Resolving Time Inconsistency of Decision Problem with Non-expectation Operator: From Internal Conflict to Internal Harmony by Strategy of Self-Coordination

XIANGYU CUI ^{*}, DUAN LI [†] AND YUN SHI [‡]

Abstract

When a stochastic decision problem is time inconsistent, the decision maker would be puzzled by his conflicting decisions optimally derived from his time-varying preferences at different time instants (with different time horizons). While the long-run self (LR) of the decision maker pursues the long-term optimality, the short-run selves (SRs) of the decision maker at different time instants bow to short-term temptations. While the literature began to recognize the importance to strike a balance between LR's and SRs' interests, the existing results are not applicable to situations where the decision maker's preferences involve non-expectation operators. We propose an operable unified two-tier dual-self game model with commitment by punishment, which can cope with general time inconsistent stochastic decision problems with both expectation and non-expectation operators in the objective function. By attaching punishment terms to both the preferences of LR and SRs which quantitatively evaluate the internal conflict among different selves, our game model aligns the interests of the LR and SRs to a certain degree. The equilibrium strategy, termed strategy of self-coordination, achieves some degree of internal harmony among various selves. We successfully apply the model to the investment

^{*}School of Statistics and Management, Shanghai Institute of International Finance and Economics, Shanghai University of Finance and Economics.

[†]School of Data Science, City University of Hong Kong.

[‡]Institute for Statistics and Interdisciplinary Sciences, Faculty of Economics and Management, East China Normal University.

and consumption problem with quasi-hyperbolic discounting and the dynamic mean-variance portfolio selection problem.

Key Words: time inconsistency, dual-self game model, commitment by punishment, strategy of self-coordination

1 Introduction

When a decision maker faces a dynamic stochastic decision problem, his preferences at different time instants for the corresponding tail parts of the time horizon could be time-varying and/or state dependent. Such kind of inconsistency among the preferences at different time instants would yield conflicting decisions optimally derived for different time horizons. This phenomena is termed *time inconsistency* in the literature. Mathematically, when the preferences at different time instants do not satisfy the *Bellman's principle of optimality* of dynamic programming, the global preference does not conform with local preferences, which yields conflict among long term optimal decision and short term optimal decisions. In such situations, the decision maker would be always puzzled by his paradoxical behavior patterns induced by short-term temptations.

Dynamic investment and consumption problem with quasi-hyperbolic discounting and dynamic mean-variance portfolio selection problem are two salient time inconsistent stochastic decision problems in management science and finance.¹ For dynamic investment and consumption problem with quasi-hyperbolic discounting, the decision maker's time preference, which represents the value at time t of \$1 received at future time s , is described by

$$D_i(t, s) = \begin{cases} e^{-r(s-t)}, & \text{if } s \in [t_i, t_{i+1}), \\ \beta e^{-r(s-t)}, & \text{if } s \in [t_{i+1}, +\infty), \end{cases}$$

where $t_i \leq t < t_{i+1}$ and the short-run discounting parameter $\beta < 1$ (see Grenadier and Wang, 2007; Harris and Laibson, 2013).² The decision maker tends to underestimate the value of payoff

¹More real life and economical time inconsistent problems can be found in Thaler (1981), Gul and Pesendorfer (2001), Grenadier and Wang (2007)

²Extensions of quasi-hyperbolic discounting can be found in Bleichrodt et al. (2009), Takeuchi (2011), Pan et al. (2015).

in far future and his preference at time t is the expected sum of discounted utilities of the future consumptions and the terminal wealth,

$$\mathbb{E}_{t,x_t} \left[\int_t^T D_i(t,s) U(c(s)) ds \right] + D_i(t,T) \mathbb{E}_{t,x_t} [U(X^{u,c}(T))].$$

Due to the existence of the short-run discounting parameter β , the preferences at different time instants are inconsistent and do not satisfy the Bellman's principle of optimality. After the decision maker comes up with his long term optimal investment and consumption plan at time 0, he may switch his mind at a later time t and prefer to consume more than his original plan to maximize his short term preference (see Thaler and Shefrin, 1981). The smaller the parameter β , the larger the conflict between the long term optimal plan and short term optimal plans.

For dynamic mean-variance portfolio selection problem, the decision maker's preference at time t is the weighted sum of the conditional expected value and the conditional variance of the terminal wealth,

$$\mathbb{E}_{t,x_t} [X^u(T)] - \lambda \text{Var}_{t,x_t} (X^u(T)),$$

where $\lambda > 0$ is the risk aversion parameter. As the variance term does not satisfy the smooth property, i.e., $\text{Var}_{t,x_t} (X^u(T)) \neq \text{Var}_{t,x_t} (\text{Var}_{s,x_s} (X^u(T)))$ for $s > t$, the preferences at different time instants would be definitely inconsistent. After a mean-variance investor derives his long term optimal investment strategy at time 0, he could be tempted to adopt a different strategy at a later time t in order to achieve a short-term mean-variance efficiency (see Basak and Chabakauri, 2010, Cui et al., 2012). One significant difference between the above two problems is that the objectives of the dynamic investment and consumption problem with quasi-hyperbolic discounting at different time instants only involve the expectation operator (which can be represented as the expected sum of the future discounted performance measures), while the objectives of the dynamic mean-variance portfolio selection problem at different time instants involve a non-expectation operator: Variance. Almost all widely adopted risk measures involving non-expectation operators in the literature, including VaR (Value-at-Risk) and CVaR (Conditional Value-at-risk), are not time consistent (see Boda and Filar (2006), Artzner et al. (2007)).

Using the terms in the dual-self model of Fudenberg and Levine (2006), we name the decision maker’s abstract desire to maximize the long term preference as Long-run self (LR), who possesses foresight, and name the decision maker’s abstract desire to maximize short term preferences at different time instants as Short-run selves (SRs), who lack foresight and bow to short term benefits and temptations. The essence of time inconsistency is the internal conflict among long-run self and short-run selves.

The first approach to deal with the conflict among LR and SRs proposed in Strotz (1956) is to completely ignore SRs’ incentives and strictly adhere to the long term optimal decision strategy of LR over the entire time horizon, which is called *pre-committed strategy* in this paper.³ Although the pre-committed strategy achieves the long term optimality, it is hard for any decision maker to adhere to when time goes. To ensure the successful implementation of the pre-committed strategy, the decision maker needs to eliminate the effect of short term temptations, for example, by making the pre-committed strategy the only feasible or economically reasonable strategy via strictly self-control commitment or external contractual commitment. The Save More Tomorrow plan of 401(k) is such an example in the reality that forces the employee to save more before retirement through an external contractual scheme (see Thaler and Benartzi, 2004).

The second approach to cope with the conflict among LR and SRs proposed in Strotz (1956) is to completely abandon LR’s incentive and assume that the SR at time t is *sophisticated*, who is aware of the time inconsistency but unable to find a way of committing to his initial plan. Then, the dynamic decision problem can be reformulated as a sequential game among SRs, in which SR at time t optimally chooses the equilibrium strategy on the premise that the future SRs have already decided their equilibrium strategies in a similar fashion. The subgame perfect Nash equilibrium strategy is the carried out strategy of the decision maker, which is called *time consistent strategy* in this paper.⁴ Although the time consistent strategy achieves equilibrium among SRs, it could lead

³It is termed “strategy of pre commitment” in Strotz (1956), and “pre-committed strategy” in Basak and Chabakauri (2010) and Cui et al. (2012).

⁴It is termed “strategy of consistent planning” in Strotz (1956), and “time consistent strategy” in Basak and Chabakauri (2010) and Björk et al. (2014). A decision maker with quasi-hyperbolic discounting who adopts time

to a bad long term performance.

