INFINITELY REPEATED GAMES WITH SELF-CONTROL: A DUAL-SELF INTERPRETATION OF THE MONKS STORY

Wei Wang, Jie Zheng
School of Economics and Management, Tsinghua University
Beijing
P.R. China
wangw.10@sem.tsinghua.edu.cn: zhengjie@sem.tsinghua.edu.cn

1. Introduction

What is dual-self? Who are those two selves? Why exercising self-control? Is it necessary? When one reads the title of this chapter, these might be some of the questions that arise. Dual-self refers to the game between a long-run self and a sequence of short-run selves within an individual. Fudenberg and Levine [2006] state that each individual has two selves: the long-run self is a patient self who tries to maximize his total utility across time; however, the short-run self is made of a sequence of selves that are impulsive who exist for only a brief time, thus each of these short-run selves cares only about their immediate experience and rewards. Therefore self-control must be established in order for the individual to make decisions that are overall better off for him.

The concept of dual-self is not an invention of this century. Adam Smith [1758, 1776] was the first to consider two selves within an individual: the impartial spectator vs. the passion-driven. Shefrin and Thaler [1981] were the first to systematically and formally treating a two-self economic man. According to their theory in 1981, at any point in time an individual is viewed as an organization that consists of a planner and a doer to reflect the conflict between short-run and long-run preferences. The planner role is concerned with lifetime utility whereas the doer is completely selfish or myopic and exists only for one period. The concept of hot/cold states mentioned in Loewenstein [2000] is later extended as a hot/cold model of addiction in Bernheim and Rangel [2004]. The latter model describes that an individual operates in a cold mode when he considers all alternatives and contemplates all consequences, and that he gets in a hot decision-making mode where he always "consumes the substance". Consistently, Loewenstein and O'Donoghue [2004] develop a two-system model in which a person's behavior is the outcome of an interaction between affective system and deliberative system. The former desires immediate gratification; the latter assesses options with a goal-based perspective and considers longer-term effects. Ashraf, Camerer and Loewenstein [2005] state that "behavior was determined by the struggle between what Smith termed the 'passions' and the 'impartial spectator'".

We adopt the theoretical framework of dual-self established by Fudenberg and Levine [2006] not only because the concepts of two-selves used are consistent with the literature mentioned above, but also the dual-self model (DSM) of Fudenberg and Levine [2006] provides advancements to the model with quasi-hyperbolic utility (QHM) done by O'Donoghue and Rabin [2001]. One key difference between DSM and QHM that is worth noting is that QHM assumes that a conflict of interests exist between present and future selves while DSM states that both the long-run self and short-run selves share the same preferences. We believe that the latter more closely depicts the fact that both the long-run and short-run selves belong to the same individual.

Our model extends from individual decision-making problems to games involving strategic situations among multiple players in order to better capture human behaviors in decision-making process in reality. Our model reflects two-dimensioned games: one dimension refers to a game played between a long-run self and short-run selves within the same individual; the other dimension refers to a game played among players. We believe that such an extension can help us better understand how individual makes decisions in reality as human beings constantly interact with each other.

When discussing multi-period games with strategic interactions, one may wonder whether our model is any different from the repeated game with history-dependent strategy (HDSM). There are two key differences that are worth noting. First, every player in our model is assumed to be a dual-self individual whereas players in HDSM are assumed to be one single-self. Second, in our model, given that each of the short-run selves lives only for one period, they only care about immediate payoffs, and thus their strategies may not depend on incidences which happened in the past. However, in HDSM, the decisions made by players at every stage game are based on the previous outcomes.

In our theoretical model, we follow the same assumptions of DSM established by Fudenberg and Levine [2006]: Assumption of Costly Self-control assumes that payoff without self-control is higher than that with self-control; Assumption of Unlimited Self-Control shows that for any player, for any strategy choice that the player takes, we can always find an optimal self-control action that maximizes his payoff; and a technical assumption, Assumption of Continuity. In addition to the assumptions mentioned above, we introduce Assumption of Independent Self-Control to our model in order to depict interactions among multiple players. This assumption is necessary as our model expands the single-player decision problem, which is the focus of current literature, to a multi-player decision problem. This assumption means that a player's payoff only depends on his own self-control action and actions of all the players, and that the other players' self-control actions have no impact on this player's payoff function. Furthermore, it is worth noting that the introduction of such an assumption makes our model more general and keeps our conclusion consistent with that of Fudenberg and Levine [2006].

In this paper, we use The Monks Story, a famous traditional Chinese proverb, as an example to illustrate the practicality of our model on the matters in reality. The Monks Story is typically used in literature that is related to marketing, human resource, and management aspects. In those research areas, they are more concerned with possible changes in players' behaviors as the number of players increase. Therefore, the main focus of those studies is on coordination and cooperation issues. However, in our paper, we look at the Monks Story from a different perspective: we assume that the number of players does not change over time and examine how players' behaviors evolve over time. Our analysis of this s-tory is focused on how individual player as a dual-self individual makes decisions when interacting with other players. In order to make this example more interesting, we analyze the story and compare the results under three scenarios: history-independent strategy case, history-dependent strategy case, and dual-self approach. Furthermore, we show two cases of analysis under the dual-self approach: 1) when the short-run selves cannot observe the previous outcomes; 2) when the short-run selves can observe the previous outcomes. We hope our analysis will provide an interesting perspective for the Monks Story.

This chapter is organized as follows: Section 2 shows the basic model; Section 3 applies our model to an example The Monks Story; Section 4 is the conclusion and discussion.

2. The Model

We adopt the framework of Fudenberg and Levine [2006] and focus on the infinite-horizon multi-player case.

