

Robust Time-Inconsistent Stochastic Control Problems^{*}

Chi Seng Pun^a

^a*School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore*

Abstract

This paper establishes a general analytical framework for continuous-time stochastic control problems for an ambiguity-averse agent (AAA) with time-inconsistent preference, where the control problems do not satisfy Bellman's principle of optimality. The AAA is concerned about model uncertainty in the sense that she is not completely confident in the reference model of the controlled Markov state process and rather considers some similar alternative models. The problems of interest are studied within a set of dominated models and the AAA seeks for an optimal decision that is robust with respect to model risks. We adopt a game-theoretic framework and the concept of subgame perfect Nash equilibrium to derive an extended dynamic programming equation and extended Hamilton–Jacobi–Bellman–Isaacs equations for characterizing the robust dynamically optimal control of the problem. We also prove a verification theorem to theoretically support our construction of robust control. To illustrate the tractability of the proposed framework, we study an example of robust dynamic mean-variance portfolio selection under two cases: 1. constant risk aversion; and 2. state-dependent risk aversion.

Key words: Stochastic control; Robustness; Time-inconsistent preference; State-dependence; Hamilton–Jacobi–Bellman–Isaacs equations; Robust dynamic mean-variance portfolio

1 Introduction

Stochastic control theory has achieved great success in modeling and providing solutions to lots of physical, biological, economical, and financial problems, to name a few. Stochastic optimal control is the serial control variables that accomplish a desired goal for the controlled state process with minimum cost or with maximum reward in the presence of noises (risks). The most commonly used approaches in solving stochastic optimal control problems are Pontryagin's maximum principle and Bellman's dynamic programming. These two principal approaches and their relationship are documented in many classic reference books such as [Yong and Zhou, 1999].

While stochastic control deals with the existence of risk, robust stochastic control deals with the existence of ambiguity as well. Here, the ambiguity refers to the Knightian (model) uncertainty originating with [Knight, 1921], who clarified the subtle difference between risk and uncertainty. [Ellsberg, 1961] reveals the inadequacy of utility theory and argues that human beings are ambiguity-averse; thus our rational decisions should be made under conditions of ambiguity. Our lack of knowledge about the actual state process or estimation error unavoidably introduces ambiguity into the control problem and it has important implications for many critical aspects such as risk quantification. We call the agent fearing ambiguity as *ambiguity-averse agent (AAA)*. The AAA has certain confidence in a reference model but rather considers some alternative models. A key to dealing with ambiguity is to quantify the model misspecification given the AAA's historical data record; see [Anderson *et al.*, 2003] for a statistical method. Following the robust decision rule in [Wald, 1945] and [Anderson *et al.*, 2003], [Maenhout, 2004] derives robust portfolio rules in the context of [Merton, 1971]'s portfolios. Recently, [Pun and Wong, 2015] extends the analysis to a general time-consistent objective functional.

^{*} Corresponding author C. S. Pun. Tel. +65-65137468.
Email address: `cspun@ntu.edu.sg` (Chi Seng Pun).

In recent years, there is a growing literature investigating time-inconsistent stochastic control problems, where the objective functional contains time-inconsistent terms such that Pontryagin’s and Bellman’s optimality principles are not applicable. A famous example is the financial mean-variance (MV) portfolio selection, pioneered by [Markowitz, 1952] and further studied in [Li and Ng, 2000] and [Zhou and Li, 2000] for dynamic settings. Some other examples include endogenous habit formulation in economics and equilibrium production economy; see [Björk and Murgoci, 2014] and [Björk *et al.*, 2017]. For a general time-inconsistent objective functional (see Equation (4) below), Lemma 2 below reveals the sources of time-inconsistency that violate Bellman’s principle of optimality. In this paper, we contribute to incorporate the model uncertainty with time-inconsistency to study a general class of robust time-inconsistent stochastic control problems. To the best of our knowledge, we are the first to establish a general, analytical, and tractable framework for such problems. Our model and objective settings nest many classes of well-known continuous-time problems, such as [Merton, 1971], [Heston, 1993], and [Zhou and Li, 2000], as special cases.

In the existing literature, several approaches to handling the time-inconsistency derive reasonable policies with different features, which include but are not limited to the followings:

- *Precommitment* policy: that optimizes the objective functional anticipated at the very beginning time point and the controller sticks with this policy over the whole control period. For example, [Li and Ng, 2000] and [Zhou and Li, 2000] introduced an embedding technique to solve for a precommitment MV portfolio.
- *Equilibrium (time-consistent)* policy: that consistently optimizes the objective functional anticipated at every time point in the similar manner of dynamic programming but using the concept of subgame perfect equilibrium. This idea is initiated in [Strotz, 1955] and [Goldman, 1980]. Related papers include [Basak and Chabakauri, 2010] for time-consistent MV portfolio and [Björk *et al.*, 2014], [Björk and Murgoci, 2014], [Bannister *et al.*, 2016], and [Björk *et al.*, 2017] for more concrete examples.

Some recent alternatives are proposed in [Cui *et al.*, 2017], [Karnam *et al.*, 2017], and [Pedersen and Peskir, 2017] with different views on dynamic objectives. Loosely speaking, from the perspective of decision making, precommitment policy emphasizes on global optimality while time-consistent policy emphasizes on local optimality. However, precommitment policy is sometimes inferior because its strong commitment leads to time-inconsistency in efficiency (see [Cui *et al.*, 2010] for details and a remedy) and its error-accumulation property brings huge estimation error (see [Chiu *et al.*, 2017]). Moreover, finding precommitment policy poses analytical challenges for general time-inconsistent stochastic control problems. In this paper, we attack the time-inconsistency with the second approach, together with robustness, resulting in robust time-consistent policy. A related works is [Zeng *et al.*, 2016], which considered similar mean-variance optimization in continuous-time setting and it will be compared with this paper in Section 2.

In this paper, we assume a general continuous-time Markov stochastic process for the state under the reference model. Moreover, we define alternative models, which are equivalent to the reference model in terms of probability measure. By the Girsanov theorem, the link between the reference and alternative models is characterized by a stochastic process that acts as an adverse control variate. The agent has a time-inconsistent preference of a general form and aims to find a time-consistent policy that is robust with respect to the choice of the alternative model. We use a maximin formulation as in [Wald, 1945] to construct robust decision rule and use the concept of subgame perfect equilibrium to characterize the time-consistent policy. With an extended dynamic programming approach derived in this paper, we characterize the robust time-consistent policy using an extended Hamilton–Jacobi–Bellman–Isaacs (HJBI) system. Moreover, we introduce and discuss the choice of ambiguity preference function, which completes a general analytical framework for a large class of time-inconsistent stochastic control problems with model uncertainty. To illustrate the tractability of the proposed framework, we apply it to solve for robust dynamic mean-variance (MV) portfolio selection under two cases: 1. constant risk aversion; and 2. state-dependent risk aversion. For the latter case, we introduce nonhomogeneous Abel’s differential equations to characterize the robust MV portfolio.

The contribution of this paper is threefold: first, we provide a rigorous mathematical definition of robust time-consistent policy and reveal its nature as the perfect equilibrium of subgames of maximin control problems, i.e. “games in subgames.” Second, we prove a verification theorem that solving the proposed extended HJBI system is a sufficient condition for robust optimality. The extended dynamic programming approach and the verification theorem, derived in this paper, are innovative to the literature on robust control. Third, we apply the proposed framework to solve an open problem of robust dynamic mean-variance portfolio selection under the robustness rule of [Anderson *et al.*, 2003]. The analyses cover two economically meaningful cases, which extend the results in [Björk *et al.*, 2014] to robust counterparts. Through this study, our discussion about the ambiguity preference function for the general case extends the results in [Maenhout, 2004] and [Pun and Wong, 2015].

The remainder of this paper is organized as follows. Section 2 introduces the reference and alternative models and the

robust time-inconsistent stochastic control problems of our interest. Moreover, Section 2.3 defines the robust time-consistent control-measure policy and robust value function. The main results are in Section 3, where we present an extended dynamic programming equation for the robust value function and the extended HJBI system while a generalized result is put in Appendix A. The verification theorem is provided in Section 3.2. To facilitate the analysis, Section 3.3 provides a simplification of the HJBI system to an HJB system. Section 4 is devoted to robust MV portfolio analysis and through this study, we discuss the choice of ambiguity preference function for general problems in Section 4.3. Finally, Section 5 concludes and discusses the future research directions.

2 Problem Formulation

We study the problem of our interest with a set of candidate models, which stem from a reference model with agent's preliminary knowledge. Specifically, we suppose the agent does not have complete confidence in the reference model, and she prefers to consider alternative models perturbed around the reference model and to make a robust decision with respect to the model risk.

2.1 The Reference and Alternative Models

The *reference model* is defined over a filtered physical probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t^{\mathbb{P}}\}_{t \geq 0}, \mathbb{P})$, where the filtration $\{\mathcal{F}_t^{\mathbb{P}}\}_{t \geq 0}$ is generated by an m -dimensional standard \mathbb{P} -Brownian motion $W_t^{\mathbb{P}} = (W_{1t}^{\mathbb{P}}, \dots, W_{mt}^{\mathbb{P}})'$. Hereafter, the transpose of a vector or matrix a is denoted by a' . We consider the p -dimensional controlled state process within a time horizon T , $\{X_t\}_{t \in [0, T]}$, driven by a stochastic differential equation (SDE)

$$dX_t = \mu(t, X_t, u_t)dt + \sigma(t, X_t, u_t)dW_t^{\mathbb{P}}, \quad (1)$$

where $\{u_t\}_{t \in [0, T]}$ is a k -dimensional control process with the constraint $u_t \in U(t, X_t)$ and $\mu \in \mathbb{R}^p$ and $\sigma \in \mathbb{R}^{p \times m}$ are the drift and diffusion coefficient functions in $(t, X_t, u_t) \in [0, T] \times \mathbb{R}^p \times U(t, X_t)$. The model (1) is general as m, p, k can be arbitrary natural numbers and it nests the stochastic multi-factor models ([Pun *et al.*, 2015]) and stochastic volatility models in finance (see [Heston, 1993]) as special cases.

The *alternative models* are conceptually defined as the models that are “similar” to the reference model. In this paper, we employ the mathematical concept of measure equivalence to characterize the “similarity” between models; see [Anderson *et al.*, 2003]. Specifically, the alternative measures (models) are induced by a class of probability measures equivalent to \mathbb{P} : $\mathcal{Q} := \{\mathbb{Q} | \mathbb{Q} \sim \mathbb{P}\}$. By the Girsanov theorem, for each $\mathbb{Q} \in \mathcal{Q}$, there is a \mathbb{R}^m -valued stochastic process $\{q_t\}_{t \in [0, T]}$ such that for $t \in [0, T]$, we have the Radon-Nikodym derivative of \mathbb{Q} with respect to \mathbb{P} :

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} = \exp \left(\int_0^t q'_s dW_s^{\mathbb{P}} - \frac{1}{2} \int_0^t q'_s q_s ds \right) =: \nu_t.$$

Moreover, if $\{q_t\}_{t \in [0, T]}$ satisfies the Novikov condition:

$$\mathbb{E}^{\mathbb{P}} \left[\exp \left(\frac{1}{2} \int_0^T q'_s q_s ds \right) \right] < \infty, \quad (2)$$

then the process $\{\nu_t\}_{t \in [0, T]}$ is a positive \mathbb{P} -martingale and $W_t^{\mathbb{Q}}$, driven by $dW_t^{\mathbb{Q}} = dW_t^{\mathbb{P}} - q_t dt$, is an m -dimensional standard \mathbb{Q} -Brownian motion. For each $\mathbb{Q} \in \mathcal{Q}$, we can rewrite the SDE for the state process in (1) under the measure \mathbb{Q} as

$$dX_t = [\mu(t, X_t, u_t) + \sigma(t, X_t, u_t)q_t]dt + \sigma(t, X_t, u_t)dW_t^{\mathbb{Q}}, \quad (3)$$

where q_t can be dependent on (t, X_t, u_t) , i.e. $q_t = q(t, X_t, u_t)$, as long as it satisfies (2). Therefore, the alternative model essentially alters the probability measure of the reference model and the dissimilarity between an alternative model \mathbb{Q} and the reference model \mathbb{P} is characterized by q_t . Choosing \mathbb{Q} can be converted to the determination of the process $\{q_t\}_{t \in [0, T]}$, which can be viewed as a control variate by “nature” from the viewpoint of game theory. Specifically, a robust optimization problem can be viewed a game played by the agent against the “nature” who can alter the probability space. By (3), the ambiguity, introduced by q_t , imposes a direct effect on the state's drift coefficient and an indirect effect on the state's diffusion coefficient through its state-dependence.