The third approach to resolve the conflict among LR and SRs is to strike a balance between LR's and SRs' incentives. Several successful models have been reported in the last decade in the literature along this direction. O'Donoghue and Rabin (1999, 2001), DellaVigna and Malmendier (2004, 2006) suggested that the decision maker with quasi-hyperbolic discounting may be "partially naive", who believes that he has a larger short-run discounting parameter $\hat{\beta}(> \beta)$ in the future. Gul and Pesendorfer (2001, 2004) proposed that the decision maker could construct rational and stable "preferences with self-control", which reconciles LR and SRs' incentives. Fudenberg and Levine (2006, 2012) proposed that the LR can influence the SRs' preferences via some self-control schemes, and the decision maker adopts the "SR-Perfect equilibrium strategy", which is the equilibrium strategy among LR and SRs.⁵ Benhabib and Bisin (2005) studied a new consumption-saving decision model, which allows for internal commitment mechanisms and self-control by introducing an additional control process during the decision. Although these models are different in many aspects, they all implicitly require some type of self-control behavior of the decision maker to resist local temptations and to restrain the deviations from the long term optimal strategy. The major deficiency of the current models is that they all require the preferences at different time instants be myopic (only depending on current state and current decision) and/or involve only expectation operator in the objective functions. While this class of strategies is called *strategy of self-control* in the literature, we name it *strategy of self-coordination* in this paper. Adopting the term of self-control implicitly implies a "correct" attitude to achieve the best possible long-term interest by resisting short-term temptations to some extent. Essentially, we almost have no base to judge which one of these paradoxical behavior patterns is the correct one. We thus use the term of self-coordination to reflect a necessary self-adjust process to strike a balance among conflicting interests at different time instants. The second reason to introduce the term of self-coordination is to distinguish our approach which is applicable to non-expectation operators from the self-control consistent strategy is called "sophisticated" decision maker in O'Donoghue and Rabin (1999, 2001), DellaVigna and Malmendier (2004, 2006) and Grenadier and Wang (2007).

⁵Please refer to Rachlin (2004) for more detailed discussion on self-control schemes.

approach in the literature.

Stochastic decision problems often involve risk measures which are mostly built up with a non-expectation operator. The dynamic mean-variance portfolio selection problem is indeed such a problem. Unfortunately, the existing framework of strategy of self-coordination is unable to be applied to problems with a non-expectation operator. For example, almost all studies in tackling the time inconsistency issue in the dynamic mean-variance portfolio selection problem are confined to the class of time consistent strategies (see Basak and Chabakauri, 2010, Wang and Forsyth, 2011, Björk et al., 2014, Cui et al., 2015).

In this paper, we extend the concept of the dual-self model in Fudenberg and Levine (2006, 2012) and develop an operable unified two-tier dual-self game model for generalized time inconsistent decision problems which could involve non-expectation operator. The proposed game model involves a particular self-control scheme, i.e., commitment by punishment. Once the carried out strategy of a SR is different from the LR's announce strategy, a type of punishment committed is imposed on the decision maker (i.e., on both LR and SRs). By attaching punishment terms to the preferences of LR and SRs, the unified two-tier dual-self game model aligns the interests of the long-run and short-run selves. Most importantly, we show that, when some mild conditions are satisfied, the two-tier dual-self game model can be reduced to a static optimization problem. The derived strategy of self-coordination (i.e., equilibrium strategy) enables the decision maker to achieve some degree of internal harmony among various selves. Following a general framework with its important properties revealed, we successfully apply the proposed model to dynamic investment and consumption problem with quasi-hyperbolic discounting and dynamic mean-variance portfolio selection problem, and derive their strategies of self-coordination. We believe that our work represents the first result on the strategy of self-coordination for time inconsistent decision problems with risk measure of a non-expectation type, including the continuous-time mean-variance portfolio selection problem. Our work pushes forward the study on resolving the long-standing issue of time inconsistency.

The remaining of this paper is organized as follows. In Section 2, we propose a unified two-tier dual-self game model with commitment by punishment, establish the required axiom system and

study its important properties. In Section 3, we apply the proposed game model to continuous-time investment and consumption problem with quasi-hyperbolic discounting, and derive its strategy of self-coordination. In Section 4, we show the proposed game model is also applicable to continuous-time mean-variance portfolio selection problem, and derive an explicit strategy of self-coordination. Finally, we conclude our paper in Section 5. To clearly present the main argument and key results, we have provided all the proofs in the online appendix.

2 Two-Tier Dual-Self Game Model

We consider a controlled Markovian process X^u in continuous time within interval $[0, T]$, which is defined on a filtered probability space $(\Omega, \mathcal{F}_T, \{\mathcal{F}_t\}, P)$. The process X^u lives on a measurable state space $\{\mathcal{X}, \mathcal{G}_X\}$, with control strategy taking values in a measurable control space $\{\mathcal{U}, \mathcal{G}_U\}$. For a strategy u , we denote by $\mathbf{A}^u$ the corresponding infinitesimal generator, which describes how strategy u influences the dynamics of the process. We use $\mathcal{L}^2(\mathcal{F}_t, P)$ to denote the set of all $\mathcal{F}_t$ -measurable square integrable random variables.

At time t , after knowing $X_t^u = x$, the decision maker wants to maximize the following preference over time horizon $[t, T]$,

$$J(t, x, u(t, x), u_f(t, x)) = \mathbb{E}_{t,x} \left[\int_t^T F(t, x, s, X_s^u, u(s, X_s^u)) ds \right] + H(t, x, T, X_T^u),$$

where $F : [0, T] \times \mathcal{X} \times [0, T] \times \mathcal{L}^2(\mathcal{F}_s, P) \times \mathcal{L}^2(\mathcal{F}_s, P) \rightarrow \mathcal{L}^2(\mathcal{F}_s, P)$, $H : [0, T] \times \mathcal{X} \times [0, T] \times \mathcal{L}^2(\mathcal{F}_T, P) \rightarrow \mathbb{R}$, $u(t, x)$ is the strategy at time t , $u_f(t, x) = \{u(s, X_s^u)\}_{s>t}^T$ is the set of future strategies and $u_f(T, x)$ is empty set. In general, these preferences at different time instants do not satisfy Bellman's principle of optimality due to a t -dependence in the F and H terms, and an unrestricted generality of the H term (e.g., the H term may involve a non-expectation operator), which makes the stochastic decision problem under consideration time inconsistent.

We call the abstract desire of the decision maker to achieve optimality over time horizon $[t, T]$ as self t , who desires to achieve the best performance over $[t, T]$ by solving the following stochastic

decision problem,

$$(SC_t) \quad \max_{\{u(s, X_s^u)\}_{|s \geq t}^T} J(t, x, u(t, x), u_f(t, x)).$$

We further classify self ts into two classes. When t is zero, self 0 is actually the *long-run self* (LR), who tries to achieve the best long term performance over $[0, T]$ by solving the following stochastic decision problem,

$$(SC_0) \quad \max_{\{u(t, X_t^u)\}_{|t \geq 0}^T} J(0, x, u(0, x), u_f(0, x)).$$

We denote the optimal solution to (SC_0) by $\{u^{Com}(t, X_t^u)\}_{|t \geq 0}^T$, which is the *pre-committed strategy*.

On the other side, we call all self ts with $t > 0$ as *short-run selves* (SRs) due to their limited time span on $[t, T]$. While we often view SRs as short-sighted decision makers who bow to short term temptation, we view the LR with a global interest over the entire time horizon as a willpower to resist such temptations. Note that the strategy at time s will be different in general when it is derived either from (SC_{t_1}) or from (SC_{t_2}) with $t_1 \leq s, t_2 \leq s$ and $t_1 \neq t_2$.

