A FEEDBACK APPROACH OF ENDOGENOUS PREFERENCES

WEI ZHANG

Job Market Paper

November, 2008

Abstract. We study intertemporal decision making problems with endogenously changing preferences. A preference state is defined, which affects the decision maker's instantaneous utility. Each period this state is, in turn, determined by past actions chosen by the decision maker. We formulate the dynamic process as a simple closed-loop feedback system: current state as an input is fed into the decision maker's preference, which determines the action chosen in the current period, this then produces the new state and starts a new round of looping. In particular, we examine *positive feedback systems* in which instantaneous optimization always drives the state further and further away from the initial state. Outcomes are investigated according to the decision maker's different levels of self-knowledge of his own endogenous preference. For welfare comparisons, we use the initial preference, and show that under a mild sufficient condition a sophisticated person is always better off than a naive person. However the payoff is in general not monotonic in the level of self-knowledge.

KEYWORDS: Endogenous Preferences, Changing Tastes, Time-inconsistency, Rational Addiction, Self-control, Positive Feedbacks, Sophisticated and Naive Decision Makers

JEL classifications: D03, D60, C73

Department of Economics, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, WI 53706-1393 (wzhang3@wisc.edu). I am grateful to William Sandholm, Larry Samuelson, Ricardo Serrano-Padial, and Daniel Quint for advice and comments.

1. Introduction

1.1. **Motivation.** Economists have long been paying attention to "changing tastes" in intertemporal choices. At the broadest level, there are two distinct points of view on changing tastes. Best illustrated by Becker and Murphy (1988), the "rational addiction" literature (Stigler and Becker (1977), Iannaccone (1986), Loewenstein, O'Donoghue, and Rabin (2003)) studied choices that have intertemporal consequences in that they alter future tastes, e.g. the phenomena of habit-formation and addiction. A fundamental assumption in these models is that an individual holds the same objective across periods as in the "rational choice framework". That is, an individual at an earlier stage of the intertemporal decision maximizes the (discounted) sum of future instantaneous utility as evaluated *instantaneously*. In other words, at any point of time the individual approves of the anticipated changing tastes and makes forward-looking choices according to a stable preference. Hence in a rational addiction model, there is no conflict between the selves in different periods, and preferences are time consistent. Conceptually, acquired tastes for wine or classical music fit the premise of the rational addiction approach to changing tastes very well. For instance, one may force himself to listen to Bach or Shostavokich today knowing that in time he will learn to appreciate the music and benefit a great deal from the enjoyment in the future. However, in situations that changing tastes are usually considered bad by an individual, e.g. addiction to cigarette, it is hardly convincing that one should agree with the taste of the future addicted self and optimize accordingly.

The second line of work on changing taste is the "inconsistent selves" approach. These models hypothesizes changing tastes as changes of preferences and conflicts between the selves of different periods, as well as explanations of some observed inconsistent behaviors in intertemporal choices. In such a model, at each point of time, an individual needs to play *strategically* against the future selves with potentially different tastes. Hence changing tastes pose a "self-control problem" on an earlier period self. In this paper we develop a model of changing preferences on the same ground.

Most of the literature of inconsistent multi-selves has focused on *time-inconsistency*, where the change of preference is the result of placing an individual at different points of time. In particular, present-biased preferences, developed by Strotz (1956), Phelps and Pollak (1968) and epitomized in Laibson (1997), O'Donoghue and Rabin (1999), lead people to gravitate towards immediate awards and to avoid immediate cost more than a constant discount factor warrants. In the simpliest formulation, at any time t the utility function takes the form of

$$U^t(u_t, u_{t+1}, \cdots, u_T) = \delta^t u_t + \beta \sum_{\tau=t+1}^{T} \delta^\tau u_\tau$$

where $0 < \delta \leq 1$ represents the normal discount factor and $0 < \beta \leq 1$ measures the bias for the present. The smaller β is, the bigger the bias is. And whenever $\beta < 1$, the decision process is a game between multiple selves, one in each period whose distinctive preference is captured by the above equation. In these models, the way preference changes is exogenously given and depends solely on the point of time the decision maker (DM) evaluates events. The DM can do nothing to prevent his preference from changing although the "sophisticated" ones (as opposed to "naive") know about this change. Secondly, inconsistent behavior in these models is not only driven by the inconsistent preference alone. It also depends on the payoff structure. If the payoff, costs and rewards, of an action is realized within the same period as will be assumed in our model, the "$\beta - \delta$ preference" would not predict any behavior inconsistency for sophisticated or naive DMs. Since then the action that maximizes the instantaneous payoff has nothing to do with the size of β and the DM will simply repeatedly choose such an action.

In contrast to the $\beta - \delta$ preference, a formulation of endogenously changing preferences in intertemporal decisions is proposed here. The premise of the model is that people's past decisions often play a significant role in shaping their future preferences. This makes the preference change an endogenous component of the decision process. For example, the objective of winning an item for under ten dollars in an auction, sometimes turns into winning the bidding competition itself, regardless of the cost. A short period of "retail therapy" might turn into chronic shopping sprees. A teenager experimenting with cigarettes, alcohol, or drugs later on finds it impossible to stop the substance abuse. These phenomena take many forms we are familiar of: emotional investment of the bidder, habit forming of the shopper, addiction of the teenager, etc. What is common in the above observations is that, all the decision makers later find themselves doing things different from what they set out to achieve earlier because their earlier actions have endogenously changed their preferences. For instance, in the online auction example, it is the choice of entering the bidding earlier that changes the bidder's objective from buying the item within a certain budget to "winning the competition" itself (Ariely and Simonson (2003)). Had the individual not entered the bidding and got emotionally involved, his objective would still be buying the item for under ten dollars.

To model endogenous preferences, we introduce a "preference state", which helps capture how the history of actions influences the DM's instantaneous payoff evaluation. A deterministic updating rule maps history into the current state. In each period, the DM chooses an action from the same set of possible actions, and payoff is realized before the new state updating process. Assumptions on the agent's self-knowledge of his future preferences is necessary to any study of behaviors under inconsistent preferences. [1] Following O'Donoghue and Rabin (1999), we call a person *sophisticated* if he knows perfectly how his future selves' preferences will be endogenously affected by past and current choices. A DM is *naive*, on the other hand, if he is completely oblivious of this possible change and believes that all future selves possess the same preference as he has at the moment. Besides the two extreme types, we will also look into DMs with *limited self-knowledge*.

Compare to the $\beta - \delta$ preference, time alone plays no role in the change of preference here. An intertemporal setting simply provides the arena for the state, hence the preference, to change over time. Moreover, our model does not mandate preference change. By choosing proper current action, a DM can potentially keep his next period self from having a different objective. Only such "control" is exercised at the cost of choosing an action that might not be instantaneously optimal. As will be demonstrated later, the crux of a sophisticated person's decision is balancing the cost and benefit of this control.

[1] In fact, in both the rational addiction and inconsistent selves approach, one can postulate different levels of self-knowledge of the changing taste on the DM. Loewenstein, O'Donoghue, and Rabin (2003), for example, assumed a "projection bias" in the prediction of future tastes in the rational addiction framework.

Another conceptual difference between our model and time-inconsistent preferences lies in that time-inconsistency is a global, not decision specific, assumption on a person. Hence one of the implications of this theory is that the DM has to exhibit the same inconsistent tendency on every decision he makes that involves multiple periods of decisions. In reality, we might observe that a "naive" person yields to immediate gratification on one thing (cheese cake vs. diet) but not the other (watching a movie vs. cleaning the house). More importantly, the time-inconsistency assumption precludes endogenous reasons of a multi-period decision as a driving force for the observed outcome of inconsistent behaviors. Our approach to preference change is decision specific, and provides an alternative explanation of inconsistent behaviors especially in the context of repeated choices of actions and preference or habit development, where the status quo of a person's preference is very much likely to be partly the result of past choices. In a *strategic* setting with repeated interactions, the model can also be used to study emotional responses triggered by endogenous play between players and the corresponding modification of players' preferences.

In addition to the above mentioned papers, Gruber and Koszegi (2001) incorporated time-inconsistent preferences in the rational addiction model. Fudenberg and Levine (2006) provided a dual-self model where a long run self faces a series of short run selves in each period. A paper by Ozdenoren, Salant, and Silverman (2008) modeled "willpower", a limited resource in helping execute self-control, which is allocated over time and between different activities. It is also worth mentioning that in a series of papers, Gul and Pesendorfer (2001), (2004), (2005), (2007) provided a "revealed preference" approach of changing tastes by taking as the model primitive the preference on the entire intertemporal choice problem and axiomatizing the conditions on the preference under which the observed inconsistent choices arise as optimal. They stressed that the preference ranking applies to the selves of all periods and therefore is time-consistent. However, preferences are defined over current and future actions as well as actions not chosen in the past, and Noor (2005) pointed out a conceptual challenge in the primitive of their model.

1.2. **Outline.** In the next section, we start by presenting the model primitives. This involves introducing the aforementioned preference state and its role in determining instantaneous utility. A deterministic state update function is defined to incorporate how the states are endogenously changed by past choices.