2.2 Stochastic Control Problems with Time-Inconsistent Objectives and Model Uncertainty

Suppose that the agent adopts a time-inconsistent reward functional of the form

$$\mathbb{E}_{t,x}^{\mathbb{Q}}[F(x, X_T)] + G(x, \mathbb{E}_{t,x}^{\mathbb{Q}}[X_T]) + D_{t,x}(\mathbb{Q} \parallel \mathbb{P}) \quad (4)$$

for any point $(t, x) \in [0, T] \times \mathbb{R}^p$, where $\mathbb{E}_{t,x}^{\mathbb{Q}}[\cdot] := \mathbb{E}^{\mathbb{Q}}[\cdot | \mathcal{F}_t]$ while the subscript $x = X_t$ emphasizes the Markovian property of (3), F and G are arbitrary functions valued in \mathbb{R} , and $D_{t,x}(\mathbb{Q} \parallel \mathbb{P}) \geq 0$ is a generalized Kullback–Leibler (KL) divergence between \mathbb{Q} and \mathbb{P} , which will be specified below. The introduction of $D_{t,x}(\mathbb{Q} \parallel \mathbb{P})$ is to measure the (state-dependent) ambiguity aversion of the agent and to regularize the choices of \mathbb{Q} .

We denote by $(\mathbf{u}, \mathbf{q}) := (\{u_t\}_{t \in [0, T]}, \{q_t\}_{t \in [0, T]})$ the *control-measure policy*, noting that a measure \mathbb{Q} is fully characterized by \mathbf{q} and the reference model. The agent aims to find a robust optimal control policy that maximizes (4) under the worst-case scenario among \mathcal{Q} for every (t, x) :

$$\sup_{\mathbf{u}} \inf_{\mathbf{q}} J(t, x, \mathbf{u}, \mathbf{q}) := \sup_{\mathbf{u}} \inf_{\mathbf{q}} \left\{ \mathbb{E}_{t,x}^{\mathbb{Q}}[F(x, X_T^{\mathbf{u}})] + G(x, \mathbb{E}_{t,x}^{\mathbb{Q}}[X_T^{\mathbf{u}}]) + D_{t,x}(\mathbb{Q} \parallel \mathbb{P}) \right\}, \quad (5)$$

where $\{X_t^{\mathbf{u}}\}_{t \in [0, T]}$ is a strong solution to (3) with \mathbf{u} . It is noteworthy that q could depend on u as a result of minimizing an objective involved u . The dependence is interpreted as follows. For each strategy u , the agent presumes the controlled state process will evolve as the worst possible; thus the model misspecification is allowed to depend on u . Under such worst-case scenarios, she finds the optimal (robust) control. To ease the notational burden, we always suppress the arguments of u_t and q_t when no confusion can arise.

The formulation (5) belongs to the Wald’s maximin paradigm and is popular in robust decision making; see [Wald, 1945]. However, the state-dependence of F , G , and $D_{t,x}$ in (4) introduces time-inconsistency to the (robust) control problem and makes the conventional dynamic programming principle invalid; see the explanation below and [Yong and Zhou, 1999] for the classic approaches for stochastic control problems. Since \mathbf{q} depends on \mathbf{u} in general, the time-inconsistency is also inherited to the adverse control \mathbf{q} . To find a time-consistent robust control-measure policy, this paper contributes to the formal definition of robust optimality using the concept of intrapersonal equilibrium games under model uncertainty.

Before we introduce the robust optimality, we first complete the mathematical details about the problem (5). We first specify the form of $D_{t,x}(\mathbb{Q} \parallel \mathbb{P})$ that measures the divergence between \mathbb{Q} and \mathbb{P} . Inspired by [Maenhout, 2004] and [Pun and Wong, 2015], we consider $D_{t,x}(\mathbb{Q} \parallel \mathbb{P})$ of the form

$$D_{t,x}(\mathbb{Q} \parallel \mathbb{P}) := \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^T C(x, s, X_s^{\mathbf{u}}; \mathbf{q}) ds \right] := \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\frac{1}{2} \int_t^T \frac{q'_s \Xi(x)^{-1} q_s}{\Phi(s, X_s^{\mathbf{u}})} ds \right], \quad (6)$$

where $\Xi(x) = \text{diag}(\xi_1(x), \dots, \xi_m(x))$ with $\xi_i(x) \geq 0$ measuring the state-dependent ambiguity aversion to the uncertainty source ($W_{it}^{\mathbb{P}}$) of the state process for $i = 1, \dots, m$, and $\Phi(\cdot, \cdot)$ is a deterministic \mathbb{R}^+ -valued function representing agent’s preference on ambiguity. We immediately have two remarks:

- (1) If $\xi_1(x) = \dots = \xi_m(x) =: \xi$ and $\Phi(\cdot, \cdot) \equiv 1$, then

$$D_{t,x}(\mathbb{Q} \parallel \mathbb{P}) = \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\frac{1}{2\xi} \int_t^T q'_s q_s ds \right] = \frac{1}{\xi} D_{KL}(\mathbb{Q} \parallel \mathbb{P}),$$

where $D_{KL}(\mathbb{Q} \parallel \mathbb{P})$ is the KL divergence between \mathbb{Q} and \mathbb{P} and ξ is an aggregate measure of ambiguity aversion. If $\xi \downarrow 0$, the agent picks the reference measure \mathbb{P} and the robust control problem is reduced to an ordinary control problem because $D_{KL}(\mathbb{Q} \parallel \mathbb{P}) \equiv 0$ in this case.

- (2) If $\xi_i(x) \downarrow 0$ for some i , we can similarly argue that $q_{it} \equiv 0$, where $q_t = (q_{1t}, \dots, q_{mt})'$. For this case, the ambiguity aversion about the uncertainty source $W_{it}^{\mathbb{P}}$ is vanished.

The specifications of ξ ’s and Φ provide flexibility to the agent in terms of modeling ambiguity. However, the key to the robust control formulation is to find a suitable preference function Φ to facilitate analysis. For the general (time-consistent) utility maximization, [Pun and Wong, 2015] proposed a specific form for Φ . In this paper, we extend

their results to time-inconsistent control problems with our illustrative example. Nevertheless, our mathematical framework allows a general Φ .

Remark: The formulation in (5) with the penalty $D_{t,x}$ in (6) is highly related to the robust problem formulation with an ellipsoidal uncertainty set as in [Biagini and Pinar, 2017] by similar arguments of the Lagrangian multiplier theorem; see [Hansen *et al.*, 2006]. By choosing the value of ξ (embed in Ξ), as with tuning the radii of ellipsoidal uncertainty set, we can control the balance between efficiency and robustness of the problem (5).

Next, we define the admissible set of control-measure policies as follows.

Definition 1 A control-measure policy $(\mathbf{u}, \mathbf{q}) = (\{u(t, X_t)\}_{t \in [0, T]}, \{q(t, X_t, u(t, X_t))\}_{t \in [0, T]})$ is admissible if the following conditions are satisfied:

- (1) (Control constraint) For every $(t, x) \in [0, T] \times \mathbb{R}^p$, we have $u(t, x) \in U(t, x)$.
- (2) (Change-of-measure condition) \mathbf{q} satisfies the Novikov condition (2).
- (3) (Unique strong solvability) For every initial point $(t, x) \in [0, T] \times \mathbb{R}^p$, the SDE

$$dX_s = [\mu(s, X_s, u(s, X_s)) + \sigma(s, X_s, u(s, X_s))q(s, X_s, u(s, X_s))]ds + \sigma(s, X_s, u(s, X_s))dW_s^{\mathbb{Q}}$$

for $s \in [t, T]$ with $X_t = x$ admits a unique strong solution.

- (4) (Uniform integrability) For every initial point $(t, x) \in [0, T] \times \mathbb{R}^p$, we have

$$\sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^T |C(x, s, X_s^u; \mathbf{q})| ds + |F(x, X_T^u)| \right] < \infty, \quad \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{t,x}^{\mathbb{Q}} [|X_T^u|_1] < \infty, \quad \text{and} \quad \sup_{\mathbb{Q} \in \mathcal{Q}} G(x, \mathbb{E}_{t,x}^{\mathbb{Q}} [X_T^u]) < \infty,$$

where $|a|_1 = \sum_{i=1}^p |a_i|$ for $a = (a_1, \dots, a_p)' \in \mathbb{R}^p$.

We denote the set of admissible control-measure policies as $\mathbf{U} \times \mathbf{Q}$.

Finally, we explore the sources of time-inconsistency in (4). Conceptually, the state-dependence in all three terms in (4) leads to time-changing preferences of the agent; thus decisions being made at different time points can be inconsistent with each other. If we view the decision making as a game played by many incarnations of the agent at different time points, then the inconsistency is primarily about commitment and credible threats. The cause of inconsistency is primarily due to the violation of *Bellman's principle of optimality* by (4). To see this, we first derive the recursive equation for $J(t, x; \mathbf{u}, \mathbf{q})$ in (5), whose proof can be referred to Appendix A.

Lemma 2 For $s > t$ and any admissible control-measure policy $(\mathbf{u}, \mathbf{q}) \in \mathbf{U} \times \mathbf{Q}$, we have

$$J(t, x; \mathbf{u}, \mathbf{q}) = \mathbb{E}_{t,x}^{\mathbb{Q}} \left[J(s, X_s^u; \mathbf{q}) + \int_t^s C(X_\tau^u, \tau, X_\tau^u; \mathbf{u}, \mathbf{q}) d\tau \right] + L^{\mathbf{u}, \mathbf{q}}(t, x, s), \quad (7)$$

where $L^{\mathbf{u}, \mathbf{q}}(t, x, s) = L_C^{\mathbf{u}, \mathbf{q}}(t, x, s) + L_H^{\mathbf{u}, \mathbf{q}}(t, x, s) + L_G^{\mathbf{u}, \mathbf{q}}(t, x, s)$ and

$$\begin{aligned} L_C^{\mathbf{u}, \mathbf{q}}(t, x, s) &= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^s [C(x, \tau, X_\tau^u) - C(X_\tau^u, \tau, X_\tau^u)] d\tau \right], \\ L_H^{\mathbf{u}, \mathbf{q}}(t, x, s) &= \mathbb{E}_{t,x}^{\mathbb{Q}} [h^{\mathbf{u}, \mathbf{q}}(s, X_s^u, x)] - \mathbb{E}_{t,x}^{\mathbb{Q}} [h^{\mathbf{u}, \mathbf{q}}(s, X_s^u, X_s^u)], \\ L_G^{\mathbf{u}, \mathbf{q}}(t, x, s) &= G(x, \mathbb{E}_{t,x}^{\mathbb{Q}} [g^{\mathbf{u}, \mathbf{q}}(s, X_s^u)]) - \mathbb{E}_{t,x}^{\mathbb{Q}} [G(X_s^u, g^{\mathbf{u}, \mathbf{q}}(s, X_s^u))], \\ h^{\mathbf{u}, \mathbf{q}}(t, x, y) &:= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^T C(y, \tau, X_\tau^u) d\tau + F(y, X_T^u) \right], \quad g^{\mathbf{u}, \mathbf{q}}(t, x) := \mathbb{E}_{t,x}^{\mathbb{Q}} [X_T^u]. \end{aligned} \quad (8)$$

If either one of F and $D_{t,x}$ depends on x , i.e. state-dependent, or $G(x, y)$ is nonlinear in y , then $L^{\mathbf{u}, \mathbf{q}}$ will be nonzero in general and it quantifies the agent's incentives to deviate from the time t optimal policy over the period $[t, s]$, leading to the violation of Bellman's Principle of Optimality. For the case of nonzero $L^{\mathbf{u}, \mathbf{q}}$, we call (5) the *robust time-inconsistent stochastic control problem*. Lemma 2 also reveals that if we determine the (robust) optimal

decision by backward induction in the same manner of dynamic programming, we should take into account the adjustment term $L^{\mathbf{u}, \mathbf{q}}$ in the derivation of a recursive relationship for the (robust) optimal objective function, i.e. the Hamilton–Jacobi–Bellman(–Isaacs) equation for the value function.