When all self ts come to make decision together, then self t can only make his decision on $u(t, x)$, while future decision $u_f(t, x)$ is determined by future selves. Thus, self t will choose his strategy on the premise that the future selves have already decided their strategies, resulting in the following sequential game,

$$\begin{aligned} (NSC_t) \quad & \max_{u(t, x)} J(t, x, u(t, x), u_f^{TC}(t, x)) \\ \text{s.t.} \quad & u^{TC}(s, X_s^u) \text{ solves } (NSC_s), \quad t < s \leq T, \end{aligned}$$

where $u_f^{TC}(t, x)$ is the set of future time consistent strategies and $u_f^{TC}(T, x)$ is an empty set. The *time consistent strategy*, $\{u^{TC}(t, X_t^u)\}_{|t \geq 0}^T$, is the subgame perfect Nash equilibrium strategy of the above game, which completely ignores the LR's interests over the entire time span $[0, T]$. The conflicting interests among different SRs and between the LR and SRs are the source of time inconsistency inherent in the generalized stochastic decision problem under consideration.

We now propose a two-tier dual-self game model for this general class of time inconsistent stochastic decision problems. We first quantify the degree of conflict by a family of distance measures for the difference between the strategies of the LR and SR at different t . Our proposed game model aims to induce a coordination process among LR and all SRs such that the degree of conflict

is under control, which can be achieved by enforcing a self-control process through commitment by punishment. More specifically, the LR decides a planned strategy denoted by $\{\nu(t, X_t^u)\}_{t \geq 0}^T$, announces it to the SRs, and commits that any derivation by a SR from the planned control strategy leads the SR himself a punishment. At the same time, the total amount of punishment applied by the LR will in turn penalize the LR himself. Under this dual-punishment scheme, LR and all SRs will revise their decision patterns accordingly and achieve some degree of internal harmony.

Under such a punishment scheme, the modified preference of SR at time t would become the following,

$$\begin{aligned} & \tilde{J}(t, x, u(t, x), \nu(t, x), u_f(t, x)) \\ &= J(t, x, u(t, x), u_f(t, x)) - G(t, x, u(t, x), \nu(t, x), u_f(t, x)), \end{aligned}$$

where $G(t, x, u(t, x), \nu(t, x), u_f(t, x))$ is a concrete mathematical representation of commitment by the imposed punishment, i.e., the punishment inflicted if the SR at time t does not follow the planned strategy $\nu(t, x)$. $G(t, \cdot, \cdot, \cdot, \cdot)$ is also a quantitative representation of internal conflict between LR and SR at time t . This type of punishment term has been widely witnessed in many real world applications, such as in disruption management (see Yu and Qi, 2004) where the punishment term is quantitatively represented by a distance measure between the schedules before the disruption and after. As the new preference depends on the future strategies, for a given planned strategy $\{\nu(t, X_t^u)\}_{t \geq 0}^T$ of LR, a sequential game among different SRs is involved. The subgame perfect Nash equilibrium strategy $\{\bar{u}(t, X_t^u)\}_{t \geq 0}^T$ can be found through solving the following sequential game problem:

$$\begin{aligned} (NSC_t^{SR}) \quad & \max_{u(t, x)} \quad \tilde{J}(t, x, u(t, x), \nu(t, x), \bar{u}_f(t, x)) \\ \text{s.t.} \quad & \bar{u}(s, X_s^u) \text{ solves } (NSC_s^{SR}), \quad t < s \leq T, \end{aligned}$$

where $\bar{u}_f(t, x) = \{\bar{u}(s, X_s^u)\}_{s > t}^T$ is the set of future subgame perfect Nash equilibrium strategies and $\bar{u}_f(T, x)$ is an empty set. In fact, the best responsive strategy of SRs, $\{\bar{u}(t, X_t^u)\}_{t \geq 0}^T$, is just the time consistent strategy of the new preferences $\tilde{J}(\cdot)$.

After solving the sequential game among SRs, the game between LR and SRs becomes a leader-follower game. Then, according to SRs' responsive subgame perfect Nash equilibrium strategy, the

LR chooses the best planned strategy to balance between the global preference and the expected total punishment exercised under self-coordination,

$$(SC^{LR}(\omega)) \quad \max_{\{\nu(t, X_t^u)\}_{t \geq 0}^T} \quad J(0, x, \bar{u}(0, x), \bar{u}_f(0, x)) \\ - \omega \mathbb{E}_{0, x} \left[\int_0^T G(t, X_t^u, \bar{u}(t, X_t^u), \nu(t, X_t^u), \bar{u}_f(t, X_t^u)) dt \right],$$

where ω is the trade-off parameter. Parameter ω is also a quantitative representation of the decision maker's determination to control the internal conflict. The larger the ω , the stronger the determination and the smaller the total punishment. We call the equilibrium strategy of this two-tier dual-self game model as *strategy of self-coordination* of LR and SRs, denoted by $\{\nu^{Coor}(t, X_t^u, \omega)\}_{t \geq 0}^T$ and $\{u^{Coor}(t, X_t^u, \omega)\}_{t \geq 0}^T$, respectively.

Remark 2.1 *We can also formulate the LR's decision model as the following constrained optimization problem,*

$$(SC^{LR}(\epsilon)) \quad \max_{\{\nu(t, X_t^u)\}_{t \geq 0}^T} \quad J(0, x, \bar{u}(0, x), \bar{u}_f(0, x)) \\ \text{s.t.} \quad \mathbb{E}_{0, x} \left[\int_0^T G(t, X_t^u, \bar{u}(t, X_t^u), \nu(t, X_t^u), \bar{u}_f(t, X_t^u)) dt \right] \leq \epsilon,$$

where the inequality represents that the expected total punishment due to self-coordination should not exceed a pre-given resource level ϵ . This formulation is consistent with the finding that the cognitive resources used for self-control is a limited resource (see Shiv and Fedorikhin, 1999; Ward and Mann, 2000; Muraven and Baumeister, 2000). It is clear that formulation $(SC^{LR}(\omega))$ is actually the Lagrangian form of problem $(SC^{LR}(\epsilon))$ and parameter ω is essentially the shadow price of the limited exogenous cognitive resource.

Next, we introduce some mild conditions to endow certain desired properties for the proposed self-coordination scheme. Then, we show that under these conditions, the strategies of self-coordination of LR and SRs can be obtained by solving an optimization problem, and the equilibrium strategy achieves a degree of internal harmony among LR and SRs.

Assumption 2.1 (Costly Self-Coordination) *If $u(t, x) = \nu(t, x)$, then*

$$\tilde{J}(t, x, u(t, x), \nu(t, x), u_f(t, x)) = J(t, x, u(t, x), u_f(t, x)).$$

If $u(t, x) \neq \nu(t, x)$, then

$$\tilde{J}(t, x, u(t, x), \nu(t, x), u_f(t, x)) < J(t, x, u(t, x), u_f(t, x)).$$

Assumption 2.2 (Unique Self-coordination) For any quadruple $\{t, x, u(t, x), u_f(t, x)\}$, there exists a unique $\nu^u(t, x)$ such that for all $u'(t, x) \neq u(t, x)$,

$$\tilde{J}(t, x, u(t, x), \nu^u(t, x), u_f(t, x)) > \tilde{J}(t, x, u'(t, x), \nu^u(t, x), u_f(t, x)).$$

Assumption 1 and Assumption 2 are extensions of Assumptions 1, 2, 3, 4 in Fudenberg and Levine (2006). Assumption 1 suggests that when $u(t, x) \neq \nu(t, x)$, there is a strictly positive punishment. Assumption 2 ensures that, for any admissible strategy chosen by self t , there exists a unique planned strategy $\nu^u(t, x)$ to make this admissible strategy an optimizer under new preference $\tilde{J}$.