To give a tighter structure to the endogenous state updating process, we formulate the decision problem as a *closed-loop feedback control system*: each period, the state is considered an input that is fed into the "controller" of the system – the DM's instantaneous utility; the controller then guides (or biases) the DM's current choice of action; next payoff is realized and state updated and fed into another round of looping. This feedback process is a closed loop [2] in the sense that, early period DM cannot employ any device other than the periodic action itself to influence later states. Hence the state update function is also the instrument a sophisticated person utilizes to exercise control. For this reason, we shall also call the state update function the "control function" when this meaning is invoked.

Such a formulation helps in characterizing the nature of the endogenous preference and answers questions like, does a temporary departure of the state from the initial state tends to lead it back to the starting point or further astray? The former case can be characterized as a "negative feedback". The latter, an amplification of the departure, is a "positive feedback". In this paper, we study intertemporal decision problems that are positive feedbacks by nature in that instantaneous optimization causes the states to go further and further away from the initial state. We focus on positive feedbacks because negative feedback systems tend to re-establish or oscillate around the initial state, therefore does not pose a significant challenge to investigations approximated with fixed preferences in economics. Decisions that are naturally positive feedbacks, for example habit formation and addictions, requires examination without the fixed preference assumption.

We also formally define the two types of decision makers and their decision problems in this section. A sophisticated person knows perfectly the changeability of his preference and the endogenous process, so his problem involves playing strategically against future selves and costly self-control. A naive person, on the other hand, believes that he has a fixed preference

[2]A *closed-loop* control system is one in which the control action is somewhat dependent on the output. In contrast, an *open-loop* control system is one in which the control action is independent of the output. (DiStefano, Stubberud, and Williams (1995))

therefore engages in simple myopic optimization. We then move on to dissect a sophisticated person's endogenous incentives.

In Section 3, we introduce our main result: a sufficient condition for the welfare comparison that a sophisticated person is always better off than his naive counterpart using the initial preference state as the payoff criterion. This condition states an increasing marginal cost of control. Under a simplifying assumption of the control function, the outcomes of the two types of decision makers are compared. It is demonstrated that a sophisticated person exerts some self-control at every period of the decision and this makes him better off than a naive person in a positive feedback decision. Section 4 follows with some discussions on generalizations and future research. In particular, we briefly talk about the decision maker with limited self-knowledge and some novel issues of endogenous preference in strategic games. Section 5 concludes and proofs are gathered in the appendix.

2. The feedback system

2.1. **The Decision Maker's Problem.** In each period t, of total T periods with no discounting, a decision maker (DM) chooses an action $p_t \in \mathbf{R}$. The instantaneous utility of the DM is determined by both the action chosen and the current state $e_t \in \mathbf{E}$ according to the continuous function $U(e_t, p_t)$. For example, in a smoking decision, p_t represents how much a person chooses to smoke at period t and e_t measures his current addiction level. We date the periods backwards for expositional convenience. So period T is chronically the first period, and period 1 the last. For each period $t < T$, the state e_t is endogenously determined by the time-invariant function $e_t = E(e_{t+1}, p_{t+1})$ and e_T is exogenously given. [3] [4] The function $E(\cdot, \cdot)$ captures the intertemporal consequences of actions as well as the dependence of the status quo utility on the past. One can imagine that the current addiction level of some substance relies on the addiction level in the past and the history of consumption of that substance.

[3]Alternatively, one can let e_t depend on the entire history of past actions. Our setup treats that the effects of actions before time $t+1$ on e_t are lumped into e_{t+1}.

[4]We use e to denote the preference state here, having in mind "emotions" (or emotional involvement) as a very important reason to cause preference to change (see for example Elster (1998)).

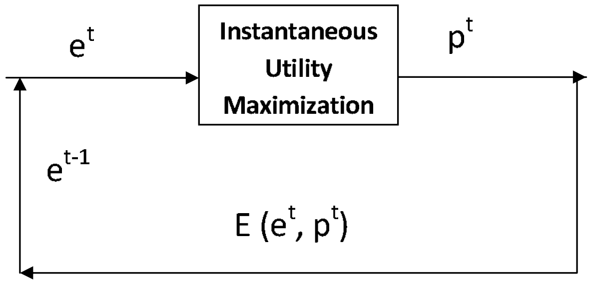

FIGURE 1. THE FEEDBACK LOOP

To characterize such an intertemporal decision making that features endogenous preference, we first describe it as a "feedback loop" illustrated in Figure 1. Given a period t state e_t, a "controller" chooses an action determining the next period state, which is fed back into the controller in the next round of recursion or looping. In the current context, the natural benchmark of the "controller", in order to characterize the nature of the DM's tradeoffs, is the instantaneous utility maximization. This controller ensures the most immediate benefit and the cost in the future is measured by how much the newly produced state differs from the current. With this definition of the controller, a "positive feedback" is one that each round of instantaneous utility maximization brings the state further and further away from the initial state e_T.

We make the following assumptions to construct a positive feedback system. Let $\mathbf{E} = [\underline{e}, \bar{e}] \in \mathbf{R}$. To avoid payoff ties, some assumptions are stated in strict inequalities and strict monotonicity.

Assumption 1. There is a unique action $P^*(e) \in \mathbf{R}$ that maximizes $U(e, p)$ for each $e \in \mathbf{E}$. $U(e, p)$ is strictly increasing in p to the left of $P^*(e)$, and strictly decreasing in p to the right of $P^*(e)$.

So there is a unique action that maximizes the instantaneous payoff evaluated at a state e. Deviating from the optimum is always costly to this instantaneous utility. The following assumption helps determine the effect of the state e on the optimal action $P^*(e)$.

Assumption 2. $U(e,p)$ is strictly supermodular. That is, given any $e > e' \in \mathbf{E}$ and $p > p' \in \mathbf{R}$, we have

$$U(e,p) + U(e',p') > U(e,p') + U(e',p).$$

This implies that the marginal utility of choosing a higher action increases with the state. The higher the current state is, the higher the action the DM of the period is inclined to choose. In the smoking example, the assumption implies that the more addicted a person is, the more cigarettes are needed to satisfy the current craving. By Topkis (1978), the optimal action $P^*(e)$ is increasing in e. [5]

Assumption 3. $E(e,p)$ is continuous and strictly increasing in each argument and $E(e,P^*(e)) \geq e$ for any $e \in \mathbf{E}$.

The operation of $E(e,P^*(e))$ finishes a feedback loop depicted earlier. Define the *feedback operator* $F : \mathbf{E} \mapsto \mathbf{E}$ by $Fe = E(e,P^*(e))$. The operator applied k times is denoted by F^k. So $F^k e = F(F^{k-1}e) = \cdots = F^{k-1}(Fe)$. By Assumption 2 and 3, it is apparent that $Fe' \geq Fe$ if $e' > e$, and $F^k e \geq F^{k-1}e \geq \cdots \geq Fe \geq e$. Hence we constructed a positive feedback system. Formally, let $F^0 e = e$.

Definition 1.[6] A decision making problem $\{U(e,p); E(e,p)\}_{e\in\mathbf{E},p\in\mathbf{R}}$ is a *positive feedback system* if $F^k e \geq F^{k-1}e$ for any $e \in E$, $k \geq 1$.

A few comments are to be made for the state update function $E(e,p)$. First, by the assumptions, for any $e \in \mathbf{E}$, there always exists a $p \in \mathbf{R}$ such that $E(e,p) = e$. This ensures that the primitives of the model never mandate the preference to change from period to period. Change of states through time, hence change of preference, is induced by the endogenous choice of actions in previous periods through $E(e,p)$. Second, as mentioned in the Introduction, the state update function is the sole mechanism through which early actions

[5]With differentiability, the assumption can be simplified as

$$\frac{\partial^2 U(e,p)}{\partial e \partial p} > 0 \text{ and } \frac{\partial^2 U(e,p)}{\partial p^2} < 0$$

However, differentiability is not necessary in any part of the analysis.

[6]Alternatively, a positive feedback can be defined as $E(e,P^*(e)) \leq e$, for any $e \in \mathbf{E}$. Without loss of generality, we assume a round of instantaneous utility maximization always brings the next period state larger than the current state, but this direction has no impact on any results.

can have any effect on later states. This implies that $E(e, p)$ is also the control function an earlier DM can use to implement some "control" on later selves. So self-control is always exercised with some cost.

Lastly, in the setup, the most general and realistic functional form $e_{t-1} = E(e_t, p_t)$ is assumed so that the next period state depends on both the current state and the current action. For example, a person's current level of addiction to alcohol and current year's average consumption of alcohol are both relevant in his next year's addiction level. However, for reasons that will be discussed in detail in the next section, it is very hard to make definitive predictions for the general setup. Results in Section 4 are, instead, derived for the special "state-independent" update function $e_{t-1} = E(p_t)$. This is, without any question, a serious restriction in the current context. Later on, this restriction will be relaxed. Nevertheless, we think that the state-independent update function still captures the fundamental nature of the problem. Although only p_t directly determines e_{t-1}, actions before time t also have indirect effect on it through the following endogenous mechanism. p_t is chosen by time t DM under the momentary objective indexed by e_t. By the model assumptions, actions other than $P^*(e_t)$ is not instantaneously optimal, so p_t must be chosen by the tradeoffs of e_t. In the same way, e_t is directly determined by p_{t+1}, which is chosen according to the interest of state e_{t+1} DM, and so on and so forth.

In terms of the "control-function" perspective of $E(p)$, state-independence does not imply that DMs at all states are equally effective in changing the next period state. Although the same action p leads to the same immediate future state $E(p)$, the same action comes at different costs for DMs of different states, measured by the difference $U(e, P^*(e)) - U(e, p)$. The higher the cost is, the less effectively the control can be exercised.