2.3 Game-Theoretic Formulation of Robust Time-Consistent Stochastic Controls

To deal with time-inconsistent stochastic control problems, there are several approaches with different perspectives proposed in the context of mean-variance (MV) portfolio selection, which involves nonlinear G in (4). [Li and Ng, 2000] and [Zhou and Li, 2000] applied embedding techniques to the MV problems and the resulting investment policy is *precommitment*, where the investor maximizes her objective at time 0 and thereafter does not deviate from that policy. [Cui *et al.*, 2010] pointed out the inefficiency of the precommitment MV portfolio and proposed semi-self-financing strategies to improve it, where the investor can withdraw money. However, the embedding technique cannot be applied to the general time-inconsistent problems in (4) and the resulting policies are not time-consistent from the economic perspective; see [Strotz, 1955]. According to [Basak and Chabakauri, 2010], a time-consistent policy is optimally chosen to take into account that the agent will also act optimally in the future while she is not restricted from revising her policy at all times. [Björk and Murgoci, 2014] and [Björk *et al.*, 2017] generalize the results of [Basak and Chabakauri, 2010] to time-inconsistent stochastic control problem with an interpretation of intrapersonal game equilibrium. To the best of our knowledge, no one attacked the time-inconsistent stochastic control problems with model uncertainty. In this paper, we adopt the game-theoretic formulation to characterize robust time-consistent stochastic controls.

A robust stochastic control problem over $[0, T]$ can be viewed as a non-cooperative game played by infinite Players indexed by $t \in [0, T]$, who can only control the state process X at time t by choosing $u(t, X_t)$ and simultaneously plays another Stackelberg game against the “nature” controlling $q(t, X_t, u(t, X_t))$ at time t . Hence, the whole robust stochastic control problem can be regarded as “games in subgames.” Following the approach of [Strotz, 1955] and [Basak and Chabakauri, 2010], a time-consistent policy $(\mathbf{u}^*, \mathbf{q}^*)$ is a subgame perfect Nash equilibrium, where for any $t \in [0, T]$, if Players $s \in (t, T]$ choose the robust control $\{(u^*(s, X_s), q^*(s, X_s, u^*(s, X_s)))\}_{s \in (t, T]}$ respectively, then it is robust optimal for Player t to choose $(u^*(t, X_t), q^*(t, X_t, u^*(t, X_t)))$. However, for each Player t , one conceptually needs to solve a robust control problem at time t , i.e. on a time set of Lebesgue measure zero; thus the control $u(t, X_t)$ has no impact. Hence for each time t and a minimal time elapse $\epsilon > 0$, we instead consider a robust optimal control problem over $[t, t + \epsilon]$ given that Players $s \in [t + \epsilon, T]$ have chosen robust optimal controls. The formal definition of the subgame perfect Nash equilibrium is provided as follows.

Definition 3 Consider an admissible control-measure policy $(\mathbf{u}^*, \mathbf{q}^*)$. For any fixed point $(t, x) \in [0, T] \times \mathbb{R}^p$ and a fixed real number $\epsilon > 0$, define an ϵ -policy $(\mathbf{u}_\epsilon, \mathbf{q}_\epsilon)$: for $y \in \mathbb{R}^p$,

$$\begin{cases} u_\epsilon(s, y) = u(s, y)\mathbf{1}_{\{s \in [t, t + \epsilon)\}} + u^*(s, y)\mathbf{1}_{\{s \in [t + \epsilon, T]\}}, \\ q_\epsilon(s, y, u_\epsilon(s, y)) = q(s, y, u)\mathbf{1}_{\{s \in [t, t + \epsilon)\}} + q^*(s, y, u^*)\mathbf{1}_{\{s \in [t + \epsilon, T]\}}, \end{cases} \quad (9)$$

where (\mathbf{u}, \mathbf{q}) is an arbitrary admissible control-measure policy and $\mathbf{1}_A$ is the indicator function of a subset A . If

(1) for all $(\mathbf{u}, \mathbf{q}) \in \mathbf{U} \times \mathbf{Q}$,

$$\liminf_{\epsilon \downarrow 0} \frac{J(t, x; \mathbf{u}_\epsilon, \mathbf{q}_\epsilon) - J(t, x; \mathbf{u}_\epsilon, \mathbf{q}^*)}{\epsilon} \geq 0;$$

(2) and for all $\mathbf{u} \in \mathbf{U}$,

$$\limsup_{\epsilon \downarrow 0} \frac{J(t, x; \mathbf{u}_\epsilon, \mathbf{q}^*) - J(t, x; \mathbf{u}^*, \mathbf{q}^*)}{\epsilon} \leq 0,$$

then $(\mathbf{u}^*, \mathbf{q}^*)$ is called an equilibrium control-measure policy. The equilibrium value function is defined as the corresponding objective function:

$$V(t, x) := J(t, x; \mathbf{u}^*, \mathbf{q}^*) \quad \text{for } t \in [0, T].$$

We remark here that $u(T, \cdot)$ and $q(T, \cdot, \cdot)$ have no impact on the equilibrium or the robust control problem because by definition, $V(T, x) = J(T, x; \mathbf{u}^*, \mathbf{q}^*) = F(x, x) + G(x, x)$. Notice that the two inequalities in Definition 3 are consistent with the martingale conditions in formulating the robust control problem as in [Biagini and Pinar, 2017].

By its definition, the equilibrium control-measure policy is time-consistent and robust optimal in the sense of (5); thus we also refer the equilibrium control-measure policy and equilibrium value function as *robust time-consistent control* and *robust value function*, respectively. The use of two inequalities in Definition 3 preserves the preference ordering (supinf) as the robustness rule suggested in [Anderson *et al.*, 2003] and [Hansen *et al.*, 2006]. Moreover, the proposed framework in this paper does not assume exchangeability of sup and inf in (5).

It is noteworthy that similar robust control formulation was considered in [Zeng *et al.*, 2016] for the special case of mean-variance optimization. [Zeng *et al.*, 2016] identifies the worst-case measure as a prior step and regards the worst-case objective as an ordinary objective. However, due to the time-inconsistent preference, the infimum problem with respect to measure also requires time-consistent treatment on \mathbb{Q} (or q) that is advocated in this paper. Definition 3 clarifies the time-consistency on both control and measure in order to reflect the consistent planning of the agent.

3 The Extended Dynamic Programming Approach

This section derives an extended dynamic programming equation (EDPE) for the robust value function V . Loosely speaking, the equation is to connect all the subgames played by the incarnation of the agent at time t for $t \in [0, T]$ in the same manner of dynamic programming. The derivation relies on the recursive equation for J and the equilibrium concept stated in Section 2.3. Motivated by the EDPE, we then define the extended Hamilton–Jacobi–Bellman–Isaacs (HJBI) equations (system). In addition, we provide a verification theorem that the solution to the HJBI equations is equal to the robust value function V and the corresponding policy is the robust time-consistent control $(\mathbf{u}^*, \mathbf{q}^*)$. In the end of this section, we deduce the expression of \mathbf{q}^* and reduce the extended HJBI system to an extended HJB system. The results derived in this section for the reward functional (4) can be generalized to a broader class of reward functionals depending on the current time t . The general results and their proofs are placed in Appendix A. Due to the page limit, we omit the proofs in the main context, which can be referred to Appendix A.

3.1 EDPE and Extended HJBI System

The derivation of EDPE and extended HJBI equations in this subsection, relying on the equilibrium concept, is motivational while its mathematical integrity is provided with the verification theorem in the next subsection.

We suppose that the robust time-consistent policy $(\mathbf{u}^*, \mathbf{q}^*)$ exists and the corresponding robust value function at time t is $V(t, x)$. For any $(t, x) \in [0, T] \times \mathbb{R}^p$ and $\epsilon > 0$, according to the definition of time-consistent policy, we consider a robust optimal control problem over $[t, t + \epsilon]$ given that the controls, u and q , over $[t + \epsilon, T]$ are fixed as u^* and q^* , respectively, while the controls over $[t, t + \epsilon]$ are free. By (7) and using the notation of ϵ -policy, we have

$$\begin{aligned} V(t, x) &= \sup_{\mathbf{u}} \inf_{\mathbf{q}} \left\{ \mathbb{E}_{t,x}^{\mathbb{Q}} \left[J(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}\epsilon}; \mathbf{u}\epsilon, \mathbf{q}\epsilon) + \int_t^{t+\epsilon} C(X_{\tau}^{\mathbf{u}\epsilon}, \tau, X_{\tau}^{\mathbf{u}\epsilon}, \mathbf{q}\epsilon) d\tau \right] + L^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon) \right\} \\ &= \sup_{\mathbf{u}} \inf_{\mathbf{q}} \left\{ \mathbb{E}_{t,x}^{\mathbb{Q}} \left[V(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}) + \int_t^{t+\epsilon} C(X_{\tau}^{\mathbf{u}}, \tau, X_{\tau}^{\mathbf{u}}; \mathbf{q}) d\tau \right] + L^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon) \right\}, \end{aligned} \quad (10)$$

where $L^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon) = L_C^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon) + L_H^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon) + L_G^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon)$,

$$\begin{aligned} L_C^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon) &= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^{t+\epsilon} [C(x, \tau, X_{\tau}^{\mathbf{u}}) - C(X_{\tau}^{\mathbf{u}}, \tau, X_{\tau}^{\mathbf{u}})] d\tau \right], \\ L_H^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon) &= \mathbb{E}_{t,x}^{\mathbb{Q}} [h(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}, x)] - \mathbb{E}_{t,x}^{\mathbb{Q}} [h(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}, X_{t+\epsilon}^{\mathbf{u}})], \\ L_G^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}(t, x, t + \epsilon) &= G(x, \mathbb{E}_{t,x}^{\mathbb{Q}} [g(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}})]) - \mathbb{E}_{t,x}^{\mathbb{Q}} [G(X_{t+\epsilon}^{\mathbf{u}}, g(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}))], \end{aligned}$$

in which $h(\cdot, \cdot, \cdot) := h^{\mathbf{u}^*, \mathbf{q}^*}(\cdot, \cdot, \cdot)$ and $g(\cdot, \cdot) := g^{\mathbf{u}^*, \mathbf{q}^*}(\cdot, \cdot)$. The equation (10) is analogous to the ordinary dynamic programming equation except for the additional term of $L^{\mathbf{u}\epsilon, \mathbf{q}\epsilon}$. Notice that if the sources of time-inconsistency in (4) do not exist, $L^{\mathbf{u}\epsilon, \mathbf{q}\epsilon} \equiv 0$. Hence, we call (10) the *extended dynamic programming equation (EDPE)*.

To define a new HJBI equation for V , it can be expected there is an extra term to adjust the Hamiltonian. Let $\mathcal{A}^{\mathbf{u}, \mathbf{q}}$ be the infinitesimal operator associated with the state process in (3). Then for any multivariate function of t and x ,

denoted by $\phi(t, x) \in \mathbb{R}^n$ with any n , we have

$$\mathcal{A}^{\mathbf{u}, \mathbf{q}} \phi(t, x) := \frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial x} [\mu(t, x, u) + \sigma(t, x, u)q] + \frac{1}{2} \sum_{i,j=1}^p \Sigma_{ij}(t, x, u) \frac{\partial^2 \phi}{\partial x_i \partial x_j},$$

where $x = (x_1, \dots, x_p)'$, $\Sigma(t, x, u) = \sigma(t, x, u)\sigma(t, x, u)' =: (\Sigma_{ij}(t, x, u))_{i,j=1,\dots,p} \in \mathbb{R}^{p \times p}$, and $\partial \phi / \partial x = (\partial \phi_i / \partial x_j)_{i=1,\dots,n; j=1,\dots,p} \in \mathbb{R}^{n \times p}$. Recall that the *Hamiltonian* is defined as

$$\mathcal{H}(t, x, u, q, M, N) = M'[\mu(t, x, u) + \sigma(t, x, u)q] + \frac{1}{2} \text{tr}(\sigma(t, x, u)' N \sigma(t, x, u)) + C(x, t, x; q) \quad (11)$$

for $(t, x, u, q, M, N) \in [0, T] \times \mathbb{R}^p \times \mathbb{R}^k \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{p \times p}$. By Appendix A, the adjustment term is given by

$$\mathcal{L}(t, x, u, q, h, g) := (\mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, y))|_{y=x} - \mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, x) + \frac{\partial G(x, y)}{\partial y} \Big|_{y=g(t, x)} \mathcal{A}^{\mathbf{u}, \mathbf{q}} g(t, x) - \mathcal{A}^{\mathbf{u}, \mathbf{q}} G(x, g(t, x)). \quad (12)$$

Then we are ready to define the HJBI equations.

Definition 4 *The extended HJBI equations for V , h , and g are defined as follows.*

(1) *The function $V(t, x)$ is determined by the HJBI equation:*

$$\frac{\partial V}{\partial t} + \sup_u \inf_q \left\{ \mathcal{H} \left(t, x, u, q, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q, h, g) \right\} = 0, \text{ for } t \in [0, T] \quad (13)$$

with $V(T, x) = F(x, x) + G(x, x)$, where \mathcal{H} and \mathcal{L} are given in (11) and (12), respectively.

