s), \varphi_k^*(s,z), \vartheta_k^*(s)$ and $\psi_k^*(s,z), k = 1, ..n-d$, denote the optimal solution to the problem (35). By (32) in the proof of Lemma 3,

$$\vartheta_k^*(s)\psi_k^*(s,z) = \exp\left\{\chi_k(z,c_k^*)\right\},$$

$$\theta_k^*(s)\varphi_k^*(s,z)\vartheta_k^*(t)\psi_k^*(s,z) = \left(\frac{\gamma c_k^*}{\gamma-1}z+1\right)^{-\gamma}\exp\left\{\chi_k(z,c_k^*)\right\},$$

implying

$$\vartheta_k^*(s) = \int_{A_k}\exp\left\{\chi_k(z,c_k^*)\right\}\Phi_k(s,dz), \tag{41}$$

$$\psi_k^*(s,z) = \frac{\exp\left\{\chi_k(z,c_k^*)\right\}}{\int_{A_k}\exp\left\{\chi_k(z,c_k^*)\right\}\Phi_k(s,dz)}, \tag{42}$$

$$\theta_k^*(s) = \int_{A_k} \left(\frac{\gamma c_k^*}{\gamma - 1} z + 1\right)^{-\gamma} \Phi_k(dz), \quad \varphi_k^*(s,z) = \frac{\left(\frac{\gamma c_k^*}{\gamma-1} z + 1\right)^{-\gamma}}{\int_{A_k} \left(\frac{\gamma c_k^*}{\gamma-1} z + 1\right)^{-\gamma} \Phi_k(dz)}.$$

And furthermore, by Theorems T10 and T11 of Chapter VIII in Bremaud [8], $\widetilde{\xi}_{\delta^*}(t)$ is a martingale, implying

$$\mathbf{E}\left[\widetilde{\xi}_{\delta^*}(T)|\mathcal{F}_T^X\right] = 1.$$

Therefore, by (40),

$$\begin{aligned}
& \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds} \eta^{\frac{1}{\gamma}} \zeta(T) \left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right] \\
\leq\ & \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds} \left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}} E[\widetilde{\xi}_{\delta^*}(T)|\mathcal{F}_T^X] e^{\left(\sum\limits_{k=1}^{n-d} \lambda_k \int_0^T \int_{A_k} f_k(z,\theta_k^*(t),\varphi_k^*(z),\vartheta_k^*(t),\psi_k^*(z))\Phi_k(dz)dt\right)}\right] \\
=\ & \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds} \left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}} \widetilde{\xi}_{\delta^*}(T) e^{\left(\sum\limits_{k=1}^{n-d} \lambda_k \int_0^T \int_{A_k} f_k(z,\theta_k^*(t),\varphi_k^*(z),\vartheta_k^*(t),\psi_k^*(z))\Phi_k(dz)dt\right)}\right] \\
=\ & \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds} \eta^{\frac{1}{\gamma}} \zeta(T) \left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}} \left(\xi_{\delta^*}^q(T)\right)^{1-\frac{1}{\gamma}}\right] \qquad (43) \\
=\ & \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds} \eta^{\frac{1}{\gamma}} \zeta(T) \left(\xi_{\delta^*}(T)\right)^{1-\frac{1}{\gamma}}\right],
\end{aligned}$$

for each $\delta \in \Gamma_X$. Here we have used the fact that $\xi_\delta^b(T) = \xi_{\delta^*}^b(T)$ since $v(t) = \theta^b$ by (34). Hence (37) is proved and this completes the proof of the lemma. ■

Note by virtue of Lemma 5 (see (40)), we have

$$\begin{aligned}
& \sup_\zeta \sup_{\xi \in \mathbb{Q}_\zeta} \mathbf{E}^\zeta\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds} \eta^{\frac{1}{\gamma}} \left(\xi_\delta(T)\right)^{1-\frac{1}{\gamma}}\right] \\
=\ & \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds} \left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}} \exp\left(\sum_{k=1}^{n-d} \lambda_k \int_0^T \sup_{\delta \in \Gamma^d} \int_{A_k} f_k(z, \theta_k(t), \varphi_k(z), \vartheta_k(t), \psi_k(z))\Phi_k(dz)dt\right)\right] \\
=\ & \mathbf{E}\left[e^{-\left(1-\frac{1}{\gamma}\right)\int_0^T r ds} \left(\xi_\delta^b(T)\right)^{1-\frac{1}{\gamma}} \exp\left(\int_0^T \sum_{k=1}^{n-d} \inf_{c_k \in \tilde{A}_k} G_k(s, c_k) ds\right)\right],
\end{aligned}$$

where the last equality is by Lemma 4 and G_k is defined as that in Theorem 1.

At last, the worst case intensity and density are obtained by (41) and (42). This completes the proof of Theorem 1.

4 Optimal Investment Strategy in the Worst Case

By Theorem 1, we obtain the indirect utility function, the jump distribution, and intensity in the worst case by the duality approach. Substituting them into the corresponding Hamilton-Jacobi-Bellman equation, we can directly find the optimal portfolio under ambiguity as follows. However, it is also possible to employ the classic martingale and duality approach to find the optimal portfolio, instead of solving the HJB equation which relies on the dynamic programming principle. For simplicity, we take the HJB equation approach in this section. The following method is based on an orthogonal decomposition technique, which is different from the decomposition approach in Jin and Zhang [19], or Aït-Sahalia et al. [1].

Note that $rank(\Sigma) = rank([\Sigma_b, \Sigma_q]) = m$ and we consider the case $m = n$. Then Σ is an invertible square matrix. Denote $\hat{\Sigma}_q = \Sigma^{-1}\Sigma_q$ and $\hat{\Sigma}_b = \Sigma^{-1}\Sigma_b$. We have two orthogonal sub-spaces of $\mathcal{R}^n$, generated by the columns of $\hat{\Sigma}_q$ and $\hat{\Sigma}_b$. Decomposed into the two orthogonal spaces, the optimal portfolio π^* can be written as

$$\pi^* = (\Sigma^{-1})^\top(\hat{\Sigma}_q\bar{\pi}^* + \hat{\Sigma}_b\pi^*_\perp), \tag{44}$$

where $\bar{\pi}^*$ is a $(n-d)\times 1$ column vector, and $\pi^*_\perp$ is a $d \times 1$ vector. Note that $\hat{\Sigma}_q^\top\hat{\Sigma}_q = \mathbf{I}_{(n-d)\times(n-d)}$, $\hat{\Sigma}_b^\top\hat{\Sigma}_b = \mathbf{I}_{d\times d}$, $\hat{\Sigma}_b^\top\hat{\Sigma}_q = 0$, and $\hat{\Sigma}_q^\top\hat{\Sigma}_b = 0$.

Then the following proposition describes two parts $\bar{\pi}^*$ and $\pi^*_\perp$.

Proposition 2

$$
\begin{aligned}
\pi_\perp^* &= \frac{1}{\gamma}\Sigma_b^\top(\Sigma\Sigma^\top)^{-1}(b - r\mathbf{1}_m) + \sigma^{X\top}\frac{f_X}{f}, \\
\bar{\pi}^* &= \arg\max_{\bar{\pi}}\ \bar{\pi}\Sigma_q^\top(\Sigma\Sigma^\top)^{-1}(b - r\mathbf{1}_m) + \frac{1}{1-\gamma}\sum_{k=1}^{n-d}\lambda_k^*\int_{A_k}(1+\bar{\pi}_k z)^{1-\gamma}\Phi_k^*(dz)
\end{aligned}
\tag{45}
$$

where the worst case density λ_k^* *and distribution* $\Phi_k^*(dz)$ *are obtained by Theorem 1.*

Proof. Given the worst probability by Theorem 1, by using the standard dynamic programming approach to stochastic control and an appropriate Ito's lemma for jump-diffusion processes, we can derive the corresponding indirect value function, J, of the investor's problem solving the HJB equation below:

$$
\begin{aligned}
0 \;=\; \max_{\pi}\Bigg\{ & J_t + \frac{1}{2}W^2\pi^\top\Sigma_b\Sigma_b^T\pi J_{WW} + W[\pi^\top(b - r\mathbf{1}_m) + r]J_W \\
& + b^X J_X + W\pi^\top\Sigma_b\sigma^{X\top}J_{WX} + \frac{1}{2}Tr(\sigma^X\sigma^{X\top}J_{XX^\top}) \\
& + \sum_{k=1}^{n-d}\lambda_k^*\int_{A_k}[J(W + W\pi^\top\Sigma_{qk}z) - J(W)]\Phi_k^*(dz) + H(\zeta_t^*)J \Bigg\},
\end{aligned}
\tag{46}
$$

where Σ_{qk} denotes the k-th column of Σ_q.

From Theorem 1, $J(t, W, X) = \frac{W^{1-\gamma}}{1-\gamma}(f(t, X))^\gamma$. Substituting J into (46), we find the optimal portfolio π^* by solving

$$
\begin{aligned}
\max_{\pi}\ & \pi^\top(b - r\mathbf{1}_n) - \frac{\gamma}{2}\pi^\top\Sigma_b\Sigma_b^\top\pi + \gamma\pi^\top\Sigma_b\sigma^{X\top}f_X/f \\
& + \frac{1}{1-\gamma}\sum_{k=1}^{n-d}\lambda_k^*\int_{A_k}(1 + \pi^\top\boldsymbol{\Sigma}_{qk}z)^{1-\gamma}\Phi_k^*(dz).
\end{aligned}
\tag{47}
$$

Substituting the decomposition $\pi = (\Sigma^{-1})^\top(\hat{\Sigma}_q\bar{\pi} + \hat{\Sigma}_b\pi_\perp)$ into the above problem, we obtain

$$
\begin{aligned}
&\max_{\pi_\perp,\bar{\pi}} (\bar{\pi}^\top \hat{\Sigma}_q^\top + \pi_\perp^\top \hat{\Sigma}_b^\top)\Sigma^{-1}(b - r\mathbf{1}_m) - \frac{\gamma}{2}\pi_\perp^\top \pi_\perp + \gamma \pi_\perp^\top \sigma^{X\top} \frac{f_X}{f} \\
&+ \frac{1}{1-\gamma}\sum_{k=1}^{n-d} \lambda_k^* \int_{A_k} (1 + \bar{\pi}^\top (\hat{\Sigma}_q^\top \hat{\Sigma}_q)_k z)^{1-\gamma} \Phi_k^*(dz)
\end{aligned}
$$

where $(\hat{\Sigma}_q^\top \hat{\Sigma}_q)_k$ denotes the k^{th} column of the matrix. As we know, $(\hat{\Sigma}_q^\top \hat{\Sigma}_q) = \mathbf{I}_{(n-d)\times(n-d)}$, so $\bar{\pi}^\top (\hat{\Sigma}_q^\top \hat{\Sigma}_q)_k = \bar{\pi}_k$.

The maximization problem can be solved separately for $\bar{\pi}$ and $\pi_\perp$. Note that $\pi_\perp^*$ solves the first order condition with respect to $\pi_\perp$:

$$
\hat{\Sigma}_b^\top \Sigma^{-1}(b - \mathbf{1}_m) - \gamma \pi_\perp + \gamma \sigma^{X\top} \frac{f_X}{f} = 0.
$$

It follows from the above equation that

$$
\pi_\perp^* = \frac{1}{\gamma}\Sigma_b^\top (\Sigma\Sigma^\top)^{-1}(b - r\mathbf{1}_m) + \sigma^{X\top}\frac{f_X}{f}.
$$

Similarly,

$$
\bar{\pi}^* = \arg\max_{\bar{\pi}} \ \bar{\pi}^\top \Sigma_q^\top (\Sigma\Sigma^\top)^{-1}(b - r\mathbf{1}_m) + \frac{1}{1-\gamma}\sum_{k=1}^{n-d} \lambda_k^* \int_{A_k} (1 + \bar{\pi}_k z)^{1-\gamma} \Phi_k^*(dz).
$$

■

Note that $\bar{\pi}^*$ may not be achieved at an interior point of its admissible region. For example, for a mixed jump size of $(-1, \infty)$, $\bar{\pi}_k^*$ must be in $[0, 1]$ and it is possible to achieve the maximum at 0 or 1. However, when the maximum is achieved at an interior point, the first order condition gives an individual equation of $\bar{\pi}_k^*$ as follows.

$$
\Sigma_{qk}^\top (\Sigma\Sigma^\top)^{-1}(b - r\mathbf{1}_m) + \lambda_k^* \int_{A_k} (1 + \bar{\pi}_k^* z)^{-\gamma} z \Phi_k^*(dz) = 0.
$$

Since $[\hat{\Sigma}_b, \hat{\Sigma}_q]$ is an $n \times n$ identity matrix, the optimal portfolio choice π can be re-written as

$$\pi = (\Sigma^{-1})^\top [\hat{\Sigma}_b, \hat{\Sigma}_q] \begin{pmatrix} \pi_\perp \\ \bar{\pi} \end{pmatrix} = (\Sigma^{-1})^\top \begin{pmatrix} \pi_\perp \\ \bar{\pi} \end{pmatrix}.$$

Compared to Proposition 1 of Jin and Zhang [19], $\bar{\pi}$ and $\pi_\perp$ here are corresponding to π_q and π_b there, respectively, given the worst probability. It worths mentioning that our decomposition approach is different from theirs and it may be extended to the case $m < n$ or $m > n$.