2.2. **Types of Decision Maker.** In decision problems involving inconsistent preferences, a person's knowledge of this inconsistency needs to be specified before outcome criterion can be properly defined. In this section, we introduce the two extreme types: the *sophisticated* person who fully understand the endogenous changeability of the future states, and the *naive* person who at any moment of time believes that all future selves will have the same state regardless of actions chosen. We use the solution concept of *perception-perfect strategy* by O'Donoghue and Rabin (1999): "a strategy that in all periods a person chooses the optimal action given his current preference and his perceptions of future behavior." Notation-wise, for presentational convenience, we shall use $DM_t^A(e)$ (similarly $DM_t^B(e)$) to stand for the sophisticated (naive) decision maker who is given the state e and t periods of actions left to choose.

For the sophisticated DM, denote by $\mathbf{p}_T^A(e) = (p_T^A(e), p_{T-1}^A(e), \cdots, p_t^A(e), \cdots, p_1^A(e))$ the equilibrium outcome of actions of the $DM_T^A(e)$'s problem, and by $\mathbf{e}_T^A(e) = (e, e_{T-1}^A(e), \cdots, e_t^A(e), \cdots, e_1^A(e))$ with given e the outcome of states. $\mathbf{p}_T^B(e)$ and $\mathbf{e}_T^B(e)$ are the counterparts for the naive person. (Figure 2 provides a summary of notations.)

The naive person's decision problem is rather simple. If a DM is not aware of the preference change, at every period the choice of action is made to maximize the instantaneous payoff. The outcome $\mathbf{p}_T^B(e)$ and and $\mathbf{e}_T^B(e)$ are uniquely defined by the following algorithm:

$$
\begin{aligned}
p_T^B(e) &= P^*(e) \\
&\quad \text{and} \;\; e_{T-1}^B(e) = E(e, P^*(e)) = Fe \\
p_{T-1}^B(e) &= P^*(e_{T-1}^B(e)) = P^*(Fe) \\
&\quad \text{and} \;\; e_{T-2}^B(e) = E(Fe, P^*(Fe)) = F^2 e \\
&\vdots \\
&\quad \text{and} \;\; e_1^B(e) = F^{T-1} e \\
p_1^B(e) &= P^*(e_1^B(e)) = P^*(F^{T-1} e)
\end{aligned}
$$

The naive person's T period problem coincides with the feedback operation with $T-1$ times of looping. The following observation is implied by the model assumptions:

Lemma 1. The naive person chooses increasing actions with time and has increasing states over time. Or equivalently, $p_T^B(e) \leq p_{T-1}^B(e) \leq \cdots \leq p_1^B(e)$ and $e \leq e_{T-1}^B(e) \leq \cdots \leq e_1^B(e)$. Also a naive person with a higher initial state chooses a path of higher actions: $\mathbf{p}_T^B(e) \leq \mathbf{p}_T^B(e')$ if $e < e'$. [7]

[7] In all vector comparisons, $(a_n, a_{n-1}, \cdots, a_1) \leq (\geq)(b_n, b_{n-1}, \cdots, b_1)$ if and only if $a_k \leq (\geq) b_k$, for all $k = 1, \cdots, n$.

A sophisticated person's problem is more complicated, because he is aware that his preference is constantly evolving and at each future period, a potentially different objective function can be used to make decisions. Take the DM with two remaining actions $DM_2^A(e)$. Once an action p_2 is chosen, the last period state will be determined by $e_1 = E(e, p_2)$ and he should expect his future self $DM_1^A(e_1)$ to choose $p_1 = P^*(e_1)$. Such a continuation (p_2, p_1) taking into consideration the optimal choice of the future will be referred to as a *feasible* continuation for $DM_2^A(e)$. Similarly, for a $DM_3^A(e)$, choosing an action p_3 updates the state to $e_2 = E(e, p_3)$ and he should expect $DM_2^A(e_2)$ to choose nothing but the (p_2, p_1) that is both feasible and optimal in the point of view of $DM_2^A(e_2)$. Hence such a vector of actions (p_3, p_2, p_1) is called a feasible continuation for $DM_3^A(e)$. A feasible continuation for $DM_t^A(e)$ is defined recursively in this way. Therefore, the problem of a sophisticated $DM_T^A(e)$ is to choose an optimal continuation of actions that is feasible, or equivalently, he needs to solve the following maximization by $\mathbf{p}_T^A(e)$ and $\mathbf{e}_T^A(e)$, ensuring that the incentives of all future selves are satisfied.

$$
\begin{aligned}
& \max_{(p_T, p_{T-1}, \cdots, p_1)} \sum_{\tau=T}^{1} U(e, p_\tau) \\
\text{s.t.} \quad & e_t = E(e_{t+1}, p_{t+1}) \quad (t = T-1, T-2, \cdots, 1) \\
(IC_{T-1}) \quad & \sum_{\tau=T-1}^{1} U(e_{T-1}; p_\tau) \geq \sum_{\tau=T-1}^{1} U(e_{T-1}; q_\tau) \text{ for any feasible } (q_{T-1}, \cdots, q_1) \text{ for } DM_t^A(e_{T-1}) \\
& \vdots \\
(IC_t) \quad & \sum_{\tau=t}^{1} U(e_t; p_\tau) \geq \sum_{\tau=t}^{1} U(e_t; q_\tau) \text{ for any feasible } (q_t, \cdots, q_1) \text{ for } DM_t^A(e_t) \\
& \vdots \\
(IC_1) \quad & p_1 = P^*(e_1)
\end{aligned}
$$

For a general update function $E(e, p)$, the feasibility of a certain continuation $(p_t, p_{t-1}, \cdots, p_1)$ is specific to each time t state e_t, because different e_t updates to different $e_{t-1} = E(e, p_t)$ and implies distinct incentives for $DM_{t-1}^A(e_{t-1})$, and so on. The problem with state-independent update function $E(p)$, on the other hand, is greatly simplified since the feasibility of a continuation is independent of the current state. The next subsection is devoted to dissecting the sophisticated person's decision problem defined above.

Besides characterizing and comparing the outcomes of the two types of the decision maker, another important task of the project is the welfare comparison. As Rubinstein (2006) mentioned, "if an agent is a collection of selves, ..., behavioral economics makes it clear that one must make assumptions about the relation between the preferences that explain behavior and the preferences used in the welfare criterion." In this model, we use the DM's initial preference as the basis of welfare analysis. Given the setup of the endogenous preference change, at the onset of the entire decision process, only the initial state determined preference is guaranteed to be carried out and all future selves of the DM are subordinate to the initial period DM's interest and choice. Notations related to total payoffs are now introduced. Define

$$\pi(e', \mathbf{p}_T^X(e)) = \sum_{\tau=T}^{1} U(e', p_\tau^X(e))$$

where $X = A, B$ as the total payoff of each type of DM's equilibrium continuation of T periods evaluated at the state e'. They are simplified to $\pi(\mathbf{p}_T^A(e))$ and $\pi(\mathbf{p}_T^B(e))$ if $e' = e$. So in welfare comparisons, the main focus will be the relationship between $\pi(\mathbf{p}_T^A(e))$ and $\pi(\mathbf{p}_T^B(e))$.

2.3. **The Conundrum of the Sophisticated.** In this part, we look at the sophisticated person's dynamic and endogenous incentive problem defined above. First, consider the easiest case when a sophisticated DM has two periods of actions to choose. The tradeoff of $DM_2^A(e_2)$ lies in the benefit and cost of one round of instantaneous utility maximization. The benefit is clearly to ensure the best payoff at the current period. The nature of the positive feedback dictates that the cost of doing so is to have a future self with a higher state $e_1 = Fe_2 \geq e_2$. The future self $DM_1^A(e_1)$ is also the last to make a choice with no further future to consider, so it is obvious that his action will be $P^*(e_1)$. Facing this tradeoff and having the control function $e_1 = E(e_2, p)$, at the initial period $DM_2^A(e_2)$ will choose an action equal to or lower

Notation	Explanation
$DM_T^A(e)$ (or $(DM_T^B(e))$)	a sophisticated (or naive) decision maker endowed with state e and T remaining actions to choose
$P^*(e) = \arg\max_p U(e,p)$	instantaneously optimal action
$Fe = E(e, P^*(e)) \geq e$	the feedback operation of state e by an instantaneous utility maximization
$\mathbf{p}_T^A(e) = (p_T^A(e), p_{T-1}^A(e), \cdots, p_t^A(e), \cdots, p_1^A(e))$	equilibrium action path of the remaining T periods of a sophisticated DM's problem with state $e_T = e$
$\mathbf{p}_T^B(e) = (p_T^B(e), p_{T-1}^B(e), \cdots, p_t^B(e), \cdots, p_1^B(e))$	the counterpart for a naive DM
$\mathbf{e}_T^A(e) = (e, e_{T-1}^A(e), \cdots, e_t^A(e), \cdots, e_1^A(e))$	equilibrium path of states for a sophisticated DM with state $e_T = e$
$\mathbf{e}_T^B(e) = (e, e_{T-1}^B(e), \cdots, e_t^B(e), \cdots, e_1^B(e))$	the counterpart for a naive DM
$\pi(e', \mathbf{p}_T^A(e)) = \sum_{\tau=T}^1 U(e', p_\tau^A(e))$	simplified as $\pi(\mathbf{p}_T^A(e))$ if $e' = e$
$\pi(e', \mathbf{p}_T^A(e)) = \sum_{\tau=T}^1 U(e', p_\tau^B(e))$	simplified as $\pi(\mathbf{p}_T^B(e))$ if $e' = e$

FIGURE 2. Summary of Notations

than $P^*(e_2)$ since the objective is to make future state smaller and the control function is increasing in action.

