Conditional Optimal Stopping:
A Time-Inconsistent Optimization

Marcel Nutz*   Yuchong Zhang†

October 13, 2019

Abstract

Inspired by recent work of P.-L. Lions on conditional optimal control, we introduce a problem of optimal stopping under bounded rationality: the objective is the expected payoff at the time of stopping, conditioned on another event. For instance, an agent may care only about states where she is still alive at the time of stopping, or a company may condition on not being bankrupt. We observe that conditional optimization is time-inconsistent due to the dynamic change of the conditioning probability and develop an equilibrium approach in the spirit of R. H. Strotz' work for sophisticated agents in discrete time. Equilibria are found to be essentially unique in the case of a finite time horizon whereas an infinite horizon gives rise to non-uniqueness and other interesting phenomena. We also introduce a theory which generalizes the classical Snell envelope approach for optimal stopping by considering a pair of processes with Snell-type properties.

1 Introduction

The classical optimal stopping problem is to maximize the expected payoff $E[G_\tau]$ over all stopping times τ, where $G = (G_t)$ is a given adapted process. In this paper, we propose to study a criterion that conditions on a given stopping time σ not being reached at the time τ:

$$\sup_{\tau} \frac{E[G_\tau 1_{\{\tau \lhd \sigma\}}]}{P(\tau \lhd \sigma)} \quad \text{where} \quad \tau \lhd \sigma \Leftrightarrow \tau < \sigma \text{ or } \sigma = \infty. \qquad (1.1)$$

When the model is based on a Markov chain X, a natural choice of σ is the first exit time from a given set B. If, for instance, the stopping decision is made by a company, one application is that X being in B indicates solvency so that σ is the time of bankruptcy. Indeed, the company may only care about states where the stopping payoff happens before σ as the company no longer exists in the other states. Or, for an individual making a financial decision, σ may be the time of death, then the model expresses that she only cares about states where the payoff happens while she is alive.

It is typically not possible to model such a conditional problem as a classical optimal stopping problem, except in the trivial case where the conditioning event does not depend on the stopping time τ. The classical framework would require us to model this as an exit time problem where a specific payoff is assigned to the exit event (that is, a value G_t for $t \geq \sigma$). E.g., for the individual facing possible death, we are unable to simply say, "I don't care what happens after I die." Instead, we have to assign a specific payoff at death. Even if the modeler were willing to fix some value in order to be "pragmatic," it may be hard to make a justifiable choice and the solution of the optimization will typically depend on it.

This paper is inspired by recent work of P.-L. Lions which introduces the optimal control of conditioned processes [25]. There, the main example is controlling the drift of a Brownian motion and the payoff is conditioned on the process staying inside a given domain. The problem is cast as an optimal control problem of Fokker–Planck equations, a particular type of mean field game problem with coupling through the final condition. The limit towards the classical case, where the domain tends to $\mathbb{R}^d$, is given particular attention. While it is observed that optimal controls depend on the starting point, the question of time-consistency is not raised.

In the present paper, we introduce optimal stopping with conditioning, a novel problem to the best of our knowledge. One of our first observations is that the problem is time-inconsistent in the sense of Strotz [29]: if an agent determines an optimal strategy at time $t = 0$ and reconsiders her decision at a later time taking into account her present state, she may contradict her previous decision and find that her strategy is no longer optimal. In this setting where the dynamic programming principle does not hold, there is more than one notion of optimization. The *precommitted* problem is to optimize the expected payoff at $t = 0$, assuming that the decision will not be challenged later on; i.e., the agent "commits" to the initial choice. (The theory of [25] corresponds to this notion.) In Strotz' terminology, a *sophisticated* agent without a commitment device is aware of the fact that her "future selves" may overturn her current plan. Thus, she takes this as a constraint for a

"strategy of consistent planning": she chooses her behavior ignoring plans that she knows her future selves will not carry out; that is, she selects an action such that her future incarnations have no incentive to deviate. The resulting time-consistent strategy is called subgame perfect Nash equilibrium, and this is the notion that we will focus on. A different interpretation follows the literature on intergenerational models or overlapping-generations models (see [28] and the work thereafter) where future decisions are taken by subsequent generations rather than other selves. For instance, a government agency may want to take into account future presidential terms and opt for policies which will not be reversed after the next election.

Beyond being interesting in and of itself, conditional stopping may also help to shed more light on the conditioned control of processes, since optimal stopping is often more tractable than control.

1.1   Literature

Following the early work of [29], a rich literature involving time-inconsistency has emerged in economics. For instance, [27] reconsiders Strotz' concept in a setting with non-exponential discounting when the number of decision points changes, and [26] studies preferences that change over time. Non-standard discounting (in particular hyperbolic) and time preferences (such as habit formation) are the most frequent reasons for time-inconsistency in this literature; see [14] for an overview. The models are mostly formulated in discrete time with finite or infinite time horizon. Time-inconsistency also arises when the optimization objective involves a nonlinear function of an expectation, such as the mean-variance criterion in [2], or a probability distortion as in [1, 15, 23]. (A probability distortion corresponds to an optimization objective that over- or underemphasizes events relative to their objective probability.)

The pioneering work of [10, 11] has initiated the study on how to define and obtain equilibrium strategies for the optimal control of continuous-time processes, using the example of Ramsey's problem when the planner uses non-exponential discounting. In the continuous setting, varying a control at a single instance in time is meaningless since it does not affect the diffusion. The authors develop a first-order criterion which corresponds to variations of the control over a short time interval, meaning that agents can commit for a short period. This has led to a number of works, including portfolio optimization with non-exponential discounting [12, 13], mean-variance portfolio selection [5, 8] and general linear–quadratic control [16, 17]. Nevertheless, this concept of equilibrium is not the only one possible; in particular,

first-order conditions are not sufficient for optimality in general. The recent study [21] introduces a stronger concept of optimality and highlights the differences. In [3, 4] the authors study time-inconsistent control in discrete and continuous time, respectively, and the relation between them, for a general class of objectives that are a sum of an expected utility and a nonlinear function of an expected utility with possible dependence on the initial condition. See also [31] for a continuous-time framework with dependence on the initial condition.

The closest reference for the present work is [22] where the authors study optimal stopping in discrete time under non-exponential discounting in a Markovian context. In the finite horizon case, a backward recursion yields the unique equilibrium. In the infinite horizon case, the authors focus on a time-homogeneous Markov chain. Under the assumption of decreasing impatience (including hyperbolic discounting), a time-homogeneous equilibrium is constructed by iterating the "strategic reasoning" or "fictitious play" map (cf. Φ in Section 2.1); that is, every agent optimizes her decision between continuing and stopping while taking as given the decisions of all other agents. Remarkably, an equilibrium which is optimal for all agents can be obtained. We remark that [22] is predated by [18] where the iterative approach was first implemented in continuous time. In [18], time-homogeneous equilibria are obtained for time-homogeneous diffusions and inhomogeneous equilibria for time-inhomogeneous diffusions. See also [20] for a discussion of optimal equilibria in continuous time and [30] for a recent study of optimal stopping with non-exponential discounting where equilibria may not exist and this fact is related to a failure of smooth pasting. Optimal stopping under probability distortion is studied in [19] with a particular focus on equilibria that are obtained by iterating from naïve strategies.

The mentioned works on optimal stopping in continuous time use a direct analogy to the discrete-time case to define equilibria: each agent may stop or continue, without any commitment device. Indeed, for optimal stopping, the first-order approach of [10] is not a necessity: the decision to stop at a single instance in time immediately affects the process. On the other hand, as highlighted by [9] in the context of prospect theory, the definition in continuous time may include unreasonable equilibria based on the fact that continuation and stopping for a time-t agent produce the same payoff if the subsequent agents stop and G is continuous. In particular, "always stopping" is an equilibrium even if, say, G is increasing. In a homogeneous diffusion model, [6, 7] use a first-order condition to define equilibria for two problems with time-inconsistency, and then "always stopping" is not necessarily an equilibrium. The relation between the two definitions has not been clarified

so far.

To the best of our knowledge, the present paper is the first investigation of conditional optimal stopping. Regarding the control of conditioned processes, we would like to mention ongoing work of R. Carmona and M. Laurière where the problem of [25] is studied as a mean field control problem for open and closed loop controls as well as ongoing work of Y. Achdou and M. Laurière on the numerical resolution.

1.2 Synopsis

We study the conditional optimal stopping problem in (1.1) in a discrete-time setting with finite or infinite time horizon. While a continuous-time setting may certainly be of interest, our choice avoids some of the difficulties mentioned in the preceding section and leads to an uncontroversial definition of an equilibrium: at every time and state (t, ω), an agent makes a binary choice—stopping or continuing—without committing future agents. We analyze such equilibria in a general stochastic framework while paying particular attention to the Markovian setting.

In the case of a finite time horizon T, there is a natural terminal condition (stopping is mandatory at T) and we shall see that there is an equilibrium which can be constructed by a backward recursion. This recursion computes two processes, a value process like in the classical case and an additional "survival process" that keeps track of the conditioning probability induced by the future selves' decisions. The equilibrium is essentially unique, and if the stochastic framework is Markovian, then so is the equilibrium. These findings are in line with the results for other time-inconsistent problem as described in Section 1.1.

In the case of an infinite horizon, we provide a fairly general existence result by passing to the limit of finite horizon problems. (Note that for non-exponential discounting, existence may fail if the discounting does not satisfy decreasing impatience; cf. [22, Example 3.1].) On the other hand, we also provide examples showing that this case is more subtle than the previous one. We shall see that there can exist non-Markovian equilibria in addition to Markovian ones in a Markovian setting, which disproves a conjecture of [4] for our problem. Moreover, equilibria need not be unique even within the class of Markovian equilibria. Even more surprisingly, we detail a time-homogeneous Markovian example which does not admit a time-homogeneous equilibrium while time-inhomogeneous equilibria do exist. This is in sharp contrast to the results of [18, 22] and illustrates that for our problem, in general, iterating the "strategic reasoning" map of [22] does not converge. At a technical level,

one reason is that non-exponential discounting with decreasing impatience as in [22] preserves one inequality of the dynamic programming principle whereas in our problem, the rescaling due to the conditioning probability can cause deviations in both directions.

