



# Who Are I: Time Inconsistency and Intrapersonal Conflict and Reconciliation\*

Xue Dong He and Xun Yu Zhou

**Abstract** Time inconsistency is prevalent in dynamic choice problems: a plan of actions to be taken in the future that is optimal for an agent today may not be optimal for the same agent in the future. If the agent is aware of this intra-personal conflict but unable to commit herself in the future to following the optimal plan today, the rational strategy for her today is to reconcile with her future selves, namely to correctly anticipate her actions in the future and then act today accordingly. Such a strategy is named intra-personal equilibrium and has been studied since as early as in the 1950s. A rigorous treatment in continuous-time settings, however, had not been available until a decade ago. Since then, the study on intra-personal equilibrium for time-inconsistent problems in continuous time has grown rapidly. In this chapter, we review the classical results and some recent development in this literature.

## 1 Introduction

When making dynamic decisions, the decision criteria of an agent at different times may not align with each other, leading to time-inconsistent behavior: an action that is optimal under the decision criterion today may no longer be optimal under the decision criterion at certain future time. A variety of preference models can lead to time inconsistent behaviors, such as those involving present-bias, mean-variance criterion, and probability weighting.

---

Xue Dong He  
The Chinese University of Hong Kong,  
e-mail: [xdhe@se.cuhk.edu.hk](mailto:xdhe@se.cuhk.edu.hk)

Xun Yu Zhou  
Columbia University,  
e-mail: [xz2574@columbia.edu](mailto:xz2574@columbia.edu)

\* There is no grammatical error in this title. The word “I” is *plural*; it refers to many different selves over time, which is the premise of this article.

In his seminal paper, Strotz (1955-1956) describes three types of agents when facing time inconsistency. Type 1, a “spendthrift” (or a naïveté as in the more recent literature), does not recognize the time-inconsistency and at any given time seeks an optimal solution from the vantage point of that moment only. As a result, his strategies are always myopic and change all the times. The next two types are aware of the time inconsistency but act differently. Type 2 is a “precommitter” who solves the optimization problem only once at time 0 and then commits to the resulting strategy throughout, even though she knows that the original solution may no longer be optimal at later times. Type 3 is a “thrift” (or a sophisticated agent) who is unable to precommit and realizes that her future selves may disobey whatever plans she makes now. Her resolution is to compromise and choose *consistent planning* in the sense that she optimizes taking the future disobedience as a *constraint*. In this resolution, the agent’s selves at different times are considered to be the players of a game, and a consistent plan chosen by the agent becomes an equilibrium of the game from which no selves are willing to deviate. Such a plan or strategy is referred to as an *intra-personal equilibrium*.

To illustrate the above three types of behavior under time inconsistency, consider an agent who has a planning horizon with a finite end date  $T$  and makes decisions at *discrete* times  $t \in \{0, 1, \dots, T - 1\}$ . The agent’s decision drives a Markov state process and the agent’s decision criterion at time  $t$  is to maximize an objective function  $J(t, x; \mathbf{u})$ , where  $x$  stands for the Markovian state at that time and  $\mathbf{u}$  represents the agent’s strategy. The agent considers Markovian strategies, so  $\mathbf{u}$  is a function of time  $s \in \{0, 1, \dots, T - 1\}$  and the Markovian state at that time. If the agent, at certain time  $t$  with state  $x$ , is a “pre-committer”, she is committed to implementing throughout the remaining horizon the strategy  $\mathbf{u}_{(t,x)}^{pc} = \{\mathbf{u}_{(t,x)}^{pc}(s, \cdot) | s = t, t + 1, \dots, T - 1\}$  that maximizes  $J(t, x; \mathbf{u})$ , and this strategy is referred to as the *pre-committed strategy* of the agent at time  $t$  with state  $x$ . If the agent is a “spendthrift”, at *every* time  $t$  with state  $x$ , she is able to implement the pre-committed strategy at that moment only and will change at the next moment; so the strategy that is actually implemented by the agent throughout the horizon is  $\mathbf{u}^n = \{\mathbf{u}_{(s, X^{\mathbf{u}^n}(s))}^{pc}(s, X^{\mathbf{u}^n}(s)) | s = 0, 1, \dots, T - 1\}$ , where  $X^{\mathbf{u}^n}$  denotes the state process under  $\mathbf{u}^n$ . This strategy is referred to as the *naïve strategy*. If the agent is a “thrift”, she chooses an intra-personal equilibrium strategy  $\hat{\mathbf{u}}$ : At any time  $t \in \{0, 1, \dots, T - 1\}$  with any state  $x$  at that time,  $\hat{\mathbf{u}}(t, x)$  is the optimal action of the agent given that her future selves follow  $\hat{\mathbf{u}}$ ; i.e.,

$$\hat{\mathbf{u}}(t, x) \in \arg \max_u J(t, x; \mathbf{u}_{t,u}), \tag{1}$$

where  $\mathbf{u}_{t,u}(t, x) := u$  and  $\mathbf{u}_{t,u}(s, \cdot) := \hat{\mathbf{u}}(s, \cdot)$  for  $s = t + 1, \dots, T - 1$ .

All the three types of behavior are important from an economic perspective. First, field and experimental studies reveal the popularity of commitment devices to help individuals to fulfill plans that would otherwise be difficult to implement due to lack of self control; see for instance Bryan et al. (2010). The demand for commitment devices implies that some individuals seek for pre-committed strategies in the presence of time inconsistency. Second, empirically observed decision-making behavior

implies that some individuals are naïvetés. For example, Barberis (2012) shows that a naïve agent would take on a series independent, unfavorable bets and take a gain-exit strategy, and this gambling behavior is commonly observed in casinos. Finally, when an agent foresees the time-inconsistency and a commitment device is either unhelpful or unavailable, the intra-personal equilibrium strategy becomes a rational choice of the agent.

It is important to note that it is hard or perhaps not meaningful to determine which type is superior than the others, simply because there is no uniform criteria to evaluate and compare them. So a naïve strategy, despite its name, is not necessarily inferior to an intra-personal equilibrium in terms of an agent's long-run utility. Indeed, O'Donoghue and Rabin (1999) show that in an optimal stopping problem with an immediate reward and present-biased preferences, a sophisticate agent has a larger tendency to preproperate than a naïveté and thus leads to a lower long-run utility. In this sense, studying the different behaviors under time inconsistency sometimes falls into the realm of being "descriptive" as in behavioral science, rather than being "normative" as in classical decision-making theory.

In this survey article, we focus on reviewing the studies on intra-personal equilibrium of a sophisticated agent in *continuous time*. Intra-personal equilibrium for time-inconsistent problems in discrete time, which is defined through the equilibrium condition (1), has been extensively studied in the literature and generated various economic implications. The extension to the continuous-time setting, however, is nontrivial because in this setting, taking a different action from a given strategy at only one time instant does not change the state process and thus has no impact on the objective function value. As a result, it becomes meaningless to examine whether the agent is willing to deviate from a given strategy at a particular moment by just comparing the objective function values before and after the deviation. To address this issue and to formalize the idea of Strotz (1955-1956), Ekeland and Pirvu (2008), Ekeland and Lazrak (2006), and Björk and Murgoci (2010) assume that the agent's self at each time can implement her strategy in an infinitesimally small, but positive, time period; consequently, her action has an impact on the state process and thus on the objective function. In Section 2 below, we follow the framework of Björk and Murgoci (2014) to define intra-personal equilibria, show a sufficient and necessary condition for an equilibrium, and present the so-called *extended HJB equation* that characterizes the intra-personal equilibrium strategy and the value under this strategy. In Section 3, we further discuss various issues related to intra-personal equilibria.

A close-loop strategy for a control system is a mapping from the historical path of the system state and control to the space of controls. So at each time, the control taken by the agent is obtained by plugging the historical path into this mapping. For example, a Markovian strategy for a Markovian control system is a closed-loop strategy. An open-loop strategy is a collection of controls across time (and across scenarios in case of stochastic control), and at each time the control in this collection is taken, regardless of the historical path of the system state and control. For a classical, time-consistent controlled Markov decision problem, the optimal close-loop strategy and the optimal open-loop strategy yield the same state-control

path. For time-inconsistent problems, however, closed-loop and open-loop intra-personal equilibria can be vastly different. In Section 4, we review the study of open-loop intra-personal equilibrium and discuss its connection with closed-loop intra-personal equilibrium.

Optimal stopping problems can be viewed as a special case of control problems, so intra-personal equilibria can be defined similarly for time-inconsistent stopping problems. These problems, however, have very special structures, and by exploiting these structures new notions of intra-personal equilibria have been proposed in the literature. We discuss these in Section 5.

If we discretize a continuous horizon of time and assume that the agent has full self control in each subperiod under the discretization, we can define and derive intra-personal equilibria as in the discrete-time setting. The limits of the intra-personal equilibria as discretization becomes infinitely finer are used by some authors to define intra-personal equilibria for continuous-time problems. In Section 6, we review this thread of research.

Time-inconsistency arises in various economic problems, and for many of them, intra-personal equilibria have been studied and their implications discussed in the literature. In Section 7, we review this literature.

Finally, in Section 8, we review the studies on dynamic consistency preferences. In these studies, starting from a preference model for an agent at certain initial time, the authors attempt to find certain preference models for the agent's future selves such that the pre-committed strategy for the agent at the initial time is also optimal for the agent at any future time and thus can be implemented consistently over time.

## 2 Extended HJB Equation

Strotz (1955-1956) is the first to study the behavior of a sophisticated agent in the presence of time-inconsistency in a continuous-time model. Without formally defining the notion of intra-personal equilibrium, the author derives a consistent plan of the sophisticated agent. Barro (1999) and Luttmer and Mariotti (2003) also investigate, for certain continuous-time models, consistent plans of sophisticated agents, again without their formal definitions. In a series of papers, Ekeland and Lazrak (2006), Ekeland and Lazrak (2008), and Ekeland and Lazrak (2010) study the classical Ramsey model with a nonexponential discount function and propose for the first time a formal notion of intra-personal equilibrium for deterministic control problems in continuous time. Such a notion is proposed in a stochastic context by Björk and Murgoci (2010), which is later split into two papers, Björk and Murgoci (2014) and Björk et al. (2017), discussing the discrete-time and continuous-time settings, respectively. In this section, we follow the framework of Björk et al. (2017) to define an intra-personal equilibrium strategy and present a sufficient and necessary condition for such a strategy.

## 2.1 Notations

We first introduce some notations. By convention,  $x \in \mathbb{R}^n$  is always a column vector. When a vector  $x$  is a row vector, we write it as  $x \in \mathbb{R}^{1 \times n}$ . Denote by  $A^\top$  the transpose of a matrix  $A$ , and by  $\text{tr}(A)$  the trace of a square matrix  $A$ . For a differentiable function  $\xi$  that maps  $x \in \mathbb{R}^m$  to  $\xi(x) \in \mathbb{R}^n$ , its derivative, denoted as  $\xi_x(x)$ , is an  $n \times m$  matrix with the entry in the  $i$ -th row and  $j$ -th column denoting the derivative of the  $i$ -th component of  $\xi$  with respect to the  $j$ -th component of  $x$ . In particular, for a mapping  $\xi$  from  $\mathbb{R}^m$  to  $\mathbb{R}$ ,  $\xi_x(x)$  is an  $m$ -dimensional row vector, and we further denote by  $\xi_{xx}$  the Hessian matrix.

Consider  $\xi$  that maps  $(z, x) \in \mathbb{Z} \times \mathbb{X}$  to  $\xi(z, x) \in \mathbb{R}^l$ , where  $\mathbb{Z}$  is a certain set and  $\mathbb{X}$ , which represents the state space throughout, is either  $\mathbb{R}^n$  or  $(0, +\infty)$ .  $\xi$  is *locally Lipschitz in  $x \in \mathbb{X}$ , uniformly in  $z \in \mathbb{Z}$*  if there exists a sequence of compact sets  $\{\mathbb{X}_k\}_{k \geq 1}$  with  $\cup_{k \geq 1} \mathbb{X}_k = \mathbb{X}$  and a sequence of positive numbers  $\{L_k\}_{k \geq 1}$  such that for any  $k \geq 1$ ,  $\|\xi(z, x) - \xi(z, x')\| \leq L_k \|x - x'\|, \forall z \in \mathbb{Z}, x, x' \in \mathbb{X}_k$ .  $\xi$  is *global Lipschitz in  $x \in \mathbb{X}$ , uniformly in  $z \in \mathbb{Z}$*  if there exists constant  $L > 0$  such that  $\|\xi(z, x) - \xi(z, x')\| \leq L \|x - x'\|, \forall z \in \mathbb{Z}, x, x' \in \mathbb{X}$ . In the case  $\mathbb{X} = \mathbb{R}^n$ ,  $\xi$  is of *linear growth in  $x \in \mathbb{X}$ , uniformly in  $z \in \mathbb{Z}$*  if there exists  $L > 0$  such that  $\|\xi(z, x)\| \leq L(1 + \|x\|), \forall z \in \mathbb{Z}, x \in \mathbb{X}$ . In the case  $\mathbb{X} = (0, +\infty)$ ,  $\xi$  has a *bounded norm in  $x \in \mathbb{X}$ , uniformly in  $z \in \mathbb{Z}$* , if there exists  $L > 0$  such that  $\|\xi(z, x)\| \leq Lx, \forall z \in \mathbb{Z}, x \in \mathbb{X}$ .  $\xi$  is of *polynomial growth in  $x \in \mathbb{X}$ , uniformly in  $z \in \mathbb{Z}$*  if there exists  $L > 0$  and integer  $\gamma \geq 1$  such that  $\|\xi(z, x)\| \leq L(1 + \varphi_{2\gamma}(x)), \forall z \in \mathbb{Z}, x \in \mathbb{X}$ , where  $\varphi_{2\gamma}(x) = \|x\|^{2\gamma}$  when  $\mathbb{X} = \mathbb{R}^n$  and  $\varphi_{2\gamma}(x) = x^{2\gamma} + x^{-2\gamma}$  when  $\mathbb{X} = (0, +\infty)$ .