Compare the total payoff of the sophisticated $DM_2^A(e_2)$ versus the naive $DM_2^B(e_2)$. One can see that the sophisticated person can secure a better payoff than the naive counterpart because the outcome $\mathbf{p}_T^B(e_2) = (P^*(e_2), P^*(Fe_2))$ is among the feasible continuations for $DM_2^A(e_2)$. All the above arguments also apply to the last two periods of any sophisticated person's problem and the result is summarized in Lemma 2.

Lemma 2. For the last two periods of a positive feedback control problem ($T \geq 2$),

(1) a sophisticated person chooses $p_2^A(e_2) \leq P^*(e_2^A) \leq p_1^A(e_2)$, and $e_2^A \leq e_1^A = E(e_2^A, p_2^A(e_2))$ with $p_1^A(e_2) = P^*(e_1^A)$.

(2) This implies that if a sophisticated and a naive person have the same state $e_2^A = e_2^B = e_2$ at period 2, the choices satisfy $\mathbf{p}_2^A(e) \leq \mathbf{p}_2^B(e)$, and the last period states $e_1^A \leq e_1^B$.

(3) In a two-period decision, a sophisticated person is always better off than a naive person, or $\pi(\mathbf{p}_2^A(e)) \geq \pi(\mathbf{p}_2^B(e))$.

That is when only two periods are left to be played out, a sophisticated DM chooses an action that is less than instantaneously optimal so that the the last period state will be less than what would have been if the DM were naive. However the comparison $\mathbf{p}_T^A(e) \leq \mathbf{p}_T^B(e)$ is in general not true for $T > 2$.

Also needed to be pointed out for the problem of $DM_2^A(e_2)$ is the initial period DM's perception of his future self. When there are only two periods and the DM is aware of the inconsistent preference, the future self for $DM_2^A(e_2)$, represents only conflict of interest that he should play against. This makes a two-period problem a degenerate case in terms of the intertemporal incentives and the welfare comparison between $DM_2^A(e_2)$ and $DM_2^B(e_2)$ a trivial one.

Now take a three-period problem of a sophisticated person. At $t = 3$ the DM has both a conflict of interest and a common interest with the next period self, and is perfectly aware of both. On the one hand, DM_3^A and DM_2^A do not share the same objective functions by changing preference. On the other hand, they share the same point of view on DM_1^A. Because of the nature of a positive feedback, if DM_3^A and DM_2^A do not exercise any control, DM_1^A will choose an action deemed too large by the DMs of both previous two periods, hence the mergence of interest of DM_3^A and DM_2^A. Moreover, each period's decision maker can only make a one-time choice of action to exert control on all future selves. DM_3^A's control on the last period self is implemented indirectly through the endogenous choice of e_2 and $DM_2^A(e_2)$'s objective function. Lastly, the feedback system we defined is a strict closed-loop control, so all of the above has to be done while balancing the cost to the instantaneous utility of DM_3^A.

Therefore, having two future selves, versus only one future self, has both benefit and cost for DM_3^A. There are both common interest with $DM_2^A(e_2)$ the initial period decision maker needs to exploit, as well as persistent conflict of interest with both future selves to play against. For decisions with more and more periods, higher and higher orders of such incentive considerations are involved. For this reason, definitive predictions for the sophisticated person are hard to obtain without more restrictive assumptions. State-independent updating turns out to be a useful restriction to simplify the sophisticated person's problem.

3. Results of outcome and welfare comparisons

In this section, we first characterize the sophisticated person's outcome and welfare as well as to compare it with that of a naive person, with the help of assuming a state-independent update (control) function $E(p)$. A mild sufficient condition is then provided for the same welfare comparison with general control functions.

3.1. **Outcomes.** First, we look at the outcome paths of a sophisticated person with different starting states. Intuitively, a person with a lower starting state should choose a series of actions that are all lower than a person with a higher states. Because of the supermodular instantaneous utility, the benefit of a lower action is higher for the lower state. Secondly, a lower action this period will lead to a lower state in the next by the monotonicity of $E(e, p)$, which also predicts the same behavior tendency for future periods. However, this rationale does not go through under a general control function. First, we provide a contrived counterexample below.

Example 1. Assume the instantaneous utility function $U(e, p) = -(e-p)^2$ with two periods of actions to take. So the instantaneously optimal choice is $P^*(e) = e$. Consider two initial states $e = 1$ or $7/6$ and let the state-dependent update functions be

$$E(e,p) = \begin{cases} p+1, & \text{if } \ e = 1 \\ 10\sqrt{p/3} - 2, & \text{if } \ e = 7/6 \text{ and } 1/3 \leq p \leq 27 \end{cases}$$

One can check that these functions satisfy all the assumptions required and the outcomes of the two-period decision of a sophisticated person for each state are $\mathbf{p}_2^A(1) = (1/2, 3/2)$ and

$\mathbf{p}_2^A(7/6) = (1/3, 4/3)$, respectively. This shows that it is possible for a sophisticated DM with a lower initial state to choose both actions higher than those of a DM with a higher initial state when the update functions are state-dependent. The reason lies in that in the example the DM with the lower state $DM_2^A(1)$ is given a less effective control function than the DM with the higher state. This can be seen by looking at the partial $\partial E(e, p)/\partial p$ around each DM's instantaneously optimal action. $DM_2^A(7/6)$ has a control function that is more effective in changing the next period state, so choosing a lower action in the initial period is more beneficial for $DM_2^A(7/6)$.

With state-independent control functions, the lemma below shows that a sophisticated person with a lower initial state always chooses a lower action path. Using smoking as the example, the result implies that a person less addicted should choose to smoke less in *every* period than a person who is more addicted, *ceteris paribus.*

Lemma 3. If a positive feedback system has state-independent update (control) function, then for any $e_0 \leq e < e'$

(1) $\mathbf{p}_T^A(e) \leq \mathbf{p}_T^A(e')$, and

(2) $\pi(e_0; \mathbf{p}_T^A(e)) \geq \pi(e_0; \mathbf{p}_T^A(e'))$

By the assumption of state-independent control function, the feasibility of an action continuation is state independent also. Hence by definition, out of all the equilibrium paths chosen by DMs with different states, a $DM_T^A(e)$ likes the path he chooses $\mathbf{p}_T^A(e)$ the best. The second part of the above result shows when evaluating the two paths, not only $DM_T^A(e)$ prefers $\mathbf{p}_T^A(e)$ over $\mathbf{p}_T^A(e')$ (where $e < e'$), but also all DMs with lower states than e prefer the former outcome over the latter. The next proposition shows that when a positive feedback system has state-independent update function, a sophisticated DM cuts back the action of *every* period from what a naive DM would choose (i.e. instantaneous maximization). Because the self-knowledge of the future consequences of current indulgence leads a sophisticated person to exercise some self-control at every period.

Proposition 1. If a positive feedback system has state-independent update (control) function, then $p_t^A(e) \leq P^*(e_t^A(e))$ for any $T \geq t \geq 1$. Hence $\mathbf{p}_T^A(e) \leq \mathbf{p}_T^B(e)$ and $\mathbf{e}_T^A(e) \leq \mathbf{e}_T^B(e)$.

With the above preliminary results, the outcome of a sophisticated person's problem can be readily characterized in the following proposition. Although a sophisticated person exerts some self-control at every period, because that the control always comes at a cost, as time goes by, a sophisticated person also observes his state to increase in the positive feedback system as established in Lemma 1 for a naive person.

Proposition 2. If a positive feedback system has state-independent update (control) function, then both the outcomes of actions and states of a sophisticated $DM_T^A(e_T^A)$ are monotonically increasing. That is $e_T^A \leq e_{T-1}^A \leq \cdots \leq e_1^A$, and $p_T^A(e_T^A) \leq p_{T-1}^A(e_T^A) \leq \cdots \leq p_1^A(e_T^A)$.

3.2. **Welfare.** Now we turn our attention into the welfare comparison part of the current inquiry. As discussed before, with any two-period problem, a sophisticated person always ensures a better payoff than the naive person. Because the sophisticated can at least choose the same action path of the naive person. For decisions of three periods or longer, the naive DM's choice path is typically not feasible for the sophisticated person with the same state. Because it might violate the incentives of future selves. Hence the "curse" of the sophisticated is that future selves are also sophisticated and their incentives have to be properly met. In the following proposition and example, we demonstrate that a sophisticated with 3 periods of actions can get around this "curse" and for longer periods we rely on, in this section, the restriction of state-independent control function to show that being sophisticated is always better off than being naive.

Proposition 3. If a positive feedback system has state-independent update (control) function, then a sophisticated DM is always better off than a naive one with the same initial state, or $\pi(\mathbf{p}_T^A(e)) \geq \pi(\mathbf{p}_T^B(e))$ for any T. With a general control function $E(e,p)$, a sophisticated person is better off for a decision of 3 periods. That is, $\pi(\mathbf{p}_3^A(e)) \geq \pi(\mathbf{p}_3^B(e))$.