There are I $(2 \leq I < \infty)$ players, $i = 1, 2, \cdots, I$, and infinite number of periods, $t = 1, 2, \cdots$. Player i's discount factor between any two consecutive periods is constant and denoted by δ_i, where $\delta_i \in [0,1]$.

Each player i is considered a dual-self agent: a sequence of short-run selves $\{SRS_i^t\}_{t=1}^{\infty}$, each of whom lives only for one period, and a long-run self LRS_i, who lives forever. SRS_i^t plays a one-period strategy to maximize his short-run payoff while alive, and LRS_i plays a series of self-control strategies over time to maximize his long-run payoff.

SRS_i^t's choice at period t is denoted by s_i^t, $s_i^t \in S_i^t \subseteq \mathbb{R}$. We write $s_i = \left(s_i^1, s_i^2, \cdots\right)$, $s^t = (s_1^t, s_2^t, \cdots, s_I^t)$, and $\mathbf{s} = (s_1, s_2, \cdots, s_I) = \left(s^1, s^2, \cdots\right)$. For simplicity, we assume $S_i^t = S_i^{t'}$ $\forall i \, \forall t, t'$. LRS_i's self-control action at period t is denoted by r_i^t, $r_i^t \in R_i^t \subseteq \mathbb{R}$. Similarly, we have $r_i = \left(r_i^1, r_i^2, \cdots\right)$, $r^t = (r_1^t, r_2^t, \cdots, r_I^t)$, and $\mathbf{r} = (r_1, r_2, \cdots, r_I) = \left(r^1, r^2, \cdots\right)$. We also assume $R_i^t = R_i^{t'}$ $\forall i \, \forall t, t'$.[1]

A finite history of play at period t, denoted by h_t, consists of all the players' past actions, $h_t = \begin{cases} \left(r^1, s^1, r^2, s^2, \cdots, r^{t-1}, s^{t-1}\right) & \text{if } t \geq 2 \\ \emptyset & \text{if } t = 1 \end{cases}$. Let $\{r^t\} = r^1, r^2, \cdots, r^t$, and let $\{s^t\} = s^1, s^2, \cdots, s^t$. Then we have $h_t = \left(\left\{r^{t-1}\right\}, \left\{s^{t-1}\right\}\right)$ if $t \geq 2$.

In general, player i's lifetime payoff (i.e. the payoff of LRS_i) depends on both all the long-run selves' self-control choices and all the short-run selves' strategies over time, denoted by $U_i(\mathbf{r}, \mathbf{s}) : \prod_{i=1}^{I} \prod_{t=1}^{\infty} R_i^t \times \prod_{i=1}^{I} \prod_{t=1}^{\infty} S_i^t \longrightarrow \mathbb{R}$. When there is no self control for any player i at any time t, that is $r_i^t = 0$ $\forall i \, \forall t$, player i's lifetime payoff becomes $U_i(\mathbf{0}, \mathbf{s})$, where $\mathbf{0}$ is the

[1] These restrictions $R_i^t = R_i^{t'}$ $\forall i \, \forall t, t'$ and $S_i^t = S_i^{t'}$ $\forall i \, \forall t, t'$ are for modeling convenience only, and they can be relaxed without significant change of the results.

null vector in $\mathbb{R}^I \times \mathbb{R}^\infty$.

In general, player i's period-t payoff (i.e. the payoff of SRS_i^t) depends on both the past actions and the current actions, denoted by $u_i^t(h_{t+1}) = u_i^t(h_t, r^t, s^t) = u_i^t(\{r^t\}, \{s^t\}) : \prod_{i=1}^{I}\prod_{\tau=1}^{t} R_i^\tau \times \prod_{i=1}^{I}\prod_{\tau=1}^{t} S_i^\tau \longrightarrow \mathbb{R}$. When there is no self control for any player i at any time $\tau \le t$, that is $r_i^\tau = 0\ \forall i\ \forall \tau \le t$, player i's period-t payoff becomes $u_i^t(\{\mathbf{0}^t\}, \{s^t\})$, where $\mathbf{0}^t$ is the null vector in $\mathbb{R}^{It}$. For simplicity, we assume from now on that a player's per-period payoff only depends on the current actions of all the players, that is,

$$u_i^t\left(r^t, s^t\right) : \prod_{i=1}^{I} R_i^t \times \prod_{i=1}^{I} S_i^t \longrightarrow \mathbb{R}.$$

Under the dual-self framework, it is natural to assume that the lifetime payoff is time-additive. Thus, player i's lifetime payoff (i.e. the payoff for LRS_i) is

$$U_i(\mathbf{r}, \mathbf{s}) = \sum_{t=1}^{\infty} \delta_i^{t-1} u_i^t\left(r^t, s^t\right).$$

At each period, for every player, his long-run self chooses a self-control action. Given that, every player's short-run self plays the specified game to maximize his utility. The mixed strategies of LRS_i at period t are maps from histories to current self-control actions, $\sigma_i^{LR,t} : H \longrightarrow \Delta R_i$. Denote the set of t-length histories by H_t. A strategy for SRS_i^t is a map $\sigma_i^t : H_{t-1} \times \prod_{i=1}^{I} R_i^t \longrightarrow \Delta S_i$. We denote the collection of all of these strategies by σ_i^{SR}. The strategy profiles $\sigma^{LR} = \left(\sigma_1^{LR}, \cdots, \sigma_I^{LR}\right)$ and $\sigma^{SR} = \left(\sigma_1^{SR}, \cdots, \sigma_I^{LR}\right)$ together give rise to probability distributions π^t over histories of length t for every t. Therefore, the lifetime utility function is given by

$$U_i\left(\sigma^{LR}, \sigma^{SR}\right) = \sum_{t=1}^{\infty} \delta_i^{t-1} \int_{h_t^A} u_i^t\left(\mathbf{r}^t(h_t), \mathbf{s}^t(h_t)\right) d\pi^t(h_t).$$

Now we impose some assumptions on the payoff functions. For simplicity, we omit the superscript t for all the notations. For example, all the payoff functions in this section are one-period payoff functions.