5 Extension to HARA utility functions

In Bajeux-Besnainou et al. [3], they obtain closed-form solutions for HARA optimal dynamic portfolios in pure-diffusion models. Specifically, they employ the duality results developed by Karatzas et al. [20] in complete markets. A key assumption for the applicability of the duality results in Karatzas et al. [20] is that there is an unique equivalent martingale measure in a complete market. By contrast, the markets in the present paper are incomplete due to random jump size and thus there exist infinitely many equivalent martingale measures. To solve a HARA optimal dynamic portfolio problem, we resort to the duality results for incomplete market recently developed by Bellini and Frittelli [5] in combination with the results developed in the last section. In the model, we assume that the dynamics of the bond and stock prices remain unchanged and an investor has a HARA utility function given by

$$U(x) = \begin{cases} \frac{1}{1-\gamma}(x-a)^{1-\gamma}, & \forall x > a \\ -\infty, & \forall x \le a \end{cases}.$$

When $a = 0$, $U(x)$ reduces to a CRRA utility function. Here we consider a realistic case with $a > 0$, that is, the relative risk aversion is decreasing with x. In Bajeux-Besnainou et al. [3], they interpret the constant a as "subsistence level".

Proposition 3 *Under the HARA utility function, for the portfolio choice problem under ambiguity and ambiguity aversion (11), we have the following duality result*

$$\begin{aligned} J(t, W_t, X_t) &= \frac{(W_t - a\kappa_t)^{1-\gamma}}{1-\gamma} \left(\sup_{\zeta} \sup_{\xi \in \mathbb{Q}_\zeta} \mathbf{E}_t^\zeta \left[\exp\left(\frac{1}{\gamma} \int_t^T (H_s + (1-\gamma) r) ds \right) \xi_\delta(t,T)^{1-\frac{1}{\gamma}} \right] \right)^\gamma \\ &= \frac{(W_t - a\kappa_t)^{1-\gamma}}{1-\gamma} \left(f(t, X_t) \right)^\gamma, \end{aligned}$$

where $\kappa_t = \mathbf{E}_t \left[\exp\left(-\int_t^T r(X_s) ds \right) \xi_\delta^b(t,T) \right]$, *and* $f(t, X_t)$ *is given by (13),* $\xi_\delta(t,T)$ *and* $\xi_\delta^b(t,T)$ *are given in Section 3.1.*

Proof. We now use the results obtained for the CRRA utility function and the results in Bellini and Frittelli [5] to solve the optimal portfolio choice problem with a HARA utility function. First we derive duality result for the model without ambiguity aversion, and then we obtain duality result for the model with ambiguity aversion by using the same idea as before. Note that

$$U\left(U'^{-1}(y)\right) \geq U(x) + y\left(U'^{-1}(y) - x\right), \ \forall x > 0, \ y > 0,$$

where $U'^{-1}(y) = I(y) = y^{-\frac{1}{\gamma}} + a$. For simplicity, we consider $t = 0$ and let $\beta_t = \exp\left(-\int_t^T r(X_s) ds\right)$. Let $\mathbb{Q}$ denote the set of all equivalent martingale measures. Thus, for any $\xi \in \mathbb{Q}$ and terminal wealth W_T, we have

$$U\left(U'^{-1}(y\beta_0\xi_T)\right) \geq U(W_T) + y\beta_0\xi_T\left(U'^{-1}(y\beta_0\xi_T) - W_T\right),$$

and

$$\mathbf{E}\left[U\left(U'^{-1}(y\beta_0\xi_T)\right)\right] \geq \mathbf{E}\left[U(W_T)\right], \tag{48}$$

where y satisfies

$$\mathbf{E}\left[\beta_0\xi_T U'^{-1}(y\beta_0\xi_T)\right] = W_0, \tag{49}$$

giving

$$y = \frac{\mathbf{E}\left[(\beta_0\xi_T)^{1-\frac{1}{\gamma}}\right]^{\gamma}}{(W_0 - a\mathbf{E}(\beta_0\xi_T))^{\gamma}}.$$

We now prove that there exists a $\xi \in \mathbb{Q}$ such that $(\beta_0\xi_T)^{-\frac{1}{\gamma}}$ can be replicated and hence $I(y\beta_0\xi_T) = y^{-\frac{1}{\gamma}}(\beta_0\xi_T)^{-\frac{1}{\gamma}} + a$ can be replicated. By Kramkov and Schachermayer [21] and by considering the utility function $\frac{1}{1-\gamma}x^{1-\gamma}$, we have that there exists a $\xi \in \mathbb{Q}$ such that $(\beta_0\xi_T)^{-\frac{1}{\gamma}}$ can be replicated. Furthermore, according to (48), we have

$$u(W_0) = \mathbf{E}\left[U(U'^{-1}(y\xi_T))\right] = \frac{(W_0 - a\mathbf{E}(\beta_0\xi_T))^{1-\gamma}}{1-\gamma}\mathbf{E}\left[(\beta_0\xi_T)^{1-\frac{1}{\gamma}}\right]^{\gamma}, \tag{50}$$

with y satisfying (49).

In the following, we use some results in Bellini and Frittelli [5] to prove the following

$$u(W_0) = \inf_{\varsigma\in\mathbb{Q}} \frac{(W_0 - a\mathbf{E}(\beta_0\varsigma_T))^{1-\gamma}}{1-\gamma}\mathbf{E}\left[(\beta_0\varsigma_T)^{1-\frac{1}{\gamma}}\right]^{\gamma}.$$

To this end, we denote with L^∞ the space of essentially bounded random variables and define

$$M_0 = \{W \in L^\infty : \mathbf{E}[\beta_0\varsigma_T W] \le W_0 \ \forall\varsigma \in \mathbb{Q}\}.$$

According to Lemma 1.1 and 1.2 of Bellini and Frittelli [5] (note we do not need Assumption 1.3), we have

$$u(W_0) = \sup_{W\in M_0} \mathbf{E}[U(W)]. \tag{51}$$

By following (1.8) in Bellini and Frittelli [5], we define

$$U(W_0;\varsigma,P) = \sup_{W\in M_0^{\varsigma}} \mathbf{E}[U(W)],$$

where $M_0^\varsigma = \{W \in L^\infty : \mathbf{E}[\beta_0 \varsigma_T W] \leq W_0\}$. It is easy to see from (51)

$$u(W_0) = \sup_{W \in M_0} \mathbf{E}[U(W)] \leq \inf_{\varsigma \in \mathbb{Q}} U(W_0; \varsigma, P), \tag{52}$$

since $M_0 \subseteq M_0^\varsigma$. As in Section 2.1 of Bellini and Frittelli [5], we define the concave conjugate $U^*(x^*)$ of the utility function $U(x)$ as:

$$U^*(x^*) = \inf_x \{xx^* - U(x)\}.$$

In particular, for the HARA utility function $U(x)$, we have

$$U^*(x^*) = \frac{\gamma}{\gamma - 1}(x^*)^{1-\frac{1}{\gamma}} + ax^*.$$

Hence, using Corollary 2.1 of Bellini and Frittelli [5], we have

$$\begin{aligned} U(W_0; \varsigma, P) &= \min_{\lambda \in (0,\infty)} \{\lambda W_0 - \mathbf{E}_P[U^*(\lambda \beta_0 \varsigma_T)]\} \\ &= \frac{(W_0 - a\mathbf{E}(\beta_0 \varsigma_T))^{1-\gamma}}{1-\gamma} \mathbf{E}\left[(\beta_0 \varsigma_T)^{1-\frac{1}{\gamma}}\right]^\gamma, \end{aligned}$$

and, by (52),

$$u(W_0) \leq \inf_{\varsigma \in \mathbb{Q}} \frac{(W_0 - a\mathbf{E}(\beta_0 \varsigma_T))^{1-\gamma}}{1-\gamma} \mathbf{E}\left[(\beta_0 \varsigma_T)^{1-\frac{1}{\gamma}}\right]^\gamma.$$

From (50), we have

$$u(W_0) = \inf_{\varsigma \in \mathbb{Q}} \frac{(W_0 - a\mathbf{E}(\beta_0 \varsigma_T))^{1-\gamma}}{1-\gamma} \mathbf{E}\left[(\beta_0 \varsigma_T)^{1-\frac{1}{\gamma}}\right]^\gamma.$$

We now turn to the model with ambiguity aversion. By following the same approach as that for Theorem 1, we can derive the indirect value function as

$$
\begin{aligned}
J(0, W_0, X_0) &= \frac{1}{1-\gamma}\left(\sup_{\zeta}\sup_{\xi\in\mathbb{Q}_\zeta}(W_0 - a\mathbf{E}^{\zeta}(\beta_0\xi_\delta(T)))^{1-\gamma}\right.\\
&\quad \left.\times\, \mathbf{E}^{\zeta}\left[\exp\left(\frac{1}{\gamma}\int_0^T (H_s + (1-\gamma)r)ds\right)\xi_\delta(T)^{1-\frac{1}{\gamma}}\right]\right)^{\gamma}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\mathbf{E}^{\zeta}\left[\beta_0\xi_\delta(T)\right] &= \mathbf{E}^{\zeta}\left[\beta_0\xi_\delta^b(T)\xi_\delta^q(T)\right] = \mathbf{E}^{\zeta}\left[\mathbf{E}^{\zeta}\left[\beta_0\xi_\delta^b(T)\xi_\delta^q(T)|\mathcal{F}_T^X\right]\right]\\
&= \mathbf{E}^{\zeta}\left[\beta_0\xi_\delta^b(T)\mathbf{E}^{\zeta}\left[\xi_\delta^q(T)|\mathcal{F}_T^X\right]\right] = \mathbf{E}^{\zeta}\left[\beta_0\xi_\delta^b(T)\right],
\end{aligned}
$$

since $\mathbf{E}^{\zeta}\left[\xi_\delta^q(T)|\mathcal{F}_T^X\right] = 1$, implying that $\mathbf{E}^{\zeta}\left[\beta_0\xi_\delta(T)\right]$ is independent of ζ. And therefore,

$$
\begin{aligned}
J(0, W_0, X_0) &= \frac{(W_0 - a\mathbf{E}\left[\beta_0\xi_\delta^b(T)\right])^{1-\gamma}}{1-\gamma}\\
&\quad \times\left(\sup_{\zeta}\sup_{\xi\in\mathbb{Q}_\zeta}\mathbf{E}^{\zeta}\left[\exp\left(\frac{1}{\gamma}\int_0^T (H_s + (1-\gamma)r)ds\right)\xi_\delta(T)^{1-\frac{1}{\gamma}}\right]\right)^{\gamma}\\
&= \frac{(W_0 - a\mathbf{E}\left[\beta_0\xi_\delta^b(T)\right])^{1-\gamma}}{1-\gamma}(f(0, X_0))^{\gamma},
\end{aligned}
$$

where $f(0, X_0)$ is given by Theorem 1. Likewise, we can show

$$
J(t, W_t, X_t) = \frac{(W_t - a\mathbf{E}\left[\beta_t\xi_\delta^b(t, T)\right])^{1-\gamma}}{1-\gamma}(f(t, X_t))^{\gamma}.
$$

■

The above proposition suggests that the worst case probability ζ is independent of the wealth W_t. In other words, the wealth of an investor with a HARA utility function does not affect her effective ambiguity aversion coefficient. The reason for this is that the special form of the normalization factor $\Lambda(x) = (1-\gamma)x$ is used in the utility function (7), which is proposed by Maenhout [27]. But it is worth mentioning that this property still holds in the framework of ambiguity aversion in Drechsler [12].