Fix integers  $r \geq 0, q \geq 2r$ , and real numbers  $a < b$ . Consider  $\xi$  that maps  $(t, x) \in [a, b] \times \mathbb{X}$  to  $\xi(t, x) \in \mathbb{R}^l$ . We say  $\xi \in \mathcal{C}^{r,q}([a, b] \times \mathbb{X})$  if for any derivative index  $\alpha$  with  $|\alpha| \leq q - 2j$  and  $j = 0, \dots, r$ , the partial derivative  $\frac{\partial^{j+\alpha} \xi(t, x)}{\partial t^j \partial x^\alpha} := \frac{\partial^{j+\alpha_1+\dots+\alpha_n} \xi(t, x)}{\partial t^j \partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$  exists for any  $(t, x) \in (a, b) \times \mathbb{X}$  and can be extended to and continuous on  $[a, b] \times \mathbb{X}$ . We say  $\xi \in \tilde{\mathcal{C}}^{r,q}([a, b] \times \mathbb{X})$  if  $\xi \in \mathcal{C}^{r,q}([a, b] \times \mathbb{X})$  and  $\frac{\partial^{j+\alpha} \xi(t, x)}{\partial t^j \partial x^\alpha}$  is of polynomial growth in  $x \in \mathbb{X}$ , uniformly in  $t \in [a, b]$ , for any derivative index  $\alpha$  with  $|\alpha| \leq q - 2j$  and  $j = 0, \dots, r$ .

## 2.2 Time-Inconsistent Stochastic Control Problems

Let be given a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with a standard  $d$ -dimensional Brownian motion  $W(t) := (W_1(t), \dots, W_d(t))^\top, t \geq 0$ , on the space, along with the filtration  $(\mathcal{F}_t)_{t \geq 0}$  generated by the Brownian motion and augmented by the  $\mathbb{P}$ -null sets. Consider an agent who makes dynamic decisions in a given period  $[0, T]$ , and for any  $(t, x) \in [0, T] \times \mathbb{X}$ , the agent faces the following stochastic control problem:

$$\left\{ \begin{array}{l} \max_{\mathbf{u}} J(t, x; \mathbf{u}) \\ \text{subject to } dX^{\mathbf{u}}(s) = \mu(s, X^{\mathbf{u}}(s), \mathbf{u}(s, X^{\mathbf{u}}(s)))ds \\ \quad + \sigma(s, X^{\mathbf{u}}(s), \mathbf{u}(s, X^{\mathbf{u}}(s)))dW(s), s \in [t, T] \\ X^{\mathbf{u}}(t) = x. \end{array} \right. \quad (2)$$

The agent's dynamic decisions are represented by a Markov strategy  $\mathbf{u}$ , which maps  $(s, y) \in [0, T] \times \mathbb{X}$  to  $\mathbf{u}(s, y) \in \mathbb{U} \subset \mathbb{R}^m$ . The controlled diffusion process  $X^{\mathbf{u}}$  under  $\mathbf{u}$  takes values in  $\mathbb{X}$ , which as aforementioned is set to be either  $(0, +\infty)$  or  $\mathbb{R}^n$ .  $\mu$  and  $\sigma$  are measurable mappings from  $[0, T] \times \mathbb{X} \times \mathbb{U}$  to  $\mathbb{R}^n$  and to  $\mathbb{R}^{n \times d}$ , respectively, where  $n$  stands for the dimension of  $\mathbb{X}$ .

The agent's goal at  $(t, x) \in [0, T] \times \mathbb{X}$  is to maximize the following objective function:

$$\begin{aligned} J(t, x; \mathbf{u}) = & \mathbb{E}_{t,x} \left[ \int_t^T C(t, x, s, X^{\mathbf{u}}(s), \mathbf{u}(s, X^{\mathbf{u}}(s))) ds + F(t, x, X^{\mathbf{u}}(T)) \right] \\ & + G(t, x, \mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]), \end{aligned} \quad (3)$$

where  $C$  is a measurable mapping from  $[0, T] \times \mathbb{X} \times [0, T] \times \mathbb{X} \times \mathbb{U}$  to  $\mathbb{R}$ , and  $F$  and  $G$  are measurable mappings from  $[0, T] \times \mathbb{X} \times \mathbb{X}$  to  $\mathbb{R}$ . Here and hereafter,  $\mathbb{E}_{t,x}[Z]$  denotes the expectation of  $Z$  conditional on  $X^{\mathbf{u}}(t) = x$ . If  $C$ ,  $F$ , and  $G$  are independent of  $(t, x)$  and  $G(t, x, \mathbb{E}_{t,x}[X^{\mathbf{u}}(T)])$  is linear in  $\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]$ , then  $J(t, x; \mathbf{u})$  becomes a standard objective function in classical stochastic control where time consistency holds. Thus, with objective function (3), time inconsistency arises from the dependence of  $C$ ,  $F$ , and  $G$  on  $(t, x)$  as well as from the nonlinearity of  $G(t, x, \mathbb{E}_{t,x}[X^{\mathbf{u}}(T)])$  in  $\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]$ .

For any feedback strategy  $\mathbf{u}$ , denote

$$\begin{aligned} \mu^{\mathbf{u}}(t, x) &:= \mu(t, x, \mathbf{u}(t, x)), \quad \sigma^{\mathbf{u}}(t, x) := \sigma(t, x, \mathbf{u}(t, x)), \\ \Upsilon^{\mathbf{u}}(t, x) &:= \sigma(t, x, \mathbf{u}(t, x))\sigma(t, x, \mathbf{u}(t, x))^{\top}, \quad C^{\tau, y, \mathbf{u}}(t, x) := C(\tau, y, t, x, \mathbf{u}(t, x)). \end{aligned}$$

With a slight abuse of notation,  $u \in \mathbb{U}$  also denotes the feedback strategy  $\mathbf{u}$  such that  $\mathbf{u}(t, x) = u, \forall (t, x) \in [0, T] \times \mathbb{X}$ ; so  $\mathbb{U}$  also stands for the set of all *constant* strategies when no ambiguity arises.

We need to impose conditions on a strategy  $\mathbf{u}$  to ensure the existence and uniqueness of the SDE in (2) and the well-posedness of the objective function  $J(t, x; \mathbf{u})$ . This consideration leads to the following definition of feasibility:

**Definition 1** A feedback strategy  $\mathbf{u}$  is *feasible* if the following hold:

- (i)  $\mu^{\mathbf{u}}$ ,  $\sigma^{\mathbf{u}}$  are locally Lipschitz in  $x \in \mathbb{X}$ , uniformly in  $t \in [0, T]$ .
- (ii)  $\mu^{\mathbf{u}}$  and  $\sigma^{\mathbf{u}}$  are of linear growth in  $x \in \mathbb{X}$ , uniformly in  $t \in [0, T]$ , when  $\mathbb{X} = \mathbb{R}^n$  and have bounded norm in  $x \in \mathbb{X}$ , uniformly in  $t \in [0, T]$ , when  $\mathbb{X} = (0, +\infty)$ .
- (iii) For each fixed  $(\tau, y) \in [0, T] \times \mathbb{X}$ ,  $C^{\tau, y, \mathbf{u}}(t, x)$  and  $F(\tau, y, x)$  are of polynomial growth in  $x \in \mathbb{X}$ , uniformly in  $t \in [0, T]$ .
- (iv) For each fixed  $(\tau, y) \in [0, T] \times \mathbb{X}$  and  $x \in \mathbb{X}$ ,  $\mu^{\mathbf{u}}(t, x)$  and  $\sigma^{\mathbf{u}}(t, x)$  are right-continuous in  $t \in [0, T]$  and  $\lim_{t' \geq t, (t', x') \rightarrow (t, x)} C^{\tau, y, \mathbf{u}}(t', x') = C^{\tau, y, \mathbf{u}}(t, x)$  for any  $t \in [0, T)$ .

Denote the set of feasible strategies as  $\mathbf{U}$ .

We impose the following assumption:

**Assumption 1** Any  $u \in \mathbf{U}$  is feasible.

### 2.3 Intra-Personal Equilibrium

Here and hereafter,  $\hat{\mathbf{u}} \in \mathbf{U}$  denotes a given strategy and we examine whether it is an equilibrium strategy. For given  $t \in [0, T)$ ,  $\varepsilon \in (0, T - t)$  and  $\mathbf{a} \in \mathbf{U}$ , define

$$\mathbf{u}_{t,\varepsilon,\mathbf{a}}(s,y) := \begin{cases} \mathbf{a}(s,y), & s \in [t, t + \varepsilon), y \in \mathbb{X} \\ \hat{\mathbf{u}}(s,y), & s \notin [t, t + \varepsilon), y \in \mathbb{X}. \end{cases} \quad (4)$$

Imagine that the agent at time  $t$  chooses strategy  $\mathbf{a}$  and is able to commit herself to this strategy in the period  $[t, t + \varepsilon)$ . The agent, however, is unable to control her future selves beyond this small time period, namely in the period  $[t + \varepsilon, T)$  and believes that her future selves will take strategy  $\hat{\mathbf{u}}$ . Then,  $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$  is the strategy that the agent at time  $t$  expects herself to implement throughout the entire horizon. Note that  $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$  is feasible because both  $\hat{\mathbf{u}}$  and  $\mathbf{a}$  are feasible.

#### Definition 2 (Intra-Personal Equilibrium)

$\hat{\mathbf{u}} \in \mathbf{U}$  is an *intra-personal equilibrium* if for any  $x \in \mathbb{X}$ ,  $t \in [0, T)$ , and  $\mathbf{a} \in \mathbf{U}$ , we have

$$\limsup_{\varepsilon \downarrow 0} \frac{J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}}) - J(t,x;\hat{\mathbf{u}})}{\varepsilon} \leq 0. \quad (5)$$

For each positive  $\varepsilon$ ,  $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$  leads to a possibly different state process and thus to a different objective function value from those of  $\hat{\mathbf{u}}$ , so it is meaningful to compare the objective function values of  $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$  and  $\hat{\mathbf{u}}$  to examine whether the agent is willing to deviate from  $\hat{\mathbf{u}}$  to  $\mathbf{a}$  in the period of time  $[t, t + \varepsilon)$ . Due to the continuous-time nature, the length of the period,  $\varepsilon$ , during which the agent at  $t$  exerts full self control, must be set to be infinitesimally small. Then,  $J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}})$  and  $J(t,x;\hat{\mathbf{u}})$  become arbitrarily close to each other; so instead of evaluating their difference, we consider the *rate of increment* in the objective function value, i.e., the limit on the left-hand side of (5). Thus, under Definition 2, a strategy  $\hat{\mathbf{u}}$  is an intra-personal equilibrium if at any given time and state, the rate of increment in the objective value when the agent deviates from  $\hat{\mathbf{u}}$  to any alternative strategy is nonpositive. As a result, the agent has little incentive to deviate from  $\hat{\mathbf{u}}$ .

## 2.4 Sufficient and Necessary Condition

We first introduce the generator of the controlled state process. Given  $\mathbf{u} \in \mathbf{U}$  and interval  $[a, b] \subseteq [0, T]$ , consider  $\xi$  that maps  $(t, x) \in [a, b] \times \mathbb{X}$  to  $\xi(t, x) \in \mathbb{R}$ . Suppose  $\xi \in \mathcal{C}^{1,2}([a, b] \times \mathbb{X})$ , and denote by  $\xi_t$ ,  $\xi_x$ , and  $\xi_{xx}$  respectively its first-order partial derivative in  $t$ , first-order partial derivative in  $x$ , and second-order partial derivative in  $x$ . Define the following generator:

$$\mathcal{A}^{\mathbf{u}}\xi(t, x) = \xi_t(t, x) + \xi_x(t, x)\mu^{\mathbf{u}}(t, x) + \frac{1}{2}\text{tr}\left(\xi_{xx}(t, x)^{\top}\Upsilon^{\mathbf{u}}(t, x)\right), \quad t \in [a, b], x \in \mathbb{X}. \quad (6)$$

For each fixed  $(\tau, y) \in [0, T] \times \mathbb{X}$ , denote

$$f^{\tau, y}(t, x) := \mathbb{E}_{t, x}[F(\tau, y, X^{\hat{\mathbf{u}}}(T))], \quad (7)$$

$$g(t, x) := \mathbb{E}_{t, x}[X^{\hat{\mathbf{u}}}(T)], \quad t \in [0, T], x \in \mathbb{X}. \quad (8)$$

In addition, for fixed  $(\tau, y) \in [0, T] \times \mathbb{X}$  and  $s \in [0, T]$ , denote

$$c^{\tau, y, s}(t, x) := \mathbb{E}_{t, x}[C^{\tau, y, \hat{\mathbf{u}}}(s, X^{\hat{\mathbf{u}}}(s))], \quad t \in [0, s], x \in \mathbb{X}. \quad (9)$$

In the following,  $\mathcal{A}^{\mathbf{u}}f^{\tau, y}$  denotes the function that is obtained by applying the operator  $\mathcal{A}^{\mathbf{u}}$  to  $f^{\tau, y}(t, x)$  as a function of  $(t, x)$  while fixing  $(\tau, y)$ . Then,  $\mathcal{A}^{\mathbf{u}}f^{t, x}(t, x)$  denotes the value of  $\mathcal{A}^{\mathbf{u}}f^{\tau, y}$  at  $(t, x)$  while  $(\tau, y)$  is also set at  $(t, x)$ . The above notations also apply to  $C^{\tau, y, \mathbf{u}}$  and  $c^{\tau, y, s}$ .