Assumption 2.1 (Costly Self-Control)**.** *For any player i, for any $s_i, \mathbf{r}_{-i}, \mathbf{s}_{-i}$, if $r_i \neq 0$, then $u_i(r_i, \mathbf{r}_{-i}, s_i, \mathbf{s}_{-i}) < u_i(0, \mathbf{r}_{-i}, s_i, \mathbf{s}_{-i})$.*

Assumption 2.2 (Unlimited Self-Control)**.** *For any player i, for any $s_i, \mathbf{r}_{-i}, \mathbf{s}_{-i}$, there exists r_i, such that for any s_i', $u_i(r_i, \mathbf{r}_{-i}, s_i', \mathbf{s}_{-i}) \le u_i(r_i, \mathbf{r}_{-i}, s_i, \mathbf{s}_{-i})$.*

Assumption 2.3 (Independent Self-Control)**.** *For any player i, for any $r_i, s_i, \mathbf{s}_{-i}$, for any $\mathbf{r}_{-i}$, $u_i(r_i, \mathbf{r}_{-i}, s_i, \mathbf{s}_{-i}) = u_i(r_i, \mathbf{0}, s_i, \mathbf{s}_{-i})$.*

Assumptions 2.1 and 2.2 are borrowed from Fudenberg and Levine [2006] whereas Assumption 2.3 was first introduced in Wang and Zheng [2012]. As the key difference between our model and Fudenberg and Levine [2006]'s model, Assumption 2.3 describes how one player's self-control action affects other players' payoffs, and hence makes it possible to extend the analysis of single-agent decision-making problems to multi-player games of strategic interactions in the context of self-control.

Definition 2.1 (Self-Control Cost)**.** *Given any short-run selves' strategy choice profile* $\mathbf{s}$*, let* $R_i(\mathbf{s}) = \{r_i \in R_i | u_i(r_i, \mathbf{0}, s_i, \mathbf{s}_{-i}) \geq u_i(r_i, \mathbf{0}, \cdot, \mathbf{s}_{-i})\}$*, then player i's self-control cost is defined as*

$$C_i(\mathbf{s}) = C_i(s_i, \mathbf{s}_{-i}) \equiv u_i(\mathbf{0}, \mathbf{s}) - \sup_{r_i \in R_i(\mathbf{s})} u_i(r_i, \mathbf{0}, \mathbf{s}).$$

For detailed discussions regarding the assumptions above, please refer to Wang and Zheng [2012].

Assumption 2.4 (Continuity)**.** *For any player i,* $u_i(r_i, \mathbf{r}_{-i}, s_i, \mathbf{s}_{-i})$ *is continuous in* r_i, s_i*.*

We have the following property regarding self control cost.

Property 2.1 (Strict Cost of Self-Control)**.** *Under Assumptions 2.1-2.4,*

$$s_i \in \arg\max_{s_i' \in S_i} u_i\left(0, \mathbf{0}, s_i', \mathbf{s}_{-i}\right) \Leftrightarrow C_i(s_i, \mathbf{s}_{-i}) = 0.$$

Please refer to Wang and Zheng [2012] for the proof of the above property.

Definition 2.2 (Optimal Self-Control)**.** *Given any player i and any strategy choice profile* $\mathbf{s}$*, an optimal self-control action* $r_{\mathbf{s}}^i$ *imposed by* LRS_i *satisfies the following two conditions: (1)* $C_i(s_i, \mathbf{s}_{-i}) = u_i(0, \mathbf{0}, s_i, \mathbf{s}_{-i}) - u_i\left(r_{\mathbf{s}}^i, \mathbf{0}, s_i, \mathbf{s}_{-i}\right)$ *and (2)* $s_i \in \arg\max_{s \in S_i} u_i\left(r_{\mathbf{s}}^i, \mathbf{0}, s, \mathbf{s}_{-i}\right)$*.*

It is easy to see from the above definition that $r_{\mathbf{s}}^i$ is such a self-control action for LRS_i that can ensure that SRS_i has no incentive to unilaterally deviate from $(s_i, \mathbf{s}_{-i})$ at the lowest possible cost. In this sense we call $r_{\mathbf{s}}^i$ an "optimal" self-control.

3. The Monks Story

Suppose that two monks, A and B, are playing an infinitely-repeated normal-form game. In each period, each monk has two actions available: E (Making Effort) and N (Making No Effort). Monk i's payoff function in period t is denoted by $u_i^t(s_A^t, s_B^t) : \{E, N\} \times \{E, N\} \longrightarrow \mathbb{R}$, where $i = A, B$, $t = 1, 2, \cdots$, and s_i^t is monk i's strategy in period t. Monk i's discount factor between any two consecutive periods is constant and denoted by δ_i, $\delta_i \in [0, 1]$, $i = A, B$. Hence monk i's total payoff is $\sum_{t=1}^{\infty} [\delta_i]^{t-1} u_i^t(s_A^t, s_B^t)$. The payoff matrix for the stage game is shown below, and we assume $c > a > 0$, $b > 0$, $2a > c - b$.