Given the indirect value function, the corresponding optimal portfolio strategy can be obtained by using the same orthogonal decomposition method in Section 4. To save space, we omit the derivation.

6 Numerical examples

To illustrate our approach and results, we specialize in a simplified jump-diffusion model with only one jump as follows. For the purpose of illustration, we do not consider the state variable X_t in this model.

$$\frac{dS_n}{S_n} = b_n dt + \sum_{m=1}^{M} \sigma_{nm} dB_m + \sigma_n^q Y dN, \quad n = 1, 2, ..., M, \tag{53}$$

where $Y = \exp(\mu_J + \sigma_J \varepsilon) - 1$ and ε is a standard normal random variable; $\mathbf{E}(dN) = \lambda dt$; B_1 to B_M are standard independent Brownian motions and independent of Y; M is the total number of stocks.

Theorem 1 cannot be applied to this case directly. As discussed in Section 3, we may add one fictitious risky asset with an undetermined drift term. Then we have eleven risky assets, ten Brownian motions and one jump process in the fictitious market. Theorem 1 is therefore able to apply. The optimal portfolio is obtained when the investment is restricted to the first ten assets by adjusting the drift term of the fictitious asset. For details of this treatment, we refer to Jin and Zhang [17], Karatzas et al. [18], and Cvitanic and Karatzas [8].

6.1 Model calibration

We calibrate the model to the monthly continuously compounded returns on the equity indices of 10 developed and 10 emerging countries/regions, respectively. The developed countries include the United States (US), United Kingdom (UK), Switzerland (SW), Germany (GE), France (FR), Australia (AU), Canada (CA), Sweden (SD), Japan (JP), and Netherlands (NE). The emerging countries/regions include Argentina (ARG), Brazil (BRA), Hong Kong (HKG), India (IND), Indonesia (IDO), South Korea (KO-R), Malaysia (MAL), Mexico (MEX), Singapore (SNG), and Taiwan (TWN). To avoid confusion, we abbreviate the developed countries with two characters and the emerging countries/regions with three characters. We collect beginning-of-month equity index levels from finance.yahoo.com. Due to data availability, our sample period is January 1993 to December 2015 for the developed group and July 1997 to January 2016 for the emerging group.

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 lead to the negative skewness of the returns. Pairwise correlations among the equity index returns are unanimously higher than 43% within the developed group and higher than 36% within the emerging group. This result indicates the close linkage of the international equity markets.

We estimate the jump-diffusion model using the method of moments approach provided by Das and Uppal [10] and Jin and Zhang [19]. The first four unconditional moments of the multivariate return series are considered. Following Das and Uppal [10], we derive in closed form the characteristic function of the continuously compounded stock returns. We then differentiate the characteristic function to obtain the moments.

Let $\bar{Y}_n = \ln(\sigma_n^q Y + 1)$. For $n, m = 1, 2, ..., M$ $(M = 10)$,

$$
\begin{aligned}
&\text{mean} && = t(b_n - 0.5\sum_{k=1}^{M}\sigma_{nk}^2 + \lambda \mathrm{E}[\bar{Y}_n]), \\
&\text{covariance} && = t(\sum_{k=1}^{M}\sigma_{nk}\sigma_{mk} + \lambda \mathrm{E}[\bar{Y}_n\bar{Y}_m]), \\
&\text{coskewness} && = \frac{t\lambda \mathrm{E}[(\bar{Y}_n)^2\bar{Y}_m]}{\text{variance}_n(\text{variance}_m)^{0.5}}, \\
&\text{excess kurtosis} && = \frac{t\lambda \mathrm{E}[(\bar{Y}_n)^4]}{(\text{variance}_n)^2},
\end{aligned}
\tag{54}
$$

where $\mathbf{E}[(\bar{Y}_n)^i(\bar{Y}_m)^j] = \int_{-\infty}^{+\infty}(\bar{Y}_n)^i(\bar{Y}_m)^j f(x)dx$ with $i = 1, 2, ...; j = 0, 1, ...;$ and $f(\cdot)$ is the standard normal density. This integral can be evaluated easily using the numeric quadrature method. We first use the 10×10 co-skewness conditions and 10×1 kurtosis conditions to estimate the 13 jump parameters (σ_n^q, μ_J, σ_J, λ) by minimizing the sum of squared deviations of the model moments from those in data. We then derive b_n and σ_{nm} by exactly matching the 10×1 mean conditions and 10×10 covariance conditions, respectively.

Table 1 presents the parameter estimates[6] and return moments on monthly basis. From Panel A, we see that the average jump size is -12.0% for the developed countries. This result is consistent with the negative skewness of the return series. The standard deviation of jump size is 6.2%. Thus a 95% confidence interval for the jump size is (-24.4%, 0.4%). As shown in the moment condition in equation (54), large-sized jumps are crucial to match the high excess kurtosis of the data. The jump intensity is estimated to be 0.073. Simultaneous jumps among the ten markets are expected to occur about once every 14 months, or once every 1.1 years. This is consistent with the literature which finds that equity indices jump about once or twice a year. Turn to Panel B, the average jump size is -8.0% with a standard deviation of jump size 19.7% for the emerging countries/regions. This much higher standard deviation of jump size helps to match the largely inflated excess kurtosis of the return series observed for the emerging countries/regions compared to the developed ones. The jump intensity is estimated to

[6]To save space, b_n and σ_{nm} are not listed but available upon request.

Table 1: Parameter Estimates

This table reports parameter estimates of the multivariate jump-diffusion model of the stock index returns. We estimate the parameters by minimizing the sum of squared deviations of the return moments implied by model from those in data. We provide the higher moments reconstructed from model and those in data. *Skew* and *ExKurt* denote return skewness and excess kurtosis, respectively. All the parameter estimates and moments are on the monthly basis. Panel A gives the results for the developed countries for the sample period January 1993 to December 2015. Panel B gives the results for the emerging countries/regions for the period July 1997 to January 2016.

Panel A: Developed countries										
	U.S.	U.K.	SW	GE	FR	AU	CA	SD	JP	NE
μ_J	-0.118									
σ_J	0.062									
λ	0.073									
σ^q	0.654	0.538	0.781	0.986	0.705	0.500	0.784	0.780	0.664	0.514
Skew: model	-0.708	-0.459	-0.472	-0.843	-0.740	-0.419	-1.202	-0.653	-0.347	-0.442
Skew: data	-0.855	-0.698	-0.557	-0.867	-0.855	-0.741	-1.226	-0.674	-0.514	-0.527
ExKurt: model	1.932	1.080	1.131	2.464	2.052	0.956	3.929	1.741	0.748	1.027
ExKurt: data	1.761	0.735	1.141	2.626	2.120	0.581	4.093	1.755	0.436	1.084
Panel B: Emerging countries/regions										
	ARG	BRA	HKG	IND	INS	KOR	MAL	MEX	SNG	TWN
μ_J	-0.080									
σ_J	0.197									
λ	0.080									
σ^q	0.998	0.903	0.666	0.531	0.850	0.752	0.695	0.654	0.658	0.483
Skew: model	-0.552	-0.759	-0.401	-0.197	-0.703	-0.438	-0.613	-0.558	-0.455	-0.149
Skew: data	-0.418	-1.142	-0.463	-0.370	-1.138	0.167	-0.061	-0.918	-0.353	-0.129
ExKurt: model	3.003	4.869	2.515	1.131	4.561	2.614	4.307	3.950	3.002	0.827
ExKurt: data	2.996	4.687	2.586	0.988	4.482	2.838	4.511	3.786	3.071	0.803

be 0.080. Simultaneous jumps among the ten markets are expected to occur once a year. For both the developed and emerging groups, the theoretic moments reconstructed using the model parameter estimates are close in magnitude to hence do a reasonably good job in fitting the moments of the return data.

In the following, we will discuss portfolio choice and the worst probabilities implied from the model.

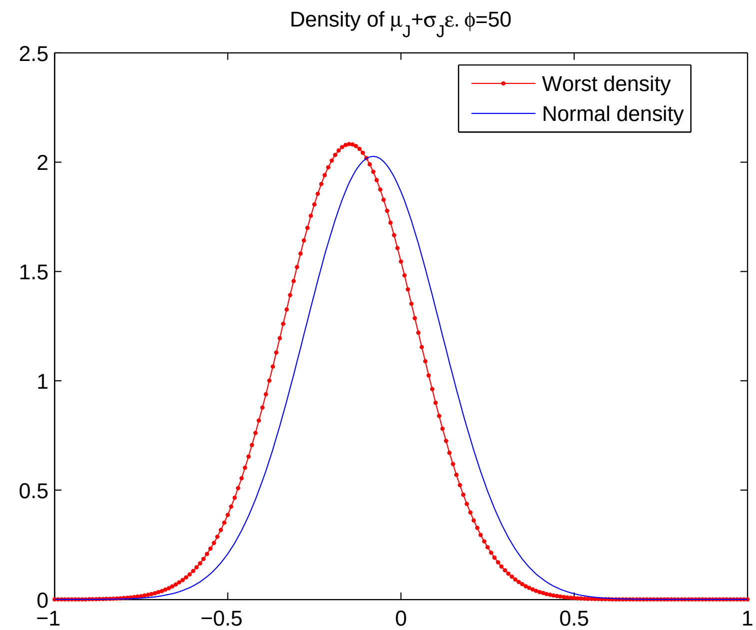

Figure 1: The worst-case density and the reference (normal) density.

6.2 International asset allocation with ambiguity aversion

We compute optimal asset allocations among the two groups of ten countries/regions across varying ambiguity aversion coefficient ϕ. The jump-diffusion model without ambiguity is used as the benchmark for comparison. The risk-free interest rate is set at 5% per year and the risk aversion coefficient γ is set to be 5.

In Table 2, we present the optimal portfolios for varying degrees of ambiguity aversion. As ambiguity aversion (ϕ) gets higher, the investor becomes more ambiguity averse. The total risky investments get abated (since the mean jump size is negative), so do the exposures to jump risks ($\tilde{\pi}_q$), reflecting the investor's fear of jump uncertainty. Undoubtedly, the total risky investments under ambiguity aversion are less than that without ambiguity aversion ($\phi = 0$). Note that ϑ^* is larger for a higher level of ambiguity aversion. Hence (negative) jumps occur more frequently in the worst case for an investor with a higher level of ambiguity aversion. Consistent with the portfolio results, the worst density shifts to the negative side as shown in Fig. 1.

Table 2: Optimal Portfolios at Different Degrees of Ambiguity Aversion

This table reports optimal portfolio positions among the two groups of 10 countries/regions at five ambiguity aversion values of ϕ. The total portfolio weights in each 10 indices given by $\sum_i \pi_i$ are listed in the row "Total". Exposure to jump risk in the risky assets is given by $\tilde{\pi}_q = \pi^\top \Sigma_q$. The worst jump intensity is $\lambda\vartheta^*$ where ϑ^* is also reported in the table. We show in the last column the optimal portfolios and the exposures to jump risks without ambiguity and ambiguity aversion ($\phi = 0$).