To illustrate how to evaluate  $J(t, x; \mathbf{u}_{t, \varepsilon, \mathbf{a}}) - J(t, x; \hat{\mathbf{u}})$  and thus the rate of increment, let us consider the second term in the objective function (3). An informal calculation yields

$$\begin{aligned} & \mathbb{E}_{t, x}[F(t, x, X^{\mathbf{u}_{t, \varepsilon, \mathbf{a}}}(T))] - \mathbb{E}_{t, x}[F(t, x, X^{\hat{\mathbf{u}}}(T))] \\ &= \mathbb{E}_{t, x}\left[\mathbb{E}_{t+\varepsilon, X^{\mathbf{u}_{t, \varepsilon, \mathbf{a}}}(t+\varepsilon)}[F(t, x, X^{\mathbf{u}_{t, \varepsilon, \mathbf{a}}}(T))]\right] - \mathbb{E}_{t, x}[F(t, x, X^{\hat{\mathbf{u}}}(T))] \\ &= \mathbb{E}_{t, x}\left[f^{(t, x)}(t+\varepsilon, X^{\mathbf{a}}(t+\varepsilon))\right] - f^{(t, x)}(t, x) \\ &\approx \mathcal{A}^{\mathbf{a}}f^{t, x}(t, x)\varepsilon, \end{aligned}$$

where the second equality holds because  $\mathbf{u}_{t, \varepsilon, \mathbf{a}}(s, \cdot) = \mathbf{a}(s, \cdot)$  for  $s \in [t, t+\varepsilon)$  and  $\mathbf{u}_{t, \varepsilon, \mathbf{a}}(s, \cdot) = \hat{\mathbf{u}}(s, \cdot)$  for  $s \in [t+\varepsilon, T]$  in addition to the definition of  $f^{t, y}$  in (7). The change of the other terms in the objective function when the agent deviates from  $\hat{\mathbf{u}}$  to  $\mathbf{a}$  in the period  $[t, t+\varepsilon)$  can be evaluated similarly. As a result, we can derive the rate of increment in the objective value, namely the limit on the left-hand side of (5), which in turn enables us to derive a sufficient and necessary condition for  $\hat{\mathbf{u}}$  to be an intra-personal equilibrium.

To formalize the above heuristic argument, we need to impose the following assumption:



**Assumption 2** For any fixed  $(\tau, y) \in [0, T] \times \mathbb{X}$  and  $t \in [0, T]$ , there exists  $\tilde{t} \in (t, T]$  such that (i)  $f^{\tau, y}, g \in \tilde{\mathcal{C}}^{1,2}([t, \tilde{t}] \times \mathbb{X})$ ; (ii)  $c^{\tau, y, s} \in \mathcal{C}^{1,2}([t, \tilde{t} \wedge s] \times \mathbb{X})$  for each fixed  $s \in (t, T]$  and  $\frac{\partial^{j+\alpha} c^{\tau, y, s}(t', x')}{\partial t^j \partial x'^\alpha}$  is of polynomial growth in  $x' \in \mathbb{X}$ , uniformly in  $t' \in [t, \tilde{t} \wedge s]$  and  $s \in (t, T]$ , for any  $\alpha$  with  $|\alpha| \leq 2 - 2j$  and  $j = 0, 1$ ; and (iii)  $G(\tau, y, z)$  is continuously differentiable in  $z$ , with the partial derivative denoted as  $G_z(\tau, y, z)$ .

**Theorem 1** Suppose Assumptions 1 and 2 hold. Then, for any  $(t, x) \in [0, T] \times \mathbb{X}$  and  $\mathbf{a} \in \mathbf{U}$ , we have

$$\lim_{\varepsilon \downarrow 0} \frac{J(t, x; \mathbf{u}_{t, \varepsilon, \mathbf{a}}) - J(t, x; \hat{\mathbf{u}})}{\varepsilon} = \Gamma^{t, x, \hat{\mathbf{u}}}(t, x; \mathbf{a}), \quad (10)$$

where for any  $(\tau, y) \in [0, T] \times \mathbb{X}$ ,

$$\begin{aligned} \Gamma^{\tau, y, \hat{\mathbf{u}}}(t, x; \mathbf{a}) &:= C^{\tau, y, \mathbf{a}}(t, x) - C^{\tau, y, \hat{\mathbf{u}}}(t, x) + \int_t^T \mathcal{A}^{\mathbf{a}} c^{\tau, y, s}(t, x) ds \\ &\quad + \mathcal{A}^{\mathbf{a}} f^{\tau, y}(t, x) + G_z(\tau, y, g(t, x)) \mathcal{A}^{\mathbf{a}} g(t, x). \end{aligned} \quad (11)$$

Moreover,  $\Gamma^{\tau, y, \hat{\mathbf{u}}}(t, x; \mathbf{a}) = \Gamma^{\tau, y, \hat{\mathbf{u}}}(t, x; \tilde{\mathbf{a}})$  for any  $\mathbf{a}, \tilde{\mathbf{a}} \in \mathbf{U}$  with  $\mathbf{a}(t, x) = \tilde{\mathbf{a}}(t, x)$  and  $\Gamma^{\tau, y, \hat{\mathbf{u}}}(t, x; \mathbf{a}) = 0$  if  $\mathbf{a}(t, x) = \hat{\mathbf{u}}(t, x)$ . Consequently,  $\hat{\mathbf{u}}$  is an intra-personal equilibrium if and only if

$$\Gamma^{t, x, \hat{\mathbf{u}}}(t, x; u) \leq 0, \quad \forall u \in \mathbf{U}, x \in \mathbb{X}, t \in [0, T]. \quad (12)$$

Theorem 1 presents a sufficient and necessary condition (12) for an intra-personal equilibrium  $\hat{\mathbf{u}}$ . Because  $\Gamma^{\tau, y, \hat{\mathbf{u}}}(\tau, y; \hat{\mathbf{u}}(t, x)) = 0$ , we have

$$\Gamma^{\tau, y, \hat{\mathbf{u}}}(t, x; \mathbf{a}) = \Pi^{\tau, y}(t, x; \mathbf{a}) - \Pi^{\tau, y}(t, x; \hat{\mathbf{u}}),$$

where

$$\begin{aligned} \Pi^{\tau, y}(t, x; \mathbf{a}) &:= C^{\tau, y, \mathbf{a}}(t, x) + \int_t^T \mathcal{A}^{\mathbf{a}} c^{\tau, y, s}(t, x) ds + \mathcal{A}^{\mathbf{a}} f^{\tau, y}(t, x) \\ &\quad + G_z(\tau, y, g(t, x)) \mathcal{A}^{\mathbf{a}} g(t, x). \end{aligned} \quad (13)$$

As a result, condition (12) is equivalent to

$$\max_{u \in \mathbf{U}} \Gamma^{t, x, \hat{\mathbf{u}}}(t, x; u) = 0, x \in \mathbb{X}, t \in [0, T] \quad (14)$$

or

$$\hat{\mathbf{u}}(t, x) \in \arg \max_{u \in \mathbf{U}} \Pi^{t, x}(t, x; u), x \in \mathbb{X}, t \in [0, T]. \quad (15)$$

This can be regarded as a time-inconsistent version of the verification theorem in (classical) stochastic control.

The proof of Theorem 1 can be found in Björk et al. (2017) and He and Jiang (2019). Assumption 1 is easy to verify because it involves only the model parame-

ters, i.e.,  $\mu$ ,  $\sigma$ ,  $C$ ,  $F$ , and  $G$ . Assumption 2 imposes some regularity conditions on  $\hat{\mathbf{u}}$ , which usually requires  $\hat{\mathbf{u}}$  to be smooth to a certain degree; see He and Jiang (2019) for a sufficient condition for this assumption. As a result, the sufficient and necessary condition (12) cannot tell us whether there exists any intra-personal equilibrium among the strategies that do not satisfy Assumption 2. This condition, however, is still very useful for us to find intra-personal equilibria for specific problems. Indeed, in most time-inconsistent problems in the literature, intra-personal equilibrium can be found and verified using (12); see Section 7.

## 2.5 Extended HJB

Define the *continuation value* of a strategy  $\hat{\mathbf{u}}$ , denoted as  $V^{\hat{\mathbf{u}}}(t, x)$ ,  $(t, x) \in [0, T] \times \mathbb{X}$ , to be the objective value over time and state under this strategy, i.e.,

$$V^{\hat{\mathbf{u}}}(t, x) := J(t, x; \hat{\mathbf{u}}) = H^{t,x}(t, x) + G(t, x, g(t, x)), \quad (16)$$

where

$$\begin{aligned} H^{\tau,y}(t, x) &:= \mathbb{E}_{t,x} \left[ \int_t^T C^{\tau,y,\hat{\mathbf{u}}}(s, X^{\hat{\mathbf{u}}}(s)) ds + F(\tau, y, X^{\hat{\mathbf{u}}}(T)) \right] \\ &= \int_t^T c^{\tau,y,s}(t, x) ds + f^{\tau,y}(t, x). \end{aligned} \quad (17)$$

Assuming certain regularity conditions and applying the operator  $\mathcal{A}^u$  to  $V^{\hat{\mathbf{u}}}(t, x)$ , we derive

$$\begin{aligned} \mathcal{A}^u V^{\hat{\mathbf{u}}}(t, x) &= -C^{t,x,\hat{\mathbf{u}}}(t, x) + \int_t^T \mathcal{A}^u c^{t,x,s}(t, x) ds + \mathcal{A}^u f^{t,x}(t, x) \\ &\quad + G_z(t, x, g(t, x)) \mathcal{A}^u g(t, x) + \mathcal{A}_{\tau,y}^u H^{t,x}(t, x) + \mathcal{A}_{\tau,y}^u G(t, x, g(t, x)) \\ &\quad + \text{tr} \left( \left( H_{xy}^{t,x}(t, x) + G_{zy}(t, x, g(t, x))^\top g_x(t, x) \right)^\top \Upsilon^u(t, x) \right) \\ &\quad + \frac{1}{2} G_{zz}(t, x, g(t, x)) \text{tr} \left( g_x(t, x) g_x(t, x)^\top \Upsilon^u(t, x) \right) \end{aligned}$$

where  $H_{xy}^{\tau,y}(t, x)$  denotes the cross partial derivative of  $H^{\tau,y}(t, x)$  in  $x$  and  $y$ ,  $G_{zy}(\tau, y, z)$  the cross partial derivative of  $G(\tau, y, z)$  in  $z$  and  $y$ , and  $G_{zz}(\tau, y, z)$  the second-order derivative of  $G(\tau, y, z)$  in  $z$ . For each *fixed*  $(t, x)$ ,  $\mathcal{A}_{\tau,y}^u H^{\tau,y}(t, x)$  denotes the generator of  $\mathcal{A}^u$  applied to  $H^{\tau,y}(t, x)$  as a function of  $(\tau, y)$ , i.e.,  $\mathcal{A}_{\tau,y}^u H^{\tau,y}(t, x) := \mathcal{A}^u \ell(\tau, y)$ , where  $\ell(\tau, y) := H^{\tau,y}(t, x)$ ,  $(\tau, y) \in [0, T] \times \mathbb{X}$ , and  $\mathcal{A}_{\tau,y}^u G(\tau, y, g(t, x))$  is defined similarly.

Now, suppose  $\hat{\mathbf{u}}$  is an intra-personal equilibrium. Recalling (11) and the sufficient and necessary condition (14), we derive the following equation satisfied by the continuation value of an intra-personal equilibrium  $\hat{\mathbf{u}}$ :

$$\begin{aligned}
& \max_{u \in \mathbb{U}} \left[ \mathcal{A}^u V^{\hat{\mathbf{u}}}(t, x) + C^{t, x, u}(t, x) - \left( \mathcal{A}_{\tau, y}^u H^{t, x}(t, x) + \mathcal{A}_{\tau, y}^u G(t, x, g(t, x)) \right) \right. \\
& \quad - \text{tr} \left( \left( H_{xy}^{t, x}(t, x) + G_{zy}(t, x, g(t, x)) \right)^\top g_x(t, x) \right)^\top \Upsilon^u(t, x) \Big) \\
& \quad - \frac{1}{2} G_{zz}(t, x, g(t, x)) \text{tr} \left( g_x(t, x) g_x(t, x)^\top \Upsilon^u(t, x) \right) \Big] = 0, (t, x) \in [0, T] \times \mathbb{X}, \\
& V^{\hat{\mathbf{u}}}(T, x) = F(T, x, x) + G(T, x, x), x \in \mathbb{X}.
\end{aligned} \tag{18}$$

By (17), the definitions of  $c^{\tau, y, s}(t, x)$  and  $f^{\tau, y}(t, x)$ , and the Feymann-Kac formula, we derive the following equation for  $H^{\tau, y}(t, x)$ :

$$\begin{aligned}
& \mathcal{A}^{\hat{\mathbf{u}}} H^{\tau, y}(t, x) + C^{\tau, y, \hat{\mathbf{u}}}(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{X}, (\tau, y) \in [0, T] \times \mathbb{X}, \\
& H^{\tau, y}(T, x) = F(\tau, y, x), \quad x \in \mathbb{X}, (\tau, y) \in [0, T] \times \mathbb{X}.
\end{aligned} \tag{19}$$

Similarly, we derive the following equation for  $g$ :

$$\begin{aligned}
& \mathcal{A}^{\hat{\mathbf{u}}} g(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{X}, \\
& g(T, x) = x, \quad x \in \mathbb{X}.
\end{aligned} \tag{20}$$

Some remarks are in order. First, instead of a single equation for the value function of a time-consistent problem, the intra-personal equilibrium and its continuation value satisfy a system of equations (18)–(20), which is referred to as the *extended HJB equation* by Björk et al. (2017).