Stage Game

Monk B

Monk A		E	N
	E	(a, a)	$(-b, c)$
	N	$(c, -b)$	$(0, 0)$

We consider the following 3 scenarios and compare the results under different scenarios.

3.1. Scenario 1: History-Independent Strategy

In Scenario 1, we assume that the game in each period is considered independent. In other words, the monks' decisions in period $t+1$ are independent of the outcomes in periods $1,\cdots,t$. It is easy to know that in any period t, the equilibrium strategy profile is (N,N) and the equilibrium payoff profile is $(0,0)$. So the total payoffs are $(0,0)$.

3.2. Scenario 2: History-Dependent Strategy

In Scenario 2, we assume that monks' strategies can depend on the history of the outcomes. In this case, with the stage game repeated infinitely, it is possible for monks to cooperate in order to achieve higher payoffs. We are interested in the following equilibrium strategy profile, which achieves the highest equilibrium payoffs for both monks.

Monk i's strategy is as follows:

1. In period 1 he plays E;
2. In period $t\,(\geq 2)$ he plays $s_i^t = \begin{cases} E & \text{if } \left(s_A^{t'}, s_B^{t'}\right) = (E,E)\ \forall t' < t \\ N & \text{otherwise} \end{cases}$, $i = A,B$.

It is easy to verify that monk i has no incentive to deviate if and only if $\delta_i \in \left[\frac{c-a}{c}, 1\right]$, $i = A,B$. When $\min\{\delta_A, \delta_B\} \in \left[\frac{c-a}{c}, 1\right]$, monk i can achieve his highest payoff $\frac{a}{1-\delta_i}$, $i = A,B$.

3.3. Scenario 3: Dual-Self Approach

In Scenario 3, instead of adopting the single-self decision making model which is used in the first 2 scenarios, we apply the dual-self approach to our example. Suppose that each monk has a long-run self and a sequence of short-run selves (each of whom lives only for one period). The long-run self of monk i (LRS_i) can impose a costly self-control action $r_i^t \in R_i^t \subseteq \mathbb{R}$ in each period t, which may vary across different periods and will affect his own payoff (but not the other monk's payoff) at the current period. Under the above settings, monk i's payoff function in period t is denoted by $u_i^t\left(r_i^t, s_A^t, s_B^t\right) : R_i^t \times \{E,N\} \times \{E,N\} \longrightarrow \mathbb{R}$, where $i = A,B$, $t = 1,2,\cdots$, r_i^t is LRS_i's self-control in period t, and s_i^t is the strategy choice of the short-run self of monk i in period t (SRS_i^t).[2] Monk i's total payoff is thus $\sum\limits_{t=1}^{\infty} [\delta_i]^{t-1} u_i^t\left(r_i^t, s_A^t, s_B^t\right)$.

Let $C_i^t\left(s_A^t, s_B^t\right)$ be the self-control cost of monk i in period t, when monks' strategy profile in that period is $\left(s_A^t, s_B^t\right)$. Note that $r^{i,t}_{s_A^t, s_B^t}$ is the optimal self-control imposed by LRS_i in period t, given $i \in \{A,B\}$ and any strategy choice profile $\left(s_A^t, s_B^t\right)$ by SRS_A^t and SRS_B^t.

The Structure of Self-Control Cost Consider payoff structure of the stage game.

According to Assumption 2.1 and Property 2.1, in any period t, we have

$$\begin{aligned} C_A^t(E,E) &> C_A^t(N,E) = 0, C_A^t(E,N) > C_A^t(N,N) = 0, \\ C_B^t(E,E) &> C_B^t(E,N) = 0, C_B^t(N,E) > C_B^t(N,N) = 0. \end{aligned}$$

[2]For simplicity, we took the redundent term $\mathbf{r}_{-i}$ out of the expression of the payoff function, by Assumption 2.3.

By Definition 2.2, in any period t, we have

$$a > u_A^t\left(r_{EE}^A, E, E\right) \geq u_A^t\left(r_{EE}^A, N, E\right) \tag{3.1}$$
$$a > u_B^t\left(r_{EE}^B, E, E\right) \geq u_B^t\left(r_{EE}^B, E, N\right) \tag{3.2}$$
$$0 = u_A^t(0, N, N) \geq u_A^t(0, E, N) \tag{3.3}$$
$$0 = u_B^t(0, N, N) \geq u_B^t(0, N, E) \tag{3.4}$$

Inequality (3.1) means that in period t there exists a nonzero optimal self-control r_{EE}^A so that SRS_A^t chooses E over N when SRS_B^t chooses E. Similarly, inequality (3.2) means that in period t there exists a nonzero optimal self-control r_{EE}^B so that SRS_B^t chooses E over N when SRS_A^t chooses E.

Inequality (3.3) means that in period t there exists an optimal self-control $r_{NN}^{At} = 0$ so that SRS_A^t chooses N over E when SRS_B^t chooses N. Similarly, inequality (3.4) means that in period t there exists an optimal self-control $r_{NN}^{Bt} = 0$ so that SRS_B^t chooses N over E when SRS_A^t chooses N.

Note that when the self-control is too costly ($C_i^t(E,E) \geq a$, $i = A, B$), the long-run selves have no incentive to cooperate, because in that case each monk's per-period payoff is non-positive if they choose to be cooperative while their payoff is zero if they defect. In order to focus on the interesting case, we make the following assumption.

Assumption 3.1 (Gains from Cooperation)**.** *For $i = A, B$, monk i's self-control cost for strategy choice profile (E, E) in any period t must satisfy the following condition:*

$$C_i^t(E, E) < a.$$

3.3.1. Case 1: the short-run selves can observe the previous results

An Equilibrium Strategy We are interested in the following equilibrium strategy profile, which achieves the highest equilibrium payoffs for both monks.