Panel A: Developed countries					
ϕ	200	100	50	10	0
US	0.740	0.738	0.733	0.711	0.667
UK	-1.269	-1.271	-1.275	-1.294	-1.333
SW	-0.509	-0.506	-0.501	-0.478	-0.431
GE	0.482	0.491	0.507	0.584	0.740
FR	-0.246	-0.240	-0.229	-0.175	-0.066
AU	-0.147	-0.151	-0.157	-0.188	-0.252
CA	-0.395	-0.384	-0.363	-0.265	-0.066
SD	0.585	0.588	0.592	0.611	0.649
JP	-0.345	-0.356	-0.377	-0.474	-0.670
NE	0.908	0.912	0.918	0.950	1.014
Total	-0.195	-0.179	-0.151	-0.018	0.253
$\tilde{\pi}_q$	0.017	0.033	0.061	0.197	0.471
ϑ^*	1.486	1.464	1.429	1.268	1
Panel B: Emerging countries/regions					
ϕ	200	100	50	10	0
ARG	0.229	0.230	0.233	0.248	0.276
BRA	-0.131	-0.128	-0.123	-0.101	-0.058
HKG	-0.111	-0.110	-0.109	-0.106	-0.099
IND	0.293	0.293	0.294	0.297	0.304
INS	0.114	0.116	0.120	0.139	0.176
KOR	0.062	0.064	0.068	0.087	0.123
MAL	-0.168	-0.164	-0.155	-0.116	-0.040
MEX	0.462	0.464	0.468	0.485	0.517
SNG	-0.487	-0.489	-0.494	-0.514	-0.553
TWN	-0.392	-0.394	-0.398	-0.416	-0.451
Total	-0.129	-0.117	-0.096	0.004	0.195
$\tilde{\pi}_q$	0.011	0.021	0.039	0.125	0.290
ϑ^*	1.237	1.226	1.207	1.124	1

We now examine the economic significance of the differences in the optimal portfolio weights between the two models with and without ambiguity aversion. To this end, we let $\pi^{(1)}$ and $\pi^{(2)}$ be the optimal portfolios for the investor with ambiguity aversion and the one without ambiguity aversion, respectively. In particular, certainty equivalent loss (CEL) is defined as the percentage of initial wealth an investor is willing to give up to switch from portfolio strategy $\pi^{(2)}$ to portfolio strategy $\pi^{(1)}$. Equivalently, CEL solves the following equation:

$$J^{(2)}(W,t) = J^{(1)}(W(1-CEL),t), \tag{55}$$

where the value function $J^{(1)}(W,t)$ associated with $\pi^{(1)}$ is calculated by Theorem 1. We evaluate $J^{(2)}(W,t)$ following Flor and Larsen [12]. Specifically,

$$J^{(2)}(W,t) = \mathbf{E}_t^{\zeta^*}\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right], \tag{56}$$

where W_T is the wealth process associated with $\pi^{(2)}$, and the worst case Radon-Nikodym derivative ζ^* is associated with the ambiguity aversion portfolio $\pi^{(1)}$. That is, $J^{(2)}$ is the value function when applying $\pi^{(2)}$ in the model with the (worst case) jump distribution corresponding to $\pi^{(1)}$. Since $\pi^{(1)}$ is optimal to maximize among all possible worst values, we know $J^{(1)}(W,t) \geq J^{(2)}(W,t)$ and the CEL defined above is non-negative.

The results are listed in Table 3. As we can see from the table, the certainty equivalent loss is significant. It can be as large as 95% in a time horizon of 20 years in the emerging markets. This indicates that a huge loss may be caused by ignoring uncertainty of jumps. It is interesting to note that the certainty equivalent loss is much larger in the emerging markets. The reasons may be due to more volatile jumps in the emerging markets (i.e. jumps size has a larger variance), hence there are more ambiguity in jumps for the emerging markets and it is more important to consider the ambiguity in the optimal portfolio in the emerging markets. If we artificially change the jump volatility σ_J from

Table 3: Certainty Equivalent Loss

This table reports the certainty equivalent loss when the investor fails to account for jump ambiguity and takes the portfolio without ambiguity aversion. The certainty equivalent loss is possibly incurred when the model encounters the worst case jumps and the investor applies the (suboptimal) portfolio strategy ignoring jump ambiguity.

	Investment horizon (in years)		
ϕ	1	10	20
Panel A: Developed countries			
10	0.0017	0.0164	0.0326
50	0.0166	0.1538	0.2840
100	0.0522	0.4152	0.6580
Panel B: Emerging countries/regions			
10	0.0092	0.0886	0.1693
50	0.0754	0.5434	0.7915
100	0.1482	0.7989	0.9595

0.197 to 0.06, we find the CEL under $\phi = 100$ becomes 0.0367, 0.3117, 0.5262 in 1, 10, and 20 years' investment horizons respectively. Compared to Panel B of Table 3, these CELs are significantly reduced. The result confirms that jump volatility affects certainty equivalent loss and explains the difference of CELs between the emerging markets and the developed markets.

6.3 Comparison with the Parameterized Model

Liu et al. [25] use a parametric approach to choose the worst jump size and jump intensity. They consider a single stock model with one jump and assume that the jump intensity is changed to $e^a\lambda$ and the jump size density is changed to $e^{bz-b\mu_J-\frac{1}{2}b^2\sigma_J^2}\Phi(dz)$ for $a, b \in \mathcal{R}$ in an alternative model, where $\Phi(dz)$ is the density of $Z \sim N(\mu_J, \sigma_J)$ and $Y = e^Z - 1$ is the jump size of the stock price. Instead of minimizing over all valid probability measures, they minimize the objective function over the real sets of a and b. We note that for many values of model parameters this parameterized approach can generate results close to those by our non-parameterized approach. Panel B of Table 4

Table 4: Comparison with the Parameterized Model

This table reports optimal portfolios under our non-parameterized model(NonP), the parameterized model(Para) of Liu et al. [25], and the model without ambiguity (W/A). The jump intensity in the worst case becomes $\lambda\vartheta^*$ in NonP, and becomes λe^{a^*} in Para.

	ARG	BRA	HKG	IND	INS	KOR	MAL	MEX	SNG	TWN
Panel A: $\phi = 250$, $\gamma = 1.1$										
NonP	1.0318	-0.6048	-0.5043	1.3296	0.5082	0.2732	-0.7835	2.0915	-2.2037	-1.7716
Para	1.0496	-0.5774	-0.4998	1.3337	0.5316	0.2960	-0.7352	2.1123	-2.2287	-1.7940
W/A	1.2005	-0.3452	-0.4621	1.3687	0.7297	0.4894	-0.3270	2.2883	-2.4398	-1.9838
$\vartheta^* = 1.2477$, $\exp(a^*) = 1.5275$, $b^* = -1.1255$										
Panel B: $\phi = 50$, $\gamma = 5$										
NonP	0.2333	-0.1233	-0.1093	0.2939	0.1201	0.0682	-0.1552	0.4675	-0.4937	-0.3977
Para	0.2333	-0.1232	-0.1093	0.2940	0.1202	0.0682	-0.1550	0.4675	-0.4937	-0.3977
W/A	0.2760	-0.0576	-0.0986	0.3038	0.1761	0.1229	-0.0397	0.5173	-0.5534	-0.4514
$\vartheta^* = 1.2066$, $\exp(a^*) = 1.2076$, $b^* = -1.7692$										

lists an example where the two approaches generate almost the same optimal portfolio.[7] In fact, we find that the worst densities under the two models are almost identical and $\exp(a^*)$ is almost equal to ϑ^* in our model. Hence the worst cases under the two models are almost identical and not surprised to see almost identical optimal portfolios under ambiguity in this example.

However, it is not always true that the two models give close results. When γ is small and ϕ is large, the two approaches may generate results with large difference. Panel A of Table 4 lists the optimal portfolios under the two models when $\phi = 250$ and $\gamma = 1.1$. There are clear differences between the allocations of the two optimal portfolios. Meanwhile, Fig. 2 shows a sharp difference between the two worst densities under the models. Furthermore, we find that the certainty equivalent losses are significant, namely, 1.99%, 18.18%, and 33.06% for 1, 10, and 20 years' investments respectively, if the worst case of the parameterized model is applied when the true model is the worst case of the non-parameterized model. These results illustrate the essential importance of our non-parameterized approach.

[7] We focus on the group of 10 emerging countries/regions for illustration. The other group of 10 developed countries produces qualitatively similar results.

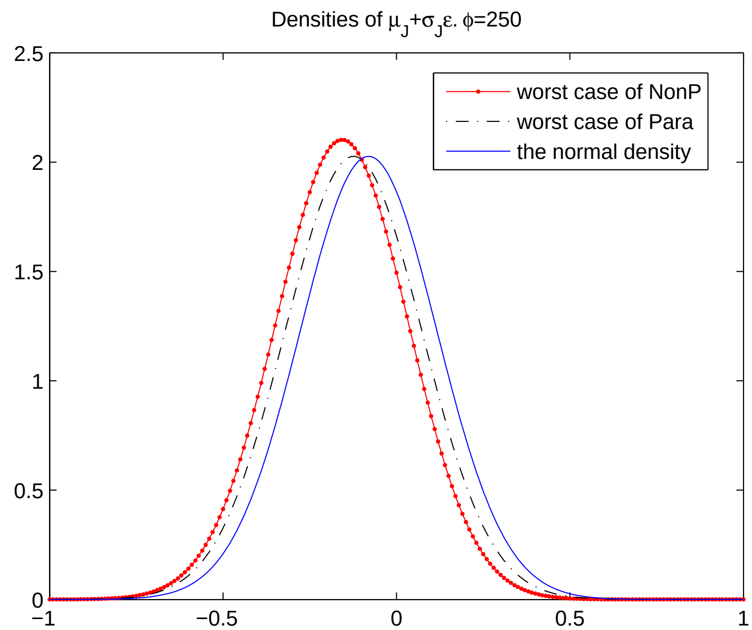

Figure 2: The worst-case density of our non-parameterized model (NonP), the worst-case density under the parameterized model (Para), and the reference (normal) density. $\phi = 250$, $\gamma = 1.1$.

7 Conclusion

Solving the optimal dynamic portfolio selection problem for a multi-asset incomplete market with or without model uncertainty is a daunting task due to the curse of dimensionality. This paper proposes a novel approach to the intertemporal portfolio selection problem in jump-diffusion models where the investor is averse not only to risk of loss but also to model uncertainty. More specifically, based on the duality-martingale methods and the minimax theorem, we evaluate the probability of the worst case scenario and the indirect value function by solving a pathwise optimization problem. We also develop an orthogonal decomposition method to obtain the optimal portfolio in the worst case. One appealing feature of our approach is that our method can deal with a large number of assets and state variables in a model with ambiguity aversion to jump risks. Our approach also circumvents the problem of dimensionality.

The theoretical results show how an ambiguity averse investor fears ambiguous jumps by attaching more weights to the likelihood of adverse events. Our calibration exercise in international markets illustrates that jump uncertainty significantly affects the optimal portfolio weights and the certainty equivalent loss may be large if the uncertainty is ignored or the alternative probability laws are confined to be parametric.

Appendices

A Proof of Proposition 1

We separate the proof of Proposition 1 into several lemmas.