Example 2. Assume that the instantaneous payoff of coffee drinking is given by the function $U(e,p) = -(e-p)^2$ where e represents the current "addictive state" and p is the number of cups the DM chooses to drink. The instantaneous optimization function is then $P^*(e) = e$. The next period addictive state is endogenously determined by the function $e_{t-1} = E(p_t) = 1 + p_{t-1}$. Assume the initial state is $e = 1$. It is easy to see when $T = 2$, $\mathbf{p}_2^B(1) = (1,2)$ and $\pi(\mathbf{p}_2^B(1)) = -1$. When the DM is aware of the changing preference, first period optimization problem changes to:

$$\begin{aligned} &\max_{p_2} -(1-p_2)^2 - (1-p_1)^2 \\ \text{s.t.} \qquad & p_1(=e_1) = 1 + p_2 \end{aligned}$$

The outcome $\mathbf{p}_2^B(1)$ satisfies the constraint of the maximization problem, so the sophisticated person can ensure at least a payoff of -1. In fact, the solution of the above is $\mathbf{p}_2^A(1) = (1/2, 3/2)$ and $\pi(\mathbf{p}_2^A(1)) = -0.5 > \pi(\mathbf{p}_2^B(1)) = -1$

When $T = 3$, we have $\mathbf{p}_3^B(1) = (1,2,3)$ and $\pi(\mathbf{p}_3^B(1)) = -5$. An optimal coffee drinking path for a sophisticated DM needs to solve the following problem:

$$\begin{aligned} &\max_{p_3} -(1-p_3)^2 - (1-p_2)^2 - (1-p_1)^2 \\ \text{s.t.} \qquad & e_2 = 1 + p_3 \\ & p_1 = e_1 = 1 + p_2 \\ & p_2 = \arg\max_{p_2} -(e_2-p_2)^2 - (e_2-p_1)^2 \end{aligned}$$

The solution is $\mathbf{p}_3^A(1) = (1/3, 5/6, 11/6)$. The comparison $\pi(\mathbf{p}_3^A(1)) = -7/6 > \pi(\mathbf{p}_3^B(1)) = -5$ still holds, although no longer trivially from the definition of the two types of DM when $T = 3$. Without the last constraint, the argument provided for $T = 2$ still applies. The last constraint is a generic one of inconsistent preference which requires the optimality of every time t DM. The result is driven by the positive feedback structure of the coffee drinking problem.

The intuition is that the outcome path $\mathbf{p}_3^B(1)$ has spiralling coffee drinking and preference further and further away from the initial state. When a sophisticated DM is aware of this tendency, he will try to control the change of his later preferences balancing the cost of doing so according to his current preference. However, for $DM_3^A(1)$, this control can be executed only once by the choice of p_3. After p_3 is chosen and e_2 is updated, choice will be made according to $DM_2^A(e_2)$'s preference, and so on. Proposition 3 holds because even if $DM_3^A(e)$ chooses an action $p_3 = P^*(e) = p_3^B(e)$ and counts on $DM_2^A(Fe)$ to exercise some kind of control, it will still yield a better outcome than $\mathbf{p}_3^B(e)$. Because $DM_2^A(Fe)$ is the second last to make a choice and has only one future DM to play against, his choice is a clear-cut "controlling action" ($p_2^A(Fe) \leq P^*(Fe)$), which for sure is beneficial to $DM_3^A(e)$, compared to the choice that would have been made by a naive $DM_2^B(Fe)$. For a general positive feedback system, this argument, however, does not apply to a sophisticated DM with 4 or more actions to choose. Take the problem of $DM_4^A(e)$ for instance. If the DM starts by choosing an instantaneously optimal action and update the state to Fe. The next period DM will be $DM_3^A(Fe)$. However, by the reason we provided in Section 3, when there are 3 periods involved in a decision, there are both conflict and common interest between DM_3^A and DM_2^A. So DM_4^A cannot make sure that DM_3^A acts more towards his interest than the naive counterpart DM_3^B.

Another interesting observation can be made here is the payoff comparison of the two outcomes in the perspective of a DM with a state other than the initial state. The following corollary shows than there exists a cutoff state so that a DM will a lower than the cutoff state will prefer the sophisticated outcome with self-control, and vice versa.

Corollary 1. If a positive feedback system has state-independent update (control) function, then there exists a $\bar{f} \geq e \in \mathbf{E}$ such that for any $f \in \mathbf{E}$:

(1) $\pi(f; \mathbf{p}_T^A(e)) \geq \pi(f; \mathbf{p}_T^B(e))$ if $f \leq \bar{f}$, and

(1) $\pi(f; \mathbf{p}_T^A(e)) \leq \pi(f; \mathbf{p}_T^B(e))$ if $f \geq \bar{f}$.

It can be seen from the proof of Proposition 3 that the reason it cannot extend beyond three periods for the general functional form $E(e, p)$ is a phenomenon illustrated by Example 1. With different control functions at different states, a sophisticated person with a higher state might find it optimal to choose a higher action path. This invalidates the conclusions of

Lemma 3 and Proposition 1 that backward induction of the proof depends on. If the outcome comparison in Lemma 3 holds, one can make the same conclusion that a DM is better off being sophisticated than being naive.

Corollary 2. A sufficient condition for $\pi(\mathbf{p}_T^A(e)) \geq \pi(\mathbf{p}_T^B(e))$ to hold with a positive feedback system is $\mathbf{p}_{t-2}^A(f) \leq \mathbf{p}_{t-2}^A(f')$ for any $f < f' \in \mathbf{E}$ and $3 \leq t \leq T$.

The corollary makes it apparent why in Proposition 3 we have $\pi(\mathbf{p}_3^A(e)) \geq \pi(\mathbf{p}_3^B(e))$ for a general positive feedback system, because then the condition is reduced to $\mathbf{p}_1^A(f) \leq \mathbf{p}_1^A(f')$ for $f < f' \in \mathbf{E}$, which is true by $p_1^A(f) = P^*(f) \leq P^*(f') = p_1^A(f')$. The next example shows that the assumption of positive feedback is a necessary one in driving the result of Proposition 3, and by itself the self-knowledge of a sophisticated person does not guarantee a better position relative to a naive counterpart for a decision longer than two periods.

Example 3. Consider a similar scenario as in the previous example. Assume again the instantaneous utility function $U(e,p) = -(e-p)^2$ with initial state $e = 1$ and 3 periods of actions to take. Let the state update function be

$$E(e,p) = \begin{cases} p+1, & \text{if } \ p \leq 3/2 \\ -3p+7, & \text{if } \ p \geq 3/2 \end{cases}$$

This decision problem is not a positive feedback system because $Fe \geq e$ is only satisfied when $e \leq 7/4$, while for $e > 7/4$, we have $Fe = -3e + 7 < e$. A naive DM chooses $\mathbf{p}_3^B(1) = (1,2,1)$ which induces states $\mathbf{e}_3^B(1) = (1,2,1)$ and his payoff is $\pi(\mathbf{p}_3^B(1)) = -1$. One can show that a sophisticated DM, in the same problem, will choose $\mathbf{p}_3^A = (5/7, 123/70, 121/70)$ with $\mathbf{e}_3^A(1) = (1, 12/7, 121/70)$. And this only arrives at a payoff of $\pi(\mathbf{p}_3^A(1)) = -83/70 < -1 = \pi(\mathbf{p}_3^B(1))$.

In this example, the sophisticated $DM_3^A(1)$ falls into the "curse" that DM_2^A will also be sophisticated. In the naive outcome $\mathbf{p}_3^B(1) = (1,2,1)$, although $DM_3^B(1)$ blindly chooses a myopic optimization and leads to a higher state in the second period, since the feedback system is not positive, the instantaneous optimization of $DM_2^B(2)$ actually benefits $DM_3^B(1)$ because it results in the last period state to be the same as the state of $DM_3^B(1)$. And as it turns out, when the middle period self is sophisticated and tries to exert self-control, it is detrimental to $DM_3^A(1)$. Because of this curse, the sophisticated $DM_3^A(1)$ cannot obtain a payoff as much as in a naive outcome, even with the perfect knowledge of the behaviors of the future selves.

Lastly, we compare the payoffs of a sophisticated DM endowed with the same instantaneous utility functions but distinct update functions $E_2(p) \geq E_1(p)$ for any p. We arrive at the intuitive result that the DM obtains a higher total payoff with the slower updating $E_1(p)$.

Proposition 4. Two positive feedback systems, with the same $U(e, p)$ and two different state-independent control functions $E_1(p)$ and $E_2(p)$, if $E_2(p) \geq E_1(p)$ for any $p \in \mathbf{R}$, then $\pi(\mathbf{p}^A_{T,E_1}(e)) \geq \pi(\mathbf{p}^A_{T,E_2}(e))$, for any $e \in \mathbf{E}$.

Corollary 3. There exists a $\bar{f} \geq e \in \mathbf{E}$ such that for any $f \in \mathbf{E}$:

(1) $\pi(f; \mathbf{p}^A_{T,E_1}(e)) \geq \pi(f; \mathbf{p}^A_{T,E_2}(e))$ if $f \leq \bar{f}$, and

(1) $\pi(f; \mathbf{p}^A_{T,E_1}(e)) \leq \pi(f; \mathbf{p}^A_{T,E_2}(e))$ if $f \geq \bar{f}$.