Monk A's strategy is as follows:

A1 In period 1, LRS_A imposes a self-control $r_A^1 = r_{EE}^A$; SRS_A^1 chooses $s_A^1\left(r_A^1\right) = \begin{cases} E & \text{if } u_A^1\left(r_A^1, E, E\right) \geq u_A^1\left(r_A^1, N, E\right) \\ N & \text{otherwise} \end{cases}$.

A2 In period $t\,(\geq 2)$, LRS_A imposes $r_A^t(h_t) = \begin{cases} r_{EE}^A & \text{if } \forall t' < t, \left(s_A^{t'}, s_B^{t'}\right) = (E, E) \\ 0 & \text{otherwise} \end{cases}$; SRS_A^t chooses

$$s_A^t\left(h_t, r_A^t\right) = \begin{cases} E & \text{if} \begin{cases} u_A^t\left(r_A^t, E, E\right) \geq u_A^t\left(r_A^t, N, E\right), \forall t' < t, \left(s_A^{t'}, s_B^{t'}\right) = (E, E), \text{ or} \\ u_A^t\left(r_A^t, N, N\right) < u_A^t\left(r_A^t, E, N\right), \exists t' < t, \left(s_A^{t'}, s_B^{t'}\right) = (E, E) \end{cases} \\ N & \text{otherwise} \end{cases}.$$

Similarly, monk B's strategy is as follows:

B1 In period 1, LRS_B imposes a self-control $r_B^1 = r_{EE}^B$; SRS_B^1 chooses $s_B^1\left(r_B^1\right) = \begin{cases} E & \text{if } u_B^1\left(r_B^1, E, E\right) \geq u_B^1\left(r_B^1, E, N\right) \\ N & \text{otherwise} \end{cases}$.

B2 In period $t\,(\geq 2)$, LRS_B imposes $r_B^t(h_t) = \begin{cases} r_{EE}^B & \text{if } \forall t' < t, \left(s_A^{t'}, s_B^{t'}\right) = (E,E) \\ 0 & \text{otherwise} \end{cases}$;

SRS_B^t chooses

$$s_B^t(h_t, r_B^t) = \begin{cases} E & \text{if} \begin{cases} u_B^t(r_B^t, E, E) \geq u_B^t(r_B^t, E, N), \forall t' < t, \left(s_A^{t'}, s_B^{t'}\right) = (E,E), \text{ or} \\ u_B^t(r_B^t, N, N) < u_B^t(r_B^t, N, E), \exists t' < t, \left(s_A^{t'}, s_B^{t'}\right) = (E,E) \end{cases} \\ N & \text{otherwise} \end{cases}.$$

Analysis To see why the strategy profile described above forms a Subgame Perfect Nash Equilibrium, let us check whether monks have incentive to deviate or not.

It suffices to only consider monk A since the game is symmetric.

Also note that monk A's dual-self problem when $t = 1$ is exactly the same as the case when $t \geq 2$ and $\forall t' < t, \left(s_A^{t'}, s_B^{t'}\right) = (E,E)$. Thus the following analysis only focuses on the $t \geq 2$ case.

Consider SRS_A^t in period t:

History On the Equilibrium Path: If the outcome is $\left(s_A^{t'}, s_B^{t'}\right) = (E,E)$ for any $t' < t$:

Assuming that monk B (LRS_B from period t on and $\{SRS_B^s\}_{s=t}^{\infty}$), $\{SRS_A^s\}_{s=t+1}^{\infty}$ and LRS_A from period t on are playing the equilibrium strategies described above, LRS_B will impose self-control $r_B^t = r_{EE}^B$. By (3.2), $u_B^t\left(r_{EE}^B, E, E\right) \geq u_B^t\left(r_{EE}^B, E, N\right)$, so SRS_B^t plays E in period t. Given that LRS_A will impose self-control $r_A^t = r_{EE}^A$ in period t, the payoff of SRS_A^t is $u_A^t\left(r_{EE}^A, E, E\right)$ by playing E and his payoff is $u_A^t\left(r_{EE}^A, N, E\right)$ by playing N.

By (3.1), $u_A^t\left(r_{EE}^A, E, E\right) \geq u_A^t\left(r_{EE}^A, N, E\right)$, so SRS_A^t in period t has no incentive to deviate from playing E.

History Off the Equilibrium Path: If the outcome is $\left(s_A^{t'}, s_B^{t'}\right) \neq (E,E)$ for some $t' < t$:

Assuming that monk B (LRS_B from period t on and $\{SRS_B^s\}_{s=t}^{\infty}$), $\{SRS_A^s\}_{s=t+1}^{\infty}$ and LRS_A from period t on are playing the equilibrium strategies described above from period t on, LRS_B will impose a zero self-control $r_B^t = 0$. By (3.4), $u_B^t(0, N, N) \geq u_B^t(0, N, E)$, so SRS_B^t plays N in period t. Given that LRS_A will impose a zero self-control $r_A^t = 0$ in period t, the payoff of SRS_A^t is $u_A^t(0, E, N)$ by playing E and his payoff is $u_A^t(0, N, N)$ by playing N.

By (3.3), $u_A^t(0, N, N) \geq u_A^t(0, E, N)$, so SRS_A^t in period t has no incentive to deviate from playing N.