(2) *For any $y \in \mathbb{R}^p$, the function $h(t, x, y)$ solves the partial differential equation (PDE):*

$$\mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} h(t, x, y) + C(y, t, x; \mathbf{q}^*) = 0, \text{ for } t \in [0, T] \quad \text{with } h(T, x, y) = F(y, x), \quad (14)$$

where $(\mathbf{u}^*, \mathbf{q}^*) = (\{u^*(t, x)\}_{t \in [0, T]}, \{q^*(t, x, u^*)\}_{t \in [0, T]})$ in which $q^*(t, x, u)$ realizes the infimum and $u^*(t, x)$ realizes the supremum in (13).

(3) *The function $g(t, x)$ solves the PDE:*

$$\mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} g(t, x) = 0, \text{ for } t \in [0, T] \quad \text{with } g(T, x) = x. \quad (15)$$

It is noteworthy that (13)-(15) have to be solved simultaneously because of their entangled relationship.

3.2 Verification Theorem

In this subsection, we provide a verification theorem that the function V and the optimal policy $(\mathbf{u}^*, \mathbf{q}^*)$ solved from (13) coincide with the robust value function and the robust time-consistent control, respectively. In other words, they satisfy the conditions in Definition 3.

We first introduce some mathematical notations. We let $\mathbb{C}^{1,2}([0, T] \times \mathbb{R}^p)$ be the set of all continuous functions $f : [0, T] \times \mathbb{R}^p \rightarrow \mathbb{R}$ such that $\partial f / \partial t$, $\partial f / \partial x'$ and $\partial^2 f / \partial x \partial x'$ are all continuous in (t, x) . Similarly, we let $\mathbb{C}^{1,2,2}([0, T] \times \mathbb{R}^p \times \mathbb{R}^p)$ be the set of all continuous functions $f : [0, T] \times \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$ such that $\partial f / \partial t$, $\partial f / \partial x'$, $\partial^2 f / \partial x \partial x'$, $\partial f / \partial y'$, and $\partial^2 f / \partial y \partial y'$ are all continuous in (t, x, y) . For $i, j \in \mathbb{N}$, we define

$$\mathbb{L}_{\mathcal{F}^{\mathbb{Q}}}^{i, loc}(0, T; \mathbb{R}^j) = \left\{ Y : [0, T] \times \Omega \rightarrow \mathbb{R}^j \mid Y(\cdot) \text{ is } \{\mathcal{F}_t\}_{t \in [0, T]} \text{-adapted and } \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}} \left[\int_0^T \|Y(t)\|_i dt \right] < \infty \right\}.$$

Theorem 5 Assume that $V(t, x)$, $h(t, x, y)$, and $g(t, x)$ solve the equations (13)-(15), $V, g \in \mathbb{C}^{1,2}([0, T] \times \mathbb{R}^p)$, $h \in \mathbb{C}^{1,2,2}([0, T] \times \mathbb{R}^p \times \mathbb{R}^p)$ and for $f(t, x) \in \{V(t, x), h(t, x, \cdot), g(t, x), G(x, g(t, x))\}$, we have

$$\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}[\mu + \sigma q], \operatorname{tr} \left(\sigma' \frac{\partial^2 f}{\partial x \partial x'} \sigma \right) \in \mathbb{L}_{\mathcal{F}^2}^{1,loc}(0, T; \mathbb{R}), \quad \sigma' \frac{\partial f}{\partial x'} \in \mathbb{L}_{\mathcal{F}^2}^{2,loc}(0, T; \mathbb{R}^m).$$

Moreover, assume that there is an admissible policy $(\mathbf{u}^*, \mathbf{q}^*)$, which realizes the supremum and infimum in (13):

$$\begin{aligned} & \mathcal{H} \left(t, x, u^*, q^*(u^*), \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u^*, q^*(u^*), h, g) \\ &= \sup_u \left\{ \mathcal{H} \left(t, x, u, q^*(u), \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q^*(u), h, g) \right\} \\ &= \sup_u \inf_{q(u)} \left\{ \mathcal{H} \left(t, x, u, q(u), \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q(u), h, g) \right\} \end{aligned}$$

for any (t, x) . Then $(\mathbf{u}^*, \mathbf{q}^*)$ is the robust time-consistent control and V is the robust value function in Definition 3.

Remark: This paper provides the verification theorem by assuming the smoothness of the value function, which is true for the examples we considered. The non-smoothness will unavoidably arise from the problem with generic control constraints and in this case, the viscosity solution technique should be adopted for the verification. However, finding conditions guaranteeing the smoothness is a difficult open problem even for the case without model uncertainty; see the conclusion and open problems discussed in [Björk *et al.*, 2017]. The pain point is the systemic complexity of equations (13)-(15). We refer to [Strulovici and Szydlowski, 2015] for the study on the smoothness of value functions and [Yong and Zhou, 1999] for the viscosity solution technique. Another related open problem is to prove that the value function is a unique solution to the extended HJBI equations. Due to its complexity, there is very little progress in proving the existence and uniqueness of the equilibrium solution. We refer to [Hu *et al.*, 2017] for proving the uniqueness of the equilibrium for linear-quadratic objectives without model uncertainty, where the open-loop control approach (that is more restricted compared to the framework of this paper) and the maximum principle approach are used. However, these problems are beyond the scope of this paper and remain open.

3.3 Reduction to Extended HJB System

To solve the HJBI equation (13), we note that $V(t, x) = h(t, x, x) + G(x, g(t, x))$. Then, we have

$$\begin{aligned} & \frac{\partial V}{\partial t} + \mathcal{H} \left(t, x, u, q, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q, h, g) = \mathcal{A}^{\mathbf{u}, \mathbf{q}} V(t, x) + C(x, t, x; q) + \mathcal{L}(t, x, u, q, h, g) \\ &= \mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, x) + \mathcal{A}^{\mathbf{u}, \mathbf{q}} G(x, g(t, x)) + C(x, t, x; q) + \mathcal{L}(t, x, u, q, h, g) \\ &= (\mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, y))|_{y=x} + \frac{\partial G(x, y)}{\partial y} \Big|_{y=g(t, x)} \mathcal{A}^{\mathbf{u}, \mathbf{q}} g(t, x) + C(x, t, x; q) \end{aligned}$$

Hence, (13) becomes

$$\sup_u \inf_q \left\{ (\mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, y))|_{y=x} + \frac{\partial G(x, y)}{\partial y} \Big|_{y=g(t, x)} \mathcal{A}^{\mathbf{u}, \mathbf{q}} g(t, x) + C(x, t, x; q) \right\} = 0. \quad (16)$$

Recall that C takes the form (6). By minimizing the quadratic form of q in (16), we can obtain the expression of q^* :

$$q^*(t, x, u) = -\Phi(t, x) \Xi(x) \sigma(t, x, u)' \left[\frac{\partial h(t, x, y)}{\partial x} \Big|_{y=x} + \frac{\partial G(x, y)}{\partial y} \Big|_{y=g(t, x)} \frac{\partial g(t, x)}{\partial x} \right]'. \quad (17)$$

Let $\mathbf{q}^* = \mathbf{q}^*(\mathbf{u}) = \{q^*(t, x, u)\}_{t \in [0, T]}$. The HJBI equation (16) becomes a highly nonlinear HJB equation:

$$\sup_u \left\{ (\mathcal{A}^{\mathbf{u}, \mathbf{q}^*} h(t, x, y))|_{y=x} + \frac{\partial G(x, y)}{\partial y} \Big|_{y=g(t, x)} \mathcal{A}^{\mathbf{u}, \mathbf{q}^*} g(t, x) + C(x, t, x; \mathbf{q}^*) \right\} = 0. \quad (18)$$

Therefore, the extended HJBI system of (13)-(15) is reduced to an extended HJB system of (18), (14), and (15). If one can solve for \mathbf{u}^* from (18) and h and g from (14) and (15), respectively, and verify the conditions in Theorem 5, then \mathbf{u}^* is the robust time-consistent control.

From the previous discussion, it is easy to see that we can also generalize the ambiguity preference function $\Phi(s, X_s^{\mathbf{u}})\Xi(x)$ in (6) to any positive-definite $\mathbb{R}^{m \times m}$ -valued function $\Phi^*(t, x, s, X_s^{\mathbf{u}})$; in this case, q^* takes the form (17) with $\Phi(t, x)\Xi(x)$ replaced by $\Phi^*(t, x, t, x)$. It is noteworthy that the specification of the ambiguity preference function determines the solvability of the extended HJB system. In the next section, we use financial applications of robust portfolio selection to illustrate the proposed framework and the choice of the ambiguity preference function.

4 Example: Robust Dynamic Mean-Variance Portfolio Selection

We consider a frictionless financial market with p risky assets and one risk-free asset, which can be continuously traded over an investment horizon $[0, T]$. We speculate that the p risky assets prices, denoted by $\{S_t\}_{t \in [0, T]} = \{(S_{1t}, \dots, S_{pt})'\}_{t \in [0, T]}$, evolve as follows.

$$dS_t = \text{diag}(S_t)[\mu_t dt + \sigma_t dW_t^{\mathbb{P}}],$$

where $\text{diag}(S_t)$ is a diagonal matrix with the diagonal elements S_t , $\mu_t \in \mathbb{R}^p$ is the appreciate rate vector, and $\sigma_t \in \mathbb{R}^{p \times m}$ is the volatility matrix of the assets returns. The risk-free asset price S_{0t} grows exponentially as $dS_{0t} = r_t S_{0t} dt$, where r_t is the risk-free interest rate at time t .

Denote by $u_i(t)$ the money amounts invested in the i th risky asset and $X^{\mathbf{u}}(t)$ by the value of the portfolio $\mathbf{u} = \{u(t)\}_{t \in [0, T]} := \{(u_1(t), \dots, u_p(t))'\}_{t \in [0, T]}$ at time t . Then by the self-financing principle, the money invested in the risk-free asset at time t is $X^{\mathbf{u}}(t) - u(t)'\mathbf{1}_p$. Therefore, the dynamic of $X^{\mathbf{u}}$ is given by

$$dX^{\mathbf{u}}(t) = [r_t X^{\mathbf{u}}(t) + u(t)'\beta_t]dt + u(t)'\sigma_t dW_t^{\mathbb{P}},$$

with the initial wealth $X(0) = x_0$, where $\beta_t := \mu_t - r_t \mathbf{1}_p$ is the excess return vector. Hence, with $\mu(t, X_t, u_t) = r_t X_t + u_t'\beta_t$ and $\sigma(t, X_t, u_t) = u_t'\sigma_t$, an ambiguity-averse investor considers alternative dynamics of X_t as in (3), characterized by q_t . We let $\Sigma_t = \sigma_t \sigma_t'$ and $\Theta_t = \beta_t' \Sigma_t \beta_t$. We assume that the functions r_t , β_t , σ_t , and Θ_t are deterministic, Lipschitz continuous, and uniformly bounded functions in t over $[0, T]$.

A robust mean-variance (MV) investor aims to find a robust portfolio (\mathbf{u}, \mathbf{q}) that maximizes the MV reward functional under the worst-case scenario for every (t, x) :

$$\begin{aligned} \sup_{\mathbf{u}} \inf_{\mathbf{q}(\mathbf{u})} J(t, x; \mathbf{u}, \mathbf{q}) &= \sup_{\mathbf{u}} \inf_{\mathbf{q}(\mathbf{u})} \left\{ \mathbb{E}_{t,x}^{\mathbb{Q}}[X_T^{\mathbf{u}}] - \frac{\gamma(x)}{2} \text{Var}_{t,x}^{\mathbb{Q}}(X_T^{\mathbf{u}}) + D_{t,x}(\mathbb{Q} \parallel \mathbb{P}) \right\} \\ &= \sup_{\mathbf{u}} \inf_{\mathbf{q}(\mathbf{u})} \left\{ \mathbb{E}_{t,x}^{\mathbb{Q}} \left[X_T^{\mathbf{u}} - \frac{\gamma(x)}{2} (X_T^{\mathbf{u}})^2 \right] + \frac{\gamma(x)}{2} (\mathbb{E}_{t,x}^{\mathbb{Q}}[X_T^{\mathbf{u}}])^2 + D_{t,x}(\mathbb{Q} \parallel \mathbb{P}) \right\}, \end{aligned} \quad (19)$$

where $\gamma(x)$ is the risk-aversion coefficient function of the investor and $D_{t,x}$ takes the form (6). Using the notations in this paper, this reward functional is equal to (5) with

$$F(y, x) = x - \frac{\gamma(y)}{2} x^2, \quad G(x, y) = \frac{\gamma(x)}{2} y^2.$$

From the previous discussion, in order to find the robust portfolio, we need to solve for the extended HJBI system of (13)-(15) or the extended HJB system of (18), (14), and (15). Next, we discuss the robust time-inconsistent stochastic control problem above for two cases: 1. $\gamma(x) = \gamma$ that is a positive constant; 2. $\gamma(x) = \gamma/x$ that is state-dependent. The first case corresponds to a classical MV framework, originally proposed by [Markowitz, 1952] while the time-consistent (dynamic) MV portfolio is studied in [Basak and Chabakauri, 2010]. The second case has been shown to be economically meaningful by empirical evidence and is considered in [Björk *et al.*, 2014]. For simplicity, we consider $\Xi = \xi I_{m \times m}$ in (6) with an aggregate measure of ambiguity aversion $\xi > 0$.