It seems natural to ask for analogues of the classical Snell envelope theory in our setting. Indeed, the two processes described in the recursion for the finite time horizon can be characterized in more abstract terms by supermartingale properties. This leads to a notion that we call Snell pair and extends to the infinite-horizon setting. Snell pairs are (essentially) in one-to-one relation with equilibria. Similarly as in the classical case, the equilibrium policy is retrieved from the Snell pair by stopping where the value process meets the obstacle G, but the survival process is needed to adjust the classical supermartingale properties in the context of conditioning. The survival process, in turn, also enjoys a supermartingale property. We are not aware of similar notions in the prior literature.

The remainder of this paper is organized as follows. In Section 2 we detail the observation of time-inconsistency and the equilibrium concept. Section 3 presents the results on the finite-horizon case. Existence of equilibria in the infinite-horizon case is covered in Section 4 and the corresponding examples are described in Section 5. The concluding Section 6 discusses Snell pairs and their relation to equilibria.

2 Setting

Let $T \in \mathbb{N} \cup \{\infty\}$ be the time horizon. If $T < \infty$, set $\mathbb{T} = \{0, 1, 2, \ldots, T\}$; if $T = \infty$, set $\mathbb{T} = \mathbb{N}$. We will work on a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F}_t)_{t \leq T}$ such that $\mathcal{F}_0$ is trivial. Let σ be a stopping time with $P(\sigma > 0) = 1$; we think of events that happen after σ as irrelevant and call $D_t := \{t < \sigma\}$ the *domain of relevance* at time $t \in \mathbb{T}$. In the case $T = \infty$, it is convenient to set $D_\infty := \cap_{t \in \mathbb{T}} D_t = \{\sigma = \infty\}$. We may note that $\sigma(\omega) = \inf\{t \in \mathbb{T} : \omega \notin D_t\}$; indeed, specifying σ is equivalent to specifying a decreasing adapted sequence $(D_t)_{t \in \mathbb{T}}$ with $P(D_0) = 1$. Here and in what follows, the convention $\inf \emptyset = \infty$ is used. Finally, let $G = (G_t)_{t \leq T}$ be an adapted process describing the payoff for stopping at time t. The value of G_t outside D_t will not matter; we set $G_t = \Delta$ on D_t^c for notational purposes, where Δ is an auxiliary state with the convention that $0 \cdot \Delta = 0$. We assume throughout that $E[\sup_{t \leq T} |G_t| 1_{D_t}] < \infty$. Since we are interested in events that happen strictly before σ, including the case where σ never happens, it

will be useful to introduce the notation

$$s \lhd t \iff s < t \text{ or } t = \infty \qquad \text{for} \quad s, t \in [0, \infty].$$

We can then consider the *precommitted* optimal stopping problem at the initial time,

$$V_{pre} = \sup_{\tau \leq T, \, P(\tau \lhd \sigma) > 0} \frac{E[G_\tau 1_{\{\tau \lhd \sigma\}}]}{P(\tau \lhd \sigma)}. \tag{2.1}$$

Note that the supremum only runs over stopping times τ which avoid conditioning on a nullset and that the set of such times always includes $\tau \equiv 0$.

Example 2.1 (Markovian Setting). Let X be a Markov chain with values in a separable metric space $\mathbb{X}$ starting at $X_0 = x_0$, let $B \subseteq \mathbb{X}$ be a measurable subset containing x_0 and let $\sigma = \inf\{t \geq 0 : X_t \notin B\}$ be the first exit time from B. Then, our model entails that we only evaluate states of the world where the trajectory of X lies in B up to the stopping time τ. A possible specification of the payoff is $G_t = \delta^t g(t, X_t)$ for a deterministic function g and a discount factor $\delta \in (0, 1]$. More generally, the set B can be time-dependent.

The conditional optimal stopping problem (2.1) reduces to a classical optimal stopping problem when $\sigma = \infty$. But in general, the conditioning in the definition of the expected payoff for τ depends on τ itself, so that it cannot be reduced to a classical stopping problem.

2.1 Equilibria

The following example illustrates that the optimization problem (2.1) is time-inconsistent in the sense that an optimal stopping strategy for an agent today may not be optimal in the future; that is, if she reconsiders her strategy at a future time using a conditional criterion, she may contradict her previous decision.

Example 2.2. Consider a two-period binomial tree with $\Omega = \{uu, ud, du, dd\}$ as illustrated in Figure 1, where u stands for up and d for down. The conditional probabilities are $1/2$ on every edge and the numbers at each node represent the payoff G. The domain of relevance includes all states except dd; i.e., the dashed line indicates the exit from the domain. Since there are only five distinct stopping times in this model, once can easily compute all possible payoffs and observe that the unique optimizer of (2.1) is the stopping time τ_{pre} with $\tau_{pre}(uu) = \tau_{pre}(ud) = 1$ and $\tau_{pre}(du) = \tau_{pre}(dd) = 2$. To wit,
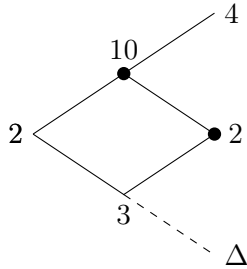
Figure 1: The binomial tree of Example 2.2.

it is optimal to stop at $t = 1$ if we have moved up in the first step and at $t = 2$ otherwise. The obtained payoffs are illustrated by the solid dots and the associated value is $V_{pre} = 10 \cdot \frac{2}{3} + 2 \cdot \frac{1}{3} = \frac{22}{3}$.

Next, consider an analogous optimization problem for an agent who solves the problem conditionally on starting in the down state at $t = 1$. This agent has only two options, either to stop immediately with payoff 3 or to wait until the horizon and receive an expected reward of 2 (since the expectation is conditioned on remaining inside the domain). Thus, this agent prefers to stop, and that is not consistent with τ_{pre}. In summary, if the first agent solves (2.1) and reconsiders her own strategy at $t = 1$ in the down state using the natural conditional criterion, she will overturn her previous decision.

For the remainder of the paper we focus on an uncommitted sophisticated agent in the sense of [29] (see [24] for a recent paper surveying other approaches). She thinks of her "future selves" at various times and states as other agents that will optimize their choices when subsequent decisions are considered as given. Thus, we look for a policy which future selves will not override. A policy is a collection of binary decisions (stop or continue), one for each time and state, and an equilibrium is a policy such that no agent is incentivized to deviate.

Before formalizing this, let us observe that each agent faces the constraint of not conditioning on a null event. That is, any agent is forced to stop if continuing would lead to exiting the domain with probability one in the next step. Thus, the problem has the (random) *effective time horizon*

$$T_e := T \wedge \inf\{0 \leq t < T : P(D_{t+1}|\mathcal{F}_t) = 0\}.$$

The following adapts the basic notions of [18, 22] to our problem of conditional stopping (instead of non-exponential discounting) and extends them

to a non-Markovian setting.

Definition 2.3. A *stopping policy* is a $\{0,1\}$-valued adapted process $\theta = (\theta_t)_{t \in \mathbb{T}}$. We interpret $\theta_t(\omega) = 1$ as the agent at (t, ω) choosing to stop and $\theta_t(\omega) = 0$ as continuing. We also introduce the *continuation stopping time*

$$\mathcal{L}_t \theta = \inf\{s > t : \theta_s = 1\};$$

this is the stopping time induced by θ for a time-t agent who decides to continue. A stopping policy θ is called *admissible* if

$$P(\mathcal{L}_t \theta \lhd \sigma | \mathcal{F}_t) > 0 \quad \text{for} \quad t < T_e \qquad \text{and} \qquad \theta_t = 1 \quad \text{for} \quad t \geq T_e.$$

We denote by Θ the set of all admissible stopping policies.

Admissibility implies that every time-t agent with $t < T_e$ has a well-defined *continuation value*

$$J_t(\theta) = \frac{E[G_{\mathcal{L}_t \theta} 1_{\{\mathcal{L}_t \theta \lhd \sigma\}} | \mathcal{F}_t]}{P(\mathcal{L}_t \theta \lhd \sigma | \mathcal{F}_t)}, \quad t < T_e.$$

Naturally, she compares $J_t(\theta)$ with her stopping value G_t and prefers the larger one, or she is invariant if they are equal. (Agents with $t = T_e$ are forced to stop, so there is no decision to be taken. The value of θ_t for $t > T_e$ is unimportant and set to 1 only for specificity.) If we start with some $\theta \in \Theta$ and all agents simultaneously update their choice according to this preference while using the convention that invariant agents stick to their preexisting decision, we are led to the updated stopping policy

$$\Phi(\theta)_t = \begin{cases} 1 & \text{if } t < T_e \text{ and } G_t > J_t(\theta), \\ \theta_t & \text{if } t < T_e \text{ and } G_t = J_t(\theta), \\ 0 & \text{if } t < T_e \text{ and } G_t < J_t(\theta), \\ 1 & \text{if } t \geq T_e. \end{cases}$$

Definition 2.4. An admissible stopping policy θ is an *equilibrium (stopping policy)* if $\Phi(\theta) = \theta$.

This notion corresponds to a subgame perfect Nash equilibrium: each agent is behaving optimally if the future agents' choices are seen as given.

Example 2.5. Consider the setting of Example 2.2. In any admissible stopping policy, the time-2 agents have to stop because of the time horizon. Both time-1 agents then prefer to stop as their stopping values (10 and 3)

9

exceed the expected continuation values (3 and 2). Given those decisions, the expected continuation value for the time-0 agent is $(10 + 3)/2$ which exceeds the stopping value of 2. It easily follows that the unique equilibrium stopping policy is given by $\theta_0 = 0$, $\theta_1 \equiv 1$ and $\theta_2 \equiv 1$. The induced stopping time for the time-0 agent is $\tau \equiv 1$. This differs from the precommitted-optimal stopping time τ_{pre} of Example 2.2, and the associated expected reward of $(10 + 3)/2$ is smaller than the precommitted value function V_{pre}.