Lemma A1

$$\mathbf{E}_t^{\zeta}\left[\ln\left(\frac{\zeta_{t+\Delta t}^{(k)}}{\zeta_t^{(k)}}\right)\right] = \lambda_k \int_t^{t+\Delta t}\int_{A_k} \mathbf{E}_t^{\zeta}[\vartheta_k(s)\psi_k(s,z)\ln(\vartheta_k(s)\psi_k(s,z))+1-\vartheta_k(s)\psi_k(s,z)]\Phi_k(dz)ds.$$

Proof. *Consider equation (4) under the probability $P(\zeta)$ with which the k-th jump intensity λ_k and density function $\Phi_k(dz)$ are changed into $\vartheta_k\lambda_k$ and $\psi_k(z)\Phi_k(dz)$ in the alternative model, respectively. We rewrite equation (4) as:*

$$\begin{aligned} d\zeta_t^{(k)} &= \int_{A_k} (\vartheta_k(t)\psi_k(t,z)-1)^2 \zeta_{t-}^{(k)} \lambda_k \Phi_k(dz)dt \\ &\quad + \int_{A_k} (\vartheta_k(t)\psi_k(t,z)-1)\zeta_{t-}^{(k)} \widetilde{q}_k(dt,dz), \end{aligned}$$

where

$$\widetilde{q}_k(dt,dz) = dN_k(t) - \lambda_k\vartheta_k(t)\psi_k(t,z)\Phi_k(dz)dt, k = 1, ..., n-d.$$

In particular, the above terms are martingale differentials under the probability $P(\zeta)$. Applying Ito's lemma to the function $f(x) = \ln x$ and the above equation gives

$$\begin{aligned} d\ln(\zeta_t^{(k)}) &= \int_{A_k} (\vartheta_k(t)\psi_k(t,z)-1)^2 \lambda_k \Phi_k(dz)dt \\ &\quad + \lambda_k\vartheta_k(t)\int_{A_k} [ln(\zeta_{t-}^{(k)} + \zeta_{t-}^{(k)}(\vartheta_k(t)\psi_k(t,z)-1)) - ln(\zeta_{t-}^{(k)}) \\ &\quad - (\vartheta_k(t)\psi_k(t,z)-1)]\psi_k(t,z)\Phi_k(dz)dt \\ &\quad + \int_{A_k} [ln(\zeta_{t-}^{(k)} + \zeta_{t-}^{(k)}(\vartheta_k(t)\psi_k(t,z)-1)) - ln(\zeta_{t-}^{(k)})]\widetilde{q}_k(dt,dz) \\ &= \lambda_k \int_{A_k} [\vartheta_k(t)\psi_k(t,z)\ln(\vartheta_k(t)\psi_k(t,z)) + 1 - \vartheta_k(t)\psi_k(t,z)]\Phi_k(dz)dt \\ &\quad + \int_{A_k} ln(\vartheta_k(t)\psi_k(t,z))\widetilde{q}_k(dt,dz). \end{aligned}$$

Thus, we have for $k = 1, 2, ..., n-d$

$$\mathbf{E}_t^{\zeta}\left[\ln\left(\frac{\zeta_{t+\Delta t}^{(k)}}{\zeta_t^{(k)}}\right)\right] = \lambda_k \int_t^{t+\Delta t}\int_{A_k} \mathbf{E}_t^{\zeta}[\vartheta_k(s)\psi_k(s,z)\ln(\vartheta_k(s)\psi_k(s,z))+1-\vartheta_k(s)\psi_k(s,z)]\Phi_k(dz)ds.$$

■

We now make the following assumption:

Assumption A For each vector $(\vartheta, \psi) \in \Theta \times \Psi$, we assume that it is continuous with respect to $t \in [0, T]$ and for the corresponding H_s, there exists a positive constant C_H such that $\sup_{t\in[0,T]} \mathbf{E}_t^{\zeta}(\sup_{s\in[0,T]} |H_s|) \leq C_H < \infty$.[8]

This assumption seems unrestrictive because the state variable vector X_t is continuous with respect to $t \in [0, T]$ and thus the optimal solution (ϑ^*, ψ^*) given by (32) is continuous with respect to $t \in [0, T]$ and bounded. This implies the corresponding H_s is bounded, leading to the second condition.

We now turn to the proof of (10). Recall $\Lambda(x) = (1 - \gamma)x$ and $\gamma > 1$. Then, from Lemma A1, we have

$$\begin{aligned}
&\frac{\Lambda\left(\mathbf{E}_t^{\zeta}(U_{t+\Delta t})\right)}{\mathbf{E}_t^{\zeta}(U_{t+\Delta t})} \sum_{k=1}^{n-d} \frac{1}{\phi_k} \mathbf{E}_t^{\zeta}\left[\ln\left(\frac{\zeta_{t+\Delta t}^{(k)}}{\zeta_t^{(k)}}\right)\right] \\
= \;& (1-\gamma) \sum_{k=1}^{n-d} \frac{\lambda_k}{\phi_k} \int_t^{t+\Delta t} \int_{A_k} \mathbf{E}_t^{\zeta}[\vartheta_k(s)\psi_k(s,z)\ln(\vartheta_k\psi_k(s,z)) + 1 - \vartheta_k(s)\psi_k(s,z)]\Phi_k(dz)ds \\
\equiv \;& \int_t^{t+\Delta t} \mathbf{E}_t^{\zeta}[H(\zeta_s)]ds \equiv \int_t^{t+\Delta t} \mathbf{E}_t^{\zeta}[H_s]ds.
\end{aligned}$$

It is evident that $H(\xi_t) \leq 0$ because $x \ln x + 1 - x \geq 0$ for $x > 0$ and $1 - \gamma < 0$. Thus, following Jin and Zhang [19], we have

$$\begin{aligned}
&\Lambda\left(\mathbf{E}_t^{\zeta}(U_{t+\Delta t})\right) \sum_{k=1}^{n-d} \frac{1}{\phi_k} \mathbf{E}_t^{\zeta}\left[\ln\left(\frac{\zeta_{t+\Delta t}^{(k)}}{\zeta_t^{(k)}}\right)\right] + \mathbf{E}_t^{\zeta}(U_{t+\Delta t}) \\
= \;& \left(1 + \int_t^{t+\Delta t} \mathbf{E}_t^{\zeta}[H_s]ds\right) \mathbf{E}_t^{\zeta}(U_{t+\Delta t}).
\end{aligned}$$

This result suggests that for a given H and a small enough Δt, $1+\int_t^{t+\Delta t} \mathbf{E}_t^{\zeta}[H_s]ds$ must be positive almost surely in order that the above function is a well-defined utility function. Or equivalently, $\int_t^{t+\Delta t} \mathbf{E}_t^{\zeta}[|H_s|]ds < 1$ for a small enough Δt. This is guaranteed by the second condition in Assumption A because

$$\int_t^{t+\Delta t} \mathbf{E}_t^{\zeta}[|H_s|]ds \leq \int_t^{t+\Delta t} \sup_{t\in[0,T]} \mathbf{E}_t^{\zeta}(\sup_{s\in[0,T]} |H_s|)ds \leq C_H \Delta t.$$

Note that by the above result, for sufficiently small Δt, $1+\int_t^{t+\Delta t} \mathbf{E}_t^{\zeta}[H_s]ds = e^{\int_t^{t+\Delta t} \mathbf{E}_t^{\zeta}[H_s]ds + O((\Delta t)^2)}$ since $ln(1+x) = x+O(x^2)$ for small x. Plugging $U_{t+\Delta t} = e^{\int_{t+\Delta t}^{t+2\Delta t} \mathbf{E}_{t+\Delta t}^{\zeta}[H_s]ds + O((\Delta t)^2)} \mathbf{E}_{t+\Delta t}^{\zeta}(U_{t+2\Delta t})$

[8] It is worth mentioning that, as shown below, the second assumption is made for the proof of (10) and the main result (11) still holds true without this assumption.

into the equation above, we have

$$\begin{aligned} U_t &= \mathbf{E}_t^{\zeta}(e^{\int_t^{t+\Delta t}\mathbf{E}_t^{\zeta}[H_s]ds+O((\Delta t)^2)}e^{\int_{t+\Delta t}^{t+2\Delta t}\mathbf{E}_{t+\Delta t}^{\zeta}[H_s]ds+O((\Delta t)^2)}\mathbf{E}_{t+\Delta t}^{\zeta}(U_{t+2\Delta t})) \\ &= \mathbf{E}_t^{\zeta}(e^{\int_t^{t+\Delta t}\mathbf{E}_t^{\zeta}[H_s]ds+\int_{t+\Delta t}^{t+2\Delta t}\mathbf{E}_{t+\Delta t}^{\zeta}[H_s]ds+2O((\Delta t)^2)}U_{t+2\Delta t}) \\ &\quad ... \\ &= \mathbf{E}_t^{\zeta}(e^{\int_t^{t+\Delta t}\mathbf{E}_t^{\zeta}[H_s]ds+...+\int_{T-\Delta t}^{T}\mathbf{E}_{T-\Delta t}^{\zeta}[H_s]ds+O(\Delta t)}U_T). \end{aligned}$$

Note that for a fixed $s \in [0,T]$, $\lim_{t\to s, t<s} E_t^{\zeta}[H_s] = E_{s-}^{\zeta}[H_s] = H_s$ since, by Assumption A, H is a continuous function. Hence, by the dominated convergence theorem on the interval $[0,T]$, we have

$$\lim_{\Delta t\to 0}\left[\int_t^{t+\Delta t}\mathbf{E}_t^{\zeta}[H_s]ds + ... + \int_{T-\Delta t}^{T}\mathbf{E}_{T-\Delta t}^{\zeta}[H_s]ds\right] = \int_t^T H_s ds,$$

almost surely since, by Assumption A, $E_t^{\zeta}[H_s] \le C_H$ for all $t,s \in [0,T]$. Thus, by the dominated convergence theorem, we can derive the continuous-time version of the utility function which is given by

$$U_t = \inf_{\zeta}\mathbf{E}_t^{\zeta}\left[e^{\int_t^T H_s ds}U_T\right],$$

since $H_s \le 0$ and $\mathbf{E}_t^{\zeta}[|U_T|] < \infty$.

For illustrative convenience, we suppose $n-d \equiv 1$, that is, there is only one type of jump in the remainder of this proof. We omit the subscript k from now on by letting, for example, $\lambda_k \equiv \lambda$, $A_k \equiv A$. Furthermore, we use the upper case Z to denote random jump size and the lower case z to denote a particular realization of Z.

In the following, we prepare for the proof of (11) by using a minimax theorem given in Lemma A5 below. For this purpose, we now verify the conditions of the lemma by presenting Lemma A2, A3 and A4. Given a constant $C > 0$, we define

$$\widetilde{\Psi}_C = \left\{\vartheta\psi\middle|(\vartheta,\psi)\in\Theta\times\Psi, \vartheta(s)\psi(s,z)\le C, \forall s\in[0,T], \forall z\in A\right\}.$$

Lemma A2 *Given a sequence $\{\vartheta_n(s)\psi_n(s,Z), n = 1,2,...\}$ in $\widetilde{\Psi}_C$, there exists a sequence of convex combination $\widetilde{\vartheta}_n(s)\widetilde{\psi}_n(s,Z) \in conv\{\vartheta_n(s)\psi_n(s,Z), \vartheta_{n+1}(s)\psi_{n+1}(s,Z), ...\} \in \widetilde{\Psi}_C$ such that $\{\widetilde{\vartheta}_n(s)\widetilde{\psi}_n(s,Z)\}$ converges $P\times l$ a.s.to a $\vartheta_0(s)\psi_0(s,Z)\in\widetilde{\Psi}_C$ and $\{\widetilde{\vartheta}_n(t+\Delta t)\widetilde{\psi}_n(t+\Delta t,Z)\}$ converges P a.s.to a $\vartheta_0(t+\Delta t)\psi_0(t+\Delta t,Z)$, where l denotes the Lebesgue measure on the interval $[0,T]$.*

Proof. *For the sequence $\{\vartheta_n(s)\psi_n(s,Z)\}$ in $\widetilde{\Psi}_C$, like the proof of Lemma 3.2 in Schied and Wu [31], by Lemma A1.1 of Delbaen and Schachermayer [11], there exists a sequence of convex combination $\bar{\vartheta}_n(s)\bar{\psi}_n(s,Z) \in conv\{\vartheta_n(s)\psi_n(s,Z), \vartheta_{n+1}(s)\psi_{n+1}(s,Z), ...\} \in \widetilde{\Psi}_C$ which converges $P\times l$ a.s.to a $\vartheta_0(s)\psi_0(s,Z)\in\widetilde{\Psi}_k$. In the same manner, considering the sequence $\{\bar{\vartheta}_n(t+\Delta t)\bar{\psi}_n(t+\Delta t,Z)\}$, there is a sequence $\widetilde{\vartheta}_n(t+\Delta t)\widetilde{\psi}_n(t+\Delta t,Z) \in conv\{\bar{\vartheta}_n(t+\Delta t)\bar{\psi}_n(t+\Delta t,Z), \bar{\vartheta}_{n+1}(t+\Delta t)\bar{\psi}_{n+1}(t+\Delta t,Z), ...\}$ which converges P a.s.to a $\vartheta_0(t+\Delta t)\psi_0(t+\Delta t,z)$. Furthermore, $\{\widetilde{\vartheta}_n(s)\widetilde{\psi}_n(s,Z)\}$ converges $P\times l$ a.s.to a $\vartheta_0(s)\psi_0(s,Z)$ since $\{\widetilde{\vartheta}_n(s)\widetilde{\psi}_n(s,Z)\}$ is a convex combination of $\{\bar{\vartheta}_n(s)\bar{\psi}_n(s,Z)\}$.* ■

We let $\Delta N(t) = N(t+\Delta t) - N(t)$ denote the number of jumps in the interval $(t, t+\Delta t]$. For $\vartheta(s)\psi(s) \in \widetilde{\Psi}_C$, we define

$$\begin{aligned}\widetilde{H}_t &= \left(1 + H_t\Delta t + \lambda\int_t^{t+\Delta t}\int_A(1-\vartheta(s)\psi(s,z))\Phi(dz)ds\right)\mathbf{1}(\mathbf{\Delta N(t) = 0}) \\ &\quad +\vartheta(t+\Delta t)\psi(t+\Delta t, Z)\mathbf{1}(\mathbf{\Delta N(t) = 1}). \qquad (A.1)\end{aligned}$$

and, for $\delta > 0$,

$$\widetilde{\Phi}_C = \widetilde{\Phi}_C(\delta) = \left\{\widetilde{H}_t \middle| \vartheta(s)\psi(s) \in \widetilde{\Psi}_C, \sup_{|t_1-t_2|\le\Delta t}|H_{t_1} - H_{t_2}| \le \delta\right\},$$

where $\mathbf{1}(\cdot)$ denotes the indicator function. And moreover, we use $\bar{\bar{\Phi}}_C$ to denote the weak closure of the set $\widetilde{\Phi}_C$ in $L^1(P)$. The following result is Theorem 7.5.10 in Yan [34].