Under Assumptions 1 and 2, we define the cost of self-coordination as

$$C(t, x, u(t, x), u_f(t, x)) = G(t, x, u(t, x), \nu^u(t, x), u_f(t, x)),$$

where $\nu^u(t, x)$ is given in Assumption 2. The cost of self-coordination equals to the punishment incurred and measures the difficulty of inducing SR t to adopt strategy $u(t, x)$ under a given commitment of punishment scheme, which quantitatively represents the degree of conflict between LR and SRs. The following assumption is about the cost of self-coordination, which is a significant extension of Assumption 5 (Opportunity-based cost of self-control) in Fudenberg and Levine (2006).

Assumption 2.3 (Zero Cost of Tempted Strategy) For a given triple $\{t, x, u_f(t, x)\}$, we define the tempted strategy as

$$u^{Tem}(t, x) = \arg \max_{u(t, x)} J(t, x, u(t, x), u_f(t, x)).$$

Then,

$$C(t, x, u^{Tem}(t, x), u_f(t, x)) = 0,$$

$$C(t, x, u(t, x), u_f(t, x)) > 0, \text{ if } u(t, x) \neq u^{Tem}(t, x).$$

Assumption 3 implies that there is no difficulty in inducing SR t to take the tempted strategy. We can derive a particular tempted strategy at time t , $u^{Tem}(t, x)$ with $u_f(t, x) = \{u^{Tem}(s, X_s^u)\}_{|s>t}^T$. It is easy to see this particular tempted strategy is just the time consistent strategy $\{u^{TC}(t, X_t^u)\}_{|t\geq 0}^T$. Then, under Assumption 3, the time consistent strategy $\{u^{TC}(t, X_t^u)\}_{|t\geq 0}^T$ is the only strategy in attaining zero expected total cost of self-coordination. We also demonstrate a specific cost of self-coordination function that satisfies Assumption 3, which is widely used in the literature (see Gul and Pesendorfer, 2001).

Assumption 3(a) (Linear Cost of Self-coordination)

$$C(t, x, u(t, x), u_f(t, x)) = \gamma [J(t, x, u^{Tem}(t, x), u_f(t, x)) - J(t, x, u(t, x), u_f(t, x))],$$

where γ is a positive constant.

Here, the specific cost of self-coordination is a linear function of the opportunity cost of adopting strategy $u(t, x)$ instead of tempted strategy $u^{Tem}(t, x)$. Next, we demonstrate that these conditions can be easily satisfied.

Proposition 2.1 *Assume that i) $J(t, x, u(t, x), u_f(t, x))$ is a strictly concave function of $u(t, x)$; ii) there are no additional constraints on the one-dimensional $u(t, x)$ and $\nu(t, x)$. Then, the self-coordination scheme with the following punishment form,*

$$G(t, x, u(t, x), \nu(t, x), u_f(t, x)) = g(t, x, u_f(t, x))(u(t, x) - \nu(t, x))^2,$$

and $g(t, x, u_f(t, x)) > 0$, satisfies Assumptions 1, 2 and 3. Furthermore, we have

$$\begin{aligned} \nu^u(t, x) &= u(t, x) - \frac{1}{2}g^{-1}(t, x, u_f(t, x)) \frac{\partial J(t, x, u(t, x), u_f(t, x))}{\partial u}, \\ C(t, x, u(t, x), u_f(t, x)) &= \frac{1}{4}g^{-1}(t, x, u_f(t, x)) \left(\frac{\partial J(t, x, u(t, x), u_f(t, x))}{\partial u} \right)^2. \end{aligned}$$

The concrete form of $g(t, x, u_f(t, x))$ can be chosen by the decision maker to accomplish different purposes. We will discuss the choices of $g(t, x, u_f(t, x))$ in two applications: continuous-time investment and consumption problem with quasi-hyperbolic discounting, and continuous-time mean-variance portfolio selection problem. From Proposition 2.1, we can see that in general, what the

LR announces would be different from what the LR would like the SRs to achieve in his mind, i.e., $\nu^u(t, x) \neq u(t, x)$.

The equilibrium strategy of the two-tier dual-self game is characterized by the following theorem. It reveals that the two-tier dual-self game formulation can be actually solved by a (static) mathematical optimization problem.

Theorem 2.1 *Under Assumptions 1 and 2, the strategy of self-coordination of SRs, $\{u^{Coor}(t, X_t^u, \omega)\}_{t \geq 0}^T$ is the optimal solution of the following optimization problem,*

$$(SC_{SR}^{LR}(\omega)) \quad \max_{\{u(t, X_t^u)\}_{t \geq 0}^T} J(0, x, u(0, x), u_f(0, x)) - \omega \mathbb{E}_{0,x} \left[\int_0^T C(t, X_t^u, u(t, X_t^u), u_f(t, X_t^u)) dt \right],$$

and the strategy of self-coordination of LR, $\{\nu^{u^{Coor}}(t, X_t^u, \omega)\}_{t \geq 0}^T$, is characterized by Assumption 2, i.e., $\nu^{u^{Coor}}(t, x, \omega)$ is the planned strategy such that

$$\begin{aligned} & \tilde{J}(t, x, u^{Coor}(t, x, \omega), \nu^{u^{Coor}}(t, x, \omega), u_f^{Coor}(t, x, \omega)) \\ & > \tilde{J}(t, x, u'(t, x), \nu^{u^{Coor}}(t, x, \omega), u_f^{Coor}(t, x, \omega)), \quad \text{for all } u'(t, x) \neq u^{Coor}(t, x), \end{aligned}$$

where $u_f^{Coor}(t, x, \omega) = \{u^{Coor}(s, X_s^u, \omega)\}_{s > t}^T$.

We need to point out here that the expected total cost of self-coordination term of problem $(SC_{SR}^{LR}(\omega))$ does not introduce additional complexity. Theoretically, to derive the strategy of self-coordination of SRs is as difficult as deriving the pre-committed strategy.

Our next theorem reveals a close relationship between the strategy of self-coordination of SRs and the other two types of strategies (pre-committed strategy and time consistent strategy).

Theorem 2.2 *Under Assumptions 1, 2 and 3, we have*

- i) $u^{Coor}(t, x, 0) = u^{Com}(t, x)$, when $\omega = 0$;
- ii) $u^{Coor}(t, x, +\infty) = u^{TC}(t, x)$, when $\omega = +\infty$.

Theorem 2.2 has shown that, the equilibrium strategy set of dual-self game model parameterized in the trade-off parameter ω includes both the pre-committed strategy and the time consistent

strategy as its extreme cases when setting ω at its zero boundary or infinite boundary, respectively. In other words, our two-tier dual-self model offers a flexibility to strike a “subjective” balance between the long term objective and a degree of internal harmony (or internal conflict) by choosing different ω . By attaching a measurement of internal conflict to the preferences of both LR and SRs and applying corresponding punishment to both LR and SRs, the dual-self game model aligns the interests of LR and SRs, and achieves a degree of internal harmony among LR and SRs.