In a Markov chain setting, a natural subset of stopping policies is also of a Markovian form. Denoting by $\sigma(Y)$ the σ-field generated by a random variable Y, this can be formalized as follows.

Definition 2.6. Consider the Markovian setting of Example 2.1. A stopping policy $\theta \in \Theta$ is called *Markovian* if θ_t is $\sigma(X_t, 1_{D_t})$-measurable for all $t \in \mathbb{T}$.

If θ is admissible, this is equivalent to the existence of measurable subsets $R_t \subseteq \mathbb{X}$ such that
$$\theta_t = 1_{\{X_t \in R_t\} \cup D_t^c}.$$
Note that such equilibria are actually path-dependent through D_t, but this is the least amount of path-dependence compatible with our general definition of admissibility. In the Markovian setting, one could assume without loss of generality that all exit states (states outside B) are absorbing. Then, we have $D_t = \{X_t \in B\}$ a.s. and one can require that θ_t is (a.s.) $\sigma(X_t)$-measurable.

3 Finite-Horizon Equilibria

In this section we discuss existence, uniqueness and construction of equilibria for the case $T < \infty$.

In the classical optimal stopping problem, the value function and the optimal decision of a time-t agent are completely determined by the value functions of the agents at time $t + 1$. This fact lies at the heart of the backward recursion of dynamic programming and the Snell envelope theory. In the problem at hand, however, the conditioning event in the computation of the continuation value $J_t(\theta)$ depends on the decisions of many future selves, not only the ones at time $t + 1$. This suggests introducing an additional process S to keep track of the probability of the conditioning event given the stopping policy of all future selves; we call S the *survival process* since it is related to survival probabilities. In Theorem 3.1 below we provide a backward recursion to construct an equilibrium; its recursive formula for $J_t(\theta)$ resembles the classical case where it would be the conditional expectation

of the value process at time $t + 1$, but now this expectation is calculated under a new measure obtained by using the normalized survival process as a density.

Just like in classical optimal stopping, one type of non-uniqueness arises when an agent is invariant; that is, when the stopping and continuation values happen to be equal: $J_t(\theta) = G_t$. Thus, an algorithm for the construction of an equilibrium necessarily comes with a specific choice. The theorem stated below uses *early stopping preference,* meaning that invariant agents choose to stop, and it yields the unique equilibrium with that preference. In the classical setting, this corresponds to the first time that the Snell envelope hits the obstacle. In general, a *stopping preference* is an adapted process with binary values, defining for each (t, ω) the choice in the case of invariance. For each such preference, one can write an algorithm similar to Theorem 3.1 and it delivers the unique equilibrium with that preference. Conversely, every finite-horizon equilibrium arises in that way.

Theorem 3.1. *Let $T < \infty$ and recall that $G_t = \Delta$ on D_t^c. Define the value process $(V_t)_{t \leq T}$ and the survival process $(S_t)_{t \leq T}$ as follows. Set $V_T = G_T$ and $S_T = 1_{D_T}$. For $t = T - 1, \ldots, 0$, set*

$$J_t = \frac{E[S_{t+1} V_{t+1} | \mathcal{F}_t]}{E[S_{t+1} | \mathcal{F}_t]} \quad if \quad t < T_e,$$

$$\begin{cases} V_t = G_t \text{ and } S_t = 1 & if \ t < T_e \text{ and } G_t \geq J_t, \\ V_t = J_t \text{ and } S_t = E[S_{t+1} | \mathcal{F}_t] & if \ t < T_e \text{ and } G_t < J_t, \\ V_t = G_t \text{ and } S_t = 1_{D_t} & if \ t \geq T_e. \end{cases}$$

Then $\theta := 1_{\{G_t \geq V_t\}}$ is the unique equilibrium with preference for early stopping.

In Section 6 we will call (V, S) a *Snell pair* and discuss its connection to Snell envelopes. A generalization including the infinite-horizon case will also be provided. We nevertheless opt to provide an elementary and self-contained treatment of the finite-horizon in the present section.

Proof of Theorem 3.1. We show in Lemma 3.2 below that θ is admissible and that J_t coincides with the continuation value $J_t(\theta)$ of θ. Once that is established, the very definition of θ shows that

$$\theta = 1_{\{G_t \geq V_t\}} = \begin{cases} 0 & \text{if } t < T_e \text{ and } G_t < J_t(\theta), \\ 1 & \text{otherwise} \end{cases}$$

11

and hence θ is an equilibrium stopping policy with early stopping preference. On the other hand, the boundary condition at T_e and a backward induction allow us to see that there is at most one such equilibrium. $\qquad\square$

Lemma 3.2. *In the setting of Theorem 3.1, θ is admissible and*

$$J_t = J_t(\theta), \quad t < T_e,$$
$$E[S_{t+1}|\mathcal{F}_t] = P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t), \quad t < T_e, \qquad (3.1)$$

and for $t \le T$ we have

$$S_t = \begin{cases} P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t) & on \ D_t \cap \{\theta_t = 0\}, \\ 1 & on \ D_t \cap \{\theta_t = 1\}, \\ 0 & on \ D_t^c. \end{cases}$$

Proof. We first check that θ is admissible. Indeed, we have $\theta_t = 1$ for $t \ge T_e$, and if $t < T_e$, backward induction shows that $P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t) > 0$.

Next, we prove the formula for S_t. The last two cases are clear from the definition. Thus, we focus on showing $S_t = P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t)$ on $D_t \cap \{\theta_t = 0\}$. For $t \ge T_e$ we have $\theta_t = 1$ so nothing needs to be proved. For $t < T_e$ we argue by induction. Indeed, using the induction hypothesis to obtain (a) below,

$$S_t = E[S_{t+1}|\mathcal{F}_t]$$
$$\overset{(a)}{=} E\big[1_{D_{t+1}}1_{\{\theta_{t+1}=0\}}P(\mathcal{L}_{t+1}\theta \lhd \sigma|\mathcal{F}_{t+1}) + 1_{D_{t+1}}1_{\{\theta_{t+1}=1\}} \cdot 1 + 1_{D_{t+1}^c} \cdot 0\big|\mathcal{F}_t\big]$$
$$\overset{(b)}{=} E[P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_{t+1})|\mathcal{F}_t] = P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t),$$

where (b) holds due to

$$P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_{t+1}) = \begin{cases} P(\mathcal{L}_{t+1}\theta \lhd \sigma|\mathcal{F}_{t+1}) & on \ D_{t+1} \cap \{\theta_{t+1} = 0\}, \\ 1 & on \ D_{t+1} \cap \{\theta_{t+1} = 1\}, \\ 0 & on \ D_{t+1}^c. \end{cases} \qquad (3.2)$$

In the last identity, the first case holds since $\theta_{t+1} = 0$ implies that $\mathcal{L}_t\theta$ and $\mathcal{L}_{t+1}\theta$ agree. The second case holds because $\theta_{t+1} = 1$ entails that $\mathcal{L}_t\theta = t+1$ and $t + 1 < \sigma$ on D_{t+1}. Finally, on D_{t+1}^c we have $\sigma \le t + 1 \le \mathcal{L}_t\theta$. This completes the proof for S_t and we note that (3.1) was obtained as part of the first display above. It remains to show that

$$J_t(\theta) \equiv \frac{E[G_{\mathcal{L}_t\theta}1_{\{\mathcal{L}_t\theta \lhd \sigma\}}|\mathcal{F}_t]}{P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t)} = \frac{E[S_{t+1}V_{t+1}|\mathcal{F}_t]}{E[S_{t+1}|\mathcal{F}_t]} \equiv J_t, \quad t < T_e.$$

Since the denominators are non-zero and agree by (3.1), it suffices to show

$$E[G_{\mathcal{L}_t\theta}1_{\{\mathcal{L}_t\theta \lhd \sigma\}}|\mathcal{F}_t] = E[S_{t+1}V_{t+1}|\mathcal{F}_t], \quad t < T. \tag{3.3}$$

Indeed, (3.3) is clear for $t \geq T_e$ since that implies $P(\mathcal{L}_t\theta \lhd \sigma) = 0$. It is also clear for $t = T - 1$. For $t < T_e \wedge (T - 1)$ we argue by backward induction. We first observe that, by similar arguments as below (3.2),

$$G_{\mathcal{L}_t\theta}1_{\{\mathcal{L}_t\theta \lhd \sigma\}} = \begin{cases} G_{\mathcal{L}_{t+1}\theta}1_{\{\mathcal{L}_{t+1}\theta \lhd \sigma\}} & \text{on } D_{t+1} \cap \{\theta_{t+1} = 0\}, \\ V_{t+1} = S_{t+1}V_{t+1} & \text{on } D_{t+1} \cap \{\theta_{t+1} = 1\}, \\ 0 = S_{t+1}V_{t+1} & \text{on } D_{t+1}^c. \end{cases} \tag{3.4}$$

On the set $D_{t+1} \cap \{\theta_{t+1} = 0\}$ occurring in the first case of (3.4) we have

$$E[G_{\mathcal{L}_{t+1}\theta}1_{\{\mathcal{L}_{t+1}\theta \lhd \sigma\}}|\mathcal{F}_{t+1}] = E[S_{t+2}V_{t+2}|\mathcal{F}_{t+1}] = S_{t+1}J_{t+1} = S_{t+1}V_{t+1},$$

where the three equalities follow from the induction hypothesis, the definitions of J_{t+1} and S_{t+1}, and $J_{t+1} = V_{t+1}$ on $\{\theta_{t+1} = 0\}$, respectively. As a result, we can take conditional expectations in (3.4) and obtain that the identity $E[G_{\mathcal{L}_t\theta}1_{\{\mathcal{L}_t\theta \lhd \sigma\}}|\mathcal{F}_{t+1}] = S_{t+1}V_{t+1}$ holds everywhere. The tower property then yields the claim (3.3) and the proof is complete. □

Corollary 3.3. *In the Markovian setting of Example 2.1 with $T < \infty$, there exists a unique equilibrium with preference for early stopping and that equilibrium is Markovian.*

Proof. We observe that G_t and V_t in Theorem 3.1 are $\sigma(X_t, 1_{D_t})$-measurable for all t, and then so is θ_t. □

One can note that the stopping preference is important in the above result: it is easy to construct examples of non-Markovian equilibria by specifying a path-dependent stopping preference and taking the reward function g to be constant.