Second, compared to the HJB equation for a time-consistent problem, which takes the form  $\max_{u \in \mathbb{U}} [\mathcal{A}^u V^{\hat{\mathbf{u}}}(t, x) + C^u(t, x)] = 0$ , equation (18) has three additional terms in the first, second, and third lines of the equation, respectively. Here and hereafter, when  $C^{\tau, y, u}(t, x)$  does not depend on  $(\tau, y)$ , we simply drop the superscript  $(\tau, y)$ . Similar notations apply to  $H^{\tau, y}(t, x)$  and to the case when there is no dependence on  $y$ . Now, recall that for the objective function (3), time inconsistency arises from (i) the dependence of  $C$ ,  $F$ , and  $G$  on  $(t, x)$  and (ii) the nonlinear dependence of  $G(t, x, \mathbb{E}_{t, x}[X^{\mathbf{u}}(T)])$  on  $\mathbb{E}_{t, x}[X^{\mathbf{u}}(T)]$ . If source (i) of time inconsistency is absent, the first and second additional terms in (18) will vanish. If source (ii) of time inconsistency is absent, the third additional term in (18) will disappear. In particular, without time inconsistency, the extended HJB equation (18) reduces to the classical HJB equation.

Third, consider the case in which  $G(t, x, \mathbb{E}_{t, x}[X^{\mathbf{u}}(T)])$  is linear in  $\mathbb{E}_{t, x}[X^{\mathbf{u}}(T)]$  and  $C$ ,  $F$ , and  $G$  do not depend on  $x$ . In this case, the second and third lines of (18) vanish and we can assume  $G \equiv 0$  without loss of generality because  $G$  can be combined with  $F$ . As a result, the extended HJB equation (18) specializes to

$$\begin{aligned}
& \max_{u \in \mathbb{U}} \left[ \mathcal{A}^u V^{\hat{\mathbf{u}}}(t, x) + C^{t, u}(t, x) \right] = h^t(t, x), (t, x) \in [0, T] \times \mathbb{X}, \\
& V^{\hat{\mathbf{u}}}(T, x) = F(T, x), x \in \mathbb{X}
\end{aligned} \tag{21}$$

where  $h^\tau(t, x) := H_\tau^\tau(t, x)$  (with the subscript  $\tau$  denoting the partial derivative with respect to  $\tau$ ) and thus satisfies

$$\begin{aligned} \mathcal{A}^{\hat{\mathbf{u}}} h^\tau(t, x) + C_\tau^{\tau, \hat{\mathbf{u}}}(t, x) &= 0, \quad (t, x) \in [0, T] \times \mathbb{X}, \tau \in [0, T], \\ h^\tau(T, x) &= F_\tau(\tau, x), \quad x \in \mathbb{X}, \tau \in [0, T]. \end{aligned} \tag{22}$$

### 3 Discussions

#### 3.1 Intra-Personal Equilibria with Fixed Initial Data

Consider an agent at time 0 with a fixed state  $x_0$  who correctly anticipates that her self at each future time  $t$  faces the problem (2) and who has no control of future selves at any time. A strategy  $\hat{\mathbf{u}}$  can be consistently implemented by the agent throughout the entire horizon  $[0, T]$  if the agent has no incentive to deviate from it at any time *along the state path*. Actions that the agent might be taking were she not on the state path are irrelevant. To be more precise, for any fixed initial data  $(0, x_0)$ , we define  $\hat{\mathbf{u}}$  to be an *intra-personal equilibrium starting from*  $(0, x_0)$  if (5) holds for any  $\mathbf{a} \in \mathbf{U}$ ,  $t \in [0, T)$ , and  $x \in \mathbb{X}_t^{0, x_0, \hat{\mathbf{u}}}$ , where  $\mathbb{X}_t^{0, x_0, \hat{\mathbf{u}}}$  denotes the set of all possible states at time  $t$  along the state path starting from  $x_0$  at the initial time and under the strategy  $\hat{\mathbf{u}}$ .

It is evident that the intra-personal equilibrium defined in Definition 2 is *universal* in that it is an equilibrium starting from *any* initial data  $(0, x_0)$ . On the other hand, starting from a fixed state  $x_0$  at time 0, the state process in the future might not be able to visit the whole state space; so an equilibrium starting from  $(0, x_0)$  is not necessarily universal, i.e., it is not necessarily an equilibrium when the agent starts from other initial data. For example, He et al. (2020) consider a continuous-time portfolio selection problem in which an agent maximizes the median of her terminal wealth. With a fixed initial wealth of the agent, the authors derive a set of intra-personal equilibrium strategies starting from this particular initial wealth level. They show that these strategies are no longer equilibria if the agent starts from some other initial wealth levels, and in particular not universal equilibria in the sense of Definition 2.

The first study of intra-personal equilibria starting from a fixed initial data dates back to Peleg and Yaari (1973). In a discrete-time setting, the authors propose that a strategy  $(s_0^*, s_1^*, \dots)$ , where  $s_t^*$  stands for the agent’s closed-loop strategy at time  $t$ , is an equilibrium strategy if for any  $t$ ,  $(s_0^*, \dots, s_{t-1}^*, s_t, s_{t+1}^*, \dots)$  is dominated by  $(s_0^*, \dots, s_{t-1}^*, s_t^*, s_{t+1}^*, \dots)$  for any  $s_t$ . They argue that the above definition is more desirable than the following one, which is based on a model in Pollak (1968):  $(s_0^*, s_1^*, \dots)$  is an equilibrium strategy if for any time  $t$ ,  $(s_0, \dots, s_{t-1}, s_t, s_{t+1}^*, \dots)$  is dominated by  $(s_0, \dots, s_{t-1}, s_t^*, s_{t+1}^*, \dots)$  for *any*  $(s_0, \dots, s_t)$ . It is clear that the equilibrium strategies considered by Peleg and Yaari (1973) are the ones starting from a fixed initial data while those studied by Pollak (1968) are universal. Recently, He

and Jiang (2019), Han and Wong (2020), and Hernández and Possamai (2020) also consider intra-personal equilibria with fixed initial data. Moreover, He and Jiang (2019) propose a formal definition of  $\mathbb{X}_t^{0,x_0,\hat{\mathbf{u}}}$ , calling it the set of reachable states.

Finally, let us comment that the sufficient and necessary condition in Theorem 1 is still valid for intra-personal equilibria starting from fixed initial data  $(0, x_0)$ , provided that we replace  $\mathbb{X}$  in this condition with the set of reachable states  $\mathbb{X}_t^{0,x_0,\hat{\mathbf{u}}}$ ; see He and Jiang (2019) for details. The extended HJB equation in Section 2.5 can be revised and applied similarly.

### 3.2 Set of Alternative Strategies

In Definition 2, the set of strategies that the agent can choose at time  $t$  to implement for the period  $[t, t + \varepsilon)$ , denoted as  $\mathbf{D}$ , is set to be the entire set of feasible strategies  $\mathbf{U}$ . This definition is used in Björk et al. (2017), Ekeland and Pirvu (2008), and Ekeland et al. (2012). In some other works, however,  $\mathbf{D}$  is set to be the set of constant strategies  $\mathbb{U}$ ; see for instance Ekeland and Lazrak (2006, 2008, 2010), Björk and Murgoci (2010), and Basak and Chabakauri (2010). He and Jiang (2019) show that the choice of  $\mathbf{D}$  is irrelevant as long as it at least contains  $\mathbb{U}$ . Indeed, this can be seen from the observation in Theorem 1 that  $\Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\mathbf{a}) = \Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\mathbf{a}(t,x))$  for any  $\mathbf{a} \in \mathbf{U}$ . He and Jiang (2019) also show that for strong intra-personal equilibrium, which will be introduced momentarily, the choice of  $\mathbf{D}$  is relevant.

### 3.3 Regular and Strong Intra-Personal Equilibrium

As noted in Remark 3.5 of Björk et al. (2017), condition (5) does not necessarily imply that  $J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}})$  is less than or equal to  $J(t,x;\hat{\mathbf{u}})$  however small  $\varepsilon > 0$  might be and thus disincentivizes the agent from deviating from  $\hat{\mathbf{u}}$ . For example, if  $J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}}) - J(t,x;\hat{\mathbf{u}}) = \varepsilon^2$ , then (5) holds, but the agent can achieve a strictly larger objective value if she deviates from  $\hat{\mathbf{u}}$  to  $\mathbf{a}$  and thus is willing to do so.

To address the above issue, Huang and Zhou (2019) and He and Jiang (2019) propose the notion of *strong intra-personal equilibrium*:

#### Definition 3 (Strong Intra-personal Equilibrium)

$\hat{\mathbf{u}} \in \mathbf{U}$  is a *strong intra-personal equilibrium strategy* if for any  $x \in \mathbb{X}$ ,  $t \in [0, T)$ , and  $\mathbf{a} \in \mathbf{D}$ , there exists  $\varepsilon_0 \in (0, T - t)$  such that

$$J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}}) - J(t,x;\hat{\mathbf{u}}) \leq 0, \quad \forall \varepsilon \in (0, \varepsilon_0]. \quad (23)$$

It is straightforward to see that a strong intra-personal equilibrium implies the one in Definition 2, which we refer to as a *weak intra-personal equilibrium* in this subsection.

Huang and Zhou (2019) consider a stochastic control problem in which an agent can control the generator of a time-homogeneous, continuous-time, finite-state Markov chain at each time to maximize expected running reward in an infinite time horizon. Assuming that at each time the agent can implement a time-homogeneous strategy only, the authors provide a characterization of a strong intra-personal equilibrium and prove its existence under certain conditions.

He and Jiang (2019) follow the framework in (2) and derive two necessary conditions for a strategy to be strong intra-personal equilibrium. Using these conditions, the authors show that strong intra-personal equilibrium does not exist for the portfolio selection and consumption problems studied in Ekeland and Pirvu (2008), Basak and Chabakauri (2010), and Björk et al. (2014). Motivated by this non-existence result, the authors propose the so-called *regular intra-personal equilibrium* and show that it exists for the above three problems and is stronger than the weak intra-personal equilibrium and weaker than the strong intra-personal equilibrium in general.

### 3.4 Existence and Uniqueness

In most studies on time-inconsistent problems in the literature, a *closed-form* strategy is constructed and verified to satisfy the sufficient and necessary condition (12) or the extended HJB equation (18)–(20). The existence of intra-personal equilibrium in general is difficult to prove because it essentially relies on a fixed point argument: For each guess of intra-personal equilibrium  $\hat{\mathbf{u}}$ , we first calculate  $\Gamma^{\tau,y;\hat{\mathbf{u}}}$  in (12) and  $H^{\tau,y}(t,x)$  and  $g$  in (19) and (20), respectively, and then derive an updated intra-personal equilibrium, denoted as  $\mathbb{T}\hat{\mathbf{u}}$ , from the condition (12) or from the equation (18). The existence of an intra-personal equilibrium then boils down to the existence of the fixed point of  $\mathbb{T}$ . The mapping  $\mathbb{T}$  is highly nonlinear; so the existence of its fixed point is hard to establish. Additional difficulty is caused by the regularity conditions that we need to pose on  $\hat{\mathbf{u}}$  to validate the sufficient and necessary condition (12) or the extended HJB equation (18)–(20).

We are only aware of very few works on the existence of intra-personal equilibria in continuous time. Yong (2012) proposes an alternative approach to defining the strategy of a sophisticated agent, which will be discussed in detail in Section 6. Assuming  $G \equiv 0$ ,  $C$  and  $F$  to be independent of  $x$  in the objective function (3), and  $\sigma(t,x,u)$  in the controlled diffusion process (2) to be independent of control  $u$  and nondegenerate, Yong (2012) proves the existence of the sophisticated agent's strategy, which is used to imply the existence of an intra-personal equilibrium under Definition 2. Wei et al. (2017) and Wang and Yong (2019) extend the result of Yong (2012) by generalizing the objective function; however for the existence of intra-personal equilibria, they need to assume the volatility  $\sigma$  to be independent of control and nondegenerate. Hernández and Possamaï (2020) study intra-personal equilibria in a non-Markovian setting, where they consider a non-Markovian version of the objective function in Yong (2012) and assume the drift  $\mu$  of the controlled process to

be in the range of the volatility matrix at each time. The authors prove the existence of intra-personal equilibria when the volatility  $\sigma$  is independent of control.

Intra-personal equilibria can be non-unique; see Ekeland and Lazrak (2010), Cao and Werning (2016), and He et al. (2020). For some problems, however, uniqueness has been established in the literature. Indeed, Yong (2012), Wei et al. (2017), Wang and Yong (2019), and Hernández and Possamaï (2020) prove the uniqueness in various settings with the common assumption that the volatility  $\sigma$  is independent of control.

### 3.5 Non-Markovian Strategies

In most studies on time-inconsistent problems, where the controlled state processes are Markovian, the search for intra-personal equilibrium is restricted to the set of Markovian strategies, i.e., strategies that are functions of time  $t$  and the *current* state value  $x$ . Motivated by some practical problems such as rough volatility models and principle-agent problems, Han and Wong (2020) and Hernández and Possamaï (2020) define and search intra-personal equilibria in the class of non-Markovian or path-dependent strategies, i.e., ones that depend on time  $t$  and the whole path of the controlled state up to  $t$ .

## 4 Closed-Loop versus Open-Loop Intra-Personal Equilibria

A *closed-loop* or *feedback* control strategy is a function  $\mathbf{u}$  that maps time  $t$  and the controlled state path  $(x_s)_{s \leq t}$  up to  $t$  to the space of actions. As a result, the action taken by an agent under such a strategy is  $\mathbf{u}(t, (x_s)_{s \leq t})$ . An *open-loop* control is a collection of actions over time and state of the nature,  $(u(t, \omega))_{t \geq 0}$ , where  $u(t, \omega)$  is the action to be taken at time  $t$  and in scenario  $\omega$ , *regardless* of the state path  $(x_s)_{s \leq t}$ . For classical time-consistent control problems and under some technical assumptions, the state-control paths under the optimal open-loop control and under the optimal closed-loop control strategy are the same if the controlled system starts from the same initial time and state; see for instance Yong and Zhou (1999).