3 Investment and Consumption with Quasi-hyperbolic Discounting

In this section, we consider a typical Black-Scholes market setting, in which two assets are traded continuously. One of the assets is a risk-free bank account whose value process $S_0(t)$ is subject to the following ordinary differential equation (ODE):

$$\begin{cases} dS_0(t) = rS_0(t)dt, & t \geq 0, \\ S_0(0) = s_0 > 0, \end{cases}$$

where $r > 0$ is the interest rate. The other asset is a risky stock, whose price process $S_1(t)$ satisfies the following stochastic differential equation (SDE):

$$\begin{cases} dS_1(t) = S_1(t)[\mu dt + \sigma dW(t)], & t \geq 0, \\ S_1(0) = s_1 > 0, \end{cases}$$

where μ is the appreciation rate, σ is the volatility or dispersion rate of the stock and $W(t)$ is standard Brownian motion defined on $(\Omega, \mathcal{F}_T, \{\mathcal{F}_t\}, P)$. We assume that all the given market parameters r, μ and σ are deterministic constants.

Assume that a quasi-hyperbolic discounting applies to investors. More specific, the value at time t of \$1 received at future time s is described by

$$D_i(t, s) = \begin{cases} e^{-r(s-t)}, & \text{if } s \in [t_i, t_{i+1}), \\ \beta e^{-r(s-t)}, & \text{if } s \in [t_{i+1}, +\infty), \end{cases}$$

where $t_i \leq t < t_{i+1}$, the short-run discounting parameter $\beta < 1$ and $i = 0, 1, 2, \dots$.

An investor, with an initial endowment $x_0 > 0$, a log utility function and a decision horizon $[0, T]$, would like to decide the investment strategy $u(t)$ and the consumption strategy $c(t)$ over $[0, T]$. Without loss of generality, we can set $t_0 = 0$, $t_n \leq T < t_{n+1}$. The wealth at time t , $X^{u,c}(t)$, then satisfies

$$\begin{cases} dX^{u,c}(t) = [rX^{u,c}(t) + (\mu - r)u(t) - c(t)]dt + \sigma u(t)dW(t), & 0 \leq t \leq T, \\ X^{u,c}(0) = x_0. \end{cases} \quad (1)$$

We define now the infinitesimal generator $\mathbf{A}^{u,c}$ as follows,

$$\mathbf{A}^{u,c}V(t, x) = V_t + [rx + (\mu - r)u - c]V_x + \frac{1}{2}\sigma^2 u^2 V_{xx},$$

where V_t , V_x and V_{xx} denote the partial derivatives of V with respect to t and x , and the Hessian of V with respect to x , respectively.

At time 0, the LR faces the following stochastic decision problem,

$$\begin{aligned} (IC_0) \quad \max \quad & \mathbb{E}_{0,x_0} \left[\int_0^T D_0(0, t) \ln(c(t)) dt \right] + D_0(0, T) \mathbb{E}_{0,x_0} [\ln(X^{u,c}(T))], \\ \text{s.t.} \quad & (X^{u,c}(\cdot), u(\cdot), c(\cdot)) \text{ satisfies (1),} \\ & 0 \leq c(t) \leq X^{u,c}(t), \end{aligned}$$

where $\ln(\cdot)$ is the log utility function of the investor. Obviously, the optimal strategy of (IC_0) forms the pre-committed strategy. The time consistent strategy can be identified by consider the following sequential game, for $t_i \leq t < t_{i+1} \wedge T$,

$$\begin{aligned} (NIC_t) \quad \max \quad & \mathbb{E}_{t,x} \left[\int_t^T D_i(t, s) \ln(c(s)) ds \right] + D_i(t, T) \mathbb{E}_{t,x} [\ln(X^{u,c}(T))], \\ \text{s.t.} \quad & (X^{u,c}(\cdot), u(\cdot), c(\cdot)) \text{ satisfies (1),} \\ & 0 \leq c(t) \leq X^{u,c}(t), \\ & u^{TC}(s, X^{u,c}(s)), c^{TC}(s, X^{u,c}(s)) \text{ solves } (NIC_s), t < s \leq T. \end{aligned}$$

Proposition 3.1 *i) The pre-committed investment strategy and consumption strategy that solve (IC_0) are given as*

$$\begin{aligned} u^{Com}(t, x) &= \frac{\mu - r}{\sigma^2} x, \quad 0 \leq t \leq T, \\ c^{Com}(t, x) &= \frac{D_0(0, t)}{\int_t^T D_0(0, s) ds + D_0(0, T)} x, \quad 0 \leq t \leq T. \end{aligned}$$

ii) The time consistent investment strategy and consumption strategy that solve (NIC_t) are given as

$$u^{TC}(t, x) = \frac{\mu - r}{\sigma^2} x, \quad 0 \leq t \leq T,$$

$$c^{TC}(t, x) = \frac{D_i(t_i, t)}{\int_t^T D_i(t_i, s) ds + D_i(t_i, T)} x, \quad t_i \leq t < t_{i+1} \wedge T.$$

Please note that when $t_0 \leq t < t_1$ or $t_n \leq t \leq T$, we have $c^{Com}(t, x) = c^{TC}(t, x)$. Only when $t_1 \leq t < t_n$, the pre-committed consumption strategy is different from the time consistent consumption strategy, which causes time inconsistency. More specifically, when $t_1 \leq t < t_n$, the time consistent investor tends to consume more and invest less.

As the two investment strategies and two consumption strategies all take linear feedback forms of the current wealth level, their corresponding coefficients can also be considered as strategies. Furthermore, as the time consistent investment strategy coincides with the pre-committed investment strategy, the investor does not need to conduct self-control over investment strategies. Denote the consumption strategy of SR as $c(t, x) = c(t)x$ and the consumption strategy of LR as $d(t, x) = d(t)x$. We propose the following punishment for the consumption strategy at time t ,

$$G(t, c(t), d(t)) = \frac{[c(t) - d(t)]^2}{2 \min\{d(t), 1 - d(t)\}}. \quad (2)$$

Notice that this punishment neither depends on future investment and consumption strategies, nor depends on the current wealth level. The reason of choosing such punishment is mainly technical consideration. The key of self-coordination scheme is to construct an one-to-one mapping between $d(t)$ and $c(t)$. As both $d(t)$ and $c(t)$ are confined in $[0, 1]$, the simple Euclidean type of punishment, such as $[c(t) - d(t)]^2$, does not work, and an additional scale parameter $\frac{1}{2 \min\{d(t), 1 - d(t)\}}$ is needed. Please refer to the proof of Proposition 3.2 for more details.

Proposition 3.2 *For the consumption strategy, the self-coordination scheme with the proposed punishment in (2) satisfies Assumptions 1, 2, and 3.*

Although the punishment scheme satisfies our three assumptions, we do not have an explicit form for the consumption strategy of self-coordination. However, as the punishment term does not depend

on future strategies and the current wealth level, we can numerically compute the cost of self-coordination function at time t and derive the numerical consumption strategy of self-coordination very efficiently. Let us demonstrate our results by the following example.