4.1 The Case of Constant Risk Aversion

We assume $\gamma(x) = \gamma$ in (19). For this case, we specify the function $\Phi(s, y) \equiv 1$, i.e. $D_{t,x}(\mathbb{Q}||\mathbb{P}) = \frac{1}{\xi} D_{KL}(\mathbb{Q}||\mathbb{P})$, a scaled KL divergence. We notice that since $F(y, x)$ is independent of y in this case, $h(t, x, y)$ is also independent of y and $(\mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, y))|_{y=x} - \mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, x) = 0$. Then (17) becomes

$$q^*(t, x, u) = -\xi \frac{\partial V(t, x)}{\partial x} \sigma'_t u$$

and the HJBI equation (13) is reduced to a HJB equation

$$\sup_u \left\{ \frac{\partial V}{\partial t} + [r_t x + u'_t \beta_t] \frac{\partial V}{\partial x} + \frac{1}{2} u'_t \Sigma_t u_t \left[\frac{\partial^2 V}{\partial x^2} - \gamma \left(\frac{\partial g}{\partial x} \right)^2 - \xi \left(\frac{\partial V}{\partial x} \right)^2 \right] \right\} = 0. \quad (20)$$

By maximizing the quadratic form of u_t in (20), we obtain the expression of the robust portfolio:

$$u^*(t, x) = -\Sigma_t^{-1} \beta_t \frac{\partial V / \partial x}{\partial^2 V / \partial x^2 - \gamma (\partial g / \partial x)^2 - \xi (\partial V / \partial x)^2},$$

where V and g solve the following coupled PDEs, which are deduced from (20) and (15) with u^* and q^* , respectively:

$$\frac{\partial V}{\partial t} + r_t x \frac{\partial V}{\partial x} - \frac{\Theta_t}{2} \frac{(\partial V / \partial x)^2}{\partial^2 V / \partial x^2 - \gamma (\partial g / \partial x)^2 - \xi (\partial V / \partial x)^2} = 0, \quad (21)$$

$$\begin{aligned} & \frac{\partial g}{\partial t} + \left[r_t x - \Theta_t \frac{\partial V / \partial x}{\partial^2 V / \partial x^2 - \gamma (\partial g / \partial x)^2 - \xi (\partial V / \partial x)^2} \right] \frac{\partial g}{\partial x} \\ & + \frac{\Theta_t}{2} \frac{(\partial V / \partial x)^2}{[\partial^2 V / \partial x^2 - \gamma (\partial g / \partial x)^2 - \xi (\partial V / \partial x)^2]^2} \left[\frac{\partial^2 g}{\partial x^2} - 2\xi \frac{\partial V}{\partial x} \right] = 0 \end{aligned} \quad (22)$$

with the boundary conditions $V(T, x) = x$ and $g(T, x) = x$.

To solve (21) and (22), we make an ansatz:

$$V(t, x) = V_1(t)x + V_0(t) \quad \text{and} \quad g(t, x) = g_1(t)x + g_0(t).$$

Substituting them into (21) and (22) yields

$$\left[\frac{\partial V_1}{\partial t} + r_t V_1(t) \right] x + \frac{\partial V_0}{\partial t} + \frac{\Theta_t}{2} \frac{V_1^2(t)}{\gamma g_1^2(t) + \xi V_1^2(t)} = 0 \quad \text{and} \quad \left[\frac{\partial g_1}{\partial t} + r_t g_1(t) \right] x + \frac{\partial g_0}{\partial t} + \Theta_t \frac{V_1(t)g_1(t) + \xi V_1^3(t)}{\gamma g_1^2(t) + \xi V_1^2(t)} = 0$$

with $V_1(T) = g_1(T) = 1$ and $V_0(T) = g_0(T) = 0$. By separating the terms by the order of x , it is easy to identify the linear ordinary differential equations (ODEs) driving V_0 , V_1 , g_0 , and g_1 . Moreover, they are given by

$$V_1(t) = e^{\int_t^T r_s ds}, \quad V_0(t) = \frac{1}{2(\gamma + \xi)} \int_t^T \Theta_s ds, \quad g_1(t) = e^{\int_t^T r_s ds}, \quad g_0(t) = \frac{1}{\gamma + \xi} \int_t^T \Theta_s (1 + \xi e^{\int_s^T r_r dr}) ds.$$

It is easy to verify the functions V and g satisfy the conditions in Theorem 5 and u^* and q^* are admissible. Based on Theorem 5, the following proposition summarizes the solution to the robust MV portfolio selection (19) with constant risk aversion.

Proposition 6 *The robust mean-variance portfolio selection problem (19) with $\gamma(x) = \gamma$ and $\Phi(\cdot, \cdot) \equiv 1$ admits the robust value function:*

$$V(t, x) = x e^{\int_t^T r_s ds} + \frac{1}{2(\gamma + \xi)} \int_t^T \Theta_s ds$$

and the robust time-consistent control $(\mathbf{u}^*, \mathbf{q}^*) = (\{u^*(t, x)\}_{t \in [0, T]}, \{q^*(t, x, u^*(t, x))\}_{t \in [0, T]})$:

$$u^*(t, x) = \frac{e^{-\int_t^T r_s ds}}{\gamma + \xi} \Sigma_t^{-1} \beta_t, \quad q^*(t, x, u^*(t, x)) = -\frac{\xi}{\xi + \gamma} \sigma'_t \Sigma_t^{-1} \beta_t.$$

4.2 The Case of State-Dependent Risk Aversion

When the risk aversion coefficient $\gamma(x)$ is state-dependent, $h(t, x, y)$ in our framework is dependent on y . It is more convenient to adopt the simplification in Section 3.3. For this case, we choose the ambiguity preference function of the form

$$\Phi(t, x) = -\frac{(\partial^2 h(t, x, y)/\partial x^2)|_{y=x} + (\partial G(x, y)/\partial y)|_{y=g(t, x)}(\partial^2 g(t, x)/\partial x^2)}{[(\partial h(t, x, y)/\partial x)|_{y=x} + (\partial G(x, y)/\partial y)|_{y=g(t, x)}(\partial g(t, x)/\partial x)]^2}. \quad (23)$$

This choice of Φ is economically meaningful and facilitates analytical tractability; we interpret it in detail in the next subsection. With (23), the q^* in (17) becomes

$$q^*(t, x, u) = \xi \frac{(\partial^2 h(t, x, y)/\partial x^2)|_{y=x} + \gamma(x)g(t, x)(\partial^2 g(t, x)/\partial x^2)}{(\partial h(t, x, y)/\partial x)|_{y=x} + \gamma(x)g(t, x)(\partial g(t, x)/\partial x)} \sigma'_t u$$

and the HJB equation (18) can be rewritten as

$$\sup_u \left\{ \cdots + u'_t \beta_t \left[\frac{\partial h(t, x, y)}{\partial x} \Big|_{y=x} + \gamma(x)g(t, x) \frac{\partial g(t, x)}{\partial x} \right] + \frac{\xi + 1}{2} u'_t \Sigma_t u_t \left[\frac{\partial^2 h(t, x, y)}{\partial x^2} \Big|_{y=x} + \gamma(x)g(t, x) \frac{\partial^2 g(t, x)}{\partial x^2} \right] + \cdots \right\} = 0,$$

where we suppressed the terms without u_t . By maximizing the quadratic form of u_t above, we obtain the expression of the robust portfolio:

$$u^*(t, x) = -\frac{1}{\xi + 1} \Sigma_t^{-1} \beta_t \frac{(\partial h(t, x, y)/\partial x)|_{y=x} + \gamma(x)g(t, x)(\partial g(t, x)/\partial x)}{(\partial^2 h(t, x, y)/\partial x^2)|_{y=x} + \gamma(x)g(t, x)(\partial^2 g(t, x)/\partial x^2)},$$

where $h(t, x, y)$ and $g(t, x)$ are solutions to PDEs (14) and (15), respectively.

We now assume $\gamma(x) = \gamma/x$ in (19) with a positive constant γ , as in [Björk *et al.*, 2014]. Inspired by the boundary conditions of h and g , i.e. $h(T, x, y) = x - \frac{\gamma}{2y}x^2$ and $g(T, x) = x$, we make an ansatz:

$$h(t, x, y) = h_1(t)x - \frac{\gamma}{2y}h_2(t)x^2 \quad \text{and} \quad g(t, x) = g_1(t)x. \quad (24)$$

Then, the robust time-consistent control can be rewritten as

$$q^*(t, x, u^*) = -\frac{\xi}{\xi + 1} \sigma'_t \Sigma_t^{-1} \beta_t, \quad u^*(t, x) = \frac{x}{\gamma(\xi + 1)} \Sigma_t^{-1} \beta_t \frac{h_1(t) + \gamma[g_1^2(t) - h_2(t)]}{h_2(t)}.$$

Substituting (24) into (14) and (15) yields the following two ODEs:

$$\frac{\partial h_1}{\partial t} x - \frac{\gamma}{2y} x^2 \frac{\partial h_2}{\partial t} + x \left[r_t + \frac{\Theta_t}{\gamma(\xi + 1)^2} \frac{h_1(t) + \gamma[g_1^2(t) - h_2(t)]}{h_2(t)} \right] \left(h_1(t) - \frac{\gamma}{y} h_2(t)x \right) \quad (25)$$

$$- \frac{\gamma}{2y} x^2 \frac{\Theta_t}{\gamma^2(\xi + 1)^2} \frac{[h_1(t) + \gamma[g_1^2(t) - h_2(t)]]^2}{h_2(t)} + x \frac{\xi \Theta_t}{2\gamma(\xi + 1)^2} \frac{[h_1(t) + \gamma[g_1^2(t) - h_2(t)]]^2}{h_2(t)} = 0,$$

$$\frac{\partial g_1}{\partial t} x + x \left[r_t + \frac{\Theta_t}{\gamma(\xi + 1)^2} \frac{h_1(t) + \gamma[g_1^2(t) - h_2(t)]}{h_2(t)} \right] g_1(t) = 0 \quad (26)$$

with the boundary conditions $h_1(T) = h_2(T) = g_1(T) = 1$. Let

$$f_1(t) := \frac{h_1(t)}{h_2(t)}, \quad f_2(t) := \frac{g_1^2(t)}{h_2(t)}, \quad \text{and } f(t) := f_1(t) + \gamma(f_2(t) - 1) = \frac{h_1(t) + \gamma[g_1^2(t) - h_2(t)]}{h_2(t)}.$$

By separating the terms in (25)-(26) by the order of x , we identify the ODEs for h_1 , h_2 , and g_1 :

$$\begin{aligned} \frac{\partial h_1}{\partial t} + \left[r_t + \frac{\Theta_t}{\gamma(\xi+1)^2} f(t) \right] h_1(t) + \frac{\xi \Theta_t}{2\gamma(\xi+1)^2} f^2(t) h_2(t) &= 0, \\ \frac{\partial h_2}{\partial t} + \left[2r_t + \frac{2\Theta_t}{\gamma(\xi+1)^2} f(t) + \frac{\Theta_t}{\gamma^2(\xi+1)^2} f^2(t) \right] h_2(t) &= 0, \\ \frac{\partial g_1}{\partial t} + \left[r_t + \frac{\Theta_t}{\gamma(\xi+1)^2} f(t) \right] g_1(t) &= 0 \end{aligned}$$

with $h_1(T) = h_2(T) = g_1(T) = 1$. Hence,

$$h_2(t) = \exp \left\{ \int_t^T \left[2r_s + \frac{2\Theta_s}{\gamma(\xi+1)^2} f(s) + \frac{\Theta_s}{\gamma^2(\xi+1)^2} f^2(s) \right] ds \right\}, \quad (27)$$

$$g_1(t) = \exp \left\{ \int_t^T \left[r_s + \frac{\Theta_s}{\gamma(\xi+1)^2} f(s) \right] ds \right\}. \quad (28)$$