In Section 2, intra-personal equilibrium is defined for closed-loop control strategies, which is also the approach taken by most studies on time-inconsistent problems in the literature. In some other works, intra-personal equilibrium is defined for open-loop controls; see for instance Hu et al. (2012), Hu et al. (2017), Li et al. (2019), and Hu et al. (2021).

Formally, under the same probabilistic framework in Section 2.2, we represent an open-loop strategy by a progressively measurable process  $(u(t))_{t \geq 0}$  that takes values in  $\mathbb{U}$ . The controlled state process  $X^u$  takes the form

$$dX^u(s) = \mu(s, X^u(s), u(s))ds + \sigma(s, X^u(s), u(s))dW(s), \quad s \in [t, T]; \quad X^u(t) = x.$$

Denote by  $\mathcal{U}$  the set of feasible open-loop controls, i.e., the set of progressively measurable processes on  $[0, T]$  satisfying certain integrability conditions. At time  $t$  with state  $x$ , suppose the agent's objective is to maximize  $J(t, x; u(\cdot))$  by choosing  $u(\cdot) \in \mathcal{U}$ . Given  $\hat{u}(\cdot) \in \mathcal{U}$ , for any  $t \in [0, T)$ ,  $x \in \mathbb{X}$ ,  $\varepsilon \in (0, T - t)$ , and  $a(\cdot) \in \mathcal{U}$ , define

$$u_{t,\varepsilon,a}(s) := \begin{cases} a(s), & s \in [t, t + \varepsilon) \\ \hat{u}(s), & s \notin [t, t + \varepsilon). \end{cases} \quad (24)$$

Suppose that at time  $t$  with state  $x$ , the agent chooses an open-loop control  $a(\cdot)$ , but is only able to implement it in the period  $[t, t + \varepsilon)$ . Anticipating that her future selves will take the given control  $\hat{u}(\cdot)$ , the agent expects herself to follow  $u_{t,\varepsilon,a}$  in the period  $[t, T]$ .

**Definition 4 (Open-Loop Intra-Personal Equilibrium)**

$\hat{u}(\cdot) \in \mathcal{U}$  is an *open-loop intra-personal equilibrium* if for any  $x \in \mathbb{X}$ ,  $t \in [0, T)$ , and  $a \in \mathcal{U}$  that is constant in a small period after  $t$ , we have

$$\limsup_{\varepsilon \downarrow 0} \frac{J(t, x; u_{t,\varepsilon,a}(\cdot)) - J(t, x; \hat{u}(\cdot))}{\varepsilon} \leq 0. \quad (25)$$

The above is analogous to the definition of an intra-personal equilibrium for closed-loop strategies. However, there is a subtle yet crucial difference between the two definitions. For the one for open-loop controls, the perturbed control  $u_{t,\varepsilon,a}(s)$  defined by (24) and the original one  $\hat{u}$  are identical on  $[t + \varepsilon, T]$  as two stochastic processes. In other words, the perturbation in the small time period  $[t, t + \varepsilon)$  will not affect the control process beyond this period. This is not the case for the closed-loop counterpart, because the perturbation (4) on  $[t, t + \varepsilon)$  changes the control in the period, which will alter the state process in  $[t, t + \varepsilon)$  and in particular the state at time  $t + \varepsilon$ . This in turn will change the control *process* on  $[t + \varepsilon, T]$  upon substituting the state process into the feedback strategy.

To characterize open-loop intra-personal equilibria, we only need to compute the limit on the left-hand side of (25). This limit can be evaluated by applying the spike variation technique that is used to derive Pontryagin's maximum principle for time-consistent control problems in continuous time (Yong and Zhou, 1999). As a result, open-loop intra-personal equilibrium can be characterized by a flow of forward-backward stochastic differential equations (SDEs); see Hu et al. (2012) for more details. In contrast, the spike variation technique no longer works for closed-loop equilibria because the perturbed control process is different from the original one beyond the small time period for perturbation, as discussed above.

This discussion suggests that closed-loop and open-loop equilibria are likely different. This is confirmed by Hu et al. (2012). The authors consider a mean-variance portfolio selection problem, where an agent decides the dollar amount invested in a stock at each time, and derive an open-loop equilibrium; see Section 5.4.1 therein. They then compare this equilibrium with the closed-loop equilibrium derived by



Björk et al. (2014) for the same portfolio selection problem, and find that the state-control path under these two equilibria are different.

It can be argued that closed-loop strategies are preferred to the open-loop ones for three reasons. First, in many problems, agents' actions naturally depend on some state variables. For example, in a consumption problem, an agent's consumption at any time is more likely to depend directly on her wealth at that time. If her wealth suddenly increases, she would probably consume more.

Second, closed-loop intra-personal equilibrium is invariant to the choice of control variables while open-loop intra-personal equilibrium might not. For example, in a portfolio selection problem where an agent decides the allocation of her wealth between a risk-free asset and a risky stock, the decision variable can be set to be the dollar amount invested in the stock or the percentage of wealth invested in the stock. Suppose  $\hat{\mathbf{u}}$  is a closed-loop intra-personal equilibrium representing the percentage of wealth invested in the stock. Then, we have

$$\limsup_{\varepsilon \downarrow 0} \frac{J(t, x; \mathbf{u}_{t, \varepsilon, \mathbf{a}}) - J(t, x; \hat{\mathbf{u}})}{\varepsilon} \leq 0. \quad (26)$$

for all  $t \in [0, T)$ ,  $x \in \mathbb{X}$ , and  $\mathbf{a} \in \mathbf{U}$ , where the state variable  $x$  represents the agent's wealth. Now, suppose we represent the agent's decision by the dollar amount invested in the risky stock, and denote a control strategy as  $\pi$ . Then, the agent's objective function is  $\tilde{J}(t, x; \pi) = J(t, x; \mathbf{u})$  with  $\mathbf{u}(s, y) = \pi(s, y)/y$ . Condition (26) implies that

$$\limsup_{\varepsilon \downarrow 0} \frac{\tilde{J}(t, x; \pi_{t, \varepsilon, \tilde{\mathbf{a}}}) - \tilde{J}(t, x; \hat{\pi})}{\varepsilon} \leq 0,$$

for any  $t \in [0, T)$ ,  $x \in \mathbb{X}$ , and strategy  $\tilde{\mathbf{a}}$  that represents the dollar amount invested in the stock, where  $\hat{\pi}(s, y) := y\hat{\mathbf{u}}(s, y)$  and  $\pi_{t, \varepsilon, \tilde{\mathbf{a}}}$  is defined similarly to  $\mathbf{u}_{t, \varepsilon, \mathbf{a}}$ . Thus,  $\hat{\pi}$ , which is the dollar amount investment strategy corresponding to the percentage investment strategy  $\hat{\mathbf{u}}$ , is also an intra-personal equilibrium. By contrast, for the mean-variance portfolio selection problem studied by Hu et al. (2012), where the agent's decision is the dollar amount invested in the stock, the open-loop intra-personal equilibrium yields a different control-state path from the one yielded by its closed-loop counterpart derived by Björk et al. (2014). If we change the agent's decision variable to the percentage of wealth invested in the stock, the open-loop intra-personal equilibrium and the closed-loop intra-personal equilibrium in Björk et al. (2014) yield the same control-state path. This implies that open-loop equilibria depend on the choice of control variables.

Third, open-loop intra-personal equilibrium may not be well-posed for some problems. Consider the discrete-time version of the consumption problem studied in Strotz (1955-1956): An agent decides the amount of consumption  $C_t$  at each time  $t = 0, 1, \dots, T$  with the total budget  $x_0$ , i.e.,  $\sum_{t=0}^T C_t = x_0$ . For this problem, any consumption plan  $(\hat{C}_t)_{t \geq 0}$  is an open-loop intra-personal equilibrium. Indeed, at each time  $t$ , anticipating her future selves will consume  $\hat{C}_s, s = t + 1, \dots, T$ , the

only amount of consumption  $C_t$  that the agent can choose at time  $t$  is  $\hat{C}_t$  due to the budget constraint  $(\sum_{s=0}^{t-1} \hat{C}_s) + C_t + (\sum_{s=t+1}^T \hat{C}_s) = x_0$ . This leads to a trivial definition of intra-personal equilibrium. The above issue can be rectified if we use closed-loop strategies. To see this, we set  $x_t$  to be the agent's remaining budget at time  $t$  before the consumption at that time. For closed-loop intra-personal equilibrium, we consider a mapping from time  $t$  and the remaining budget  $x_t$  to the consumption amount. As a result, if the agent consumes more at time  $t$ , her future selves will consume less because the remaining budget in the future becomes smaller; consequently, the budget constraint is still satisfied. To elaborate, suppose the agent's future selves' strategies are to consume  $\hat{k}_s$  fractional of wealth at time  $s$ ,  $s = t + 1, \dots, T$  with  $\hat{k}_s \in [0, 1]$ ,  $s = t + 1, \dots, T - 1$  and  $\hat{k}_T = 1$ . Then, given that the agent at time  $t$  consumes any amount  $C_t \in [0, x_t]$ , the agent's consumption in the future is  $C_s = \hat{k}_s x_s$ ,  $s = t + 1, \dots, T$ , where  $x_s = x_{s-1} - C_{s-1}$ ,  $s = t + 1, \dots, T$ . As a result, the aggregate consumption from time  $t$  to the end is  $\sum_{s=t}^T C_s = x_t$ . Recall that the aggregate consumption strictly prior to time  $t$  is  $x_0 - x_t$ ; so the aggregate consumption throughout the entire horizon is  $x_0$  satisfying the budget constraint. Thus, at each time  $t$ , the agent can consume any amount up to his wealth level at that time and her future selves will adjust their consumption according to a given strategy so that the budget constraint is still satisfied.

Finally, we establish a connection between closed-loop and open-loop intra-personal equilibria. If a closed-loop equilibrium  $\hat{\mathbf{u}}$  is independent of the state variable  $x$ , then it follows from the definition that it is also an open-loop equilibrium. For a general closed-loop equilibrium  $\hat{\mathbf{u}}$ , we can consider the following controlled state process:

$$d\hat{X}^v(s) = \hat{\mu}(s, \hat{X}^v(s), v(s))ds + \hat{\sigma}(s, \hat{X}^v(s), v(s))dW(s), \quad s \in [t, T]; \quad X^v(t) = x,$$

where  $\hat{\mu}(s, y, v) := \mu(s, y, \hat{\mathbf{u}}(s, y) + v)$ ,  $\hat{\sigma}(s, y, v) := \sigma(s, y, \hat{\mathbf{u}}(s, y) + v)$ , and  $v(\cdot)$  is a progressively measurable control process. We further consider the following objective function:

$$\hat{J}(t, x; v(\cdot)) := \mathbb{E}_{t,x} \left[ \int_t^T \hat{C}(t, x, s, \hat{X}^v(s), v(s)) ds + F(t, x, \hat{X}^v(T)) \right] + G(t, x, \mathbb{E}_{t,x}[\hat{X}^v(T)]),$$

where  $\hat{C}(t, x, s, y, v) := C(t, x, s, y, \hat{\mathbf{u}}(s, y) + v)$ . Then, by definition,  $\hat{\mathbf{u}}$  is a closed-loop equilibrium if and only if  $\hat{v}(\cdot) \equiv 0$  is an open-loop equilibrium for the problem of maximizing  $\hat{J}(t, x; v(\cdot))$  in  $v(\cdot)$  with the controlled state process  $\hat{X}^v$ . In particular, we can characterize  $\hat{\mathbf{u}}$  by a flow of forward-backward SDEs by applying the spike variation technique. In order to apply this technique, however, we need to assume that  $\hat{\mu}(s, y, v)$  and  $\hat{\sigma}(s, y, v)$  to be twice differentiable in  $y$ , which in turn requires  $\hat{\mathbf{u}}$  to be twice differentiable; see Yong and Zhou (1999) for the detailed regularity conditions needed for the spike variation technique. Thus, the spike variation technique does not seem to be advantageous over the approached reviewed in Section 2.

## 5 Optimal Stopping

An optimal stopping problem is one to search an optimal random time  $\tau$  to stop a given, *uncontrollable* process  $(X_t)_{t \geq 0}$  (taking values in a state space  $\mathbb{X}$ ) in the set of stopping times with respect to the filtration generated by the process. It is well known that if the objective function of the optimal stopping problem depends on the path of  $(X_t)_{t \geq 0}$  up to the stopping time only, this problem can be “embedded” into a general control problem with (i) a closed-loop control strategy  $\mathbf{u}$  taking binary values 0 and 1 representing the action of stopping and not stopping  $(X_t)_{t \geq 0}$  respectively; and (ii) a controlled state process  $(\tilde{X}^{\mathbf{u}})_{t \geq 0}$  that is set to be  $(X_t)_{t \geq 0}$  until the first time the control path under  $\mathbf{u}$  takes value 0 and is set to be an *absorbing state* afterwards; see for instance Section 3.4 of Bertsekas (2017). We call the control strategy  $\mathbf{u}$  associated with a stopping time  $\tau$  in the above embedding a *stopping rule*, which maps each pair of time  $t$  and a path of the process  $X$  up to time  $t$  to  $\{0, 1\}$ . A stopping time  $\tau$  is *Markovian* if the associated stopping rule is Markovian, i.e., it is a mapping from the time–state space to  $\{0, 1\}$ . With a Markovian stopping time, at each time  $t$ , given that the process has not yet been stopped, whether to stop at  $t$  depends on the value of the process at  $t$  only.