4 Infinite-Horizon Equilibria: Existence

The following result establishes the existence of infinite-horizon equilibria in a setting that includes Markov chains with a countable state space.

Theorem 4.1. *Suppose that $\mathcal{F}_t$ is a.s. discrete[1] for all $t \in \mathbb{T}$ and that $\lim_{t \to \infty} G_t = G_\infty$ a.s. Moreover, assume that*

$$P(\exists\, t \in \mathbb{T} : G_t \geq 0) > 0 \text{ and there exists } c > 1 \text{ such that} \quad (4.1)$$
$$(c^t G_t)_{t \geq 0} \text{ is uniformly bounded from above.}$$

Then an equilibrium exists.

Let us comment on the assumptions before stating the proof.

Remark 4.2. (a) Condition (4.1) covers in particular problems with discounting for a payoff function with sub-exponential growth. Consider for instance the Markov chain setting of Example 2.1 with a bounded and non-negative payoff function $g(t, x)$ and a discount factor $\delta \in (0, 1)$. Then setting $G_t = \delta^t g(t, X_t)$ for $t \in \mathbb{T}$ (and $G_\infty = 0$), we see that (4.1) is satisfied for any $c \in (1, \delta^{-1})$.

(b) The proof of Theorem 4.1 below has three steps. The construction of a limiting stopping policy θ and the verification of its optimality condition do not require (4.1) at all. The latter is used to ensure that θ is *admissible*. There are many other situations where admissibility holds, including without discounting, that can be established on a case-by-case basis, for instance the case of a Markov chain with a finite state space and a homogeneous reward $G_t = g(X_t)$. Condition (4.1) is merely one way to write a simple and fairly general result. Of course, $\sigma = \infty$ a.s. is always a sufficient condition for $P(\tau \lhd \sigma) \neq 0$, for any stopping time τ.

(c) Similarly, there are many cases where one can see directly from additional structure of G that $\mathcal{L}_t \theta < \infty$ a.s. for all $t \in \mathbb{T}$. In that case, G_∞ is irrelevant.

(d) On the other hand, existence is not guaranteed without some assumption. For instance, if $T_e = \infty$ inside the domain but $P(\sigma < \infty) = 1$ (cf. Example 5.1 below with $p_{21} > 0$), a strictly increasing reward G leads to non-existence since stopping is undesirable for any agent but $\theta \equiv 0$ is not admissible.

Proof of Theorem 4.1. For $t < \infty$, let $\mathcal{A}_t$ be the (countable) collection of atoms generating $\mathcal{F}_t$. Given $n \geq 1$, consider a modified problem with time horizon n and let $(\theta_t^n)_{0 \leq t \leq n}$ be the equilibrium stopping policy obtained by applying Theorem 3.1 with the payoff $(G_t)_{t \leq n}$. We also set $\theta_t^n \equiv 1$ for

---

[1] We call a σ-field discrete if it is generated by a countable partition of Ω. In the case of a Markov chain with countable state space one can define $\mathcal{F}_t$ as the σ-field generated by the sample paths up to time t.

$t \geq n$. Note that each θ_t^n is a binary sequence $(\theta_t^n(A))_{A \in \mathcal{A}_t}$. By a diagonal procedure we can thus find a subsequence (again denoted θ^n) which converges to a stopping policy θ in the following sense: given $t < \infty$ and $A \in \mathcal{A}_t$, we have $\theta_t^n(A) = \theta_t(A)$ for all sufficiently large n. If T_e and T_e^n denote the effective horizons, then $T_e \wedge n = T_e^n$ and thus the admissibility of θ^n for $n \geq 1$ implies that $\theta_t = 1$ for $t \geq T_e$.

To complete the proof that θ is admissible and an equilibrium, we fix arbitrary $t_0 \in \mathbb{T}$ and $A_0 \in \mathcal{F}_{t_0}$ and check the admissibility and optimality conditions at that state. For simplicity of notation, we assume that $t_0 = 0$ and $A_0 = \Omega$ (the general case differs only by writing conditional expectations and probabilities). To further simplify the notation, we set $\tau = \mathcal{L}_0 \theta$ and $\tau^n = \mathcal{L}_0 \theta^n$. The convergence of θ^n to θ implies that $\tau^n \to \tau$ a.s. More precisely, this convergence is stationary on $\{\tau < \infty\}$, yielding that $1_{\{\tau^n < \sigma < \infty\}} \to 1_{\{\tau < \sigma < \infty\}}$ a.s. Moreover, $\{\tau \lhd \sigma\} = \{\tau < \sigma < \infty\} \cup \{\sigma = \infty\}$, where the union is disjoint, and similarly for τ^n. It follows that

$$1_{\{\tau^n \lhd \sigma\}} \to 1_{\{\tau \lhd \sigma\}} \quad \text{a.s.} \tag{4.2}$$

Admissibility. We must ensure that $P(\tau \lhd \sigma) \neq 0$. In view of (4.2) it suffices to exhibit a reachable state where stopping happens for all large n, as that will imply that $P(\tau \lhd \sigma) = \lim_n P(\tau^n \lhd \sigma) > 0$. Indeed, by (4.1) we can find $t \geq 0$ and $A \in \mathcal{A}_t$ with $A \subseteq D_t$ such that $G_t(A) \geq 0$ and

$$c^t G_t(A) \geq \frac{1}{c} \sup_{s \geq 0, A' \in \mathcal{A}_s, A' \subseteq D_s} c^s G_s(A') \geq \sup_{s \geq t+1, A' \in \mathcal{A}_s, A' \subseteq D_s} c^{s-1} G_s(A')$$

and hence

$$G_t(A) \geq G_s(A') \quad \text{for all} \quad s > t, \quad A' \in \mathcal{A}_s \text{ with } A' \subseteq D_s.$$

This shows that for the agent at (t, A), stopping is optimal no matter what future selves do. In particular, $\theta_t^n(A) = 1$ for all $n \geq t$ and thus $\tau \leq t < \sigma$ on A. As a result, $P(\tau \lhd \sigma) \geq P(A) > 0$.

Optimality. It suffices to show that the continuation values converge at the fixed initial state; i.e., $J_0^n := J_0(\theta^n) \to J_0 := J_0(\theta)$. Once that is established, if $\theta_0 = 0$, then $\theta_0^n = 0$ for n large and hence $G_0 \leq J_0^n \to J_0$ shows that $\theta_0 = 0$ is optimal, and similarly for $\theta_0 = 1$. To see that

$$J_0^n = \frac{E[G_{\tau^n} 1_{\{\tau^n \lhd \sigma\}}]}{P(\tau^n \lhd \sigma)} \to \frac{E[G_\tau 1_{\{\tau \lhd \sigma\}}]}{P(\tau \lhd \sigma)} = J_0,$$

note that the denominators are non-zero by admissibility and $P(\tau^n \lhd \sigma) \to P(\tau \lhd \sigma)$ by (4.2). In view of $\tau^n \to \tau$ a.s. we have $G_{\tau^n} \to G_\tau$ a.s. on $\{\tau < \infty\}$.

As we have assumed that $G_n \to G_\infty$ a.s., this convergence holds everywhere. Using also the standing assumption that $E[\sup_{t \leq T} |G_t| 1_{D_t}] < \infty$ and (4.2), the convergence of the numerators follows by dominated convergence. $\square$

Corollary 4.3. *Consider the Markovian setting (Example 2.1) under the conditions of Theorem 4.1. Then there exists a Markovian equilibrium.*

Proof. We revisit the proof of Theorem 4.1. Each of the finite-horizon problems is Markovian, so Corollary 3.3 shows that θ^n is Markovian. Since θ_t was constructed as a pointwise limit of θ_t^n, it is again $\sigma(X_t, 1_{D_t})$-measurable. $\square$

We shall see in Example 5.3 that this corollary cannot be improved in a time-homogeneous setting: the equilibria may nevertheless be time-dependent.

5 Infinite-Horizon Equilibria: Examples

5.1 Non-Uniqueness and Non-Markovian Equilibria

The following example shows that in the infinite-horizon case, multiple equilibria may exist. In these equilibria, all agents' choices are uniquely determined; i.e., the non-uniqueness is not merely due to different choices of agents that are invariant between stopping and continuing. Moreover, the multiplicity arises even within the class of time-homogeneous Markov equilibria. The example also shows that non-Markovian equilibria may exist in a Markovian setting.

Example 5.1. Consider a homogeneous Markov chain X on the states $\{0, 1, 2\}$ with initial value $X_0 = 1$ and transition probabilities (p_{ij}) in its natural filtration. Only the states in $B = \{1, 2\}$ are relevant for the agents, meaning that $\sigma = \inf\{t \geq 0 : X_t = 0\}$ and $D_t = \{X_1, \ldots, X_t \in B\}$. The payoff $G_t = \delta^t g(X_t)$ is given by a function g of the current state and a discount factor $\delta \in (0, 1)$. Specifically,

$$p_{10} = p_{11} = p_{12} = 1/3 \quad \text{and} \quad g(0) = \Delta, \quad g(1) = 1, \quad g(2) = a,$$

where a is a constant satisfying

$$1 < \frac{3 - \delta}{2\delta} < a < \frac{2 - \delta}{\delta}.$$

We also assume that $p_{20} \neq 1$; the other transition probabilities are arbitrary. Then, there are exactly two Markovian equilibria:

(i) stop everywhere; i.e., $\theta \equiv 1$;

(ii) stop if the chain is at State 2 or has exited; i.e., $\theta_t = 1_{\{X_t=2\}\cup D_t^c}$.