Lemma A3 *Let $\mathcal{H}$ be a subset of $L^1(P)$. Then the following two conditions are equivalent:*

(1) $\mathcal{H}$ is a uniformly integrable family;

(2) For any sequence $\{X_n\}$ in $\mathcal{H}$, there exists a subsequence $\{X_{n_k}\}$ of $\{X_n\}$ and a random variable $X_0 \in L^1(P)$ such that $\{X_{n_k}\}$ weakly converges to X_0 in $L^1(P)$, that is, $\lim_{k\to\infty} E(X_{n_k}Y) = E(X_0Y)$ for any bounded random variable Y.

Proof. *See Theorem 7.5.10 in Yan [34].* ■

Lemma A4 *The set $\bar{\bar{\Phi}}_C$ is weakly compact in $L^1(P)$.*

Proof. *Like Lemma 3.2 in Schied and Wu [31], we use James' theorem to prove this result. For this, we let $F \in L^\infty(P)$ and $\{\widetilde{H}_{n,t} \in \bar{\bar{\Phi}}_C\}$ such that $\mathbf{E}[\widetilde{H}_{n,t}F]$ tends to $\sup_{\widetilde{H}_t\in\bar{\bar{\Phi}}_C}\mathbf{E}[\widetilde{H}_tF]$. Without loss of the generality, we assume that $\{\widetilde{H}_{n,t} \in \widetilde{\Phi}_C\}$. Note that the set $\widetilde{\Phi}_C$ is bounded. Hence it is uniformly integrable. Thus, by Lemma A3, there is a subsequence $\{\widetilde{H}_{n_k,t}\}$ of $\{\widetilde{H}_{n,t}\}$ such that it weakly converges to $\widetilde{H}_{0,t}$ in $L^1(P)$. This implies that $\widetilde{H}_{0,t} \in \bar{\bar{\Phi}}_C$ and $\mathbf{E}[\widetilde{H}_{0,t}F] = \sup_{\widetilde{H}_t\in\bar{\bar{\Phi}}_k}\mathbf{E}[\widetilde{H}_tF]$. That is, the continuous linear functional $\mathbf{E}[\widetilde{H}_tF]$ attains its sup on $\bar{\bar{\Phi}}_C$. As a result, by the James' theorem, the set $\bar{\bar{\Phi}}_C$ is weakly compact in $L^1(P)$.* ■

We now turn to minimax results. A function $f : X \times Y \to R$ is said to be *convex-concave like* on $X \times Y$ if, for $\alpha \in [0,1]$,

(1) for x_1 and x_2 in X, there exists x_3 in X such that

$$f(x_3, y) \le \alpha f(x_1, y) + (1-\alpha)f(x_2, y)$$

for all y in Y; and

(2) for y_1 and y_2 in Y, there exists y_3 in Y such that

$$f(x, y_3) \ge \alpha f(x, y_1) + (1-\alpha)f(x, y_2)$$

for all x in X.

The following result plays a key role in the proof of Proposition 1.

Lemma A5 *Suppose X and Y are non-empty sets with f convex-concave like on $X \times Y$. Suppose that X is compact and $f(\cdot, y)$ is lower semicontinuous on X for each y in Y. Then*

$$\min_X \sup_Y f(x,y) = \sup_Y \min_X f(x,y).$$

Proof. *See Theorem A of Borwein and Zhuang [6].* ∎

Lemma A6 *For $\epsilon > 0$, the function $\mathbf{E}_t[\widetilde{H}_t U^\epsilon_{t+\Delta t}]$ is convex-concave like on $\bar{\widetilde{\Phi}}_C \times \mathcal{W}(w_0)$, where U^ϵ_t is defined by (7) with $U^\epsilon_T = \frac{(W_T+\epsilon)^{1-\gamma}}{1-\gamma}$ and we also replace $\int_t^{t+\Delta t} H_s ds$ with $H_t \Delta t$.*

Proof. *Note that the set $\mathcal{W}(w_0)$ is convex and $\widetilde{H}_t$ is positive when Δt is small enough. Hence, the function $\mathbf{E}_t[\widetilde{H}_t U^\epsilon_{t+\Delta t}]$ is concave on $\mathcal{W}(w_0)$ for all $\widetilde{H}_t \in \bar{\widetilde{\Phi}}_C$. We now show the first condition of a convex-concave like function. For this, let $\widetilde{H}^{(1)}_t, \widetilde{H}^{(2)}_t \in \bar{\widetilde{\Phi}}_C$ and two sequences $\{\widetilde{H}^{(1)}_{n,t}\}$ and $\{\widetilde{H}^{(2)}_{n,t}\}$ in $\widetilde{\Phi}_C$ such that $\{\widetilde{H}^{(1)}_{n,t}\}$ and $\{\widetilde{H}^{(2)}_{n,t}\}$ weakly converge to $\widetilde{H}^{(1)}_t$ and $\widetilde{H}^{(2)}_t$, respectively. In particular,*

$$\begin{aligned} &\lim_{n\to\infty} \mathbf{E}_t[(\alpha \widetilde{H}^{(1)}_{n,t} + (1-\alpha)\widetilde{H}^{(2)}_{n,t})U^\epsilon_{t+\Delta t}] \\ =\ & \mathbf{E}_t[(\alpha \widetilde{H}^{(1)}_t + (1-\alpha)\widetilde{H}^{(2)}_t)U^\epsilon_{t+\Delta t}], \end{aligned}$$

since $U^\epsilon_{t+\Delta t}$ is bounded. And thus, given $\epsilon_0 > 0$, there exists n_0 such that for $n \geq n_0$,

$$\begin{aligned} & \mathbf{E}_t[(\alpha \widetilde{H}^{(1)}_{n,t} + (1-\alpha)\widetilde{H}^{(2)}_{n,t})U^\epsilon_{t+\Delta t}] \\ \leq\ & \mathbf{E}_t[(\alpha \widetilde{H}^{(1)}_t + (1-\alpha)\widetilde{H}^{(2)}_t)U^\epsilon_{t+\Delta t}] + \epsilon_0. \end{aligned}$$

We let $\vartheta^{(1)}_n \psi^{(1)}_n$ and $\vartheta^{(2)}_n \psi^{(2)}_n$ define $\widetilde{H}^{(1)}_{n,t}$ and $\widetilde{H}^{(2)}_{n,t}$, respectively. Furthermore, we let $\widetilde{H}^{(3)}_{n,t}$ be defined by $\alpha\vartheta^{(1)}_n \psi^{(1)}_n + (1-\alpha)\vartheta^{(2)}_n \psi^{(2)}_n = \vartheta^{(3)}_n \psi^{(3)}_n$. Note that $\mathbf{E}_t[\widetilde{H}_t U^\epsilon_{t+\Delta t}]$ is a convex function with respect to $\vartheta_n \psi_n$ since from the definition (A.1), $\widetilde{H}_t$ is a concave function with respect to $\vartheta_n \psi_n$ and $U^\epsilon_{t+\Delta t}$ is negative. Hence we have

$$\mathbf{E}_t[\widetilde{H}^{(3)}_{n,t} U^\epsilon_{t+\Delta t}] \leq E_t[(\alpha \widetilde{H}^{(1)}_{n,t} + (1-\alpha)\widetilde{H}^{(2)}_{n,t})U^\epsilon_{t+\Delta t}],$$

Applying Lemma A2 to the sequence $\vartheta^{(3)}_n \psi^{(3)}_n$, we can find there a sequence of convex combination $\vartheta^{(4)}_n \psi^{(4)}_n \in conv\{\vartheta^{(3)}_n \psi^{(3)}_n, \vartheta^{(3)}_{n+1} \psi^{(3)}_{n+1}, ...\} \in \widetilde{\Psi}_C$ such that $\{\vartheta^{(4)}_n \psi^{(4)}_n\}$ converges $P \times l$ a.s.to a $\vartheta^{(4)}_0 \psi^{(4)}_0$ and $\{\vartheta^{(4)}_n(t+\Delta t)\psi^{(4)}_n(t+\Delta t)\}$ converges P a.s.to a $\vartheta^{(4)}_0(t+\Delta t)\psi^{(4)}_0(t+\Delta t)$.

Finally, by noticing that the sequence $\{\vartheta^{(4)}_n \psi^{(4)}_n\}$ is bounded, the corresponding $\{\widetilde{H}^{(4)}_{n,t}\}$ is bounded. And thus, for $\vartheta^{(4)}_0(s)\psi^{(4)}_0(s)$, the corresponding $\widetilde{H}^{(4)}_t \in \bar{\widetilde{\Phi}}_C$ and

$$\mathbf{E}_t[\widetilde{H}^{(4)}_t U^\epsilon_{t+\Delta t}] \leq \mathbf{E}_t[(\alpha \widetilde{H}^{(1)}_t + (1-\alpha)\widetilde{H}^{(2)}_t)U^\epsilon_{t+\Delta t}],$$

for all $W_T \in \mathcal{W}(w_0)$, completing the proof. $\blacksquare$

Lemma A7 *Given $\Delta t > 0$, $\delta > 0$ and $\epsilon > 0$, we have*

$$\sup_{W_T \in \mathcal{W}(w_0)} \inf_{\widetilde{H}_t \in \tilde{\bar{\Phi}}_C} \mathbf{E}_t[\widetilde{H}_t U^{\epsilon}_{t+\Delta t}] = \inf_{\widetilde{H}_t \in \tilde{\bar{\Phi}}_C} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t[\widetilde{H}_t U^{\epsilon}_{t+\Delta t}].$$

Proof. *We get the result by using Lemma A4, Lemma A5 and Lemma A6.* $\blacksquare$*:*

Next we show that

$$\begin{aligned} J(t, W_t, X_t) &= \sup_{W \in \mathcal{W}(w_0)} U_t \\ &= \sup_{W_T \in \mathcal{W}(w_0)} \inf_{\zeta} \mathbf{E}^{\zeta}_t \left[e^{\int_t^T H_s ds} u(W_T) \right] = \inf_{\zeta} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}^{\zeta}_t \left[e^{\int_t^T H_s ds} u(W_T) \right]. \end{aligned}$$

As in the proof of Lemma 3.4 of Schied and Wu [31], we have

$$U(w_0 + \epsilon) \geq \sup_{W_T \in \mathcal{W}(w_0)} \inf_{\zeta} \mathbf{E}^{\zeta}_t \left[e^{\int_t^T H_s ds} u(W_T + \epsilon) \right]$$