Moreover, we can deduce the ODEs for $f_1 = h_1/h_2$ and $f_2 = g_1^2/h_2$:

$$\frac{\partial f_1}{\partial t} - r_t f_1(t) - \frac{\Theta_t}{\gamma(\xi+1)^2} f(t) f_1(t) - \frac{\Theta_t}{\gamma^2(\xi+1)^2} f^2(t) f_1(t) + \frac{\xi \Theta_t}{2\gamma(\xi+1)^2} f^2(t) = 0, \quad (29)$$

$$\frac{\partial f_2}{\partial t} - \frac{\Theta_t}{\gamma^2(\xi+1)^2} f^2(t) f_2(t) = 0 \quad (30)$$

with the boundary conditions $f_1(T) = f_2(T) = 1$, where we recall that $f(t) = f_1(t) + \gamma(f_2(t) - 1)$. The first-order system of ODEs (29)-(30) can be rewritten as

$$\frac{\partial f_1}{\partial t} = P(t, f_1, f_2), \quad \frac{\partial f_2}{\partial t} = Q(t, f_1, f_2),$$

where P and Q are cubic polynomial functions of (f_1, f_2) with bounded Lipschitz continuous coefficient functions of t . This cubic system has the Abel form; thus (29) and (30) are called *the Abel differential equations*. By the uniqueness theorem, there exists a unique solution to (29)-(30); see [Zwillinger, 1997]. However, it is hard to obtain the explicit solution to the Abel differential equations. One may rely on numerical or asymptotic methods as in [Pun, 2017]. This cubic system is also related to an integral equation for $f(t)$ when we express $f_1(t)$ and $f_2(t)$ in terms of $f(t)$:

$$\begin{aligned} f(t) &= e^{-\int_t^T \left[r_s + \frac{\Theta_s}{\gamma(\xi+1)^2} f(s) + \frac{\Theta_s}{\gamma^2(\xi+1)^2} f^2(s) \right] ds} + \gamma e^{-\int_t^T \frac{\Theta_s}{\gamma^2(\xi+1)^2} f^2(s) ds} - \gamma \\ &\quad + \frac{\xi}{2\gamma(\xi+1)^2} \int_t^T e^{-\int_t^s \left[r_\tau + \frac{\Theta_\tau}{\gamma(\xi+1)^2} f(\tau) + \frac{\Theta_\tau}{\gamma^2(\xi+1)^2} f^2(\tau) \right] d\tau} \Theta_s f^2(s) ds. \end{aligned}$$

The similar degenerate ($\xi = 0$) integral equation was studied in [Björk *et al.*, 2014].

The following proposition summarizes the discussion and the results in this subsection.

Proposition 7 *The robust mean-variance portfolio selection problem (19) with $\gamma(x) = \gamma/x$ and Φ taking the form (23) admits the robust value function:*

$$V(t, x) = \left[h_1(t) + \frac{\gamma}{2} (g_1^2(t) - h_2(t)) \right] x$$

and the robust time-consistent control $(\mathbf{u}^*, \mathbf{q}^*) = (\{u^*(t, x)\}_{t \in [0, T]}, \{q^*(t, x, u^*(t, x))\}_{t \in [0, T]})$:

$$u^*(t, x) = \frac{f(t)x}{\gamma(\xi + 1)} \Sigma_t^{-1} \beta_t, \quad q^*(t, x, u^*(t, x)) = -\frac{\xi}{\xi + 1} \sigma_t' \Sigma_t^{-1} \beta_t,$$

where $h_1(t) = f_1(t)h_2(t)$, $h_2(t)$ is given in (27), $g_1(t)$ is given in (28), $f(t) = f_1(t) + \gamma(f_2(t) - 1)$, and $f_1(t)$ and $f_2(t)$ are jointly determined by the Abel differential equations (29)-(30).

Recall that [Maenhout, 2004] found that the robustness on the power-utility-maximization portfolio just enlarges the risk aversion of the investor. In contrast, we provide an example that the robustness effect on u^* in Proposition 7 is highly nonlinear by noting the dependence of f on ξ .

4.3 Some Remarks on Ambiguity Preference Function Φ

The choice of $\Phi(t, x)$ is applicable for the general reward functional (5) when the dimension of the state process is one. When there is no G term and F is independent of y , i.e. no source of time-inconsistency, (23) is reduced to

$$\Phi(t, x) = -\frac{\partial^2 V(t, x)/\partial x^2}{(\partial V(t, x)/\partial x)^2} = \frac{1}{R(t, x)(\partial V(t, x)/\partial x)},$$

where $R(t, x) := -\frac{\partial V(t, x)/\partial x}{\partial^2 V(t, x)/\partial x^2}$ is the risk-tolerance function. This is identical to the proposal in [Pun and Wong, 2015] and consistent with the specification of Φ in [Maenhout, 2004], which considers a special case of power utility maximization. Therefore, the specification of Φ in (23) generalizes the studies of model uncertainty for time-consistent preference in [Maenhout, 2004] and [Pun and Wong, 2015] to time-inconsistent preference. Moreover, the risk-tolerance function for an agent with time-inconsistent preference can be defined similarly as

$$\tilde{R}(t, x) = -\frac{(\partial h(t, x, y)/\partial x)|_{y=x} + (\partial G(x, y)/\partial y)|_{y=g(t, x)}(\partial g(t, x)/\partial x)}{(\partial^2 h(t, x, y)/\partial x^2)|_{y=x} + (\partial G(x, y)/\partial y)|_{y=g(t, x)}(\partial^2 g(t, x)/\partial x^2)}.$$

The specification of $\Phi(t, x)$ in (23) is reasonable in the sense that Φ is monotonically decreasing with respect to the risk-tolerance function \tilde{R} ; thus (23) preserves the close relationship between risk and ambiguity aversions. Specifically, for the case of state-dependent risk aversion in the previous example, the ambiguity preference function is given by

$$\Phi(t, x) = \frac{\gamma}{x} \frac{h_2(t)}{[h_1(t) + \gamma[g_1^2(t) - h_2(t)]]^2},$$

which is also inversely proportional to the state x and is an affine map of the risk aversion coefficient function $\gamma(x) = \gamma/x$. This imposes the desired homotheticity property for robustness.

However, the specification of Φ in (23) is subject to an ansatz:

$$\frac{\partial^2 h(t, x, y)}{\partial x^2} \Big|_{y=x} + \frac{\partial G(x, y)}{\partial y} \Big|_{y=g(t, x)} \frac{\partial^2 g(t, x)}{\partial x^2} < 0$$

such that $\Phi(t, x) > 0$. For the example with $\gamma(x) = \gamma/x$, the $\Phi(t, x)$ as shown above is positive since $h_2(t) > 0$ due to (27). However, for the example with $\gamma(x) = \gamma$, using (23) leads to an unsolvable HJB system or one can see that the ansatz made in Section 4.1 leads to $\Phi(t, x) = 0$ in (23). Therefore, for this case, the suitable choice of Φ is simply a positive constant (as one) and the divergence $D_{t, x}(\mathbb{Q}||\mathbb{P})$ essentially becomes KL divergence that is interpretable. We stress that these two choices of Φ in different settings are actually consistent in the sense that the optimal q^* 's in these two cases are essentially the same. To see this, we can think of using $\Phi(t, x) = \gamma$ in Section 4.1 such that the results, especially q^* , are amended with ξ replaced by $\gamma\xi$ or we note that ξ is a free parameter and can take value of $\gamma\xi$.

5 Conclusion

We established a general tractable framework for robust time-inconsistent stochastic control problems and studied a specific example of robust dynamic mean-variance portfolio selection. Moreover, we discussed the modeling of ambiguity preference in detail. Interestingly, we found that the effect of ambiguity could be complicated (Section 4.2), in contrast with the finding of “the robustness as risk aversion multiplier” in [Maenhout, 2004]. It is noteworthy that Section 4.2 not only extends the results in [Björk *et al.*, 2014] to the robust counterparts but also provides a differential equation approach to investigate the time-inconsistent control problems. It is well-known that (matrix) Riccati equations have significant applications in linear-quadratic (time-consistent) stochastic control problems, especially portfolio selection in finance. This paper enriches the control and mathematical finance literature by introducing the Abel differential equations (or so-called extended Riccati equations) for the time-inconsistent stochastic control problems. Future research could be devoted to the properties of the solution to matrix Abel differential equations.

The model uncertainty in this paper, which accords with the rule of model misspecification in [Anderson *et al.*, 2003], is studied for a set of models dominated by the reference model. There are several studies in the context of portfolio selection considering non-dominated models as alternatives; see [Fouque *et al.*, 2016] and [Ismail and Pham, 2016]. It will be interesting to extend the framework in this paper to the non-dominated models such that one can introduce the ambiguity directly into the state’s diffusion coefficient.

Acknowledgements

The author thanks the editor, the associate editor, and three anonymous referees for their valuable comments and their efforts towards improving the manuscript. Chi Seng Pun gratefully acknowledges Nanyang Technological University (NTU), Start-Up Grant (SUG) and Ministry of Education (MOE), AcRF Tier 2 grant (Reference No: MOE2017-T2-1-044) for the funding of this research.

References

- [Anderson *et al.*, 2003] E. W. Anderson, L. P. Hansen, and T. J. Sargent (2003). A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection. *Journal of the European Economic Association* 1, 68–123.
- [Bannister *et al.*, 2016] H. Bannister, B. Goldys, S. Penev, and W. Wu (2016). Multiperiod Mean-Standard-Deviation Time Consistent Portfolio Selection. *Automatica* 73, 15–26.
- [Basak and Chabakauri, 2010] S. Basak and G. Chabakauri (2010). Dynamic Mean-Variance Asset Allocation. *The Review of Financial Studies* 23, 2970–3016.
- [Biagini and Pınar, 2017] S. Biagini and M. Ç. Pınar (2017). The Robust Merton Problem of an Ambiguity Averse Investor. *Mathematics and Financial Economics* 11, 1–24.
- [Björk and Murgoci, 2014] T. Björk and A. Murgoci (2014). A Theory of Markovian Time-Inconsistent Stochastic Control in Discrete Time. *Finance and Stochastics* 18, 545–592.
- [Björk *et al.*, 2017] T. Björk, M. Khapko, and A. Murgoci (2017). On Time-Inconsistent Stochastic Control in Continuous Time. *Finance and Stochastics* 21, 331–360.
- [Björk *et al.*, 2014] T. Björk, A. Murgoci, and X. Y. Zhou (2014). Mean-Variance Portfolio Optimization with State-Dependent Risk Aversion. *Mathematical Finance* 24, 1–24.
- [Chiu *et al.*, 2017] M. C. Chiu, C. S. Pun, and H. Y. Wong (2017). Big Data Challenges of High-Dimensional Continuous-Time Mean-Variance Portfolio Selection and a Remedy. *Risk Analysis* 37, 1532–1549.
- [Cui *et al.*, 2010] X. Y. Cui, D. Li, S. Y. Wang, and S. S. Zhu (2010). Better than Dynamic MeanVariance: Time Inconsistency and Free Cash Flow Stream. *Mathematical Finance* 22, 346–378.
- [Cui *et al.*, 2017] X. Y. Cui, D. Li, and Y. Shi (2017). Self-Coordination in Time Inconsistent Stochastic Decision Problems: A Planner-Doer Game Framework. *Journal of Economic Dynamics & Control* 75, 91–113.
- [Ellsberg, 1961] D. Ellsberg (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics* 75, 643–669.
- [Fouque *et al.*, 2016] J.-P. Fouque, C. S. Pun, and H. Y. Wong (2016). Portfolio Optimization with Ambiguous Correlation and Stochastic Volatilities. *SIAM Journal on Control and Optimization* 54, 2309–2338.
- [Goldman, 1980] S. Goldman (1980). Consistent Plans. *The Review of Economic Studies* 47, 533–537.
- [Hansen *et al.*, 2006] L. P. Hansen, T.J. Sargent, G. Turmuhambetova, and N. Williams (2006). Robust Control and Model Misspecification. *Journal of Economic Theory* 128, 45–90.
- [Heston, 1993] S. L. Heston (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency. *The Review of Financial Studies* 6, 327–343.