If $p_{21} > 0$, there are further, non-Markovian equilibria. In these equilibria, the induced stopping time for a given agent at some state (t,ω) coincides with the stopping time induced by (i) or (ii), conditionally on $\mathcal{F}_t$.

Proof. We first note that as $a > g(1)$ and $\delta < 1$, the only optimal choice for a time-t agent on $\{X_t = 2\}$ is to stop, no matter what future agents choose.

(a) To see that $\theta \equiv 1$ is an equilibrium, consider an agent at State 1, without loss of generality at $t = 0$. Then

$$J_0(\theta) = \frac{\delta(p_{11} + ap_{12})}{1 - p_{10}} = \frac{\delta(1/3 + a/3)}{2/3} = \delta\frac{1+a}{2} < 1 = G_0, \qquad (5.1)$$

showing that stopping is indeed optimal and θ is an equilibrium.

(b) The policy θ defined by $\theta_t = 1_{\{X_t=2\}\cup D_t^c}$ is admissible. To see that it defines an equilibrium, consider again the time-0 agent at State 1. Let τ_j be the first hitting time of state j, so that $\sigma = \tau_0$ and $\tau := \mathcal{L}_0\theta = \tau_0 \wedge \tau_2$. We have $\{\tau \lhd \sigma\} = \{\tau_2 < \tau_0\}$ a.s. since $P(\tau_0 \wedge \tau_2 = \infty) = 0$. As $p_{10} = p_{12}$, the symmetry between $\{\tau_2 < \tau_0\}$ and $\{\tau_0 < \tau_2\}$ yields that $P(\tau_2 < \tau_0) = P(\tau_0 < \tau_2) = 1/2$ and thus $P(\tau \lhd \sigma) = 1/2$. Moreover,

$$E[\delta^\tau g(X_\tau)1_{\tau\lhd\sigma}] = a\sum_{k\geq 1}\delta^k P(\tau_2 = k,\, k < \tau_0) = a\sum_{k\geq 1}\delta^k(1/3)^k = \frac{a\delta}{3 - \delta}$$

since $P(\tau_2 = k,\, k < \tau_0) = P(X_1 = \cdots = X_{k-1} = 1,\, X_k = 2) = p_{11}^{k-1}p_{12}$. It follows that

$$J_0(\theta) = \frac{E[\delta^\tau G_\tau 1_{\tau\lhd\sigma}]}{P(\tau \lhd \sigma)} = \frac{2a\delta}{3 - \delta} > 1 = G_0, \qquad (5.2)$$

showing that continuation is optimal. Thus θ is an equilibrium.

(c) Let θ be a Markovian equilibrium; we show that θ must be one of the two above policies. We have already observed that any agent at State 2 must stop. The same holds for any agent at State 0, by admissibility. That is, $1_{\{X_t=2\}\cup D_t^c} \leq \theta_t \leq 1$ for all $t \in \mathbb{T}$. If no other agent stops, θ is the policy of (ii). Otherwise there exists a time-t agent stopping at State 1: $\theta_t = 1$ on $\{X_t = 1\}$. But then the same calculation as in (5.1) shows that any agent at time $(t-1)$ and State 1 must also stop, etc., so that $\theta_s \equiv 1$ for all $s \leq t$. As a result, the set of all agents at State 1 that stop can be thought of as a half-line starting at $t = 0$. If this half-line is infinite, θ is the equilibrium from (i). If not, there is some maximal $t < \infty$ where the time-t agent stops,

meaning that $\theta_s \equiv 1$ for $s \leq t$ and $\theta_s = 1_{\{X_s=2\}\cup D_s^c}$ for $s > t$. But now the calculation in (5.2) shows that stopping is not optimal for any time-t agent on $\{X_t = 1\}$, a contradiction.

(d) Next, we give an example of a non-Markovian equilibrium. Indeed, set $\theta_0 = \theta_1 \equiv 1$. For $t \geq 2$, we define

$$\theta_t(\omega) = \begin{cases} 0 & \text{if } \omega \in \{X_1 = 2, \, X_t = 1\} \cap D_t, \\ 1 & \text{else.} \end{cases}$$

Simple calculations analogous to (5.1) and (5.2) show that θ is an equilibrium. If $p_{21} > 0$, both cases in the definition of θ happen with positive probability so that θ is indeed non-Markovian.

(e) Let θ be any equilibrium, possibly non-Markovian. The first argument from (c) still shows that for (t, ω) such that $X_t(\omega) = 1$ and $\theta_t(\omega) = 1$, it follows that $\theta_{t-1}(\omega) = 1$. However the second argument from (c) merely shows that for (t, ω) such that $X_{t-1}(\omega) = X_t(\omega) = 1$ and $\theta_{t-1}(\omega) = 0$, it follows that $\theta_t = 0$. (But this need not hold if $X_{t-1}(\omega) = 2$, in contrast to the Markovian case where the policy cannot depend directly on X_{t-1}). This implies that given the past up to time t, the stopping time induced by θ is either immediate stopping as in (i) or the first exit time of $\{1\}$ as in (ii). Note that, as in (d), the choice between these two may depend on ω. $\qquad\square$

Remark 5.2. (a) The finite-horizon version of Example 5.1 has a unique equilibrium, given by stopping everywhere. This follows by a backward recursion and the same calculation as in (5.1), since the time-T agents have to stop. The limit of this equilibrium as $T \to \infty$ is the infinite-horizon equilibrium (i). On the other hand, the equilibrium (ii) does not arise as a limit of finite-horizon equilibria.

(b) In this particular example the two Markovian equilibria are ordered: equilibrium (ii) has a larger value function for all agents. It is worth noting that the limit equilibrium is the inferior one.

(c) Example 5.3 shows that in general, no dominating equilibrium exists. One can also construct simple examples where the equilibrium value processes and stopping policies corresponding to different preferences are not ordered.

5.2 Non-Existence of Time-Homogeneous Equilibria

In this section we construct an example of a time-homogeneous Markov chain which admits Markovian equilibria but no time-homogeneous equilibria. In

that sense, Theorem 4.1 and Corollary 4.3 cannot be improved, and a restriction to time-homogeneous notions is not possible (or will lead to non-existence). Importantly, the example also shows that the remarkable iterative approach of [22] does not apply in our setting. Indeed, in the problem of non-exponential discounting with decreasing impatience, an iterated application of Φ (from a suitable starting point) produces a monotone sequence which converges to a time-homogeneous equilibrium. In our case however, the iteration can fail to be monotone. This can be related to a failure of both inequalities of the dynamic programming principle, whereas decreasing impatience preserves one.

Example 5.3. Consider the homogeneous Markov chain X on $\{0, 1, 2, 3, 4\}$ with transition probabilities as labeled next to the edges in Figure 2. In particular, States 0, 3 and 4 are absorbing. We set $B = \{1, 2, 3, 4\}$ so that 0 is the only exit state. The payoff process is given by $G_t = \delta^t g(X_t)$ where $\delta \in (0, 1)$ is the discount factor and $g(1) = a$, $g(2) = 2$, $g(3) = 0$, $g(4) = b$ as labeled in the boxes in Figure 2. To avoid trivialities, we assume that the initial position is one of the non-absorbing states, i.e., either $X_0 = 1$ or $X_0 = 2$, and we also restrict our attention to equilibria that stop at State 3.[2] A Markovian equilibrium $\theta_t = f(t, X_t)$ is called time-homogeneous if f does not depend on t.
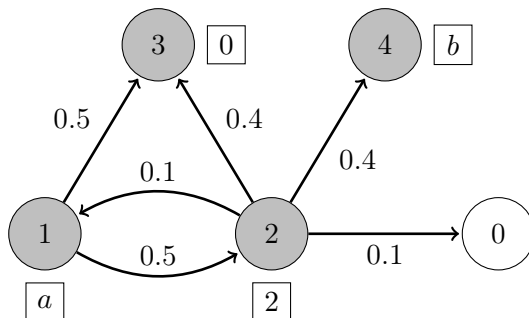

Figure 2: The Markov chain of Example 5.3, with states x labeled in circles and payoffs $g(x)$ in boxes.

We fix $0 < a < \delta < 1 < 2 < b$ such that the following inequalities are

---

[2]Since State 3 is absorbing and $g(3) = 0$, all policies have zero reward for an agent at State 3 who is therefore invariant. This leads to an infinity of (uninteresting) equilibria. If early stopping preference is assumed, stopping at State 3 is a consequence rather than a condition.

satisfied:

$$a < \delta, \quad \delta(a + 4b) < 18, \tag{5.3}$$

$$\delta(\delta + 4b) > 18, \quad 0.01\delta^3 \min(5a, b\delta^2) + 0.2b\delta^3 + 4b\delta > 17.9, \tag{5.4}$$

$$\delta^2(\max(\delta, 0.25b\delta^2) + 4b) < 18.9a. \tag{5.5}$$

One possible choice is $\delta = 0.999$, $a = 0.96$, $b = 4.257$. Then, up to a.s. equivalence:

(i) All equilibria are Markovian.

(ii) There exists no time-homogeneous equilibrium.