Define

$$\bar{H}(\Delta t) = (H_t + ... + H_{T-\Delta t})\Delta t.$$

We let ζ^* denote the optimal solution to the above optimal problem on the right hand side. Given the optimal $\vartheta^* \psi^*$ and the corresponding H^*_t, by Assumption A, for any $\delta > 0$ and $\epsilon_1 > 0$, there exists a $\Delta t^* > 0$ and $C > 0$ such that

$$P\left\{ \sup_{|t_1 - t_2| \leq \Delta t^*} |H^*_{t_1} - H^*_{t_2}| \geq \delta, \text{or} \sup_{t \in [0,T]} \vartheta^*(t)\psi^*(t) \geq C \right\} \leq \epsilon_1.$$

Then, by letting $h_1(H, \Delta t^*) = \sup_{|t_1 - t_2| \leq \Delta t^*} |H_{t_1} - H_{t_2}|$ and $h_2(\vartheta, \psi) = \sup_{t \in [0,T]} \vartheta(t)\psi(t)$, we have

$$\begin{aligned} & U(w_0 + \epsilon) \\ \geq & \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}^{\zeta^*}_t \left[e^{\int_t^T H^*_s ds} u(W_T + \epsilon) \right] \\ \geq & \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}^{\zeta^*}_t \left[e^{\int_t^T H^*_s ds} u(W_T + \epsilon) : h_1(H^*, \Delta t^*) \leq \delta, \text{and}, h_2(\vartheta^*, \psi^*) \leq C \right] - \epsilon_2 \\ \geq & \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}^{\zeta^*}_t \left[e^{\bar{H}^*(\Delta t^*)} u(W_T + \epsilon) : h_1(H^*, \Delta t^*) \leq \delta, \text{and}, h_2(\vartheta^*, \psi^*) \leq C \right] - \epsilon_3 \\ \geq & \sup_{W_T \in \mathcal{W}(w_0)} \inf_{\widetilde{H} \in \tilde{\bar{\Phi}}_C(\delta)} \mathbf{E}^{\zeta}_t \left[e^{\bar{H}(\Delta t^*)} u(W_T + \epsilon) \right] - \epsilon_3, \end{aligned}$$

where the last inequality follows from the definition of the set $\tilde{\bar{\Phi}}_C(\delta)$. The variables $\epsilon_i, i = 1, 2, 3$ above and $\epsilon_i, i = 4, 5, 6$ below can be made arbitrarily small by letting Δt^*

tend to zero. We will show that

$$\begin{aligned}
&\Big| \sup_{W_T \in \mathcal{W}(w_0)} \inf_{\widetilde{H} \in \bar{\widetilde{\Phi}}_C(\delta)} \mathbf{E}_t^{\zeta} \left[e^{\bar{H}(\Delta t^*)} u(W_T + \epsilon) \right] \\
&- \inf_{\widetilde{H} \in \bar{\widetilde{\Phi}}_C(\delta)} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\bar{H}(\Delta t^*)} u(W_T + \epsilon) \right] \Big| \le \epsilon_4,
\end{aligned} \tag{A.2}$$

and

$$\begin{aligned}
&\inf_{\widetilde{H} \in \bar{\widetilde{\Phi}}_C(\delta)} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\bar{H}(\Delta t^*)} u(W_T + \epsilon) \right] \\
= &\inf_{\widetilde{H} \in \widetilde{\Phi}_C(\delta)} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\bar{H}(\Delta t^*)} u(W_T + \epsilon) \right].
\end{aligned} \tag{A.3}$$

Thus,

$$\begin{aligned}
U(w_0 + \epsilon) &\ge \inf_{\widetilde{H} \in \bar{\widetilde{\Phi}}_C(\delta)} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\bar{H}(\Delta t^*)} u(W_T + \epsilon) \right] - \epsilon_5 \\
&= \inf_{\widetilde{H} \in \widetilde{\Phi}_C(\delta)} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\bar{H}(\Delta t^*)} u(W_T + \epsilon) \right] - \epsilon_5 \\
&\ge \inf_{\widetilde{H} \in \widetilde{\Phi}_C(\delta)} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\int_t^T H_s ds} u(W_T + \epsilon) \right] - \epsilon_6 \\
&\ge \inf_{\zeta} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\int_t^T H_s ds} u(W_T + \epsilon) \right] - \epsilon_6 \\
&\ge \inf_{\zeta} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\int_t^T H_s ds} u(W_T) \right] - \epsilon_6 \\
&\ge \sup_{W_T \in \mathcal{W}(w_0)} \inf_{\zeta} \mathbf{E}_t^{\zeta} \left[e^{\int_t^T H_s ds} u(W_T) \right] - \epsilon_6 \\
&= U(w_0) - \epsilon_6.
\end{aligned}$$

Furthermore, by using the same manner as in Theorem 2.2 and Lemma 3.1 in Schied and Wu [31], we can show that the value function $U(x)$ is concave and continuous. As a result, by letting Δt, δ and ϵ go to zero, we obtain

$$U(w_0) = \sup_{W_T \in \mathcal{W}(w_0)} \inf_{\zeta} \mathbf{E}_t^{\zeta} \left[e^{\int_t^T H_s ds} u(W_T) \right] = \inf_{\zeta} \sup_{W_T \in \mathcal{W}(w_0)} \mathbf{E}_t^{\zeta} \left[e^{\int_t^T H_s ds} u(W_T) \right].$$

The following result is used in the proof of (A.2) and (A.3).

Lemma A8 *We have*

$$\mathbf{E}_t \left(\frac{\zeta_{t+\Delta t}}{\zeta_t} U^{\epsilon}_{t+\Delta t} \mathbf{1}(\mathbf{\Delta N(t) \ge 2}) \right) \le \frac{\epsilon^{1-\gamma}}{\gamma - 1} \exp\left(\frac{(2 + C^3)\lambda}{3} \right) (\Delta t)^{\frac{4}{3}}.$$

Proof. *From the proof of Lemma A1, we have*

$$d\ln(\zeta_t) = \lambda \int_A [1 - \vartheta(t)\psi(t,z)]\Phi(dz)dt + \int_A ln(\vartheta(t)\psi(t,z))dN(t),$$

implying that

$$\frac{\zeta_{t+\Delta t}}{\zeta_t} = \exp\left(\lambda \int_A [1 - \vartheta(t)\psi(t,z)]\Phi(dz)dt + \int_A ln(\vartheta(t)\psi(t,z))dN(t)\right).$$

Note that

$$\mathbf{E}_t\left\{\exp\left(\lambda \int_A [1 - \vartheta^3(t)\psi^3(t,z)]\Phi(dz)dt + \int_A ln[\vartheta^3(t)\psi^3(t,z)]dN(t)\right)\right\} = 1.$$

Hence,

$$\begin{aligned}
& \mathbf{E}_t\left(\frac{\zeta_{t+\Delta t}}{\zeta_t}\right)^3 \\
= \;& \mathbf{E}_t\left\{\exp\left(3\lambda \int_A [1 - \vartheta(t)\psi(t,z)]\Phi(dz)dt + \int_A ln[\vartheta^3(t)\psi^3(t,z)]dN(t)\right)\right\} \\
= \;& \mathbf{E}_t\left\{\exp\left(\lambda \int_A [2 - 3\vartheta(t)\psi(t,z) + \vartheta^3(t)\psi^3(t,z)]\Phi(dz)dt\right)\right\} \\
\leq \;& \exp\left((2 + C^3)\lambda\right).
\end{aligned}$$

As a result

$$\begin{aligned}
& \left|\mathbf{E}_t\left(\frac{\zeta_{t+\Delta t}}{\zeta_t}U^{\epsilon}_{t+\Delta t}\mathbf{1}(\boldsymbol{\Delta}\mathbf{N}(\mathbf{t}) \geq \mathbf{2})\right)\right| \\
\leq \;& \frac{\epsilon^{1-\gamma}}{\gamma - 1}\left|\mathbf{E}_t\left(\frac{\zeta_{t+\Delta t}}{\zeta_t}\mathbf{1}(\boldsymbol{\Delta}\mathbf{N}(\mathbf{t}) \geq \mathbf{2})\right)\right| \\
\leq \;& \frac{\epsilon^{1-\gamma}}{\gamma - 1}\left(\mathbf{E}_t\left(\frac{\zeta_{t+\Delta t}}{\zeta_t}\right)^3\right)^{\frac{1}{3}}\left(P(\Delta N(t) \geq 2)\right)^{\frac{2}{3}} \\
\leq \;& \frac{\epsilon^{1-\gamma}}{\gamma - 1}\exp\left(\frac{(2 + C^3)\lambda}{3}\right)(\Delta t)^{\frac{4}{3}}
\end{aligned}$$

■

We now turn to proving (A.2). Note that, by dynamic programming, the problem "$\sup_{W_T \in \mathcal{W}(w_0)} \inf_{\widetilde{H} \in \bar{\tilde{\Phi}}_C(\delta)}$" in (A.2) can be approximately solved below with total approximation error $O(\Delta t)$.

$$
\begin{aligned}
& \sup_{W_T\in\mathcal{W}(w_0)} \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} \Lambda\Big(\mathbf{E}_t^\zeta(U^\epsilon_{t+\Delta t})\Big)H_t\Delta t + \mathbf{E}_t^\zeta(U^\epsilon_{t+\Delta t}) \\
= & \sup_{W_T\in\mathcal{W}(w_0)} \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} (1+H_t\Delta t)\mathbf{E}_t^\zeta(U_{t+\Delta t}) = \sup_\pi \inf_\zeta (1+H_t\Delta t)\mathbf{E}_t\left(\frac{\zeta_{t+\Delta t}}{\zeta_t}U^\epsilon_{t+\Delta t}\right) \\
= & \sup_{W_T\in\mathcal{W}(w_0)} \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} (1+H_t\Delta t)\mathbf{E}_t\left[e^{\lambda\int_t^{t+\Delta t}\int_A(1-\vartheta\psi(z))\Phi(dz)ds}\prod_{i=1}^{\Delta N(t)}\vartheta\psi(z_i)U^\epsilon_{t+\Delta t}\right] \\
= & \sup_{W_T\in\mathcal{W}(w_0)} \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} (1+H_t\Delta t)\mathbf{E}_t\left[e^{\lambda\int_t^{t+\Delta t}\int_A(1-\vartheta\psi(z))\Phi(dz)ds}U^\epsilon_{t+\Delta t}|\Delta N(t)=0\right]e^{-\lambda\Delta t} \\
& +(1+H_t\Delta t)\mathbf{E}_t\left[e^{\lambda\int_t^{t+\Delta t}\int_A(1-\vartheta\psi(z))\Phi(dz)ds}\vartheta\psi(z_1)U^\epsilon_{t+\Delta t}|\Delta N(t)=1\right](\lambda\Delta t)\,e^{-\lambda\Delta t} \\
& +(1+H_t\Delta t)\mathbf{E}_t\left(\frac{\zeta_{t+\Delta t}}{\zeta_t}U^\epsilon_{t+\Delta t}\mathbf{1}(\mathbf{\Delta N(t)}\geq\mathbf{2})\right).
\end{aligned}
$$

Thus, by Lemma A8,

$$
\begin{aligned}
& \sup_{W_T\in\mathcal{W}(w_0)} \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} \Lambda\Big(\mathbf{E}_t^\zeta(U^\epsilon_{t+\Delta t})\Big)H_t\Delta t + \mathbf{E}_t^\zeta(U^\epsilon_{t+\Delta t}) \\
= & \sup_{W_T\in\mathcal{W}(w_0)} \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} \Bigg\{\mathbf{E}_t\left[\left(1+H_t\Delta t+\lambda\int_t^{t+\Delta t}\int_A(1-\vartheta\psi(z))\Phi(dz)ds\right)U^\epsilon_{t+\Delta t}|\Delta N(t)=0\right]e^{-\lambda\Delta t} \\
& +\mathbf{E}_t\left[\vartheta\psi(z_1)U^\epsilon_{t+\Delta t}|\Delta N(t)=1\right]\lambda\Delta t e^{-\lambda\Delta t}\Bigg\}+o(\Delta t) \\
= & \sup_{W_T\in\mathcal{W}(w_0)} \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} \mathbf{E}_t\Bigg\{\Bigg[\left(1+H_t\Delta t+\lambda\int_t^{t+\Delta t}\int_A(1-\vartheta(s)\psi(s,z))\Phi(dz)ds\right)\mathbf{1}(\mathbf{\Delta N(t)}=\mathbf{0}) \\
& +\vartheta(t+\Delta t)\psi(t+\Delta t,Z)\mathbf{1}(\mathbf{\Delta N(t)}=\mathbf{1})\Bigg]U^\epsilon_{t+\Delta t}\Bigg\}+o(\Delta t) \\
= & \sup_{W_T\in\mathcal{W}(w_0)} \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} E_t[\widetilde{H}_tU^\epsilon_{t+\Delta t}]+o(\Delta t).
\end{aligned}
$$

Likewise, the problem "$\inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)}\sup_{W_T\in\mathcal{W}(w_0)}$" in (A.2) can be approximately solved by the following dynamic programing with total approximation error $O(\Delta t)$:

$$
\begin{aligned}
& \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} \sup_{W_T\in\mathcal{W}(w_0)} \Lambda\Big(\mathbf{E}_t^\zeta(U^\epsilon_{t+\Delta t})\Big)H_t\Delta t + \mathbf{E}_t^\zeta(U_{t+\Delta t}) \\
= & \inf_{\widetilde{H}\in\bar{\widetilde{\Phi}}_C(\delta)} \sup_{W_T\in\mathcal{W}(w_0)} \mathbf{E}_t[\widetilde{H}_tU^\epsilon_{t+\Delta t}]+o(\Delta t),
\end{aligned}
$$

implying (A.2) by Lemma A7. The result (A.3) follows from the facts that $\bar{\widetilde{\Phi}}_C(\delta)$ is the weak closure of $\widetilde{\Phi}_C(\delta)$ and $U_{t+\Delta t}(W_T+\epsilon)$ is bounded.