Consider LRS_A in period t:

History On the Equilibrium Path: If the outcome is $\left(s_A^{t'}, s_B^{t'}\right) = (E,E)$ for any $t' < t$:

Assuming that monk B (LRS_B from period t on and $\{SRS_B^s\}_{s=t}^{\infty}$), $\{SRS_A^s\}_{s=t}^{\infty}$ and LRS_A from period $t+1$ on are playing the equilibrium strategies described above, LRS_B in period t will impose self-control $r_B^t = r_{EE}^B$. By (3.2), SRS_B^t plays E in period t.

If LRS_A in period t follows the equilibrium strategy described above, his total payoff should be $\sum_{\tau=0}^{\infty} [\delta_A]^{\tau} u_A^t\left(r_{EE}^A, E, E\right) = \frac{1}{1-\delta_A} u_A\left(r_{EE}^A, E, E\right)$. However, if he deviates in period t by imposing a different self-control r_A^*, his total payoff would be $\begin{cases} u_A(r_A^*, E, E) + \frac{\delta_A}{1-\delta_A} u_A\left(r_{EE}^A, E, E\right) & \text{if } u_A^t(r_A^*, E, E) \geq u_A^t(r_A^*, N, E) \\ u_A(r_A^*, N, E) & \text{otherwise} \end{cases}$.

In order to make sure that LRS_A has no incentive to deviate in period t, we must have

$$\frac{1}{1-\delta_A} u_A\left(r_{EE}^A, E, E\right) \geq \max_{\substack{r_A^* \neq r_{EE}^A, \\ u_A^t(r_A^*, E, E) \geq u_A^t(r_A^*, N, E)}} u_A\left(r_A^*, E, E\right) + \frac{\delta_A}{1-\delta_A} u_A\left(r_{EE}^A, E, E\right) \quad (3.5)$$

$$\frac{1}{1-\delta_A} u_A\left(r_{EE}^A, E, E\right) \geq \max_{\substack{r_A^* \neq r_{EE}^A, \\ u_A^t(r_A^*, E, E) < u_A^t(r_A^*, N, E)}} u_A\left(r_A^*, N, E\right) \quad (3.6)$$

By Definition 2.2, we know

$C_A(E,E) = u_A(0,E,E) - u_A\left(r_{EE}^A, E, E\right) \leq u_A(0,E,E) - \max_{\substack{r_A^* \neq r_{EE}^A, \\ u_A^t(r_A^*, E, E) \geq u_A^t(r_A^*, N, E)}} u_A\left(r_A^*, E, E\right)$.

Hence $u_A\left(r_{EE}^A, E, E\right) \geq \max_{\substack{r_A^* \neq r_{EE}^A, \\ u_A^t(r_A^*, E, E) \geq u_A^t(r_A^*, N, E)}} u_A\left(r_A^*, E, E\right)$, which implies that (3.5) always holds.

Thus, in order to make sure that LRS_A has no incentive to deviate in period t, it suffices to have (3.6) hold. Since

$$\max_{\substack{r_A^* \neq r_{EE}^A, \\ u_A^t(r_A^*, E, E) < u_A^t(r_A^*, N, E)}} \left[u_A\left(r_A^*, N, E\right)\right] = u_A(0,N,E) = c \quad (3.7)$$

by Property 2.1, it suffices to have

$$\frac{1}{1-\delta_A} u_A\left(r_{EE}^A, E, E\right) \geq c \quad (3.8)$$

Solving for δ_A, we obtain the following result:

$$\delta_A \geq \frac{c - u_A\left(r_{EE}^A, E, E\right)}{c} > \frac{c-a}{c} \quad (3.9)$$

History Off the Equilibrium Path: If the outcome is $\left(s_A^{t'}, s_B^{t'}\right) \neq (E,E)$ *for some* $t' < t$*:*

Assuming that monk B (LRS_B from period t on and $\{SRS_B^s\}_{s=t}^{\infty}$), $\{SRS_A^s\}_{s=t}^{\infty}$ and LRS_A from period $t+1$ on are playing the equilibrium strategies described above, LRS_B will impose a zero self-control $r_B^t = 0$. By (3.4), SRS_B^t plays N in period t.

If LRS_A follows the strategy described above, his total payoff should be $\sum_{\tau=0}^{\infty} [\delta_A]^{\tau} u_A^t(0,N,N) = 0$. However, if he deviates in period t by imposing a different self-control r_A^*, his total payoff would be $\begin{cases} u_A\left(r_A^*, E, N\right) & \text{if } u_A^t\left(r_A^*, E, N\right) > u_A^t\left(r_A^*, N, N\right) \\ u_A\left(r_A^*, N, N\right) & \text{otherwise} \end{cases}$.

However, for any $r_A^* \in R$, we have $u_A\left(r_A^*, E, N\right) \leq u_A(0,E,N) = -b < 0$, and $u_A\left(r_A^*, N, N\right) \leq u_A(0,N,N) = 0$. Therefore in this case LRS_A has no incentive to deviate in period t.

Based on the analysis above, monk A has no incentive to deviate if and only if $\delta_A \geq \frac{c - u_A\left(r_{EE}^A, E, E\right)}{c}$.

Similarly, we can have an inequality for δ_B:

$$\delta_B \geq \frac{c - u_A\left(r_{EE}^B, E, E\right)}{c} > \frac{c-a}{c} \quad (3.10)$$

This means that monk B has no incentive to deviate if and only if $\delta_B \geq \frac{c-u_B\left(r_{EE}^B,E,E\right)}{c}$. We conclude our analysis with the following proposition.