(iii) There are exactly two equilibria and they are given by shifts of one another. Indeed, let

$$\theta_t^1 = f(t, X_t), \ \theta_t^2 = f(t+1, X_t), \ \theta_t^3 = f(t+2, X_t), \ \theta_t^4 = f(t+3, X_t)$$

where

$$f(t, x) = \begin{cases} 1_{R_4}(x), & t \equiv 0 \mod 4 \\ 1_{R_3}(x), & t \equiv 1 \mod 4 \\ 1_{R_2}(x), & t \equiv 2 \mod 4 \\ 1_{R_1}(x), & t \equiv 3 \mod 4 \end{cases}$$

for

$$R_1 = \{0, 1, 2, 3, 4\}, \ R_2 = \{0, 2, 3, 4\}, \ R_3 = \{0, 3, 4\}, \ R_4 = \{0, 1, 3, 4\}.$$

Then θ^i, $i = 1, \ldots, 4$ are Markovian equilibria, and exactly two of them are distinct up to a.s. equivalence: if $X_0 = 1$, then $\theta^1 = \theta^4$ and $\theta^2 = \theta^3$, whereas if $X_0 = 2$, then $\theta^1 = \theta^2$ and $\theta^3 = \theta^4$, a.s. [3]

That all equilibria are Markovian is related to the filtration being relatively small (a.s.) due to various states being absorbing—this fact should not be given too much weight. The proofs for the other items are rather lengthy, so let us try to summarize the key mechanics heuristically. First, the dynamics are engineered such that in any equilibrium, the decision of a time-t agent

---

[3]Recall that the initial condition is deterministic in our basic setup. If $X_0 = 1$, then State 1 can only be visited at odd t and State 2 only at even t; the reverse is true if $X_0 = 2$. This leads to the a.s. equivalence of two pairs of θ^i. Whereas if we treated the initial state as not being fixed (as may be considered natural in a Markovian framework) or if we assumed that X_0 has a distribution with support including both states, then all four equilibria would be distinct.

depends only on the agents at $t+1$. Moreover, as highlighted in Lemma 5.4 below, it embeds two types of agents that cannot agree (and cannot even agree to disagree): Call Minnie$_t$ the agent at State 2 and time t and Donald$_t$ the agent at State 1 and time t. Minnie prefers to live in harmony and always wants to agree, whereas Donald is only happy if he contradicts Minnie. Suppose that at some time t, Donald$_t$ says "1" (stop). Then Minnie$_{t-1}$ also opts for 1, but the combative Donald$_{t-2}$ immediately replies with 0, thus implying time-inhomogeneity as he is contradicting Donald$_t$. The situation is similar if Donald$_t$ starts with 0.

Conversely, there are exactly two equilibria because the above backward recursion also implies a unique forward recursion once the initial Donald$_0$ (or Minnie$_0$, depending on what the initial state X_0 is) fixes one of the two possible choices 0 or 1.

Proof of (i)–(iii). Let us first observe that any equilibrium stopping policy θ (possibly non-Markovian) must stop on $\{X_t = 0\}$, by admissibility. Furthermore, it must stop on $\{X_t = 4\}$: State 4 is absorbing and $g(4) > 0$, so that continuing is never optimal due to the discount factor $\delta < 1$. Since we have also convened that θ stops on $\{X_t = 3\}$, we may henceforth restrict our attention to equilibria satisfying $\theta_t = 1$ on $\{X_t \in \{0, 3, 4\}\}$ for all $t \in \mathbb{T}$.

(i) Let θ be any equilibrium; we show that θ is Markovian (or rather, a.s. equivalent to a Markovian equilibrium). Indeed, suppose first that the initial condition is $X_0 = 1$, and fix $t \in \mathbb{T}$. We have that $\theta_t = 1$ on $\{X_t \in \{0, 3, 4\}\}$. But since $0, 3, 4$ are absorbing states, $\{X_t \in \{0, 3, 4\}\} = \cup_{s \leq t} \{X_s \in \{0, 3, 4\}\}$. Suppose that $t \in \mathbb{T}$ is odd. Then $\{X_t = 1\}$ is a nullset, so that up to a.s. equivalence, only the value of θ_t on $\{X_t = 2\}$ has not been determined yet. But due to the absorption on $\{0, 3, 4\}$ and the fact that exactly one of the sets $\{X_s = 1\}$ and $\{X_s = 2\}$ has positive probability for every $s \leq t$, we have $\{X_t = 2\} = \{X_1 = 2, X_2 = 1, X_3 = 2, \ldots, X_t = 2\}$ which implies that $\{X_t = 2\}$ is an atom in $\mathcal{F}_t$. In particular, θ_t is a.s. constant on $\{X_t = 2\}$, and since $\theta_t = 1$ a.s. on $\{X_t = 2\}^c$, it follows that θ_t is of Markovian form. The situation is analogous if t is even, and hence θ is Markovian. The initial condition is $X_0 = 2$ is dealt with similarly.

The proof of (ii) and (iii) necessitates the following lemma which describes the Minnie–Donald relationship sketched above.

Lemma 5.4. *Let $0 < a < \delta < 1 < 2 < b$ satisfy (5.3)–(5.5) and let θ be an admissible stopping policy such that $\theta_t = 1$ on $\{X_t \in \{0, 3, 4\}\}$ for $t \in \mathbb{T}$. Then for all $t \geq 1$,*

(P1) *if $\theta_t = 1$ on $\{X_t = 1\}$, then $\Phi(\theta)_{t-1} = 1$ on $\{X_{t-1} = 2\}$;*

(P2) *if $\theta_t = 0$ on $\{X_t = 1\}$, then $\Phi(\theta)_{t-1} = 0$ on $\{X_{t-1} = 2\}$;*

(P3) *if $\theta_t = 1$ on $\{X_t = 2\}$, then $\Phi(\theta)_{t-1} = 0$ on $\{X_{t-1} = 1\}$;*

(P4) *if $\theta_t = 0$ on $\{X_t = 2\}$, then $\Phi(\theta)_{t-1} = 1$ on $\{X_{t-1} = 1\}$.*

The proof of the lemma is reported after the proof of (ii) and (iii).

(ii) Define the 4-periodic sequence (R_n) by $R_n = R_{n+4\mathbb{Z}}$ where $R_1, \ldots, R_4$ are as in (iii) above. Note that $R_1, \ldots, R_4$ exhaust all combinations of $\{0, 3, 4\}$ and the remaining states. Thus, a time-homogeneous equilibrium θ must (a.s.) be of the form $\theta_t = 1_{R_n}(X_t)$, $t \in \mathbb{T}$, for some n. On the other hand, for any $t \in \mathbb{T}$, (P1)–(P4) imply that $\Phi(\Phi(1_{R_n}(X_t))) = 1_{R_{n+2}}(X_t) \neq 1_{R_n}(X_t)$, thus ruling out the existence of a time-homogeneous equilibrium. (We iterate Φ twice to ensure that the policies differ also modulo a.s. equivalence).

(iii) Admissibility of θ^i ($i = 1, 2, 3, 4$) holds since $D_t = \{X_t = 0\}^c$ a.s. (due to $\{0\}$ being absorbing) and since from any non-absorbing state there is a positive probability of reaching $\{3, 4\}$ before reaching $\{0\}$. Moreover, $\Phi(\theta^i) = \theta^i$ follows by direct verification using (P1)–(P4). Hence, θ^i are equilibria.

To see that there are exactly two equilibria, suppose first that the initial condition is $X_0 = 1$ and let θ be a (necessarily Markovian) equilibrium. Modulo a.s. equivalence, θ is completely determined by its values on $\{X_0 = 1\}$, $\{X_1 = 2\}$, $\{X_3 = 1\}$, etc., since State 1 can only be visited at even times and State 2 only at odd times. Next, we use (P1)–(P4): Suppose that $\theta_0 = 1$ on $\{X_0 = 1\}$. This implies $\theta_1 = 0$ on $\{X_1 = 2\}$, which implies $\theta_2 = 0$ on $\{X_2 = 1\}$, which implies $\theta_3 = 1$ on $\{X_3 = 2\}$, etc. Therefore, we have $\theta = \theta^1 = \theta^4$ a.s.

Alternately, $\theta_0 = 0$ on $\{X_0 = 1\}$. This implies $\theta_1 = 1$ on $\{X_1 = 2\}$, thus $\theta_2 = 1$ on $\{X_2 = 1\}$, thus $\theta_3 = 0$ on $\{X_3 = 2\}$, etc. In particular, we have $\theta = \theta^2 = \theta^3$ a.s.

The case of the initial condition $X_0 = 2$ is similar.[4]

$\square$

Proof of Lemma 5.4. Let $t \geq 1$ and set

$$\tilde{J}_t(\theta) = \frac{E[\delta^{\mathcal{L}_t\theta - t} g(X_{\mathcal{L}_t\theta}) 1_{\{\mathcal{L}_t\theta < \sigma\}} | \mathcal{F}_t]}{P(\mathcal{L}_t\theta < \sigma | \mathcal{F}_t)}$$

---

[4]If the initial state is not considered fixed or if it is random with $P(X_0 = 1) > 0$ and $P(X_0 = 2) > 0$, then (P1)–(P4) imply that θ is a.s. equal to exactly one of the four θ^i, uniquely determined by the values of θ_0 on $\{X_0 = 1\}$ and $\{X_0 = 2\}$.

so that $J_t(\theta) = \delta^t \tilde{J}_t(\theta)$ is the continuation value at time t. Note that comparing $J_t(\theta)$ with G_t is equivalent to comparing $\tilde{J}_t(\theta)$ with $g(X_t)$. We first show (P1) and (P3).

(P1): Suppose that $X_{t-1} = 2$. This implies $X_t \in \{0, 1, 3, 4\}$ and thus the assumption of (P1) yields that $\theta_t = 1$, $\mathcal{L}_{t-1}\theta = t$ and $\tilde{J}_{t-1}(\theta) = \delta(0.1a + 0.4b)/0.9$. By the second part of (5.3), we have $\tilde{J}_{t-1}(\theta) < 2 = g(2)$ and thus $\Phi(\theta)_{t-1} = 1$ as claimed.

(P3): If $X_{t-1} = 1$, then $X_t \in \{2, 3\}$ and the assumption of (P3) imply $\theta_t = 1$, $\mathcal{L}_{t-1}\theta = t$ and $\tilde{J}_{t-1}(\theta) = \delta$. By the first part of (5.3), we have $\tilde{J}_{t-1}(\theta) > a = g(1)$ and thus $\Phi(\theta)_{t-1} = 0$.