3.3. **General Control Functions.** As mentioned before, although the assumption of state-independent control functions simplifies the prediction of outcomes for a sophisticated person and yields a definitive welfare comparison, it is very restrictive in the context of endogenous preferences. Here we explore the general update (control) function $E(e, p)$ in the positive feedback control problem of a sophisticated DM, starting by characterizing a sufficient condition so that the same welfare comparison of Proposition 3 holds. The condition lies in precluding the "pathological" case in Example 1 where the high state DM is endowed with a more effective control. Essentially, the condition requires that in achieving a certain self-control of the future state, it is always less costly at a lower state in terms of the sacrifice of the instantaneous utility. For example, quitting smoking is easier for a less addicted.

Definition 2. A positive feedback system $\{U(e, p); E(e, p)\}_{e \in \mathbf{E}, p \in \mathbf{R}}$ is said to have *increasing marginal cost of control* if for any $e < e' \in \mathbf{E}$, $f < \hat{f} \leq Fe \in \mathbf{E}$, let $p, \hat{p}, q, \hat{q} \in \mathbf{P}$ be such that $E(e,p)=f$, $E(e,\hat{p})=\hat{f}$, $E(e',q)=f$, and $E(e',\hat{q})=\hat{f}$, we have

$$U(e',\hat{q})-U(e',q)\geq U(e,\hat{p})-U(e,p)$$

One can see that under the state-independence assumption, the above condition actually is reduced to the supermodularity of the instantaneous utility $U(e,p)$. Because then, the actions will be $p=q<\hat{p}=\hat{q}$.

Proposition 5. If a positive feedback system has increasing marginal cost of control, then a sophisticated DM is always better off than a naive one, or $\pi(\mathbf{p}^A_T(e))\geq\pi(\mathbf{p}^B_T(e))$ for any T.

The proof is a combination and modification of the previous proofs and is omitted here. Intuitively, the welfare comparison is restored with the sufficient condition, because no DM of any period will choose an action above the instantaneous optimal. Since doing so will only drive the future state to be further away from the current, as well as making the future control harder hence the future self to exert less effective control. With every future self to cut back in their action, the sophisticated DM of the initial period is better off, given that the control problem has a positive feedback nature.

With a general control function $E(e,p)$, however, a "state reversal" can happen in a sophisticated outcome. That is, a sophisticated DM might exert a self-control significant enough to make the next period state lower than the current level, invalidating the result of Proposition 2. This might happen because the sophisticated DM tries to take advantage of the lower control cost of a future self with a lower state.

4. Discussions

4.1. **Limited Self-knowledge.** It has been shown that under a mild condition a sophisticated person can do better than a naive person in a positive feedback control problem. This section shows that in general the payoff of a DM is not monotonic in his "level of self-knowledge", even with state-independent control. First, we introduce the "middle" level of self-knowledge between the two extreme cases we have been studying so far. The middle ground we choose is a person who knows perfectly only L periods ($L \geq 1$) of the endogenous updating and believes that the state will not change afterwards. [8] We call this type the DM with L *self-knowledge* and use the capital letter L in place of A and B of the extreme types for symbols related.

Obviously, if L is greater than the number of times the preference is subject to change bounded by the length of the decision ($L \leq T-1$), a person with L self-knowledge behaves like a sophisticated person. If the capacity of L self-knowledge is not enough to perfectly project all the endogenous preference changing, at the first few periods of the decision the person mistakenly makes predictions about the tail of the decision. However, since the DM is still aware of the immediate preference change, his incentive is similar to a sophisticated person: to reduce the current action so to reduce the future state. When the control function is state-independent, we obtain

Proposition 6. If a positive feedback system has state-independent update (control) function, then both the outcomes of actions and states of a person of L *self-knowledge* $DM_T^L(e_T^L)$ are monotonically increasing. That is $e_T^L \leq e_{T-1}^L \leq \cdots \leq e_1^L$, and $p_T^L(e_T^L) \leq p_{T-1}^L(e_T^L) \leq \cdots \leq p_1^L(e_T^L)$.

The proof is a slight modification of Proposition 2 and is omitted here. The focus of this section is to demonstrate by examples that, contrary to what might be the initial conjecture, the payoff of the DM is not monotonic in the level of self-knowledge. Example 4 shows that it is possible for a person with L self-knowledge to obtain payoffs higher than the sophisticated.

Example 4. A person with L self-knowledge ($L = 1$) can be better off than a sophisticated person.

Assume $E = 2(p - 1/3)$ and $U = -(e-p)^2$. Compare the choice of a sophisticated DM with the choice of a DM with L-awareness ($L = 1$). Some algebra shows that $\pi(\mathbf{p}_4^L(1)) > \pi(\mathbf{p}_4^A(1))$. [9]

[8] There are apparently many forms of such limited self-knowledge one can imagine, we choose the one that is easiest to define by our model setup. The interpretation of such a limited self-knowledge can be that the cognitive load allows a person to think only L periods ahead.

[9] The outcomes are $\mathbf{p}_4^L(1) = (\frac{11}{13}, \frac{304}{351}, \frac{106}{117}, \frac{134}{117})$, $\mathbf{p}_4^A(1) = (\frac{5905}{6963}, \frac{248420}{285483}, \frac{1300198}{1427415}, \frac{1648786}{1427415})$. Payoffs are approximately $\pi(\mathbf{p}_4^L(1)) = -0.0715497 > -0.0719368 = \pi(\mathbf{p}_4^A(1))$.

The intuition is that a person with a limited self-knowledge can either over-control or under-control himself, compared to a sophisticated counterpart. And over-control by a latter period self can be beneficial to the initial period DM. The problem of the limited awareness is something we hope to explore in the future. We conclude this part with the simple observation that if a period t decision maker is the last to have any wrong predictions of future preference, then he is worse off than a sophisticated counterpart.

Corollary 4. If a person has L self-knowledge with $L = T - 2$, then $\pi(\mathbf{p}_T^A(e)) \geq \pi(\mathbf{p}_T^L(e))$.

So it is possible for $DM_T^L(e)$ to make better payoff than $DM_T^A(e)$ only if some future selves also make wrong predictions.

4.2. **Strategic Play.** In this part, with future research direction in mind, we briefly discuss endogenous preferences in strategic games and some of the novel issues that are not present in one-person decision making problems. Adding strategic play of course makes the issue more complex and is beyond the scope of current investigation. So again we opt for concrete examples.

Consider a two-player finitely repeated game. We look at the case that one player is subject to endogenous preference change, while the other is not. Also assume that the player with the fixed preference has the perfect knowledge of the endogeneity of her opponent's preference. The following examples illustrate that the payoff comparison between a naive person and a sophisticated person is no longer trivial even for $T = 2$ by the nature of strategic play.

Example 5. Consider a two period employee-employer game where the employee's preference is subject to change. In each period, the employee (he) chooses between a high or low effort to make in his work. And the employer (she) chooses whether to offer him a bonus. The above bimatrix gives the preferences in the initial period. With fixed preferences, it is obvious that both periods' outcome will be (Low, No bonus). Assume if in the initial period, a bonus is offered by the employer, the employee's preference will change and he will strictly prefer to play a high effort in the second period. On the other hand, if no bonus is offered in the first period, the employee maintains the same preference. Now consider the outcome of the game with the assumed preference change on the part of the employee. First of all,

		Employer	
		Bonus	No bonus
Employee	High	8,8	0,10
	Low	10,2	4,4

Figure 3. INITIAL PERIOD PREFERENCE

the employer is aware that if she offers the bonus in the first period, she can induce the employee to play a high effort in the second period. In that case, she would choose to pay no bonus and get a payoff of 10. If she does not offer a bonus in the initial period, the only possible outcome will be (Low, No bonus) in the second period. Hence by offering the bonus in the first period, she gains a benefit of 6 in the second period. And it is easy to see that, since offering a bonus only costs her 2 in the first period no matter what effort the employee chooses to make, indeed it is strictly optimal for the employer to offer the bonus in the first period.

Now consider the behavior of the two possible types of employee. By the above analysis, the employer offers a bonus for sure in the initial period. The employee will make a low effort no matter what type he is. The difference is that a sophisticated employee is aware that he will make a high effort in the second period, while a naive employee believes that he will make a low effort later. In both cases the unique outcome actually played out is (Low, Bonus) in the first period and (High, No bonus) in the second. The total payoff of the employee evaluated by the initial preference is 10.

Example 6. Consider a slight modification of the above example in that the employee's preference will change as described only if the first period has (Low, Bonus) played. Following any other action profile, he keeps his initial preference in the second period. Now the employee deems a bonus in the first period worthwhile only if at the same time, the employee chooses low effort in the first period. (Low, Bonus) followed by (High, No bonus) is still the actual outcome if the employee is a naive player. Because he believes that he will choose low effort in the second period no matter what he does in the first. Then in his mind, by making a low effort instead of a high effort in the first period, he gets an extra 2. Of course, with such a decision, he will only find himself making a high effort in the second period once his preference changes. Thus his total payoff evaluated by his initial preference is again 10.