- [Hu *et al.*, 2017] Y. Hu, H. Jin, and X. Y. Zhou (2017). Time-Inconsistent Stochastic Linear-Quadratic Control: Characterization and Uniqueness of Equilibrium. *SIAM Journal on Control and Optimization* 55, 1261–1279.
- [Ismail and Pham, 2016] A. Ismail and H. Pham (2016). Robust Markowitz Mean-Variance Portfolio Selection under Ambiguous Covariance Matrix. Forthcoming in *Mathematical Finance*.
- [Karnam *et al.*, 2017] C. Karnam, J. Ma, and J. Zhang (2017). Dynamic Approaches for Some Time Inconsistent Problems. *Annals of Applied Probability* 27, 3435–3477.
- [Knight, 1921] F. H. Knight (1921). *Risk, Uncertainty, and Profit*. Houghton Mifflin, New York.
- [Li and Ng, 2000] D. Li and W. L. Ng (2000). Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation. *Mathematical Finance* 10, 387–406.
- [Maenhout, 2004] P. J. Maenhout (2004). Robust Portfolio Rules and Asset Pricing. *The Review of Financial Studies* 17, 951–983.
- [Markowitz, 1952] H. M. Markowitz (1952). Portfolio Selection. *The Journal of Finance* 7, 77–91.
- [Merton, 1971] R. C. Merton (1971). Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory* 3, 373–413.
- [Pedersen and Peskir, 2017] J. L. Pedersen and G. Peskir (2017). Optimal Mean-Variance Portfolio Selection. *Mathematics and Financial Economics* 11, 137–160.
- [Pun, 2017] C. S. Pun (2017). Time-Consistent Investment Strategy with Only Risky Assets and its Link with Myopic Strategy in High Dimensions. *Working paper of Nanyang Technological University*.
- [Pun *et al.*, 2015] C. S. Pun, S. F. Chung, and H. Y. Wong (2015). Variance Swap with Mean Reversion, Multifactor Stochastic Volatility and Jumps. *European Journal of Operational Research* 245, 571–580.
- [Pun and Wong, 2015] C. S. Pun and H. Y. Wong (2015). Robust Investment-Reinsurance Optimization with Multiscale Stochastic Volatility. *Insurance: Mathematics and Economics* 62, 245–256.
- [Strotz, 1955] R. Strotz (1955). Myopia and Inconsistency in Dynamic Utility Maximization. *The Review of Economic Studies* 23, 165–180.
- [Strulovici and Szydlowski, 2015] B. Strulovici and M. Szydlowski (2015). On the Smoothness of Value Functions and the Existence of Optimal Strategies in Diffusion Models. *Journal of Economic Theory* 159, 1016–1055.
- [Wald, 1945] A. Wald (1945). Statistical Decision Functions which Minimize the Maximum Risk. *The Annals of Mathematics* 46, 265–280.
- [Yong and Zhou, 1999] J. Yong and X. Y. Zhou (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer-Verlag New York.
- [Zeng *et al.*, 2016] Y. Zeng, D. Li, and A. Gu (2016). Robust Equilibrium Reinsurance-Investment Strategy for a Mean-Variance Insurer in a Model with Jumps. *Insurance: Mathematics and Economics* 66, 138–152.
- [Zhou and Li, 2000] X. Y. Zhou and D. Li (2000). Continuous-Time Mean-Variance Portfolio Selection: A Stochastic LQ Framework. *Applied Mathematics & Optimization* 42, 19–33.
- [Zwillinger, 1997] D. Zwillinger (1997). *Handbook of Differential Equations (3rd Edition)*. Academic Press.

A General Results and Proofs

The results for the reward functional (4) in Sections 2-3 can be generalized to the following reward functional that depends on current time t :

$$J(t, x; \mathbf{u}, \mathbf{q}) = \mathbb{E}_{t,x}^{\mathbb{Q}} \left[F(t, x, X_T^{\mathbf{u}}) + \int_t^T C(t, x, s, X_s^{\mathbf{u}}; \mathbf{q}) \right] + G(t, x, \mathbb{E}_{t,x}^{\mathbb{Q}}[X_T^{\mathbf{u}}]). \quad (\text{A.1})$$

Since we generalize the reward functional, we need to first modify the uniform integrability condition for admissible policy in Definition 1 as follows while other three conditions remain unchanged.

4*. (*Uniform integrability*) For every initial point $(t, x) \in [0, T] \times \mathbb{R}^p$, we have

$$\sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^T |C(t, x, s, X_s^{\mathbf{u}}; \mathbf{q})| ds + |F(t, x, X_T^{\mathbf{u}})| \right] < \infty, \quad \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{t,x}^{\mathbb{Q}}[|X_T^{\mathbf{u}}|_1] < \infty, \quad \text{and} \quad \sup_{\mathbb{Q} \in \mathcal{Q}} G(t, x, \mathbb{E}_{t,x}^{\mathbb{Q}}[X_T^{\mathbf{u}}]) < \infty.$$

We have the following lemma for the recursive equation for J :

Lemma 8 For $s > t$ and any admissible control-measure policy $(\mathbf{u}, \mathbf{q}) \in \mathbf{U} \times \mathbf{Q}$, we have

$$J(t, x; \mathbf{u}, \mathbf{q}) = \mathbb{E}_{t,x}^{\mathbb{Q}} \left[J(s, X_s^{\mathbf{u}}; \mathbf{q}) + \int_t^s C(\tau, X_\tau^{\mathbf{u}}, \tau, X_\tau^{\mathbf{u}}; \mathbf{u}, \mathbf{q}) d\tau \right] + L^{\mathbf{u}, \mathbf{q}}(t, x, s), \quad (\text{A.2})$$

where $L^{\mathbf{u}, \mathbf{q}}(t, x, s) = L_C^{\mathbf{u}, \mathbf{q}}(t, x, s) + L_H^{\mathbf{u}, \mathbf{q}}(t, x, s) + L_G^{\mathbf{u}, \mathbf{q}}(t, x, s)$ and

$$\begin{aligned} L_C^{\mathbf{u}, \mathbf{q}}(t, x, s) &= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^s [C(t, x, \tau, X_\tau^{\mathbf{u}}) - C(\tau, X_\tau^{\mathbf{u}}, \tau, X_\tau^{\mathbf{u}})] d\tau \right], \\ L_H^{\mathbf{u}, \mathbf{q}}(t, x, s) &= \mathbb{E}_{t,x}^{\mathbb{Q}}[h^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}, t, x)] - \mathbb{E}_{t,x}^{\mathbb{Q}}[h^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}, s, X_s^{\mathbf{u}})], \\ L_G^{\mathbf{u}, \mathbf{q}}(t, x, s) &= G(t, x, \mathbb{E}_{t,x}^{\mathbb{Q}}[g^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}})]) - \mathbb{E}_{t,x}^{\mathbb{Q}}[G(s, X_s^{\mathbf{u}}, g^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}))], \\ h^{\mathbf{u}, \mathbf{q}}(t, x, s, y) &:= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^T C(s, y, \tau, X_\tau^{\mathbf{u}}) d\tau + F(s, y, X_T^{\mathbf{u}}) \right], \quad g^{\mathbf{u}, \mathbf{q}}(t, x) := \mathbb{E}_{t,x}^{\mathbb{Q}}[X_T^{\mathbf{u}}]. \end{aligned} \quad (\text{A.3})$$

PROOF. By the definition of J in (A.1) and using the notations in this lemma, we have

$$J(t, x; \mathbf{u}, \mathbf{q}) = h^{\mathbf{u}, \mathbf{q}}(t, x, t, x) + G(t, x, g^{\mathbf{u}, \mathbf{q}}(t, x)).$$

Noting that $g^{\mathbf{u}, \mathbf{q}}(t, x) := \mathbb{E}_{t,x}^{\mathbb{Q}}[g^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}})]$, we have

$$\begin{aligned} G(t, x, g^{\mathbf{u}, \mathbf{q}}(t, x)) &= \mathbb{E}_{t,x}^{\mathbb{Q}}[G(s, X_s^{\mathbf{u}}, g^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}))] + G(t, x, \mathbb{E}_{t,x}^{\mathbb{Q}}[g^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}})]) - \mathbb{E}_{t,x}^{\mathbb{Q}}[G(s, X_s^{\mathbf{u}}, g^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}))] \\ &= \mathbb{E}_{t,x}^{\mathbb{Q}}[G(s, X_s^{\mathbf{u}}, g^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}))] + L_G^{\mathbf{u}, \mathbf{q}}(t, x, s). \end{aligned}$$

Moreover, by the tower rule, we have

$$\begin{aligned} h^{\mathbf{u}, \mathbf{q}}(t, x, t, x) &= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\mathbb{E}_{s, X_s^{\mathbf{u}}}^{\mathbb{Q}} \left[\int_t^T C(t, x, \tau, X_\tau^{\mathbf{u}}) d\tau + F(t, x, X_T^{\mathbf{u}}) \right] \right] \\ &= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[h^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}, s, X_s^{\mathbf{u}}) + h^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}, t, x) - h^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}, s, X_s^{\mathbf{u}}) + \int_t^s C(t, x, \tau, X_\tau^{\mathbf{u}}) d\tau \right] \\ &= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[h^{\mathbf{u}, \mathbf{q}}(s, X_s^{\mathbf{u}}, s, X_s^{\mathbf{u}}) + \int_t^s C(\tau, X_\tau^{\mathbf{u}}, \tau, X_\tau^{\mathbf{u}}) d\tau \right] + L_C^{\mathbf{u}, \mathbf{q}}(t, x, s) + L_H^{\mathbf{u}, \mathbf{q}}(t, x, s). \end{aligned}$$

Then the result follows.

By (A.2), the EDPE for the value function $V(t, x) := \sup_{\mathbf{u}} \inf_{\mathbf{q}} J(t, x; \mathbf{u}, \mathbf{q})$ is given by

$$V(t, x) = \sup_{\mathbf{u}} \inf_{\mathbf{q}} \left\{ \mathbb{E}_{t,x}^{\mathbb{Q}} \left[V(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}) + \int_t^{t+\epsilon} C(\tau, X_{\tau}^{\mathbf{u}}, \tau, X_{\tau}^{\mathbf{u}}; \mathbf{q}) d\tau \right] + L^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) \right\},$$

where

$$\begin{aligned} L^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) &= L_C^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) + L_H^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) + L_G^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon), \\ L_C^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) &= \mathbb{E}_{t,x}^{\mathbb{Q}} \left[\int_t^{t+\epsilon} [C(t, x, \tau, X_{\tau}^{\mathbf{u}}) - C(\tau, X_{\tau}^{\mathbf{u}}, \tau, X_{\tau}^{\mathbf{u}})] d\tau \right], \\ L_H^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) &= \mathbb{E}_{t,x}^{\mathbb{Q}} [h(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}, t, x)] - \mathbb{E}_{t,x}^{\mathbb{Q}} [h(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}, t + \epsilon, X_{t+\epsilon}^{\mathbf{u}})], \\ L_G^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) &= G(t, x, \mathbb{E}_{t,x}^{\mathbb{Q}}[g(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}})]) - \mathbb{E}_{t,x}^{\mathbb{Q}} [G(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}, g(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}))], \end{aligned} \tag{A.4}$$

in which for $(t, x, s, y) \in [0, T] \times \mathbb{R}^p \times [0, T] \times \mathbb{R}^p$,

$$h(t, x, s, y) = \mathbb{E}_{t,x}^{\mathbb{Q}^*} \left[\int_t^T C(s, y, \tau, X_{\tau}^{\mathbf{u}^*}) d\tau + F(s, y, X_T^{\mathbf{u}^*}) \right], \quad g(t, x) = \mathbb{E}_{t,x}^{\mathbb{Q}^*} [X_T^{\mathbf{u}^*}].$$

To define a new HJBI equation for V , extra efforts are put into investigating the asymptotics of $L^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon)$ when ϵ is minimal. To this end, we study the adjustment terms individually as follows. We recall that (\mathbf{u}, \mathbf{q}) is admissible and satisfies the conditions (1-3, 4*) in Definition 3.