Next, we analyze (P2) and (P4). Denote by $h_t(\theta)$ and $p_t(\theta)$ the numerator and denominator of $\tilde{J}_t(\theta)$. It is clear that $h_t(\theta) \le \tilde{J}_t(\theta) \le b$ for all t, since b is the maximum possible payoff. By iterated conditioning, we have that on the set $\{X_t = 1\} \subseteq \{X_{t+1} \in \{2, 3\}\}$,

$$
\begin{aligned}
h_t(\theta) &= E[\delta^{\mathcal{L}_t\theta - t} g(X_{\mathcal{L}_t\theta}) 1_{\{\mathcal{L}_t\theta < \sigma\}} | \mathcal{F}_t] \\
&= E[1_{\{X_{t+1}=2\}} \delta^{\mathcal{L}_t\theta - t} g(X_{\mathcal{L}_t\theta}) 1_{\{\mathcal{L}_t\theta < \sigma\}} + 1_{\{X_{t+1}=3\}} \delta g(3) | \mathcal{F}_t] \\
&= E[1_{\{X_{t+1}=2, \theta_{t+1}=1\}} \delta g(2) | \mathcal{F}_t] \\
&\quad + E[1_{\{X_{t+1}=2, \theta_{t+1}=0\}} \delta^{\mathcal{L}_{t+1}\theta - (t+1)} \delta g(X_{\mathcal{L}_{t+1}\theta}) 1_{\{\mathcal{L}_{t+1}\theta < \sigma\}} | \mathcal{F}_t] \\
&= \delta E[1_{\{X_{t+1}=2, \theta_{t+1}=1\}} 2 + 1_{\{X_{t+1}=2, \theta_{t+1}=0\}} h_{t+1}(\theta) | \mathcal{F}_t], \quad (5.6)
\end{aligned}
$$

where we have used that $\mathcal{L}_t\theta = t + 1$ if $\theta_{t+1} = 1$ and $\mathcal{L}_t\theta = \mathcal{L}_{t+1}\theta$ if $\theta_{t+1} = 0$. Similarly, we deduce that on $\{X_t = 1\}$,

$$
p_t(\theta) = E[1_{\{X_{t+1}=2, \theta_{t+1}=1\}} + 1_{\{X_{t+1}=2, \theta_{t+1}=0\}} p_{t+1}(\theta) | \mathcal{F}_t] + 0.5, \quad (5.7)
$$

and on $\{X_t = 2\}$,

$$
h_t(\theta) = \delta E[1_{\{X_{t+1}=1, \theta_{t+1}=1\}} a + 1_{\{X_{t+1}=1, \theta_{t+1}=0\}} h_{t+1}(\theta) | \mathcal{F}_t] + 0.4b\delta, \quad (5.8)
$$
$$
p_t(\theta) = E[1_{\{X_{t+1}=1, \theta_{t+1}=1\}} + 1_{\{X_{t+1}=1, \theta_{t+1}=0\}} p_{t+1}(\theta) | \mathcal{F}_t] + 0.8. \quad (5.9)
$$

Equations (5.6)-(5.9) yield the following bounds: on $\{X_t = 1\}$,

$$
h_t(\theta) \le \delta E[1_{\{X_{t+1}=2\}} \max(2, h_{t+1}(\theta)) | \mathcal{F}_t], \quad (5.10)
$$
$$
h_t(\theta) \ge \delta E[1_{\{X_{t+1}=2\}} \min(2, h_{t+1}(\theta)) | \mathcal{F}_t], \quad (5.11)
$$
$$
p_t(\theta) \ge 0.5 + E[1_{\{X_{t+1}=2\}} p_{t+1}(\theta) | \mathcal{F}_t], \quad (5.12)
$$

23

and on $\{X_t = 2\}$,

$$h_t(\theta) \le \delta E[1_{\{X_{t+1}=1\}} \max(a, h_{t+1}(\theta))|\mathcal{F}_t] + 0.4b\delta, \qquad (5.13)$$

$$h_t(\theta) \ge \delta E[1_{\{X_{t+1}=1\}} \min(a, h_{t+1}(\theta))|\mathcal{F}_t] + 0.4b\delta, \qquad (5.14)$$

$$0.8 + E[1_{\{X_{t+1}=1\}}p_{t+1}(\theta)|\mathcal{F}_t] \le p_t(\theta) \le 0.9. \qquad (5.15)$$

(P2): Suppose that $\theta_t = 0$ on $\{X_t = 1\}$. Throughout the proof of (P2), we assume that we are on the set $\{X_{t-1} = 2\}$; i.e., all statements are conditional on $X_{t-1} = 2$. Then, $1_{\{X_t=1,\theta_t=0\}} = 1_{\{X_t=1\}}$ and $1_{\{X_t=1,\theta_t=1\}} = 0$. To establish that $\Phi(\theta)_{t-1} = 0$, it suffices to show that $\tilde{J}_{t-1}(\theta) > g(2) = 2$. To that end, we derive a lower bound for $h_{t-1}(\theta)$ and an upper bound for $p_{t-1}(\theta)$ (conditionally on $X_{t-1} = 2$). Let

$$\gamma_{t-1} := P(X_t = 1, X_{t+1} = 2, \theta_{t+1} = 1|\mathcal{F}_{t-1})$$

and note that $0 \le \gamma_{t-1} \le 0.05$. Starting from the fact that $h_{t+3}(\theta) \ge 0.4b\delta$ on $\{X_{t+3} = 2\}$ by (5.14), we use (5.11) to see that on $\{X_{t+2} = 1\}$,

$$\begin{aligned} h_{t+2}(\theta) &\ge \delta E[1_{\{X_{t+3}=2\}} \min(2, h_{t+3}(\theta))|\mathcal{F}_{t+2}] \\ &\ge \delta E[1_{\{X_{t+3}=2\}} \min(2, 0.4b\delta)|\mathcal{F}_{t+2}] \\ &= 0.5\delta \min(2, 0.4b\delta) = \min(\delta, 0.2b\delta^2), \end{aligned}$$

and then (5.14) and $a < \delta$ to deduce that on $\{X_{t+1} = 2\}$,

$$\begin{aligned} h_{t+1}(\theta) &\ge \delta E[1_{\{X_{t+2}=1\}} \min(a, h_{t+2}(\theta))|\mathcal{F}_{t+1}] + 0.4b\delta \\ &\ge \delta E[1_{\{X_{t+2}=1\}} \min(a, \min(\delta, 0.2b\delta^2))|\mathcal{F}_{t+1}] + 0.4b\delta \\ &= 0.1\delta \min(a, 0.2b\delta^2) + 0.4b\delta = 0.1A, \qquad (5.16) \end{aligned}$$

where

$$A := \delta \min(a, 0.2b\delta^2) + 4b\delta.$$

By (5.8), the assumption of (P2), (5.6), and iterated conditioning, we have

$$\begin{aligned} &h_{t-1}(\theta) \\ &= \delta E[1_{\{X_t=1\}}h_t(\theta)|\mathcal{F}_{t-1}] + 0.4b\delta \\ &= \delta^2 E[1_{\{X_t=1,X_{t+1}=2,\theta_{t+1}=1\}}2 + 1_{\{X_t=1,X_{t+1}=2,\theta_{t+1}=0\}}h_{t+1}(\theta)|\mathcal{F}_{t-1}] + 0.4b\delta. \end{aligned}$$

Substituting the lower bound (5.16) for $h_{t+1}(\theta)$ into the above equation,

$$h_{t-1}(\theta) \ge \delta^2(2\gamma_{t-1} + 0.1A(0.05 - \gamma_{t-1})) + 0.4b\delta.$$

Similarly, using (5.9), the assumption of (P2), (5.7), iterated conditioning and (5.15), we obtain that

$$
\begin{aligned}
p_{t-1}(\theta) &= E[1_{\{X_t=1\}} p_t(\theta) | \mathcal{F}_{t-1}] + 0.8 \\
&= E[1_{\{X_t=1, X_{t+1}=2, \theta_{t+1}=1\}} + 1_{\{X_t=1, X_{t+1}=2, \theta_{t+1}=0\}} p_{t+1}(\theta) | \mathcal{F}_{t-1}] \\
&\quad + 0.5 P(X_t = 1 | \mathcal{F}_{t-1}) + 0.8 \\
&\leq E[1_{\{X_t=1, X_{t+1}=2, \theta_{t+1}=1\}} + 1_{\{X_t=1, X_{t+1}=2, \theta_{t+1}=0\}} 0.9 | \mathcal{F}_{t-1}] + 0.85 \\
&= \gamma_{t-1} + 0.9(0.05 - \gamma_{t-1}) + 0.85 = 0.1\gamma_{t-1} + 0.895.
\end{aligned}
$$

These two bounds yield that

$$
\tilde{J}_{t-1}(\theta) = \frac{h_{t-1}(\theta)}{p_{t-1}(\theta)} \geq \frac{\delta^2(2\gamma_{t-1} + 0.1A(0.05 - \gamma_{t-1})) + 0.4b\delta}{0.1\gamma_{t-1} + 0.895}.
$$

As a consequence, a sufficient condition for $\tilde{J}_{t-1}(\theta) > 2$ is that

$$
f(y) := \delta^2(2y + 0.1A(0.05 - y)) + 0.4b\delta - (0.2y + 1.79) > 0 \quad \text{for all } y \in [0, 0.05].
$$

Since f is linear in y, this is equivalent to $f(0) > 0$ and $f(0.05) > 0$, which is precisely (5.4).

(P4): Suppose that $\theta_t = 0$ on $\{X_t = 2\}$, so that $1_{\{X_t=2, \theta_t=0\}} = 1_{\{X_t=2\}}$ and $1_{\{X_t=2, \theta_t=1\}} = 0$. We assume throughout the proof of (P4) that we are on the set $\{X_{t-1} = 1\}$, and we shall establish that $\Phi(\theta)_{t-1} = 1$ by showing the inequality $\tilde{J}_{t-1}(\theta) < g(1)$.