An interesting phenomenon happens here when the employee is aware of his potential preference change. His awareness allows him to comprehend that, when a bonus is offered in the first period, he should choose to make a high effort. Although this action incurs an immediate cost of 2, he steers his future self away from making a high effort which will only be responded by no bonus from the employer. Evaluated by the current preference, this is a benefit of 4. However, this mentality of the employee destroys the reason for the employer to offer a bonus in the first period. The unique outcome, therefore, is (Low, No bonus) in both periods. And the employee's total payoff is 8 which is lower than his payoff of 10 if he is unaware of his preference change.

Remark. In both examples, offering a bonus in the first period can be considered an act of manipulating the employee's future preference by the employer. Because as a "normal" rational player with a fixed preference, she has a dominant strategy of offering no bonus. A bonus is offered solely for the purpose of changing the employee's preference to her favor in the second period. In the first example, both types of employee end up taking the manipulation because even when he is aware of this manipulative effect, he can do nothing to prevent it. In the latter example, he can preempt the manipulation by choosing a high effort. However, this also unravels the employer's incentive to offer the bonus in the first place. Both strategic reasons and the payoff structure play a role in driving the result that being aware of the preference change actually decreases the total payoff evaluated at the initial preference.

These examples show that the payoff comparison is a non-trivial question in a strategic play even with only two periods. The future task is to identify classes of games and conditions (using the feedback structure developed by this paper) where a definitive comparison can be made, one way or the other.

5. Conclusion

In this paper, we provide a framework to model changing preferences in intertemporal decisions. The model is based on the observation that in decision that involves repeated choice of actions over time, past choices can endogenously change a person's preference, for example in the phenomena of addiction and habit-formation. We focus on decision problems that tend to lead the decision maker further and further away from the initial state: positive feedback systems. Two extreme cases of decision makers, in terms of their self-knowledge of the endogenous preference updating process, are studied and outcomes characterized. Using the initial preference as the welfare standard, we conclude that a sophisticated person is better off than a naive person in the case of state-independent control functions. An intuitive sufficient condition is also given for the general control function. Furthermore, we show by means of an example that the above payoff comparison is not monotonic in the level of self-knowledge. Future research possibilities, especially in strategic play, are discussed.

Appendix: Proofs

Lemma 3. If a positive feedback system has state-independent update (control) function, then for any $e_0 \leq e < e'$

(1) $\mathbf{p}_T^A(e) \leq \mathbf{p}_T^A(e')$, and

(2) $\pi(e_0; \mathbf{p}_T^A(e)) \geq \pi(e_0; \mathbf{p}_T^A(e'))$

Proof. (1) For $T = 2$, let $\mathbf{p}_2^A(e) = (p_2, p_1)$ and $\mathbf{p}_2^A(e') = (r_2, r_1)$. Since $p_1 = P^*(E(p_2))$ and $r_1 = P^*(E(r_2))$, we only need to show $p_2 \leq r_2$ because of the increasing $P^*(\cdot)$ and the increasing state-independent update function $E(\cdot)$. Suppose by way of contradiction that $p_2 > r_2$. This also implies $p_1 > r_1$. Because of state-independent update function, (p_2, p_1) is also a feasible continuation for $DM_2^A(e')$, optimization of $DM_2^A(e')$ ensures that

$$U(e', r_2) + U(e', r_1) \geq U(e', p_2) + U(e', p_1)$$

Since $e < e'$, the supermodularity of $U(\cdot, \cdot)$ implies

$$\begin{aligned}&[U(e, r_2) - U(e, p_2)] + [U(e, r_1) - U(e, p_1)] \\ &> [U(e', r_2) - U(e', p_2)] + [U(e', r_1) - U(e', p_1)] \geq 0\end{aligned}$$

This is a contradiction because $DM_2^A(e)$ strictly prefers the pair (r_2, r_1) which is also a feasible continuation.

The next step is to use mathematical and backward induction to show that if the condition holds for $T = K$, then it also holds for $T = K + 1$. Assume $\mathbf{p}_K^A(e) \leq \mathbf{p}_K^A(e')$ for any pair $e < e'$. And consider the choices of $DM_{K+1}^A(f)$ and $DM_{K+1}^A(f')$, where $f < f'$. Let $\mathbf{p}_{K+1}^A(f) = (p_{K+1}, \cdots, p_1)$ and $\mathbf{p}_{K+1}^A(f') = (r_{K+1}, \cdots, r_1)$. Since the condition to be proven holds for K periods, we only need to show that $p_{K+1} \leq r_{K+1}$, hence $E(p_{K+1}) \leq E(r_{K+1})$ implies the rest. Suppose not. Let $p_{K+1} > r_{K+1}$, then $p_t > r_t$ for all $t \leq K$ also. Since $(p_{K+1}, \cdots, p_1)$ is the optimal choice of $DM_{K+1}^A(f)$, it must satisfy the incentives for all future DMs. And because $E(\cdot)$ is state-independent, $\mathbf{p}_{K+1}^A(f) = (p_{K+1}, \cdots, p_1)$ is also a feasible continuation for $DM_{K+1}^A(f')$. The optimization of $DM_{K+1}^A(f')$ ensures

$$\pi(f'; \mathbf{p}_{K+1}^A(f')) \geq \pi(f'; \mathbf{p}_{K+1}^A(f)) \ \Rightarrow \ \sum_{\tau=K+1}^{1} U(f', r_\tau) \geq \sum_{\tau=K+1}^{1} U(f', p_\tau)$$

The contradiction lies in that the supermodularity of $U(\cdot,\cdot)$ then implies $DM_{K+1}^A(f)$ strictly prefers $(r_{K+1}, \cdots, r_1)$, which is feasible again by state-independent $E(\cdot)$, over $(p_{K+1}, \cdots, p_1)$.

(2) follows immediately from $\pi(e; \mathbf{p}_T^A(e)) \geq \pi(e; \mathbf{p}_T^A(e'))$, $\mathbf{p}_T^A(e) \leq \mathbf{p}_T^A(e')$), and the supermodularity of $U(\cdot,\cdot)$. □

Proposition 1. If a positive feedback system has state-independent update (control) function, then $\mathbf{p}_T^A(e) \leq \mathbf{p}_T^B(e)$.

Proof. Lemma 2 proves the condition for $T = 2$. For $T = 3$, we start by showing $p_3^A(e) \leq P^*(e) = p_3^B(e)$. Let $\mathbf{p}_3^A(e) = (r_3, r_2, r_1)$ and $\mathbf{p}_2^A(Fe) = (p_2, p_1)$. We need to show that $r_3 \leq P^*(e)$. Suppose not, assume $r_3 > P^*(e)$. Then $E(r_3) > Fe \geq e$, therefore $\mathbf{p}_2^A(E(r_3)) \geq \mathbf{p}_2^A(Fe)$ by lemma 3. The feasibility of continuation $\mathbf{p}_2^A(E(r_3))$ implies

$$\pi(Fe, \mathbf{p}_2^A(Fe)) \geq \pi(Fe, \mathbf{p}_2^A(E(r_3)))$$

And with $e \leq Fe$, we have

$$U(e, P^*(e)) + U(e, p_2) + U(e, p_1)$$

$$> U(e, r_3) + U(e, p_2) + U(e, p_1)$$

$$\geq U(e, r_3) + U(e, r_2) + U(e, r_1)$$

or equivalently $U(e, P^*(e)) + \pi(e, \mathbf{p}_2^A(Fe)) > \pi(\mathbf{p}_3^A(e))$, a contradiction. Therefore, we have shown that $p_3^A(e) \leq P^*(e)$, By both Lemma 1 and 2, this implies that

$$\mathbf{p}_3^A(e)(e) \leq (P^*(e), \mathbf{p}_2^A(Fe)) \leq (P^*(e), \mathbf{p}_2^B(Fe)) = \mathbf{p}_3^B(e)$$

The result for $T > 3$ can be shown similarly by first establishing

$$\mathbf{p}_T^A(e) \leq P^*(e)$$

and using mathematical and backward induction, with the help of Lemma 3. □

Proposition 2. If a positive feedback system has state-independent update (control) function, then both the outcomes of actions and states of a sophisticated $DM_T^A(e_T^A)$ are monotonically increasing. That is $e_T^A \leq e_{T-1}^A \leq \cdots \leq e_1^A$, and $p_T^A(e_T^A) \leq p_{T-1}^A(e_T^A) \leq \cdots \leq p_1^A(e_T^A)$.

Proof. Lemma 2 shows the result of $T = 2$. Again by way of induction assume $e_K^A \leq e_{K-1}^A \leq \cdots \leq e_1^A$ and $p_K^A(e_K^A) \leq p_{K-1}^A(e_K^A) \leq \cdots \leq p_1^A(e_K^A)$. State-independent control and Lemma 3 imply that if $e' < e''$, then

$$\pi(e; \mathbf{p}_K^A(e)) \geq \pi(e; \mathbf{p}_K^A(e')) \geq \pi(e; \mathbf{p}_K^A(e''))$$

Suppose by way of contradiction that $e_{K+1}^A > e_K^A$. Let $p, r \in \mathbf{R}$ be the actions such that $E(p) = e_K^A$, $E(r) = e_{K+1}^A$. Hence $p < r$. Because of $Fe_{K+1}^A \geq e_{K+1}^A$, we have $p < r \leq P^*(e_{K+1}^A)$ and

$$U(e_{K+1}^A, r) + \pi(e_{K+1}^A; \mathbf{p}_K^A(e_{K+1}^A)) > U(e_{K+1}^A, p) + \pi(e_{K+1}^A; \mathbf{p}_K^A(e_K^A))$$

This is the needed contradiction. The monotonicity of the state path implies the monotonicity of the action path by increasing $E(p)$. □

Proposition 3. If a positive feedback system has state-independent update (control) function, then a sophisticated DM is always better off than a naive one with the same initial state, or $\pi(\mathbf{p}_T^A(e)) \geq \pi(\mathbf{p}_T^B(e))$ for any T. With a general control function $E(e, p)$, a sophisticated person is better off for a decision of 3 periods. That is, $\pi(\mathbf{p}_3^A(e)) \geq \pi(\mathbf{p}_3^B(e))$.