Example 1 *The market parameters are assumed to take values of $r = 0.05$, $\mu = 0.12$, $\sigma = 0.15$, $T = 4$, $\beta = 0.9$ and $x_0 = 1$. We also set $t_0 = 0$, $t_1 = 1$, $t_2 = 2$ and $t_3 = 3$.*

For given consumption strategy of LR, $\{d(t)\}_{t \geq 0}^4$, the subgame perfect Nash equilibrium consumption strategy $\{\bar{c}(t)\}_{t \geq 0}^4$ is derived by solving the following sequential game,

$$\begin{aligned}
(NIC_t^{SR}) \quad & \max \quad \mathbb{E}_{t,x} \left[\int_t^4 D_i(t,s) \ln(c(s)X^{u,c}(s)) ds \right] + D_i(t,4) \mathbb{E}_{t,x} [\ln(X^{u,c}(4))] \\
& - G(t, c(t), d(t)) \\
\text{s.t.} \quad & (X^{u,c}(\cdot), u(\cdot), c(\cdot)X^{u,c}(\cdot)) \text{ satisfies (1),} \\
& 0 \leq c(t) \leq 1, \\
& u(\tau, X^{u,c}(\tau)) = \frac{\mu-r}{\sigma^2} X^{u,c}(\tau), \quad t < s \leq 4, \\
& \bar{c}(s) \text{ solves } (NIC_s^{SR}), \quad t < s \leq 4.
\end{aligned}$$

Following the analysis in the proof of Proposition 3.2, we can show that $\bar{c}(t)$ is the larger solution of the equation,

$$-\bar{c}^2(t) + d(t)\bar{c}(t) - \epsilon(t) \min\{d(t), 1 - d(t)\}\bar{c}(t) + \min\{d(t), 1 - d(t)\} = 0,$$

where $\epsilon(t) = \int_t^T D_i(t,s)ds + D_i(t,T)$. Then, the cost of self-coordination is given by

$$C(t, \bar{c}(t)) = G(t, \bar{c}(t), d(t)) = \frac{[\bar{c}(t) - d(t)]^2}{2 \min\{d(t), 1 - d(t)\}},$$

which is deterministic, dependent neither on the current wealth level, nor on the future strategies.

Both $\bar{c}(t)$ and $C(t, \bar{c}(t))$ can be easily computed. We show the SRs' subgame perfect Nash equilibrium consumption strategies $\bar{c}(t)$ at time instants t_0, t_1, t_2, t_3 in Figure 1(a), and express the cost of self-coordination functions at the four time instants in Figure 1(b) .

Based on Theorem 2.1, the consumption strategies of self-coordination adopted by SRs, $c^{Coor}(t)$,

(a) Short-run selves' consumption strategies

(b) cost of self-control functions

Figure 1: Short-run selves' consumption strategies and Cost of self-control

is the optimal solution of the following problem:

$$\begin{aligned}
 (IC_{SR}^{LR}(\omega)) \quad & \max \quad \mathbb{E}_{0,1} \left[\int_0^4 D_0(0,t) \ln(c(t)X^{u,c}(t))dt \right] - \omega \int_0^4 C(t, c(t))dt \\
 & + D_0(0,4)\mathbb{E}_{0,1}[\ln(X^{u,c}(4))], \\
 \text{s.t.} \quad & (X^{u,c}(\cdot), u(\cdot), c(\cdot)X^{u,c}(\cdot)) \text{ satisfies (1),} \\
 & u(t, X^{u,c}(t)) = \frac{\mu-r}{\sigma^2} X^{u,c}(t), \\
 & 0 < c(t) < 1.
 \end{aligned}$$

Following the analysis of problem (IC_0) in the proof of Proposition 3.1, we can show that the consumption strategy of self-coordination at time t is just the optimal solution of the following static optimization problem,

$$\max_c \left\{ D_0(0,t) \ln(c) - \omega C(t, c) - \left[\int_t^4 D_0(0,s)ds + D_0(0,4) \right] c \right\}.$$

Figure 2 illustrates consumption strategies of self-coordination at time instants t_0, t_1, t_2, t_3 as functions of ω . When $\omega \rightarrow +\infty$, all the equilibrium consumption strategies converge to time consistent consumption strategies, while all the equilibrium consumption strategies converge to the pre-committed consumption strategies when $\omega \rightarrow 0$. We observe that the investor with self-coordination consumes less than the time consistent investor in general. Please note that all the three types of consumption strategies are the same at both initial time interval and terminal time interval.

Figure 2: Consumption Strategies under Coordination v.s. ω

4 Dynamic Mean-Variance Portfolio Selection Problem

We assume that the market setting remains the same as the Black-Scholes one in the last section. We now consider an investor of a mean-variance type with an initial endowment $x_0 > 0$ and an investment horizon $[0, T]$. Assume that the trading is self-financed and takes place continuously, and that there is no transaction cost and consumptions during the investment process. The wealth process, $X^u(t)$, then satisfies

$$\begin{cases} dX^u(t) = [rX^u(t) + (\mu - r)u(t)]dt + \sigma u(t)dW(t), & 0 \leq t \leq T, \\ X^u(0) = x_0, \end{cases} \quad (3)$$

where portfolio $u(t)$ denotes the total market value of wealth in the stock, which is required to be in $\mathcal{L}^2(\mathcal{F}_t, P)$. We define the infinitesimal generator $\mathbf{A}^u$ now as follows,

$$\mathbf{A}^u V(t, x) = V_t + [rx + (\mu - r)u]V_x + \frac{1}{2}\sigma^2 u^2 V_{xx}.$$

At time 0, the long-run self of the investor is seeking the pre-committed optimal strategy such that the expected value of the terminal wealth is maximized and the variance of the terminal wealth is minimized,

$$\begin{aligned} (MV_0(\lambda)) \quad & \max \quad \mathbb{E}_{0, x_0}[X^u(T)] - \lambda \text{Var}_{0, x_0}(X^u(T)), \\ & \text{s.t.} \quad (X^u(\cdot), u(\cdot)) \text{ satisfies (3)}, \end{aligned} \quad (4)$$

where $\lambda > 0$ is risk aversion parameter, which represents the trade-off between the two conflicting objectives (the larger the λ , the larger the risk aversion). Varying parameter λ in $(0, +\infty)$ in problem formulation $(MV_0(\lambda))$ traces out the efficient frontier in the mean-variance space.

The pre-committed strategy of $(MV_0(\lambda))$ can be found in Zhou and Li (2000),

$$u^{Com}(t, x) = -\frac{\mu - r}{\sigma^2} \left[x - \left(e^{rT} x_0 + \frac{e^{\delta T}}{2\lambda} \right) e^{-r(T-t)} \right], \quad (5)$$

where $\delta = \frac{(\mu - r)^2}{\sigma^2}$. Moreover, the corresponding efficient frontier is expressed as

$$\begin{aligned} \text{Var}_{0, x_0}(X^u(T))|_{u^{Com}} &= \frac{e^{-\delta T}}{1 - e^{-\delta T}} (\mathbb{E}_{0, x_0}[X^u(T)]|_{u^{Com}} - e^{rT} x_0)^2, \\ &\text{for } \mathbb{E}_{0, x_0}[X^u(T)]|_{u^{Com}} \geq e^{rT} x_0. \end{aligned}$$

The time consistent strategy can be identified by solving the following sequential game,

$$\begin{aligned} (NMV_t(\lambda)) \quad \max \quad & \mathbb{E}_{t, x}[X^u(T)] - \lambda \text{Var}_{t, x}(X^u(T)), \\ \text{s.t.} \quad & (X^u(\cdot), u(\cdot)) \text{ satisfies (3),} \\ & u(s, X^u(s)) = u^{TC}(s, X^u(s)), \quad t < s \leq T, \\ & u^{TC}(s, X^u(s)) \text{ solves } (NMV_s(\lambda)), \quad t < s \leq T. \end{aligned}$$

Following Basak and Chabakauri (2010), we can derive the time consistent strategy as

$$u^{TC}(t, x) = \frac{1}{2\lambda} e^{-r(T-t)} \frac{\mu - r}{\sigma^2}, \quad (6)$$

with the corresponding efficient frontier expressed as

$$\begin{aligned} \text{Var}_{0, x_0}(X^u(T))|_{u^{TC}} &= \frac{1}{\delta T} (\mathbb{E}_{0, x_0}[X^u(T)]|_{u^{TC}} - e^{rT} x_0)^2, \\ &\text{for } \mathbb{E}_{0, x_0}[X^u(T)]|_{u^{TC}} \geq e^{rT} x_0. \end{aligned}$$