This completes the proof of (11). ■

B Results for Logarithm Utility Function

For the logarithm utility function, we let $\Gamma(x) = 1$ as suggested by the literature. Then we have the following proposition corresponding to Proposition 1 for the power utility function. (For illustration, we suppose $k = 1$, i.e., only one jump, and the subscriptions regarding k are ignored.)

Proposition 1' When $u(x) = \ln(x)$, we let $\Gamma(x) = 1$, then the continuous-time version of the utility in equation (7) is given by

$$U_t = \inf_{\zeta} \mathbf{E}_t^{\zeta} \left[\int_t^T H_s ds + \ln(W_T) \right],$$

where $H = \frac{\lambda}{\phi} \int_A [\vartheta\psi \ln(\vartheta\psi) + 1 - \vartheta\psi] \Phi(dz)$.

To prove (11) for $u(x) = \ln(x)$, we need to slightly modify the method of proof of the result (11) in Proposition 1 for CRRA utility functions because, unlike a CRRA utility function with $\gamma > 1$, the log utility function $\ln(W_T)$ is unbounded above. Specifically, (11) can be recovered for the log utility function $\ln(W_T)$ by using the following results:

$$\begin{aligned} \sup_{\xi \in \mathcal{Q}_\zeta} \inf_{\zeta} \mathbf{E}_t^{\zeta} \left[\int_t^T H_s ds + \ln(W_T) \right] &= \sup_n \sup_{\xi \in \mathcal{Q}_\zeta} \inf_{\zeta} \mathbf{E}_t^{\zeta} \left[\int_t^T H_s ds + \min\{\ln(W_T), n\} \right], \\ \inf_{\zeta} \sup_{\xi \in \mathcal{Q}_\zeta} \mathbf{E}_t^{\zeta} \left[\int_t^T H_s ds + \ln(W_T) \right] &= \sup_n \inf_{\zeta} \sup_{\xi \in \mathcal{Q}_\zeta} \mathbf{E}_t^{\zeta} \left[\int_t^T H_s ds + \min\{\ln(W_T), n\} \right]. \end{aligned}$$

Moreover, following the proof for CRRA utility functions with $\gamma > 1$ in Appendix A, we can show that for any n

$$\sup_{\xi \in \mathcal{Q}_\zeta} \inf_{\zeta} \mathbf{E}_t^{\zeta} \left[\int_t^T H_s ds + \min\{\ln(W_T), n\} \right] = \inf_{\zeta} \sup_{\xi \in \mathcal{Q}_\zeta} \mathbf{E}_t^{\zeta} \left[\int_t^T H_s ds + \min\{\ln(W_T), n\} \right].$$

Then we have a theorem corresponding to Theorem 1.

Theorem 1' When $u(x) = \ln(x)$, we have

$$J(t, W_t, X_t) = \ln(W_t) + \mathbf{E}_t \left[\frac{1}{2} \int_t^T ||\theta^b||^2 ds + \int_t^T r(X_s) ds - \int_t^T \inf_c G(s, c) ds \right],$$

where

$$G(s, c) = -c\theta^q - \frac{\lambda}{\phi} \int_A (1 - e^{\chi(z,c)}) \Phi(s, dz),$$

and

$$\chi(z, c) = (1 + cz)^{-\phi}.$$

The worst probability is given by the same formulas as in Theorem 1.

Proof. We prove Theorem 1' briefly. First, applying the convex conjugate approach after Lemma 1, we obtain

$$J(t, W_t, X_t) = \ln(W_t) + \inf_{\zeta} \inf_{\xi \in \mathcal{Q}_\zeta} \mathbf{E}_t^\zeta \left[\int_t^T H_s ds - \ln(\xi_\delta) + \int_t^T r(X_s) ds \right].$$

Hence, the objective function in the optimization (25) is rewritten as:

$$\begin{aligned} & f(z, \theta(t), \varphi(z), \vartheta(t), \psi(z)) \\ = \ & -x_1(z) + x_2(z)\ln(x_1(z)) + x_2(z) - x_2(z)\ln(x_2(z)) - \frac{1}{\phi}\left(x_2(z)\ln(x_2(z)) + 1 - x_2(z)\right). \end{aligned}$$

The definition of g_0 in (29) becomes:

$$g_0(x) \quad = \quad -x_1 + x_2\ln(x_1) + x_2 - x_2\ln(x_2) - \frac{1}{\phi}(x_2\ln(x_2) + 1 - x_2),$$

if $x_1 \geq 0, x_2 \geq 0$, and $-\infty$ otherwise. Then corresponding to (32), we have

$$\begin{aligned} \hat{x}_1 &= x_2(1+cz)^{-1}, \\ \hat{x}_2 &= (1+cz)^{-\phi_k}, \end{aligned}$$

and

$$\inf_{x_1, x_2} cx_1 z - g_0(x) = \frac{1}{\phi}(1 - (1+cz)^{-\phi}).$$

Corresponding to Lemma 4, we have

$$\begin{aligned} & \sup_{(x_1,x_2)} \int_A -x_1(z) + x_2(z)\ln(x_1(z)) + x_2(z) - x_2(z)\ln(x_2(z)) \\ & - \frac{1}{\phi}[x_2(z)\ln(x_2(z)) + 1 - x_2(z)]\Phi(dz) \\ & = \inf_c \ -\frac{c\theta^q}{\lambda} - \frac{1}{\phi}\int_A (1 - (1+cz)^{-\phi})\Phi(dz). \end{aligned}$$

If we define $e^{\chi(c,z)} = (1+cz)^{-\phi}$, then the worst probability for the logarithm utility is given by the same expressions as (15) and the equation before (15) in Theorem 1, where c^* solves the minimization problem in the above. This completes the proof. ∎

References

[1] Aït-Sahalia, Y., Cacho-Diaz, J., Hurd, T., 2009. Portfolio choice with jumps: A closed-form solution. Annals of Applied Probability 19, 556-584.

[2] Anderson, E., Hansen, L., Sargent, T., 2003. A quartet of semigroups for model

specification, robustness, prices of risk and model detection. Journal of the European Economic Association 1, 68-123.

[3] Bajeux-Besnainou, I., Jordan, J., Portait, R., 2003. Dynamic asset allocation for stocks, bonds and cash. Journal of Business 76, 263-287.

[4] Bardhan, I., Chao, X., 1996. On martingale measures when asset returns have unpredictable jumps. Stochastic Processes and their Applications 63, 35-54.

[5] Bellini, F., Frittelli, M., 2002. On the existence of minimax martingale measures. Mathematical Finance 12, 1-21.

[6] Borwein, J., Zhuang, D., 1986. On Fan's minimax theorem. Mathematical programming 34, 232-234.

[7] Branger, N., Larsen, L., 2013. Robust portfolio choice with uncertainty about jump and diffusion risk. Journal of Banking & Finance 37, 5036-5047

[8] Bremaud, P., 1981. Point Processes and Queues: Martingale Dynamics, Springer, Berlin.

[9] Cvitanic, J., Karatzas, I., 1992. Convex duality in constrained portfolio optimization. Annals of Applied Probability 2, 767-818.

[10] Das, S., Uppal, R., 2004. Systemic risk and international portfolio choice. Journal of Finance 59, 2809–2834.

[11] Delbaen, F., Schachermayer, W., 1994. Arbitrage and free lunch with bounded risk for unbounded continuous processes. Mathematical Finance 4, 343-348.

[12] Drechsler, I., 2013. Uncertainty, time-varying fear, and asset prices. Journal of Finance 68, 1843-1889.

[13] Epstein, L., Schneider, M., 2003. Recursive multiple priors. Journal of Economic Theory 113, 32-50.

[14] Flor, C., Larsen, L., 2014. Robust portfolio choice with stochastic interest rates. Annals of Finance 10, 243-265.

[15] Goldfarb D. and Iyengar G., 2003. Robust portfolio selection problems. Mathematics of Operations Research 28, 1-38.

[16] Guidolin, M., Rinaldi, F., 2013. Ambiguity in asset pricing and portfolio choice: A review of the literature. Theory and Decision 74, 183-217.

[17] Hansen, L. P., Sargent, T. J., 2001. Robust control and model uncertainty. Amer. Econom. Rev. 91, 60-66.

[18] Hansen, L. P., Sargent, T. J., 2007. Robustness. Princeton University Press, NJ, Princeton.

[19] Jin, X., Zhang, A., 2012. Decomposition of optimal portfolio weights in a jump-diffusion model and its applications. Review of Financial Studies 25, 2877-2919.

[20] Karatzas, I., Lehoczky, J., Shreve, S., 1987. Optimal portfolio and consumption decisions for a small investor on a finite horizon. SIAM Journal of Control and Optimization 25, 1557-1586.

[21] Kramkov, D., Schachermayer, W., 1999. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. Annals of Applied Probability 9, 904-950.

[22] Kramkov, D., Schachermayer, W., 2003. Necessary and sufficient condition in the problem of optimal investment in incomplete markets. Annals of Applied Probability 13, 1504-1516.

[23] Laeven, R., and Stadje, M., 2014. Robust portfolio choice and indifference valuation. Mathematics of Operations Research 39, 1109-1141.

[24] Liu, J., Longstaff, F., Pan, J., 2003. Dynamic asset allocation with event risk. Journal of Finance 58, 231-259.

[25] Liu, J., Pan, J., Wang, T., 2005. An equilibrium model of rare-event premia and its implication for option smirks. Review of Financial Studies 18, 131-164.

[26] Luenberger, D., 1969. Optimization by Vector Space Methods, John Wiley and Sons.

[27] Maenhout, P., 2004. Robust portfolio rules and asset pricing. Review of Financial Studies 17, 951-983.

[28] Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125-144.

[29] Pennanen, T., 2011. Convex duality in stochastic optimization and mathematical finance. Mathematics of Operations Research 36, 340-362.

[30] Ruszczynski, A, Shapiro, A., 2006. Conditional risk mappings. Mathematics of Operations Research 31, 544-561.

[31] Schied, A., Wu, C., 2005. Duality theory for optimal investments under model uncertainty. Statistical Decisions 23, 199-217.

[32] Seifried, F. T., 2010. Optimal investment for worst-case crash scenarios: A martingale approach. Mathematics of Operations Research 35, 559-579.

[33] Wang, T., 2003. Conditional preferences and updating. Journal of Economic Theory 108, 286-321.

[34] Yan, J, 1988. Measure and Integration. Shanxi Normal University Press, China.

[35] Zhao, Y., Tiemba, W. T., 2001. A stochastic programming model using an endogenously determined worst case risk measure for dynamic asset allocation. Math. Program. Ser. B 89, 293-309.