Proposition 3.1. *In the infinitely-repeated dual-self monks game described in this subsection, where the short-run selves can observe the previous results, the strategies specified above form a Subgame Perfect Nash Equilibrium if and only if*

$$\delta_A \in \left[\frac{c-u_A\left(r_{EE}^A,E,E\right)}{c},1\right] \text{ and } \delta_B \in \left[\frac{c-u_B\left(r_{EE}^B,E,E\right)}{c},1\right].$$

3.3.2. Case 2: the short-run selves do not observe the previous results

An Equilibrium Strategy We are interested in the following equilibrium strategy profile and equilibrium belief profile, which achieve the highest equilibrium payoffs for both monks.

Let $\mu_i^t\left(s_{-i}^t|r_i^t\right):\{E,N\}\to[0,1]$ denote SRS_i^t's belief about the choice of SRS_{-i}^t in period t, given the self-control action r_i^t. Obviously we should require $\mu_A^t\left(E|r_A^t\right)+\mu_A^t\left(N|r_A^t\right)=1$, $\forall r_A^t\in R_i^t$.

Monk A's strategy is as follows:

A1 In period 1, LRS_A imposes a self-control $r_A^1=r_{EE}^A$; SRS_A^1 chooses $s_B^1\left(r_B^1\right)=\begin{cases} E & \text{if } u_A^1\left(r_A^1,E,E\right)\geq u_A^1\left(r_A^1,N,E\right) \\ N & \text{otherwise}\end{cases}$.

A2 In period $t\,(\geq 2)$, LRS_A imposes $r_A^t\left(h_t\right)=\begin{cases} r_{EE}^A & \text{if } \forall t'<t, \left(s_A^{t'},s_B^{t'}\right)=(E,E) \\ 0 & \text{otherwise}\end{cases}$; SRS_A^t chooses $s_A^t\left(r_A^t\right)=\begin{cases} E & \text{if } u_A^t\left(r_A^t,E,E\right)\geq u_A^t\left(r_A^t,N,E\right), r_A^t\neq 0 \\ N & \text{otherwise}\end{cases}$. SRS_A^t's belief is such that $\mu_A^t\left(E|r_A^t\neq 0\right)=1, \mu_A^t\left(N|r_A^t=0\right)=1$.

Similarly, monk B's strategy is as follows:

B1 In period 1, LRS_B imposes a self-control $r_B^1=r_{EE}^B$; SRS_B^1 chooses $s_B^1\left(r_B^1\right)=\begin{cases} E & \text{if } u_B^1\left(r_B^1,E,E\right)\geq u_B^1\left(r_B^1,E,N\right) \\ N & \text{otherwise}\end{cases}$.

B2 In period $t\,(\geq 2)$, LRS_B imposes $r_B^t\left(h_t\right)=\begin{cases} r_{EE}^B & \text{if } \forall t'<t, \left(s_A^{t'},s_B^{t'}\right)=(E,E) \\ 0 & \text{otherwise}\end{cases}$; SRS_B^t chooses $s_B^t\left(r_B^t\right)=\begin{cases} E & \text{if } u_B^t\left(r_B^t,E,E\right)\geq u_B^t\left(r_B^t,E,N\right), r_B^t\neq 0 \\ N & \text{otherwise}\end{cases}$. SRS_B^t's belief is such that $\mu_B^t\left(E|r_B^t\neq 0\right)=1, \mu_B^t\left(N|r_B^t=0\right)=1$.

Analysis The analysis of case 2 is analogous to that of case 1. We hence skip the details and conclude our analysis with the following proposition.

Proposition 3.2. *In the infinitely-repeated dual-self monks game described in this subsection, where the short-run selves do not observe the previous results, the strategies and beliefs specified above form a Perfect Baysian Nash Equilibrium if and only if*

$$\delta_A \in \left[\frac{c-u_A\left(r_{EE}^A,E,E\right)}{c},1\right] \text{ and } \delta_B \in \left[\frac{c-u_B\left(r_{EE}^B,E,E\right)}{c},1\right].$$

Remarks: Here are a few remarks regarding the results we have found.

1. In Scenario 3, in equilibrium the range for δ_i depends on monk i's self-control cost structure, whereas in Scenario 2, the range for δ_i is the same for both monks and does not depend on the monks' self-control cost structure.

2. In Scenario 3, the highest equilibrium lifetime payoffs (with the strategies specified earlier in this subsection) are $\left(\frac{1}{1-\delta_A}u_A\left(r_{EE}^A,E,E\right),\frac{1}{1-\delta_B}u_B\left(r_{EE}^B,E,E\right)\right)\geq(c,c)$, whereas in Scenario 2, the highest equilibrium payoffs are $\left(\frac{a}{1-\delta_A},\frac{a}{1-\delta_B}\right)\geq(c,c)$. The difference in equilibrium payoffs between Scenarios 2 and 3 is $\frac{C_i^1(E,E)}{1-\delta_i}$, $i=A,B$.

3. In Scenario 3, the equilibrium range for δ_i in Proposition 3.2 (for case 2) is exactly the same as that in Proposition 3.1 (for case 1), whereas the equilibrium concepts and the equilibrium strategies are different. Based on the results in cases 1 and 2, it is worth noting that less information in the dual-self monks game does not necessarily lead to a decrease in efficiency.

4. We compare our approach with the standard coordination explanation. Although we both study efficiency issue, our approach provides different perspectives. The standard one examines how efficiency is compromised as the number of players i.e. monks increases. Our approach focuses on how cooperation is achieved through self-control actions given a fixed number of players and how efficiency depends on information: in an infinitely repeated setting, the long-run selves of monks have incentive to impose optimal self-control so that the short-run selves of monks will interact in a cooperative way.