As both the pre-committed strategy and the time consistent strategy take linear forms, we confine ourselves to consider only linear controls in our two-tier dual-self game model. Denote the SRs' carried out strategy as $u(t, x) = k^u(t)x + l^u(t)$ and the LR's planned strategy as $\nu(t, x) = k^\nu(t)x + l^\nu(t)$. We next propose a particular form of punishment as follows,

$$\begin{aligned} & G(t, x, k^u(t)x + l^u(t), k^\nu(t)x + l^\nu(t), u_f(t, x)) \\ &= \lambda \sigma^2 ([a(t)]^2 - A(t)) [k^u(t)x + l^u(t) - k^\nu(t)x - l^\nu(t)]^2, \end{aligned} \quad (7)$$

where $u_f(t, x) = \{k^u(s)X^u(s) + l^u(s)\}_{|s>t}^T$ is the set of future SRs' strategies, $a(t)$ and $A(t)$ depend on future SRs' strategies and satisfy the following ODEs,

$$\begin{cases} a_t + ra + (\mu - r)ak^u = 0, \\ a(T) = 1, \\ A_t + 2A[r + (\mu - r)k^u] - \sigma^2(a^2 - A)(k^u)^2 = 0, \\ A(T) = 0. \end{cases}$$

The reason of choosing such punishment is to ensure that the cost of self-coordination satisfies Assumption 3(a). Please refer to the proof of Proposition 4.1 for more details.

Proposition 4.1 *The self-coordination scheme with the proposed form of punishment in (7) satisfies Assumptions 1, 2 and 3(a) (with $\gamma = 1$). Moreover, the cost of self-coordination can be expressed as*

$$\begin{aligned} & C(t, x, k^u(t)x + l^u(t), u_f(t, x)) \\ &= \lambda \sigma^2([a(t)]^2 - A(t)) [k^u(t)x + l^u(t) - \hat{k}(t)x - \hat{l}(t)]^2, \end{aligned} \quad (8)$$

where

$$\begin{aligned} \hat{k}(t) &= \frac{(\mu - r)A(t)}{\sigma^2([a(t)]^2 - A(t))}, \\ \hat{l}(t) &= \frac{(\mu - r)B(t)}{2\sigma^2\lambda([a(t)]^2 - A(t))}, \\ &\begin{cases} B_t + B[r + (\mu - r)k^u] + 2\lambda(\mu - r)Al^u - 2\lambda\sigma^2(a^2 - A)k^ul^u = 0, \\ B(T) = 1. \end{cases} \end{aligned} \quad (9)$$

Based on Theorem 2.1, the equilibrium decision set of the mean-variance portfolio selection problem can be obtained by solving the following problem parameterized in ω ,

$$\begin{aligned} (MV_{SR}^{LR}(\lambda, \omega)) \quad \max \quad & \mathbb{E}_{0, x_0}[X^u(T)] - \lambda \text{Var}_{0, x_0}(X^u(T)) \\ & - \omega \mathbb{E}_{0, x_0} \left[\int_0^T C(t, X^u(t), u(t, X^u(t)), u_f(t, X^u(t))) dt \right], \\ \text{s.t.} \quad & (X^u(\cdot), u(\cdot)) \text{ satisfies (3),} \\ & u(t, x) = k^u(t)x + l^u(t). \end{aligned}$$

Due to the non smoothness of variance term, we are unable to solve problem $(MV_{SR}^{LR}(\lambda, \omega))$ directly by dynamic programming. We, instead, employ the mean-field formulation proposed by Cui et al. (2014) to derive the explicit strategy of self-coordination.

Theorem 4.1 *The strategy of self-coordination adopted by short-run selves which solves $(MV_{SR}^{LR}(\lambda, \omega))$, $\{u^{Coor}(t, x)\}_{t \geq 0}^T$, is given as*

$$u^{Coor}(t, x) = k^{Coor}(t)x + l^{Coor}(t), \quad (10)$$

with

$$k^{Coor}(t) = \hat{\beta}(t) \cdot \frac{\mu - r}{\sigma^2},$$

$$l^{Coor}(t) = \left[\lambda^{-1} \hat{\gamma}(t) - \hat{\beta}(t) \left(e^{rt} x_0 + \lambda^{-1} \delta \int_0^t e^{r(t-s)} \hat{\gamma}(s) ds \right) \right] \cdot \frac{\mu - r}{\sigma^2},$$

where the parameters $\hat{\beta}(t)$ and $\hat{\gamma}(t)$ are defined as follows,

$$\hat{\beta}(t) = \frac{\omega A(t) - \zeta(t)}{\omega([a(t)]^2 - A(t)) + \zeta(t)},$$

$$\hat{\gamma}(t) = \frac{\omega \gamma(t) + e^{r(T-t)}}{2\omega([a(t)]^2 - A(t)) + 2\zeta(t)},$$

and other parameters are governed by the following ODEs,

$$\begin{cases} \zeta_t + 2r\zeta - \delta[\omega(a^2 - A) + \zeta]\hat{\beta}^2 + \frac{\omega\delta A^2}{a^2 - A} = 0, \\ \zeta(T) = 1, \end{cases}$$

$$\begin{cases} a_t + ra + (\mu - r)ak^{Coor} = 0, \\ a(T) = 1, \end{cases}$$

$$\begin{cases} A_t + 2A[r + (\mu - r)k^{Coor}] - \sigma^2(a^2 - A)(k^{Coor})^2 = 0, \\ A(T) = 0. \end{cases}$$

$$\begin{cases} \beta_t + r\beta + [r + (\mu - r)k^{Coor}]\beta - (\mu - r)Ak^{Coor} + \sigma^2(a^2 - A)(k^{Coor})^2 = 0, \\ \beta(T) = 0, \end{cases}$$

$$\begin{cases} \gamma_t + 2\delta\beta\hat{\gamma} + [r + (\mu - r)k^{Coor}]\gamma + 2\delta A\hat{\gamma} - 2(\mu - r)(a^2 - A)k^{Coor}\hat{\gamma} = 0, \\ \gamma(T) = 1. \end{cases}$$

Finally, the efficient frontier is expressed as

$$\begin{aligned}
& \text{Var}_{0,x_0}(X^u(T))|_{u^{Coor}} \\
&= \frac{\int_0^T e^{\int_s^T [2r + 2\delta\hat{\beta}(\tau) + \delta(\hat{\beta}(\tau))^2] d\tau} (\hat{\gamma}(s))^2 ds}{\delta \left(\int_0^T e^{r(T-s)} \hat{\gamma}(s) ds \right)^2} \left[\mathbb{E}_{0,x_0}[X^u(T)]|_{u^{Coor}} - e^{rT} x_0 \right]^2, \\
& \text{for } \mathbb{E}_{0,x_0}[X^u(T)]|_{u^{Coor}} \geq e^{rT} x_0.
\end{aligned}$$

We have shown that the strategy of self-coordination takes a linear feedback form, and the efficient frontier achieved by strategy of self-coordination is also a line in mean-standard deviation plane. Furthermore, when $\omega = 0$ (or $\omega = +\infty$), the strategy under coordination of short-run selves converges to pre-committed strategy (or time consistent strategy) (see Theorem 2.2).

Although the calculation of the parameters involved seems complicated, it does not depend on risk attitude parameter λ and can be computed off-line easily by, for example, Euler method.