Proof. To show the case of $T = 3$ for a general update function $E(e, p)$, let $\mathbf{p}_3^B(e) = (r_3, r_2, r_1)$. Consider what happens if $DM_3^A(e)$ starts by choosing $P^*(e) = r_3$. The next period DM will have a state Fe and choose continuation $\mathbf{p}_2^A(Fe)$, denoted by (p_2, p_1). Since $e_2^A = e_2^B = Fe$, by Lemma 2 we have $p_2 \leq r_2$ and $p_1 \leq r_1$. Since (r_2, r_1) is feasible for $DM_2^A(Fe)$, we must have

$$U(Fe, p_2) + U(Fe, p_1) \geq U(Fe, r_2) + U(Fe, r_1)$$

or

$$[U(Fe, p_2) - U(Fe, r_2)] + [U(Fe, p_1) - U(Fe, r_1)] \geq 0$$

Also, $e \leq Fe$ and $(p_2, p_1) \leq (r_2, r_1)$ implies that

$$\begin{aligned}&[U(e, p_2) - U(e, r_2)] + [U(e, p_1) - U(e, r_1)]\\ &\geq [U(Fe, p_2) - U(Fe, r_2)] + [U(Fe, p_1) - U(Fe, r_1)] \geq 0\end{aligned}$$

So by playing one period of instantaneous optimization, $DM_3^A(e)$ can ensure a payoff

$$U(e, r_3) + U(e, p_2) + U(e, p_1) \geq U(e, r_3) + U(e, r_2) + U(e, r_1) = \pi(\mathbf{p}_3^B(e))$$

Therefore, it is shown that $\pi(\mathbf{p}_3^A(e)) \geq \pi(\mathbf{p}_3^B(e))$ for any positive feedback system. Extending the result to $T > 3$ for systems with state-independent control relies on mathematical induction and $\mathbf{p}_K^A(e) \leq \mathbf{p}_K^B(e)$ for any K by Lemma 4. Assume $\pi(\mathbf{p}_{T-1}^A(e)) \geq \pi(\mathbf{p}_{T-1}^B(e))$ for any $e \in \mathbf{E}$. Because of $e \leq Fe$ and $\mathbf{p}_{T-1}^A(e) \leq \mathbf{p}_{T-1}^B(e)$,

$$\begin{aligned}&\pi(Fe, \mathbf{p}_{T-1}^A(Fe)) \geq \pi(Fe, \mathbf{p}_{T-1}^B(Fe))\\ \Rightarrow\quad &\pi(e, \mathbf{p}_{T-1}^A(Fe)) \geq \pi(e, \mathbf{p}_{T-1}^B(Fe))\\ \Rightarrow\quad &\pi(\mathbf{p}_T^A(e)) \geq U(e, P^*(e)) + \pi(e, \mathbf{p}_{T-1}^A(Fe))\\ &\geq U(e, P^*(e)) + \pi(e, \mathbf{p}_{T-1}^B(Fe))\\ &= \pi(\mathbf{p}_T^B(e)) \qquad \square\end{aligned}$$

Corollary 1. If a positive feedback system has state-independent update (control) function, then there exists a $\bar{f} \geq e \in \mathbf{E}$ such that for any $f \in \mathbf{E}$:

(1) $\pi(f; \mathbf{p}_T^A(e)) \geq \pi(f; \mathbf{p}_T^B(e))$ if $f \leq \bar{f}$, and

(1) $\pi(f; \mathbf{p}_T^A(e)) \leq \pi(f; \mathbf{p}_T^B(e))$ if $f \geq \bar{f}$.

Proof. By Proposition 3, $\pi(e; \mathbf{p}_T^A(e)) \geq \pi(e; \mathbf{p}_T^B(e))$. And for $\bar{e}$, it must be that $\pi(\bar{e}; \mathbf{p}_T^A(e)) \leq \pi(\bar{e}; \mathbf{p}_T^B(e))$. By the intermediate value theorem, there must exist a $\bar{f}$, $e \leq \bar{f} \leq \bar{e}$ so that $\pi(\bar{f}; \mathbf{p}_T^A(e)) = \pi(\bar{f}; \mathbf{p}_T^B(e))$. The corollary is true because of the supermodularity of $U(e,p)$ and $\mathbf{p}_T^A(e) \leq \mathbf{p}_T^B(e)$.

Proposition 4. Two positive feedback systems, with the same $U(e,p)$ and two different state-independent control functions $E_1(p)$ and $E_2(p)$, if $E_2(p) \geq E_1(p)$ for any $p \in \mathbf{R}$, then $\pi(\mathbf{p}_{T,E_1}^A(e)) \geq \pi(\mathbf{p}_{T,E_2}^A(e))$, for any $e \in \mathbf{E}$.

Proof. T=2, let $\mathbf{p}_{2,E_1}^A(e) = (p_2, p_1)$ and $\mathbf{p}_{2,E_2}^A(e) = (r_2, r_1)$. Suppose that $\pi(\mathbf{p}_{2,E_1}^A(e)) < \pi(\mathbf{p}_{2,E_2}^A(e))$. By Lemma 2 $p_2 \leq P^*(e) \leq p_1$ and $r_2 \leq P^*(e) \leq r_1$. Also $E_1(r_2) \leq E_2(r_2)$. If $P^*(e) \leq P^*(E_1(r_2)) \leq P^*(E_2(r_2)) = r_1$, then

$$U(e, r_2) + U(e, P^*(E_1(r_2))) \geq U(e, r_2) + U(e, r_1) > U(e, p_2) + U(e, p_1)$$

. If $P^*(e) > P^*(E_1(r_2))$, then by the continuity of $E_1(\cdot)$ there exist a pair of actions (q_2, q_1) such that $r_2 < q_2 < P^*(e) < q_1 < r_1$, so

$$U(e, q_2) + U(e, q_1) > U(e, r_2) + U(e, r_1) > U(e, p_2) + U(e, p_1)$$

Both contradict that (p_2, p_1) is the optimal choice of the sophisticated person with control function E_1. Then induction is carried out as usual for $T = K + 1$ assuming the condition is true for $T = K$. $\square$

References

Ariely, D., and I. Simonson (2003): "Buying, Bidding, Playing, or Competing? Value Assessment and Decision Dynamics in Online Auctions," *Journal of Consumer Psychology*, 13(1), 113–123.

Becker, G., and K. Murphy (1988): "A Theory of Rational Addiction," *The Journal of Political Economy*, 96(4), 675.

DiStefano, J., A. Stubberud, and I. Williams (1995): *Schaum's Outline of Theory and Problems of Feedback and Control Systems*. McGraw-Hill.

Elster, J. (1998): "Emotions and Economic Theory," *Journal of Economic Literature*, 36(1), 47–74.

Fudenberg, D., and D. Levine (2006): "A Dual-Self Model of Impulse Control," *The American Economic Review*, 96(5), 1449–1476.

Gruber, J., and B. Koszegi (2001): "Is Addiction Rational? Theory and Evidence," *Quarterly Journal of Economics*, 116(4), 1261–1303.

Gul, F., and W. Pesendorfer (2001): "Temptation and Self-Control," *Econometrica*, 69(6), 1403–1435.

——— (2004): "Self-Control and the Theory of Consumption," *Econometrica*, 72(1), 119–158.

——— (2005): "The Revealed Preference Theory of Changing Tastes," *Review of Economic Studies*, 72(2), 429–448.

——— (2007): "Harmful Addiction," *Review of Economic Studies*, 74(1), 147–172.

Iannaccone, L. (1986): "Addiction and Satiation," *Economics Letters*, 21(1), 95–99.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112(2), 443–477.

Loewenstein, G., T. O'Donoghue, and M. Rabin (2003): "Projection Bias In Predicting Future Utility," *Quarterly Journal of Economics*, 118(4), 1209–1248.

Noor, J. (2005): "Temptation, Welfare, and Revealed Preference," *Working Paper*.

O'Donoghue, T., and M. Rabin (1999): "Doing It Now or Later," *American Economics Review*, 89(1), 103–124.

Ozdenoren, E., S. Salant, and D. Silverman (2008): "Willpower and the Optimal Control of Visceral Urges," *working paper*.

Phelps, E., and R. Pollak (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35(2), 185–199.

Rubinstein, A. (2006): "Discussion of BEHAVIORAL ECONOMICS," in *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, vol. 2, pp. 246–254.

Stigler, G., and G. Becker (1977): "De Gustibus Non Est Disputandum," *American Economic Review*, 67(2), 76–90.

Strotz, R. (1956): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23(3), 165–180.

Topkis, D. (1978): "Minimizing a Submodular Function on a Lattice," *Operations Research*, 26(2), 305–321.