In view of the above embedding, intra-personal equilibrium stopping rules can be defined naturally for time-inconsistent stopping problems; see for instance Tan et al. (2018), Christensen and Lindensjö (2018), Ebert et al. (2020), and Christensen and Lindensjö (2020). In particular, Tan et al. (2018) show that the smooth pasting principle, which is the main approach used to construct explicit solutions for classical time-consistent optimal stopping, may fail to find an equilibrium when one changes merely the exponential discounting to non-exponential one while keeping everything else the same. The authors also construct an explicit example in which no equilibrium exists. These results caution blindly extending the classical approach for time-consistent stopping to their time-inconsistent counterpart.

By exploiting special structures of stopping problems in continuous time, Huang and Nguyen-Huu (2018) propose an alternative approach to defining the optimal stopping rule for a sophisticated agent; see also applications of this approach in Huang et al. (2020), Ebert and Strack (2017), and Huang and Yu (2021). Precisely, consider a Markov state process

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

in  $\mathbb{R}^n$ , where  $(W_t)_{t \geq 0}$  is an  $d$ -dimensional standard Brownian motion and  $\mu$  and  $\sigma$  are functions of time  $t$  and state  $x$  taking values in  $\mathbb{R}^n$  and  $\mathbb{R}^{n \times d}$ , respectively. Following the settings in the above papers, we consider Markovian stopping times only in the following presentation, but the case of non-Markovian stopping times can be investigated similarly. At each time  $t$  with state  $x$ , given that the state process has not been stopped, the agent’s goal is to choose a Markovian stopping time  $\tau \in [t, T]$  to maximize an objective value  $J(t, x; \tau)$ . Here,  $J(t, x; \tau)$  can be of the form  $\mathbb{E}_{t,x} \left[ \int_t^\tau g(t, x, s, X_s) ds + h(t, x, \tau, X_\tau) \right]$  for some functions  $g$  and  $h$ , or be a functional of the distribution of  $X_\tau$  conditional on  $X_t = x$ .

Recall the embedding of optimal stopping problems into a general control framework and the stopping rule associated with each stopping time as discussed at the beginning of the present subsection. With a slight abuse of notation, we use  $\tau$  to denote both a stopping time and a stopping rule. Let us now consider a given stopping rule  $\tau$  and the current time–state pair  $(t, x)$ . If the agent decides to stop, then she has the immediate reward  $J(t, x; t)$ . If the agent decides not to stop at  $t$  but expects her future selves will still follow the original rule  $\tau$ , then she will stop at time  $\mathcal{L}^*\tau$ , the first time  $s > t$  at which  $\tau$  would stop the process. In this case the objective value is  $J(t, x; \mathcal{L}^*\tau)$ . Then, the optimal action of the agent at time  $t$  with state  $x$  is to stop if  $J(t, x; t) > J(t, x; \mathcal{L}^*\tau)$ , to continue if  $J(t, x; t) < J(t, x; \mathcal{L}^*\tau)$ , and to follow the originally assigned stopping rule  $\tau$  in the break-even case  $J(t, x; t) = J(t, x; \mathcal{L}^*\tau)$ . The above plan across all time  $t$  and state  $x$  constitutes a *new* stopping rule, denoted as  $\Theta\tau$ , which can be proved to be feasible in the sense that it can generate stopping times; see Huang and Nguyen-Huu (2018) and Huang et al. (2020).

The above game-theoretic thinking shows that for any arbitrarily given stopping rule  $\tau$ , at any time  $t$  with any state  $x$ , the agent finds  $\Theta\tau$  to be always no worse than  $\tau$ , *assuming* that her future selves will follow  $\tau$ . Hence, an equilibrium stopping rule  $\tau$  can be defined as one that can not be strictly improved by taking  $\Theta\tau$  instead. Following Bayraktar et al. (2021), we name it as a *mild intra-personal equilibrium* stopping rule:

**Definition 5** A stopping rule  $\tau$  is a mild intra-personal equilibrium if  $\Theta\tau = \tau$ .

So a mild intra-personal equilibrium is a fix-point of the operator  $\Theta$ . If  $\tau$  is to stop the process at any time and with any state, then it is straightforward to see that  $\mathcal{L}^*\tau = \tau$ . Consequently, by definition  $\Theta\tau = \tau$  and thus  $\tau$  is a mild intra-personal equilibrium. In other words, following Definition 5, immediate stop is *automatically* a (trivial) mild intra-personal equilibrium.

For a general stopping rule  $\tau$ , consider any time  $t$  and state  $x$  in the interior of the stopping region of  $\tau$ , where the stopping region refers to the set of time-state pairs at which the stopping rule  $\tau$  would stop the process. Then, it is also easy to see that  $\mathcal{L}^*\tau = \tau$  at time  $t$  and state  $x$ , so one should immediately stop under  $\Theta\tau$  as well. As a result, the stopping region of  $\Theta\tau$  is at least as large as that of  $\tau$ , if we ignore the time-state pairs that are on the boundary of the stopping region of  $\tau$ . Therefore, we expect the iterative sequence  $\Theta^n\tau$  to converge as  $n \rightarrow \infty$ , and the convergent point  $\tau^*$  satisfies  $\tau^* = \Theta\tau^*$  and thus is a mild intra-personal equilibrium. It is, however mathematically challenging to formalize the above heuristic derivation. Rigorous proofs have been established in various settings by Huang and Nguyen-Huu (2018), Huang et al. (2020), and Huang and Yu (2021). The above iterative algorithm, which generates a sequence  $\Theta^n\tau$ ,  $n = 0, 1, \dots$ , not only yields a mild intra-personal equilibrium as the limit of the sequence, but also has a clear economic interpretation: each application of  $\Theta$  corresponds to an additional level of strategic reasoning; see Huang and Nguyen-Huu (2018) and Huang et al. (2020) for elaborations.

As discussed in the above, immediate stop is always a mild equilibrium; so it is expected that there exist multiple mild intra-personal equilibrium stopping rules; see Huang and Nguyen-Huu (2018) and Huang et al. (2020). To address the issue

of multiplicity, Huang and Zhou (2020) and Huang and Wang (2020) consider, in the setting of an infinite-horizon, continuous-time optimal stopping under nonexponential discounting, the “optimal” mild intra-personal equilibrium stopping rule  $\tau^*$  which achieves the maximum of  $J(t, x; \tau)$  over  $\tau \in \mathcal{E}$  for all  $t \in [0, T)$ ,  $x \in \mathbb{X}$ , where  $\mathcal{E}$  is the set of all mild intra-personal equilibrium stopping rules.

Bayraktar et al. (2021) compare mild intra-personal equilibrium stopping rules with weak (respectively strong) intra-personal equilibrium stopping rules obtained by embedding optimal stopping into stochastic control and then applying Definition 2 (respectively Definition 3). Assuming the objective function to be a multiplication of a discount function and a Markov process taking values in a finite or countably infinite state space, the authors prove that the optimal mild intra-personal equilibrium is a strong intra-personal equilibrium.

## 6 Discretization Approach

In the discrete-time setting, an intra-personal equilibrium strategy of a sophisticated agent can be easily defined and derived in a backward manner starting from the last period. Thus, for a continuous-time problem, it is natural to discretize and then pass to the limit. Specifically, one partitions the continuous-time period  $[0, T]$  into a finite number of subperiods, assumes the agent is able to commit in each subperiod but not beyond it, and computes the strategy chosen by the agent. Sending the length of the longest subperiod in the partition to zero, the limit of the above strategy, if it exists, can be regarded as the strategy of a sophisticated agent for the continuous-time problem. This idea was first employed by Pollak (1968) to study the consumption problem of Strotz (1955-1956) and has recently been revisited and extensively studied by a series of papers; see for instance Yong (2012), Wei et al. (2017), Mei and Yong (2019), and Wang and Yong (2019).

Specifically, consider the control problem in Section 2 and assume that in the objective function in (3),  $C$  and  $F$  do not depend on  $x$  and  $G \equiv 0$ . For a partition  $\Pi$  of  $[0, T]: 0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T$ , we denote  $\|\Pi\| := \max_{k=1, \dots, N} |t_k - t_{k-1}|$ . A control strategy  $\hat{\mathbf{u}}^\Pi$  is an intra-personal equilibrium with respect to the partition  $\Pi$  if

$$J(t_k, x_k; \hat{\mathbf{u}}^\Pi) \geq J(t_k, x_k; \mathbf{u}_{k, \mathbf{a}}^\Pi) \quad (27)$$

for any  $k = 0, 1, \dots, N-1$ , reachable state  $x_k$  at time  $k$  under  $\hat{\mathbf{u}}^\Pi$ , and strategy  $\mathbf{a}$ , where  $\mathbf{u}_{k, \mathbf{a}}^\Pi(s, \cdot) := \mathbf{a}(s, \cdot)$  for  $s \in [t_k, t_{k+1})$  and  $\mathbf{u}_{k, \mathbf{a}}^\Pi(s, \cdot) = \hat{\mathbf{u}}^\Pi(s, \cdot)$  for  $s \in [t_{k+1}, T)$ . In other words,  $\hat{\mathbf{u}}(s, \cdot), s \in [t_k, t_{k+1})$ , is optimal for an agent who can commit in the period  $[t_k, t_{k+1})$  and anticipates that her future selves will take strategy  $\hat{\mathbf{u}}$  beyond time  $t_{k+1}$ . In the aforementioned literature, the authors define a strategy  $\hat{\mathbf{u}}$  to be a *limiting intra-personal equilibrium* if there exists a sequence of partition  $(\Pi_m)_{m \in \mathbb{N}}$  with  $\lim_{m \rightarrow \infty} \|\Pi_m\| = 0$  such that the state process, control process, and continuation value process under certain intra-personal equilibrium with respect to  $\Pi_m$  converge

to those under  $\hat{\mathbf{u}}$ , respectively, as  $m \rightarrow \infty$ . Assuming that the diffusion coefficient of the controlled state process is independent of control and non-degenerate and that some other conditions hold, Wei et al. (2017) prove the above convergence for any sequence of partitions with mesh size going to zero, and the limit of the continuation value function satisfies a flow of PDEs. Moreover, this flow of PDEs admits a unique solution, so the limiting intra-personal equilibrium uniquely exists. Furthermore, the limiting equilibrium is also an equilibrium under Definition 2.

Whether the equilibrium with respect to  $\Pi$  converges when  $\|\Pi\| \rightarrow 0$  for a general time-inconsistent problem, however, is still unknown. Moreover, the definition of this equilibrium relies on the assumptions that  $C$  and  $F$  do not depend on  $x$  and  $G \equiv 0$ . Otherwise, for a given partition  $\Pi$ , the optimal strategy the agent at time  $t_k$  implements in the subperiod  $[t_k, t_{k+1})$  is *semi-Markovian*: the agent's action at time  $s \in [t_k, t_{k+1})$  is a function of  $s$ , the state at  $s$ , and the state at  $t_k$ . As a result, the intra-personal equilibrium with respect to  $\Pi$  is non-Markovian; so we cannot restrict limiting equilibria to be Markov strategies.

## 7 Applications

### 7.1 Present-bias Preferences

Present-biased preferences, also known as hyperbolic discounting, refer to the following observation in intertemporal choice: when considering time preferences between two moments, individuals become more impatient when the two moments are closer to the present time. Thaler (1981) provides an illustrative example of present-biased preferences: some people may prefer an apple today to two apples tomorrow, but very few people would prefer an apple in a year to two apples in a year plus one day. Noted as early as in Strotz (1955-1956), present-biased preferences lead to time inconsistency. For example, consider an agent whose time preferences for having apples are as described in the above illustrative example by Thaler (1981). At time 0, faced with Option A of having one apple at time  $t = 365$  (days) and Option B of having two apples at time  $s = 366$  (days), the agent chooses Option B. When time  $t = 365$  arrives, however, if the agent gets to choose again, she would choose Option A. This shows that the agent in the future will change her actions planned today; hence time-inconsistency is present. For a review of the literature on present-biased preferences, see Frederick et al. (2002).

In a time-separable discounted utility model, present-biased preferences can be modeled by a non-exponential discount function. For example, consider an intertemporal consumption model in continuous time for an agent. The agent's preference value of a random consumption stream  $(C_s)_{s \in [t, T]}$  can be represented as

$$\mathbb{E}_t \left[ \int_t^T h(s-t) u(C_s) ds \right], \quad (28)$$

where  $u$  is the agent's utility function,  $h$  is the agent's discount function, and  $\mathbb{E}_t$  denotes the expectation conditional on all the information available at time  $t$ . To model present-biased preferences, we assume  $h(s+\Delta)/h(s)$  to be *strictly* increasing in  $s \geq 0$  for any fixed  $\Delta > 0$ ; hence it excludes the standard exponential discount function. An example is the generalized hyperbolic discount function proposed by Loewenstein and Prelec (1992):  $h(s) = (1 + \alpha s)^{-\beta/\alpha}$ ,  $s \geq 0$ , where  $\alpha > 0$  and  $\beta > 0$  are two parameters. Ebert et al. (2020) introduce a class of weighted discount functions that is broad enough to include most commonly used non-exponential discount functions in finance and economics.

In various continuous-time settings, Barro (1999), Ekeland and Lazrak (2006), Ekeland and Lazrak (2008), Ekeland and Lazrak (2010), Ekeland and Pirvu (2008), Marín-Solano and Navas (2010), and Ekeland et al. (2012) study intra-personal equilibria for portfolio selection and consumption problems with present-biased preferences. Ebert et al. (2020) and Tan et al. (2018) study real option problems for agents with general weighted discount functions and derive equilibrium investment strategies. Harris and Laibson (2013) and Grenadier and Wang (2007) apply a stochastic, piece-wise step discount function to a consumption problem and a real option problem, respectively, and derive intra-personal equilibrium strategies. Asset pricing for sophisticated agents with present-biased preferences and without commitment has been studied by Luttmer and Mariotti (2003) and Björk et al. (2017).