First, it is easy to see that $L_C^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) = o(\epsilon)$. Second, by Itô's lemma under the measure \mathbb{Q} , we have

$$\begin{aligned} L_H^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) &= \mathbb{E}_{t,x}^{\mathbb{Q}} [h(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}, t, x) - h(t, x, t, x)] - \mathbb{E}_{t,x}^{\mathbb{Q}} [h(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}, t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}) - h(t, x, t, x)] \\ &= [(\mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, s, y))|_{s=t, y=x} - \mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, t, x)] \epsilon + o(\epsilon), \end{aligned}$$

where $\mathcal{A}^{\mathbf{u}, \mathbf{q}}$ is the infinitesimal operator associated with the state process and given as follows. For any multivariate function of t and x , denoted by $\phi(t, x) \in \mathbb{R}^n$ with any n , we define

$$\mathcal{A}^{\mathbf{u}, \mathbf{q}} \phi(t, x) := \frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial x} [\mu(t, x, u) + \sigma(t, x, u)q] + \frac{1}{2} \sum_{i,j=1}^p \Sigma_{ij}(t, x, u) \frac{\partial^2 \phi}{\partial x_i \partial x_j},$$

where $x = (x_1, \dots, x_p)'$, $\Sigma(t, x, u) = \sigma(t, x, u)\sigma(t, x, u)' =: (\Sigma_{ij}(t, x, u))_{i,j=1, \dots, p} \in \mathbb{R}^{p \times p}$, and $\partial \phi / \partial x = (\partial \phi_i / \partial x_j)_{i=1, \dots, n; j=1, \dots, p} \in \mathbb{R}^{n \times p}$. Third, we similarly have

$$\begin{aligned} \mathbb{E}_{t,x}^{\mathbb{Q}} [g(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}})] &= g(t, x) + \mathcal{A}^{\mathbf{u}, \mathbf{q}} g(t, x) \epsilon + o(\epsilon) \mathbf{1}_p, \\ \mathbb{E}_{t,x}^{\mathbb{Q}} [G(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}, g(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}))] &= G(t, x, g(t, x)) + \mathcal{A}^{\mathbf{u}, \mathbf{q}} G(t, x, g(t, x)) \epsilon + o(\epsilon), \end{aligned}$$

where $\mathbf{1}_p = (1, \dots, 1)' \in \mathbb{R}^p$. By using the Taylor expansion on the function $G(t, x, y)$ with respect to y , we have

$$G(t, x, \mathbb{E}_{t,x}^{\mathbb{Q}} [g(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}})]) = G(t, x, g(t, x)) + \frac{\partial G(t, x, y)}{\partial y} \Big|_{y=g(t, x)} \mathcal{A}^{\mathbf{u}, \mathbf{q}} g(t, x) \epsilon + o(\epsilon).$$

Hence,

$$L_G^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) = \left[\frac{\partial G(t, x, y)}{\partial y} \Big|_{y=g(t, x)} \mathcal{A}^{\mathbf{u}, \mathbf{q}} g(t, x) - \mathcal{A}^{\mathbf{u}, \mathbf{q}} G(t, x, g(t, x)) \right] \epsilon + o(\epsilon).$$

Therefore, we have $L^{\mathbf{u}, \mathbf{q}, \epsilon}(t, x, t + \epsilon) = \mathcal{L}(t, x, u, q, h, g) \epsilon + o(\epsilon)$, where

$$\mathcal{L}(t, x, u, q, h, g) = (\mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, s, y))|_{s=t, y=x} - \mathcal{A}^{\mathbf{u}, \mathbf{q}} h(t, x, t, x) + \frac{\partial G(t, x, y)}{\partial y} \Big|_{y=g(t, x)} \mathcal{A}^{\mathbf{u}, \mathbf{q}} g(t, x) - \mathcal{A}^{\mathbf{u}, \mathbf{q}} G(t, x, g(t, x)). \tag{A.5}$$

Recall that the *Hamiltonian* is defined as

$$\mathcal{H}(t, x, u, q, M, N) = M'[\mu(t, x, u) + \sigma(t, x, u)q] + \frac{1}{2}\text{tr}(\sigma(t, x, u)'N\sigma(t, x, u)) + C(t, x, t, x; q) \quad (\text{A.6})$$

for $(t, x, u, q, M, N) \in [0, T] \times \mathbb{R}^p \times \mathbb{R}^k \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{p \times p}$. Then we are ready to define the HJBI equations.

Definition 9 *The extended HJBI equations for V , h , and g are defined as follows.*

(1) *The function $V(t, x)$ is determined by the HJBI equation:*

$$\frac{\partial V}{\partial t} + \sup_u \inf_q \left\{ \mathcal{H} \left(t, x, u, q, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q, h, g) \right\} = 0, \text{ for } t \in [0, T] \quad (\text{A.7})$$

with $V(T, x) = F(T, x, x) + G(T, x, x)$, where \mathcal{L} and \mathcal{H} are given in (A.5) and (A.6), respectively.

(2) *For any $(s, y) \in [0, T] \times \mathbb{R}^p$, the function $h(t, x, s, y)$ solves the PDE:*

$$\mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} h(t, x, s, y) + C(s, y, t, x; \mathbf{q}^*) = 0, \text{ for } t \in [0, T] \quad \text{with } h(T, x, s, y) = F(s, y, x), \quad (\text{A.8})$$

where $(\mathbf{u}^*, \mathbf{q}^*) = (\{u^*(t, x)\}_{t \in [0, T]}, \{q^*(t, x, u)\}_{t \in [0, T]})$ in which $q^*(t, x, u)$ realizes the infimum and $u^*(t, x)$ realizes the supremum in (A.7).

(3) *The function $g(t, x)$ solves the PDE:*

$$\mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} g(t, x) = 0, \text{ for } t \in [0, T] \quad \text{with } g(T, x) = x. \quad (\text{A.9})$$

Finally, we can obtain the following verification theorem.

Theorem 10 *Assume that $V(t, x)$, $h(t, x, s, y)$, and $g(t, x)$ solve the equations (A.7)-(A.9), $V, g \in \mathbb{C}^{1,2}([0, T] \times \mathbb{R}^p)$, $h \in \mathbb{C}^{1,2,1,2}([0, T] \times \mathbb{R}^p \times [0, T] \times \mathbb{R}^p)$ and for $f(t, x) = \{V(t, x), h(t, x, \cdot, \cdot), g(t, x), G(t, x, g(t, x))\}$, we have*

$$\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}[\mu + \sigma q], \text{tr} \left(\sigma' \frac{\partial^2 f}{\partial x \partial x'} \sigma \right) \in \mathbb{L}_{\mathcal{F}^{\mathcal{Q}}}^{1,loc}(0, T; \mathbb{R}), \quad \sigma' \frac{\partial f}{\partial x'} \in \mathbb{L}_{\mathcal{F}^{\mathcal{Q}}}^{2,loc}(0, T; \mathbb{R}^m).$$

Moreover, assume that there is an admissible policy $(\mathbf{u}^*, \mathbf{q}^*)$ that realizes the supremum and infimum in (A.7): for every (t, x) ,

$$\begin{aligned} & \mathcal{H} \left(t, x, u^*, q^*(u^*), \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u^*, q^*(u^*), h, g) \\ &= \sup_u \left\{ \mathcal{H} \left(t, x, u, q^*(u), \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q^*(u), h, g) \right\} \\ &= \sup_u \inf_q \left\{ \mathcal{H} \left(t, x, u, q(u), \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q(u), h, g) \right\}. \end{aligned}$$

Then $(\mathbf{u}^*, \mathbf{q}^*)$ is the robust time-consistent control and V is the robust value function in sense of Definition 3.

PROOF. We first show that $V(t, x) = J(t, x; \mathbf{u}^*, \mathbf{q}^*)$. By the Feynman-Kac formula, we have

$$h(t, x, s, y) = \mathbb{E}_{t,x}^{\mathbb{Q}^*} \left[\int_t^T C(s, y, \tau, X_\tau^{\mathbf{u}^*}; \mathbf{q}^*) d\tau + F(s, y, X_T^{\mathbf{u}^*}) \right], \quad g(t, x) = \mathbb{E}_{t,x}^{\mathbb{Q}^*} [X_T^{\mathbf{u}^*}].$$

With $(\mathbf{u}^*, \mathbf{q}^*)$, the HJBI equation (A.7) becomes a PDE:

$$\mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} V(t, x) + C(t, x, t, x) + \mathcal{L}(t, x, u^*, q^*, h, g) = 0,$$

where $\mathcal{L}(t, x, u^*, q^*, h, g) = -C(t, x, t, x) - \mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} h(t, x, t, x) - \mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} G(t, x, g(t, x))$ by using (A.8) and (A.9). By Itô's lemma, we have

$$\begin{aligned} \mathbb{E}_{t,x}^{\mathbf{Q}^*}[V(T, X_T^{\mathbf{u}^*})] &= V(t, x) + \mathbb{E}_{t,x}^{\mathbf{Q}^*} \left[\int_t^T \mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} V(\tau, X_\tau^{\mathbf{u}^*}) d\tau \right] \\ &= V(t, x) + \mathbb{E}_{t,x}^{\mathbf{Q}^*} \left[\int_t^T [\mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} h(\tau, X_\tau^{\mathbf{u}^*}, \tau, X_\tau^{\mathbf{u}^*}) + \mathcal{A}^{\mathbf{u}^*, \mathbf{q}^*} G(\tau, X_\tau^{\mathbf{u}^*}, g(\tau, X_\tau^{\mathbf{u}^*}))] d\tau \right] \\ &= V(t, x) + \mathbb{E}_{t,x}^{\mathbf{Q}^*} [h(T, X_T^{\mathbf{u}^*}, T, X_T^{\mathbf{u}^*}) - h(t, x, t, x) + G(T, X_T^{\mathbf{u}^*}, X_T^{\mathbf{u}^*}) - G(t, x, \mathbb{E}_{t,x}^{\mathbf{Q}^*}[X_T^{\mathbf{u}^*})] \\ &\Rightarrow V(t, x) = h(t, x, t, x) + G(t, x, \mathbb{E}_{t,x}^{\mathbf{Q}^*}[X_T^{\mathbf{u}^*}]) = J(t, x; \mathbf{u}^*, \mathbf{q}^*). \end{aligned}$$

Now, we prove that the ϵ -policy defined in (9) satisfies the two inequalities in Definition 3. Using the recursive relation of J , (A.2), we have

$$\begin{aligned} J(t, x; \mathbf{u}_\epsilon, \mathbf{q}_\epsilon) &= \mathbb{E}_{t,x}^{\mathbf{Q}} \left[J(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}_\epsilon}; \mathbf{u}_\epsilon, \mathbf{q}_\epsilon) + \int_t^{t+\epsilon} C(\tau, X_\tau^{\mathbf{u}_\epsilon}, \tau, X_\tau^{\mathbf{u}_\epsilon}; \mathbf{q}_\epsilon) d\tau \right] + L^{\mathbf{u}_\epsilon, \mathbf{q}_\epsilon}(t, x, t + \epsilon) \\ &= \mathbb{E}_{t,x}^{\mathbf{Q}} \left[V(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}) + \int_t^{t+\epsilon} C(\tau, X_\tau^{\mathbf{u}}, \tau, X_\tau^{\mathbf{u}}; \mathbf{q}) d\tau \right] + L^{\mathbf{u}_\epsilon, \mathbf{q}_\epsilon}(t, x, t + \epsilon), \end{aligned}$$

where $L^{\mathbf{u}_\epsilon, \mathbf{q}_\epsilon}(t, x, t + \epsilon)$ is given in (A.4). By Itô's lemma, we have

$$\begin{aligned} &\left[\frac{\partial V}{\partial t} + \mathcal{H} \left(t, x, u, q, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q, h, g) \right] \epsilon \\ &= \mathbb{E}_{t,x}^{\mathbf{Q}} \left[V(t + \epsilon, X_{t+\epsilon}^{\mathbf{u}}) - V(t, x) + \int_t^{t+\epsilon} C(\tau, X_\tau^{\mathbf{u}}, \tau, X_\tau^{\mathbf{u}}; \mathbf{q}) d\tau \right] + L^{\mathbf{u}_\epsilon, \mathbf{q}_\epsilon}(t, x, t + \epsilon) + o(\epsilon) \\ &= J(t, x; \mathbf{u}_\epsilon, \mathbf{q}_\epsilon) - V(t, x) + o(\epsilon). \end{aligned}$$

Moreover, by the construction of \mathbf{q}^* , we have

$$\begin{aligned} \mathcal{H} \left(t, x, u, q^*, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q^*, h, g) &\leq \mathcal{H} \left(t, x, u, q, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q, h, g) \\ J(t, x; \mathbf{u}_\epsilon, \mathbf{q}^*) &\leq J(t, x; \mathbf{u}_\epsilon, \mathbf{q}_\epsilon) + o(\epsilon), \end{aligned}$$

which leads to the first inequality in Definition 3. By the construction of \mathbf{u}^* , we have

$$\mathcal{H} \left(t, x, u^*, q^*, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u^*, q^*, h, g) \geq \mathcal{H} \left(t, x, u, q^*, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q^*, h, g).$$

Noting that V solves (A.7), where (u^*, q^*) realizes the supremum and infimum. Hence,

$$\begin{aligned} -\frac{\partial V}{\partial t} &\geq \mathcal{H} \left(t, x, u, q^*, \frac{\partial V}{\partial x'}, \frac{\partial^2 V}{\partial x \partial x'} \right) + \mathcal{L}(t, x, u, q^*, h, g) \\ &\geq J(t, x; \mathbf{u}_\epsilon, \mathbf{q}^*) - V(t, x) + o(\epsilon), \\ J(t, x; \mathbf{u}^*, \mathbf{q}^*) &\geq J(t, x; \mathbf{u}_\epsilon, \mathbf{q}^*) + o(\epsilon), \end{aligned}$$

which leads to the second inequality in Definition 3. Then the result follows.