4. Conclusion

The concept of dual-self refers to the game played between a long-run patient self and a sequence of short-run impulsive selves. We expand the single-player decision problem, which is the focus of current literature, into a multi-player decision problem. Such an expansion is in accordance of our belief that a multi-player dual-self model would be a step closer to depicting individuals interacting with each other in reality. We thus propose a dual-self model that adopts the theoretical framework established by Fudenberg and Levine [2006] mainly for two reasons: First, the concept of two selves used in Fudenberg and Levine

[2006] is consistent with the previous literature; second, their DSM provides advancements to O'Donoghue and Rabin [2001]'s model with quasi-hyperbolic discounting.

Our assumptions on self-control are consistent with the axioms of Fudenberg and Levine [2006], which are Assumption of Costly Self-Control, Assumption of Unlimited Self-Control, and Assumption of Continuity. We add a new assumption, Assumption of Independent Self-Control, to our model in order to show the interactions among multiple players in the context of games.

In this paper, we use an example of The Monks Story to show the interaction effects among players as well as between long-run self and short-run selves within each single player over time, which is in line with the two-dimensioned game that our model depicts: a game played among multiple players as well as a game played between short-run selves and long-run self. It is worth noting that this story is commonly discussed in marketing, human resource and management studies to show coordination and cooperation effects as the number of players increases, while we offer a different perspective to look at the same example. Moreover, we use our proposed dual-self model to analyze the Monks Story and compare the results found in three scenarios such as history-independent strategy case, history-dependent strategy case, and dual-self approach. In the dual-self scenario, we show results in two cases: 1) the short-run selves can observe the previous outcomes; and 2) the short-run selves cannot do so.

In conclusion, we would like to point out that our current work on dual-self has not taken into account the impact of social preferences such as how the degree of being selfish and altruistic can affect dual-self individual decision-making problems. This could be a direction for future work.

Acknowledgments

This research is supported by National Science Foundation of China Young Scholar Research Grant (grant number: 71203112), Tsinghua University Independent Research Project Grant (grant number: 2012Z02181) and National Institute of Fiscal Studies, Tsinghua University. We are indebted to David K. Levine and Chong-En Bai for their generous advice and continuous support. We would like to thank Jaimie W. Lien for the constructive suggestions and beneficial discussions. We are grateful to Jason Shachat, Yongchao Zhang and other session participants at the China International Conference on Game Theory and Applications for their helpful comments. All errors are our own.

References

[1] N. Ashraf, C. Camerer, G. Loewenstein, *Adam Smith, Behavioral Economist*, The Journal of Economic Perspectives **19(3)** (2005), 131-145.

[2] J. Benhabib, A. Bisin, *Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions,* Games and Economic Behavior **52(2)** (2004), 460-92.

[3] B. D. Bernheim, A. Rangel, *Addiction and Cue-Triggered Decision Processes*, American Economic Review **94(5)** (2004), 1558-1590

[4] R. Binabou, M. Pycia, *Dynamic Inconsistency and Self-Control: A Planner-Doer Interpretation*, Economics Letters **77(3)** (2002), 419-24.

[5] C. Camerer, G. Loewenstein, D. Prelic, *Neuroeconomics: Why Economics Needs Brains,* Scandinavian Journal of Economics **106(3)** (2004), 555-579.

[6] K. Chatterjee, R. V. Krishna, *A 'Dual Self' Representation for Stochastic Temptation*, American Economic Journal: Microeconomics **(1:2)** (2009), 148–167.

[7] S. DellaVigna, U. Malmendier, *Contract Design and Self-Control: Theory and Evidence*, Quarterly Journal of Economics **119(2)** (2004), 353-402.

[8] E. Dekel, D. Fudenberg, D. K. Levine, Learning to Play Bayesian Games, Games and Economic Behavior **46(2)** (2004), 282-303.

[9] E. Derel, B. L. Lipman, A. Rustichini, *Temptation-Driven Preferences*, Review of Economic Studies Limited **76** (2009), 937–971.

[10] D. Fudenberg, D. K. Levine, *The Theory of Learning in Games*, Cambridge, MA: MIT Press, 1998.

[11] D. Fudenberg, D. K. Levine, *A Dual-Self Model of Impulse Control*, American Economic Review **96(5)** (2006), 1449-1476.

[12] D. Fudenberg, D. K. Levine, *Risk, Delay, and Convex Self-Control Costs*, American Economic Journal: Microeconomics **3(3)** (2011), 34-68(35).

[13] D. Fudenberg, D. K. Levine, *Timing and Self-Control*, Econometrica **80(1)** (2012), 1–42.

[14] F. Gul, W. Pesendorfer, *Temptation and Self-Control*, Econometrica **69(6)** (2001), 1403-35.

[15] F. Gul, W. Pesendorfer, *Self-Control and the Theory of Consumption*, Econometrica **72(1)** (2004), 119-58.

[16] G. Loewenstein, *Emotions in Economic Theory and Economic Behavior*, American Economic Review **90(2)** (2000), 428.

[17] G. Loewenstein, T. O'Donoghue, *Animal Spirits: Affective and Deliberative Processes in Economic Behavior*, CAE Working Paper #04-14 (2004).

[18] H. M. Shefrin, R. H. Thaler, *An Economic Theory of Self-Control*, Journal of Political Economy **89(2)** (1981), 392-406.

[19] W. Wang, J. Zheng, *Multi-Period Complete-Information Games with Self-Control: a Dual-Self Approach*, Operations Research Transactions **16(4)** (2012), 95-104.

[20] W. Wang, J. Zheng, *Games with Dual-Self Players*, Working Paper (2012)