## 7.2 Mean-Variance

A popular decision criterion in finance is mean–variance, with which an agent minimizes the variance and maximizes the mean of certain random quantity, e.g., the wealth of a portfolio at the end of a period. Any mean–variance model is inherently time inconsistent due to the variance part. To see this, consider a two-period decision problem with dates 0, 1, and 2 for an agent. The agent is offered various options at time 1 that will yield certain payoffs at time 2. The set of options offered to the agent at time 1 depends on the outcome of a fair coin that is tossed between time 0 and 1. If the toss yields a head, the agent is offered two options at time 1: Option H1 that yields \$0 and \$200 with equal probabilities and Option H2 that yields \$50 and \$150 with equal probabilities. If the toss yields a tail, the agent is offered another two options at time 1: Option T1 that yields \$0 and \$200 with equal probabilities and Option T2 that yields \$1050 and \$1150 with equal probabilities. Suppose that at both time 0 and 1, the agent's decision criterion is to minimize the variance of the terminal payoff at time 2. At time 0, the agent has not yet observed the outcome of the toss; so she will need to make choices contingent on this outcome, i.e., she chooses between the following four plans: (H1,T1), (H1,T2), (H2,T1), and (H2,T2), where the first and second components of each of the above four plans stand for the agent's planned choice when the toss yields a head and a tail, respectively. Straight-forward calculation shows that the plan (H2,T1) yields the smallest variance of the terminal payoff; so at time 0 the agent plans to choose H2 when the toss yields a

head and choose T1 when the toss yields a tail. At time 1, after having observed the outcome of the toss, if the agent can choose again with the objective of minimizing the variance of the terminal payoff, she would choose H2 if the outcome is a head and T2 if the outcome is a tail. Consequently, what the agent plans at time 0 is different from what is optimal for the agent at time 1, resulting in time inconsistency.

The reason of having time inconsistency above can be seen from the following conditional variance formula:  $\text{var}(X) = \mathbb{E}[\text{var}(X|Y)] + \text{var}(\mathbb{E}[X|Y])$ , where  $X$  stands for the terminal payoff and  $Y$  denotes the outcome of the coin toss. At time 0, the agent's objective is to maximize  $\text{var}(X)$  and at time 1, her objective is to maximize  $\text{var}(X|Y)$ . Although the plan (H2,T2) yields small variance of  $X$  given the outcome of the toss  $Y$  and thus a small value of the average conditional variance  $\mathbb{E}[\text{var}(X|Y)]$ , it yields very different expected payoffs conditional on having a head and on having a tail, leading to a large value of  $\text{var}(\mathbb{E}[X|Y])$ . Consequently,  $\text{var}(X)$  under plan (H2,T2) is larger than under plan (H2,T1), which yields a larger value of  $\mathbb{E}[\text{var}(X|Y)]$  than the former but a much smaller value of  $\text{var}(\mathbb{E}[X|Y])$ . Consequently, (H2,T1) is preferred to (H2,T2) for the agent at time 0.

A lot of recent works study intra-personal equilibrium investment strategies for agents with mean-variance preferences. For continuous-time models, see for instance Basak and Chabakauri (2010), Björk et al. (2014), Pun (2018), Bensoussan et al. (2014), Cui et al. (2016), Sun et al. (2016), Landriault et al. (2018), Bensoussan et al. (2019), Kryger et al. (2020), and Han et al. (2021). In all these works, the mean-variance criterion is formulated as a weighted average of the mean and variance of wealth at a terminal time, i.e., at each time  $t$ , the agent's objective is to maximize  $\mathbb{E}_t[X] - \frac{\gamma}{2} \text{var}_t(X)$ , where  $\mathbb{E}_t$  and  $\text{var}_t$  stand for the conditional mean and variance of the terminal wealth  $X$ , respectively, and  $\gamma$  is a risk aversion parameter. Alternatively, He and Jiang (2020b) and He and Jiang (2020a) study intra-personal equilibria for mean-variance investors in a constrained formulation: at each time, an investor minimizes the variance of terminal wealth with a target constraint of the expected terminal wealth. Dai et al. (2021) consider a mean-variance model for log returns. Hu et al. (2012), Hu et al. (2017), Czichowsky (2013), and Yan and Wong (2020) investigate open-loop intra-personal equilibria for mean-variance portfolio selection problems. For equilibrium mean-variance insurance strategies, see for instance Zeng and Li (2011), Li et al. (2012), Zeng et al. (2013), Liang and Song (2015), and Bi and Cai (2019).

### 7.3 Non-EUT Preferences

There is abundant empirical and experimental evidence showing that when making choices under uncertainty, individuals do not maximize expected utility (EU); see for instance a survey by Starmer (2000). Various alternatives to the EU model, which are generally referred to as *non-EU* models, have been proposed in the literature. Some of these models employ probability weighting functions to describe the tendency of overweighting extreme outcomes that occur with small probabilities,



examples being prospect theory (PT) (Kahneman and Tversky, 1979, Tversky and Kahneman, 1992) and rank-dependent utility (RDU) theory (Quiggin, 1982).

It has been noted that when applied to dynamic choice problems, non-EU models can lead to time inconsistency; see Machina (1989) for a review of early works discussing this issue. For illustration, consider a casino gambling problem studied by Barberis (2012): a gambler is offered 10 independent bets with equal probabilities of winning and losing \$1, plays these bets sequentially, and decides when to stop playing. Suppose at each time, the gambler's objective is to maximize the preference value of the payoff at end of the game and the preferences are represented by a non-EU model involving a probability weighting function. We represent the cumulative payoff of playing the bets by a binomial tree with up and down movements standing for winning and losing, respectively. At time 0, the top most state (TMS) of the tree at  $t = 10$  represents the largest possible payoff achievable and the probability of reaching this state is extremely small ( $2^{-10}$ ). The gambler overweighs this state due to probability weighting and aspires to reach it. Hence, at time 0, her plan is to play the 10-th bet if and when she has won all the previous 9 bets. Now, suppose she has played and indeed won the first 9 bets. If she has a chance to re-consider her decision of whether to play the 10-th bet at *that* time, she may find it no longer favorable to play because the probability of reaching the TMS at time 10 is  $1/2$  and thus this state is not overweighed. Consequently, when deciding whether to play the 10-th bet conditioning on she has won the first 9 bets, the gambler may choose differently when she is at time 0 and when she is at time 9, showing time inconsistency.

In a continuous-time, complete market, Hu et al. (2021) study a portfolio selection problem in which an agent maximizes the following RDU of her wealth  $X$  at a terminal time:

$$\int_{\mathbb{R}} u(x)w(1 - F_X(x)), \quad (29)$$

where  $u$  is a utility function,  $w$  is a probability weighting function, and  $F_X$  is the cumulative distribution function of  $X$ . The authors derive an open-loop intra-personal equilibrium and show that it is in the same form as in the classical Merton model but with a properly scaled market price of risk. He et al. (2020) consider median and quantile maximization for portfolio selection, where the objective function, namely the quantile of the terminal wealth, can be regarded as a special case of RDU with a particular probability weighting function  $w$ . The authors study closed-loop intra-personal equilibrium and find that an affine trading strategy is an equilibrium if and only if it is a portfolio insurance strategy. Ebert and Strack (2017) consider the optimal time to stop a diffusion process with the objective to maximize the value of the process at the stopping time under a PT model. Using the notion of mild intra-personal equilibrium as previously discussed in Section 5, the authors show that under reasonable assumptions on the probability weighting functions, the only equilibrium among all two-threshold stopping rules is to immediately stop. Huang et al. (2020) study mild intra-personal equilibrium stopping rules for an agent who wants to stop a geometric Brownian motion with the objective of maximizing the RDU value at the stopping time.

Risk measures, such as value-at-risk (VaR) and conditional value-at-risk (CVaR), can also be considered to be non-EU models leading to time consistency. There are, however, few studies on intra-personal equilibria for mean-risk models in continuous time. For relevant studies in discrete-time settings, see for instance Cui et al. (2019).

Models with Knightian uncertainty or ambiguity can also result in time inconsistency. For example, the  $\alpha$ -maxmin model proposed by Ghirardato et al. (2004) is dynamically inconsistent in general; see for instance Beissner et al. (2020). Li et al. (2019) find an open-loop intra-personal equilibrium investment strategy for an agent with  $\alpha$ -maximin preferences. Huang and Yu (2021) consider a problem of stopping a one-dimensional diffusion process with preferences represented by the  $\alpha$ -maxmin model and study the mild intra-personal equilibrium stopping rule for the problem.

## 8 Dynamically Consistent Preferences

Machina (1989) notes that, in many discussions of time inconsistency in the literature, a hidden assumption is *consequentialism*: at any intermediate time  $t$  of a dynamic decision process, the agent employs the *same* preference model as used at the initial time to evaluate the choices in the *continuation* of the dynamic decision process from time  $t$ , conditional on the circumstances at time  $t$ . For example, consider a dynamic consumption problem for an agent with present-bias preferences and suppose that at the initial time 0, the agent's preference value for a consumption stream  $(C_s)_{s \geq 0}$  is represented by  $\mathbb{E}[\int_0^\infty h(s)u(C_s)ds]$ , where the discount function  $h$  models the agent's time preferences at the initial time 0 and  $u$  is the agent's utility function. The consequentialism assumption implies that at any intermediate time  $t$ , the agent's preferences for the continuation of the consumption stream, i.e.,  $(C_s)_{s \geq t}$ , are represented by the same preference model as at the initial time 0, conditional on the situations at time  $t$ , i.e., by  $\mathbb{E}_t[\int_t^\infty h(s-t)u(C_s)ds]$ , where the discount function  $h$  and  $u$  are the same as the ones in the preference model at the initial time 0. Similarly, for a dynamic choice problem with RDU preferences for the payoff at a terminal time, the consequentialism assumption stipulates that the agent uses the same utility function  $u$  and probability weighting function  $w$  at all intermediate times  $t$  when evaluating the terminal payoff at those times.

The consequentialism assumption, however, has not been broadly validated because there are few experimental or empirical studies on how individuals dynamically update their preferences. Machina (1989) consider a class of non-EU maximizers, referred to as  $\gamma$ -people, who adjust their preferences dynamically over time so as to remain time consistent. The idea in Machina (1989) was further developed by Karnam et al. (2017) who propose the notion of *time-consistent dynamic preference models*. The idea of considering time-consistent dynamic preferences is also central in the theory of forward performance criteria proposed and developed by Musiela and Zariphopoulou (2006, 2008, 2009, 2010a,b, 2011); see also He et al. (2021) for a related discussion.

Formally, consider a dynamic choice problem in a period  $[0, T]$ . A preference model at time 0 is specified for an agent, denoted as  $J_0(u(\cdot))$ , where  $(u(s))_{s \in [0, T]}$  denotes the agent's dynamic choice. A family of dynamic preference models  $J_t$ ,  $t \in (0, T)$ , are called time-consistent for the initial model  $J_0$  if the optimal strategy under  $J_0$ , namely, the pre-committed strategy for the agent at time 0, is also optimal under  $J_t$  for the agent at any future time  $t \in (0, T)$ . Note that given the pre-committed strategy at time 0, we can always find preference models at  $t > 0$  such that this strategy remains optimal. Thus, a more interesting question is whether we can find a family of time-consistent dynamic preference models that are of the same type as the initial preference model.

He et al. (2021) study portfolio selection in the Black-Scholes market for an agent whose initial preference model for wealth at a terminal time is represented by RDU. The authors show that there exists a family of time-consistent dynamic RDU models if and only if (i) the probability weighting function in the initial model belongs to a parametric class of functions proposed by Wang (1996); and (ii) the parameter of the probability weighting function, the absolute risk aversion index of the utility function, and the market price of risk must be coordinated with each other over time in a specific way. Cui et al. (2012), Karnam et al. (2017), and He and Jiang (2020a) find that mean-variance models become time consistent if the dynamic trade-off between the mean and variance over time is set properly. For mean-CVaR models, where an agent maximizes the mean and minimizes the CVaR at certain confidence level, Pflug and Pichler (2016) and Strub et al. (2019) note, in discrete-time settings, that time consistency is retained as long as the tradeoff between the mean and CVaR and the confidence level evolve dynamically in a certain way.

The problem of intra-personal equilibria and that of dynamically consistent preferences can be considered primal-dual to each other: the former finds equilibrium strategies given the time-inconsistent preferences, whereas the latter identifies preferences given the problem is time-consistent. Diving deeper into this relationship may call for innovative mathematical analysis and result in profound economic insights.

**Acknowledgements** Xue Dong He gratefully acknowledges financial support through the General Research Fund of the Research Grants Council of Hong Kong SAR (Project No. 14200917). Xun Yu Zhou gratefully acknowledges financial support through a start-up grant at Columbia University and through the Nie Center for Intelligent Asset Management. This author also thanks the hospitality of The Chinese University of Hong Kong during the summer of 2020 when the present project started.

## References

- Barberis, N. (2012). A model of casino gambling, *Management Science* **58**(1): 35–51.
- Barro, R. J. (1999). Ramsey meets Laibson in the neoclassical growth model, *Quarterly Journal of Economics* **114**(4): 1125–1152.