Example 2 We assume the parameters set as $r = 0.05$, $\mu = 0.12$, $\sigma = 0.10$ and $T = 1$. As shown in Figure 3(a), the efficient frontiers achieved by strategies of self-coordination lie in between the two achieved by the pre-committed strategy and the time consistent strategy. For different ω s, we also compute the corresponding Sharpe ratios, which are depicted in Figure 3(b). It is evident that the smaller the ω , the larger the Sharpe ratio achieved by the strategy of self-coordination. Please note that the largest Sharpe ratio and smallest Sharpe ratio are achieved by the pre-committed strategy and the time consistent strategy, respectively.

Figure 4 reveals the following findings on the expected total cost of self-coordination: 1) the stronger the willpower to conduct self-coordination (the smaller the value of ω), the better the long term investment performance, and the larger the expected total cost of self-coordination, i.e., to achieve better mean-variance efficient frontier of the terminal wealth, the investor needs to bear larger punishment; 2) the less risk aversion of the investor (the smaller the value of λ), the larger the intention of holding risky assets, and the harder to conduct self-coordination; 3) when $\lambda = +\infty$, the expected total cost of self-coordination is zero, i.e., when the investor is fully risk averse and only invests in riskless asset, there is no need to conduct self-coordination.

(a) Efficient frontiers

(b) Sharp ratio v.s. ω

Figure 3: Long term investment performances

Figure 4: Expected total cost of self-control v.s. ω

5 Conclusion

To obtain the strategy of self-coordination for time inconsistent decision problems, possibly with non-expectation operator in the objective, we have extended in this paper the existing models in the literature and propose an operable two-tier dual-self game model with commitment by punishment in continuous-time setting. In the game model, we quantitatively describe the internal conflict between long-run self and short-run selves by the expected total punishments (cost of self-coordination), and quantitatively describe the willpower of the decision maker to conduct self-control by the trade-off parameter between long-run self's performance objective and expected total punishments (cost of self-coordination). By doing so, we integrate pre-committed strategy, time consistent strategy and strategy of self-coordination in a unified two-tier dual-self game model. The decision maker can subjectively choose the most suitable strategy in accordance with the punishment he adopts and his willpower to conduct self-coordination, which achieves the most suitable level of internal harmony between long-run self and short-run selves. Successful applications of the proposed unified dual-self model in the time separable investment and consumption problem with quasi-hyperbolic discounting and in the time nonseparable mean-variance portfolio selection problem demonstrate the flexibility of the model. Our work pushes forward the study on resolving time inconsistency of time nonseparable decision problems.

Acknowledgments

This work was partially supported by National Natural Science Foundation of China under grant 71671106, 71601107, by the State Key Program in the Major Research Plan of National Natural Science Foundation of China under grant 91546202, by Shanghai Institute of International Finance and Economics.

References

- Artzner, P., Delbaen, F., Eber, J. M., Heath, D., Ku, H. 2007. Coherent multiperiod risk adjusted values and Bellman's principle. *Ann. Oper. Res.* 152, 5-22.
- Basak, S., Chabakauri, G. 2010. Dynamic mean-variance asset allocation. *Rev. Financ. Stud.* 23, 2970-3016.
- Benhabib, J., Bisin, A. 2005. Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption-saving decisions. *Games Econ. Behav.* 52, 460-492.
- Björk, T., Murgoci, A., Zhou, X. Y. 2014. Mean-variance portfolio optimization with state dependent risk aversion. *Math. Financ.* 24, 1-24.
- Bleichrodt, H., Rohde, K. I.M., Wakker, P. P. 2009. Non-hyperbolic time inconsistency. *Games Econ. Behav.* 66, 27-38.
- Boda, K., Filar, J. A. 2006. Time consistent dynamic risk measures. *Math. Method. Oper. Res.* 63, 169-186.
- Cui, X. Y., Li, X., Li, D. 2014. Unified framework of mean-field formulations for optimal multiperiod mean-variance portfolio selection, *IEEE T. Automat. Contr.* 59, 1833-1844.
- Cui, X. Y., Li, D., Wang, S. Y., Zhu, S. S. 2012. Better than dynamic mean-variance: Time inconsistency and free cash flow stream. *Math. Financ.* 22, 346-378.
- Cui, X. Y., Xu, L., Zeng, Y. 2015. Continuous Time Mean-Variance Portfolio Optimization with Piecewise State-Dependent Risk Aversion, to appear in *Optim. Lett.*
- DellaVigna, S., Malmendier, U. 2004. Contract Design and Self-Control: Theory and Evidence. *Q. J. Econ.* 119, 353-402.
- DellaVigna, S., Malmendier, U. 2006. Overestimating self-control: Evidence from the health club industry. *Am. Econ. Rev.* 96, 694-19.

- Fudenberg, D., Levine, D. K. 2006. A Dual-Self Model of Impulse Control. *Am. Econ. Rev.* 96, 1449-1476.
- Fudenberg, D., Levine, D. K. 2012. Timing and Self-Control. *Econometrica*. 80, 1-42.
- Grenadier, S.R., Wang, N. 2007. Investment under uncertainty and time-inconsistent preferences. *J. Financ. Econ.* 84, 2-39.
- Gul, F., Pesendorfer, W. 2001. Temptation and Self-Control. *Econometrica*. 69, 1403-1435.
- Gul, F., Pesendorfer, W. 2004. Self-Control and the Theory of Consumption. *Econometrica*. 71, 119-158.
- Harris, C., Laibson, D. 2013. Instantaneous gratification. *Q. J. Econ.* 128, 205-248.
- Marín-Solano, J., Navas, J. 2010. Consumption and portfolio rules for time-inconsistent investors. *Eur. J. Oper. Res.* 201, 860-872.
- Muraven, M., Baumeister, R. F. 2000. Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle? *Psychol. Bull.* 126, 247-259.
- O'Donoghue, T., Rabin, M. 1999. Doing it now or later. *Am. Econ. Rev.* 89, 103-124.
- O'Donoghue, T., Rabin, M. 2001. Choice and procrastination. *Q. J. Econ.* 116, 121-160.
- Pan, J. R., Webb, C. S., Zank, H. 2015. An extension of quasi-hyperbolic discounting to continuous time. *Games Econ. Behav.* 89, 43-55.
- Rachlin, H. *The Science of Self-Control*. Harvard University Press, Cambridge, Massachusetts, (2004).
- Shiv, B., Fedorikhin, A. 1999. Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making. *J. Consum. Res.* 26, 278-292.
- Strotz, R. H. 1956. Myopia and Inconsistency in Dynamic Utility Maximization. *Rev. Econ. Stud.* 23, 165-180.

- Takeuchi, K. 2011. Non-parametric test of time consistency: Present bias and future bias. *Games Econ. Behav.* 71, 456-478.
- Thaler, R. H. 1981. Some Empirical Evidence on Dynamic Inconsistency. *Econ. Lett.* 8, 201-207.
- Thaler, R. H., Benartzi, S. 2004. Save More Tomorrow: Using behavioral economics to increase employee saving. *J. Polit. Econ.* 114, 164-187.
- Thaler, R. H., Shefrin, H. M. 1981. An Economic Theory of Self-Control. *J. Polit. Econ.* 89, 392-406.
- Wang, J., Forsyth, P. A. 2011. Continuous time mean variance asset allocation: A time-consistent strategy. *Eur. J. Oper. Res.* 209, 184-201.
- Ward, A., Mann, T. 2000. Don't mind if I do: Disinhibited eating under cognitive load. *J. Pers. Soc. Psychol.* 78, 753-763.
- Yu, G., and X. T. Qi, *Disruption Management*, World Scientific, (2004).
- Zhou, X. Y., Li, D. 2000. Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Appl. Math. Opt.* 42, 19-33.