- Basak, S. and Chabakauri, G. (2010). Dynamic mean-variance asset allocation, *Review of Financial Studies* **23**(8): 2970–3016.
- Bayraktar, E., Zhang, J. and Zhou, Z. (2021). Equilibrium concepts for time-inconsistent stopping problems in continuous time, *Mathematical Finance* **31**: 508–530.
- Beissner, P., Lin, Q. and Riedel, F. (2020). Dynamically consistent alpha-maxmin expected utility, *Mathematical Finance* **30**(3): 1073–1102.
- Bensoussan, A., Wong, K. C. and Yam, S. C. P. (2019). A paradox in time-consistency in the mean–variance problem, *Finance and Stochastics* **23**(1): 173–207.
- Bensoussan, A., Wong, K. C., Yam, S. C. P. and Yung, S.-P. (2014). Time-consistent portfolio selection under short-selling prohibition: From discrete to continuous setting, *SIAM Journal on Financial Mathematics* **5**(1): 153–190.
- Bertsekas, D. P. (2017). *Dynamic programming and optimal control*, 4 edn, Athena Scientific, Nashua.
- Bi, J. and Cai, J. (2019). Optimal investment-reinsurance strategies with state dependent risk aversion and VaR constraints in correlated markets, *Insurance: Mathematics and Economics* **85**: 1–14.
- Björk, T., Khapko, M. and Murgoci, A. (2017). On time-inconsistent stochastic control in continuous time, *Finance and Stochastics* **21**(2): 331–360.
- Björk, T. and Murgoci, A. (2010). A general theory of Markovian time inconsistent stochastic control problems. SSRN:1694759.
- Björk, T. and Murgoci, A. (2014). A theory of Markovian time-inconsistent stochastic control in discrete time, *Finance and Stochastics* **18**(3): 545–592.
- Björk, T., Murgoci, A. and Zhou, X. Y. (2014). Mean-variance portfolio optimization with state dependent risk aversion, *Mathematical Finance* **24**(1): 1–24.
- Bryan, G., Karlan, D. and Nelson, S. (2010). Commitment devices, *Annual Review of Economics* **2**(1): 671–698.
- Cao, D. and Werning, I. (2016). Dynamic savings choices with disagreements, *Technical report*, National Bureau of Economic Research.
- Christensen, S. and Lindensjö, K. (2018). On finding equilibrium stopping times for time-inconsistent markovian problems, *SIAM Journal on Control and Optimization* **56**(6): 4228–4255.
- Christensen, S. and Lindensjö, K. (2020). On time-inconsistent stopping problems and mixed strategy stopping times, *Stochastic Processes and their Applications* **130**(5): 2886–2917.
- Cui, X., Gao, J., Shi, Y. and Zhu, S. (2019). Time-consistent and self-coordination strategies for multi-period mean-conditional value-at-risk portfolio selection, *European Journal of Operational Research* **276**(2): 781–789.
- Cui, X., Li, D., Wang, S. and Zhu, S. (2012). Better than dynamic mean-variance: Time inconsistency and free cash flow stream, *Mathematical Finance* **22**(2): 346–378.
- Cui, X., Xu, L. and Zeng, Y. (2016). Continuous time mean-variance portfolio optimization with piecewise state-dependent risk aversion, *Optimization Letters* **10**(8): 1681–1691.

- Czichowsky, C. (2013). Time-consistent mean-variance portfolio selection in discrete and continuous time, *Finance and Stochastics* **17**(2): 227–271.
- Dai, M., Jin, H., Kou, S. and Xu, Y. (2021). A dynamic mean-variance analysis for log returns, *Management Science* **67**: 661–1328.
- Ebert, S. and Strack, P. (2017). Never, ever getting started: On prospect theory without commitment. SSRN:2765550.
- Ebert, S., Wei, W. and Zhou, X. Y. (2020). Weighted discounting—on group diversity, time-inconsistency, and consequences for investment, *Journal of Economic Theory* **189**: 1–40.
- Ekeland, I. and Lazrak, A. (2006). Being serious about non-commitment: subgame perfect equilibrium in continuous time. arXiv:math/0604264.
- Ekeland, I. and Lazrak, A. (2008). Equilibrium policies when preferences are time inconsistent. arXiv:0808.3790.
- Ekeland, I. and Lazrak, A. (2010). The golden rule when preferences are time inconsistent, *Mathematics and Financial Economics* **4**(1): 29–55.
- Ekeland, I., Mbodji, O. and Pirvu, T. A. (2012). Time-consistent portfolio management, *SIAM Journal on Financial Mathematics* **3**(1): 1–32.
- Ekeland, I. and Pirvu, T. (2008). Investment and consumption without commitment, *Mathematics and Financial Economics* **2**(1): 57–86.
- Frederick, S., Loewenstein, G. and O’donoghue, T. (2002). Time discounting and time preference: A critical review, *Journal of Economic Literature* **40**(2): 351–401.
- Ghirardato, P., Maccheroni, F. and Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory* **118**(2): 133–173.
- Grenadier, S. R. and Wang, N. (2007). Investment under uncertainty and time-inconsistent preferences, *Journal of Financial Economics* **84**(1): 2–39.
- Han, B., Pun, C. S. and Wong, H. Y. (2021). Robust mean-variance portfolio selection with state-dependent ambiguity aversion and risk aversion: A closed-loop approach, *Finance and Stochastics* **forthcoming**.
- Han, B. and Wong, H. Y. (2020). Time-inconsistency with rough volatility. arXiv:1907.11378.
- Harris, C. and Laibson, D. (2013). Instantaneous gratification, *Quarterly Journal of Economics* **128**(1).
- He, X. D. and Jiang, Z. (2019). On the equilibrium strategies for time-inconsistent problems in continuous time. SSRN:3308274.
- He, X. D. and Jiang, Z. (2020a). Dynamic mean-variance efficient fractional kelly portfolios in a stochastic volatility model. SSRN:3670621.
- He, X. D. and Jiang, Z. (2020b). Mean-variance portfolio selection with dynamic targets for expected terminal wealth, *Mathematics of Operations Research* **forthcoming**. SSRN:3084657.
- He, X. D., Jiang, Z. and Kou, S. (2020). Portfolio selection under median and quantile maximization. SSRN:3657661.
- He, X. D., Strub, M. and Zariphopoulou, T. (2021). Forward rank-dependent performance criteria: Time-consistent investment under probability distortion, *Mathematical Finance* **31**: 683–721.

- Hernández, C. and Possamaï, D. (2020). Me, myself and I: a general theory of non-Markovian time-inconsistent stochastic control for sophisticated agents. arXiv:2002.12572.
- Hu, Y., Jin, H. and Zhou, X. Y. (2012). Time-inconsistent stochastic linear-quadratic control, *SIAM Journal on Control and Optimization* **50**: 1548–1572.
- Hu, Y., Jin, H. and Zhou, X. Y. (2017). Time-inconsistent stochastic linear-quadratic control: Characterization and uniqueness of equilibrium, *SIAM Journal on Control and Optimization* **55**(2): 1261–1279.
- Hu, Y., Jin, H. and Zhou, X. Y. (2021). Consistent investment of sophisticated rank-dependent utility agents in continuous time, *Mathematical Finance* **forthcoming**. arXiv:2006.01979.
- Huang, Y.-J. and Nguyen-Huu, A. (2018). Time-consistent stopping under decreasing impatience, *Finance and Stochastics* **22**(1): 69–95.
- Huang, Y.-J., Nguyen-Huu, A. and Zhou, X. Y. (2020). General stopping behaviors of naïve and non-committed sophisticated agents, with application to probability distortion, *Mathematical Finance* **30**: 310–340.
- Huang, Y.-J. and Wang, Z. (2020). Optimal equilibria for multi-dimensional time-inconsistent stopping problems. arXiv:2006.00754.
- Huang, Y.-J. and Yu, X. (2021). Optimal stopping under model ambiguity: a time-consistent equilibrium approach. arXiv:1906.01232.
- Huang, Y.-J. and Zhou, Z. (2019). Strong and weak equilibria for time-inconsistent stochastic control in continuous time, *Mathematics of Operations Research* **forthcoming**. arXiv:1809.09243.
- Huang, Y.-J. and Zhou, Z. (2020). Optimal equilibria for time-inconsistent stopping problems in continuous time, *Mathematical Finance* **30**(3): 1103–1134.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk, *Econometrica* **47**(2): 263–291.
- Karnam, C., Ma, J., Zhang, J. et al. (2017). Dynamic approaches for some time-inconsistent optimization problems, *Annals of Applied Probability* **27**(6): 3435–3477.
- Kryger, E., Nordfang, M.-B. and Steffensen, M. (2020). Optimal control of an objective functional with non-linearity between the conditional expectations: solutions to a class of time-inconsistent portfolio problems, *Mathematical Methods of Operations Research* **91**(3): 405–438.
- Landriault, D., Li, B., Li, D. and Young, V. R. (2018). Equilibrium strategies for the mean-variance investment problem over a random horizon, *SIAM Journal on Financial Mathematics* **9**(3): 1046–1073.
- Li, B., Luo, P. and Xiong, D. (2019). Equilibrium strategies for alpha-maxmin expected utility maximization, *SIAM Journal on Financial Mathematics* **10**(2): 394–429.
- Li, Z., Zeng, Y. and Lai, Y. (2012). Optimal time-consistent investment and reinsurance strategies for insurers under hestons sv model, *Insurance: Mathematics and Economics* **51**(1): 191–203.

- Liang, Z. and Song, M. (2015). Time-consistent reinsurance and investment strategies for mean–variance insurer under partial information, *Insurance: Mathematics and Economics* **65**: 66–76.
- Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation, *Quarterly Journal of Economics* **107**(2): 573–597.
- Luttmer, E. G. and Mariotti, T. (2003). Subjective discounting in an exchange economy, *Journal of Political Economy* **111**(5): 959–989.
- Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty, *Journal of Economic Literature* **27**(4): 1622–1668.
- Marín-Solano, J. and Navas, J. (2010). Consumption and portfolio rules for time-inconsistent investors, *European Journal of Operational Research* **201**(3): 860–872.
- Mei, H. and Yong, J. (2019). Equilibrium strategies for time-inconsistent stochastic switching systems, *ESAIM: Control, Optimisation and Calculus of Variations* **25**: 64.
- Musiela, M. and Zariphopoulou, T. (2006). Investments and forward utilities. <https://www.ma.utexas.edu/users/zariphop/pdfs/TZ-TechnicalReport-4.pdf>.
- Musiela, M. and Zariphopoulou, T. (2008). Optimal asset allocation under forward exponential performance criteria, in S. N. Ethier, J. Feng and R. H. Stockbridge (eds), *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*, Vol. 4, Institute of Mathematical Statistics, pp. 285–300.
- Musiela, M. and Zariphopoulou, T. (2009). Portfolio choice under dynamic investment performance criteria, *Quantitative Finance* **9**(2): 161–170.
- Musiela, M. and Zariphopoulou, T. (2010a). Portfolio choice under space-time monotone performance criteria, *SIAM Journal on Financial Mathematics* **1**(1): 326–365.
- Musiela, M. and Zariphopoulou, T. (2010b). Stochastic partial differential equation and portfolio choice, in C. Chiarella and A. Novikov (eds), *Contemporary Quantitative Finance: Essays in Honour of Eckhard Platen*, Springer, pp. 195–216.
- Musiela, M. and Zariphopoulou, T. (2011). Initial investment choice and optimal future allocations under time-monotone performance criteria, *International Journal of Theoretical and Applied Finance* **14**(1): 61–81.
- O’Donoghue, T. and Rabin, M. (1999). Doing it now or later, *American Economic Review* **89**(1): 103–124.
- Peleg, B. and Yaari, M. E. (1973). On the existence of a consistent course of action when tastes are changing, *Review of Economic Studies* **40**(3): 391–401.
- Pflug, G. C. and Pichler, A. (2016). Time-inconsistent multistage stochastic programs: Martingale bounds, *European Journal of Operational Research* **249**(1): 155–163.
- Pollak, R. A. (1968). Consistent planning, *Review of Economic Studies* **35**(2): 201–208.
- Pun, C. S. (2018). Time-consistent mean-variance portfolio selection with only risky assets, *Economic Modelling* **75**: 281–292.
- Quiggin, J. (1982). A theory of anticipated utility, *Journal of Economic Behavior and Organization* **3**: 323–343.

- Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk, *Journal of Economic Literature* **38**(2): 332–382.
- Strotz, R. H. (1955–1956). Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies* **23**(3): 165–180.
- Strub, M., Li, D., Cui, X. and Gao, J. (2019). Discrete-time mean-cvar portfolio selection and time-consistency induced term structure of the cvar, *Journal of Economic Dynamics and Control* **108**(103751).
- Sun, J., Li, Z. and Zeng, Y. (2016). Precommitment and equilibrium investment strategies for defined contribution pension plans under a jump-diffusion model, *Insurance: Mathematics and Economics* **67**: 158–172.
- Tan, K. S., Wei, W. and Zhou, X. Y. (2018). Failure of smooth pasting principle and nonexistence of equilibrium stopping rules under time-inconsistency. arXiv:1807.01785.
- Thaler, R. (1981). Some empirical evidence on dynamic inconsistency, *Economics Letters* **8**(3): 201–207.
- Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* **5**(4): 297–323.
- Wang, H. and Yong, J. (2019). Time-inconsistent stochastic optimal control problems and backward stochastic Volterra integral equations. arXiv:1911.04995.
- Wang, S. (1996). Premium calculation by transforming the layer premium density, *ASTIN Bulletin* **26**(1): 71–92.
- Wei, Q., Yong, J. and Yu, Z. (2017). Time-inconsistent recursive stochastic optimal control problems, *SIAM Journal on Control and Optimization* **55**(6): 4156–4201.
- Yan, T. and Wong, H. Y. (2020). Open-loop equilibrium reinsurance-investment strategy under mean–variance criterion with stochastic volatility, *Insurance: Mathematics and Economics* **90**: 105–119.
- Yong, J. (2012). Time-inconsistent optimal control problems and the equilibrium HJB equation, *Mathematical Control and Related Fields* **2**(3): 271–329.
- Yong, J. and Zhou, X. Y. (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer.
- Zeng, Y. and Li, Z. (2011). Optimal time-consistent investment and reinsurance policies for mean-variance insurers, *Insurance: Mathematics and Economics* **49**(1): 145–154.
- Zeng, Y., Li, Z. and Lai, Y. (2013). Time-consistent investment and reinsurance strategies for mean–variance insurers with jumps, *Insurance: Mathematics and Economics* **52**(3): 498–507.