Proceeding similarly as in the proof of (P2), we start from the fact that $h_{t+3}(\theta) \leq b$ and $p_{t+3}(\theta) \geq 0$ and apply (5.10), (5.13), (5.12) and (5.15) repeatedly to derive the following bounds on $\{X_t = 2\}$:

$$
h_t(\theta) \leq 0.1\delta \max(\delta, 0.25b\delta^2) + 0.4b\delta, \qquad p_t(\theta) \geq 0.89.
$$

Then, we use (5.6), (5.7), the assumption of (P4) and (5.5) to deduce that

$$
\begin{aligned}
\tilde{J}_{t-1}(\theta) = \frac{h_{t-1}(\theta)}{p_{t-1}(\theta)} &= \frac{\delta E[1_{\{X_t=2\}} h_t(\theta) | \mathcal{F}_{t-1}]}{E[1_{\{X_t=2\}} p_t(\theta) | \mathcal{F}_{t-1}] + 0.5} \\
&\leq \frac{0.5\delta\{0.1\delta \max(\delta, 0.25b\delta^2) + 0.4b\delta\}}{0.5 \cdot 0.89 + 0.5} \\
&= \frac{\delta^2\{\max(\delta, 0.25b\delta^2) + 4b\}}{18.9} < a = g(1).
\end{aligned}
$$

The proof is complete. □

Remark 5.5. The results in Example 5.3 extend to the undiscounted case $\delta = 1$ if we focus on equilibria that stop in the absorbing State 4 (or focus on equilibria with early stopping preference). The situation is the same as for State 3: without discounting, any agent at State 4 is invariant between stopping and continuing which leads to an infinity of equilibria.

6 Snell Pairs and Equilibria

In this section we provide a theory which extends both the Snell envelope of classical optimal stopping and the recursion from the finite-horizon case in Theorem 3.1. As mentioned in Section 3, the value process V (which is the Snell envelope of G in the classical case) needs to be complemented with the survival process S to provide a sufficient statistic for an agent's optimality criterion. We introduce the Snell pair (V, S) pragmatically in Definition 6.1 by stating the properties that will be used most often in the proofs. Alternately, both processes can be described through a more elegant Snell envelope property (Lemma 6.3), whence the terminology. The main result of this section will be a correspondence between Snell pairs and equilibria; see Theorem 6.5 and its corollary.

We focus on equilibria with early stopping preference throughout this section. Other preferences could be accommodated but lead to (even) heavier notation. For the infinite-horizon case $T = \infty$, we assume throughout that

$$G_\infty = \limsup_{t \to \infty} G_t. \tag{6.1}$$

We also recall that $\mathbb{T} = \{0, 1, \dots\}$ if $T = \infty$, so that $\mathbb{T} \cup \{T\}$ will be used when the horizon is included in the index set.

Definition 6.1. A pair (V, S) consisting of adapted processes $V = (V_t)_{t \in \mathbb{T}}$ and $S = (S_t)_{t \in \mathbb{T} \cup \{T\}}$ is said to be a *Snell pair* (with early stopping preference) if the following hold:

(i) $0 < S_t \leq 1$ on D_t and $S_t = 0$ on D_t^c for all $t \in \mathbb{T}$, and $V_t = G_t$ for all $t \geq T_e$.[5]

(ii) Given S, V is the smallest adapted process which dominates G and renders $(SV)_{\cdot \wedge T_e}$ a supermartingale.[6]

---

[5]The property that $V_t = G_t$ for $t \geq T_e$ is in fact redundant with (iii).

[6]We follow the usual convention that supermartingale properties, Snell envelopes, etc., are understood on $\mathbb{T}$ unless explicitly mentioned; that is, $t = \infty$ is not included.

(iii) Given V, S is the smallest nonnegative supermartingale on $\mathbb{T} \cup \{T\}$ satisfying $S_t = 1$ on $D_t \cap \{V_t = G_t\}$ for all $t \in \mathbb{T}$ as well as $S_\infty = 1_{\{\sigma=\infty\}}$ if $T = \infty$.

(iv) For all $t_0 < T$, the process $(S^{t_0}V)_{\cdot \wedge T_e}$ is a supermartingale, where

$$S_t^{t_0} := 1_{\{t \neq t_0\}}S_t + 1_{\{t=t_0\}}E[S_{t+1}|\mathcal{F}_t].$$

Some comments on the definition are in order before we connect Snell pairs with equilibrium stopping policies.

Lemma 6.2. *Properties (i)–(iii) imply the following "martingale properties away from the obstacle,"*

(v) *if $t < T_e$ and $V_t > G_t$, then $S_t = E[S_{t+1}|\mathcal{F}_t]$ and $S_tV_t = E[S_{t+1}V_{t+1}|\mathcal{F}_t]$.*

Proof. If the first identity fails for some t, replacing S_t by $E[S_{t+1}|\mathcal{F}_t]$ yields a smaller supermartingale with the required properties, contradicting (iii). If the second identity fails, replacing V_t by $E[S_{t+1}V_{t+1}|\mathcal{F}_t]/S_t$ yields a smaller process with the required properties, contradicting (ii). □

Lemma 6.3. *Properties (i)–(iii) are jointly equivalent to the following:*

(i') *$S_t > 0$ on D_t for all $t \in \mathbb{T}$ and $V_t = G_t$ for all $t \geq T_e$.*

(ii') *$(SV)_{\cdot \wedge T_e}$ is the Snell envelope of $(SG)_{\cdot \wedge T_e}$.*

(iii') *S is the Snell envelope of $1_{\{t<\infty\}\cap\{V_t=G_t\}\cap D_t} + 1_{\{t=\sigma=\infty\}}$ on $\mathbb{T} \cup \{T\}$.*

Proof. Clearly (i) implies (i'). To see the reverse, suppose that $S_t > 0$ on D_t. Then $S'_t := 1_{D_t}$, $t \leq T$ is a nonnegative supermartingale. Thus, (iii') yields that $0 \leq S_t \leq 1_{D_t}$ and (i) follows. Given (i'), the equivalence of (ii) and (ii') is immediate. For the equivalence of (iii) and (iii'), note that $U \wedge 1$ is a supermartingale whenever U is a supermartingale. □

Lemma 6.4. *(a) The processes $(SV)_{\cdot \wedge T_e}$ and $(S^{t_0}V)_{\cdot \wedge T_e}$ occurring in (ii') and (iv) are uniformly integrable.*
(b) Let $T = \infty$ and let (V, S) be a Snell pair. Then

$$\lim_{t \to \infty} S_t = S_\infty = 1_{\{\sigma=\infty\}}, \tag{6.2}$$

$$\lim_{t \to \infty} (SV)_{t \wedge T_e} = 1_{\{T_e \lhd \sigma\}}G_{T_e}. \tag{6.3}$$

Proof. (a) Recall that $\sup_t |G_t| 1_{D_t} \in L^1$ and that the Snell envelope of any process with an L^1-majorant is uniformly integrable. In view of of (ii'), it follows that $(SV)_{\cdot \wedge T_e}$ is uniformly integrable, and then so is $(S^{t_0}V)_{\cdot \wedge T_e}$.

(b) We have from (i) and (iii) that S is a bounded supermartingale with $S_\infty = 1_{D_\infty}$. In particular, $S_t \geq E[S_\infty | \mathcal{F}_t]$. Passing to the limit, martingale convergence yields that $\liminf_t S_t \geq S_\infty$. Conversely, (i) clearly implies that $\limsup_t S_t \leq 1$ and that $\lim S_t = 0$ on $\cup_t D_t^c = D_\infty^c$. Hence, (6.2) is proved.

Part (a), (ii') and the classical limit property of the Snell envelope yield that $\lim_{t \to \infty}(SV)_{t \wedge T_e} = \limsup_{t \to \infty}(SG)_{t \wedge T_e}$. Moreover, using (6.2) and (6.1),

$$\limsup_{t \to \infty}(SG)_{t \wedge T_e} = 1_{\{T_e < \infty\}} 1_{D_{T_e}} G_{T_e} + 1_{\{T_e = \infty\}} \limsup_{t \to \infty} G_t$$

$$= 1_{\{T_e < \sigma\}} G_{T_e} + 1_{\{T_e = \sigma = \infty\}} G_\infty = 1_{\{T_e \lhd \sigma\}} G_{T_e}$$

and thus (6.3) follows. $\qquad \square$

We can now state the main result which relates Snell pairs to equilibria, thus extending the classical Snell envelope theory to conditional optimal stopping.

Theorem 6.5. *(a) Let (V, S) be a Snell pair. Then, $\theta = 1_{\{G \geq V\}}$ defines an equilibrium stopping policy with early stopping preference. Moreover,*

$$V_t = 1_{\{t < T_e\}} \max(G_t, J_t(\theta)) + 1_{\{t \geq T_e\}} G_t, \quad t \in \mathbb{T}, \tag{6.4}$$

$$S_t = 1_{D_t} \left(1_{\{V_t = G_t\}} + 1_{\{V_t > G_t\}} P(\mathcal{L}_t \theta \lhd \sigma | \mathcal{F}_t) \right), \quad t \in \mathbb{T} \tag{6.5}$$

and $S_\infty = \lim_{t \to \infty} S_t = 1_{\{\sigma = \infty\}}$ if $T = \infty$.

(b) If θ is an equilibrium stopping policy with early stopping preference, then there exists a unique Snell pair (V, S) such that $\theta = 1_{\{G \geq V\}}$. This Snell pair is given by (6.4)–(6.5).

As mentioned above, Snell pairs reduce to the usual Snell envelope in the classical case.

Corollary 6.6. *Suppose that $D_t = \Omega$ for all $t \in \mathbb{T}$.*
(i) Any equilibrium θ corresponds to optimal stopping in the classical sense: $E[G_{\tau_t} | \mathcal{F}_t] = \operatorname{ess\,sup}_{\tau \geq t} E[G_\tau | \mathcal{F}_t]$ for $\tau_t = \inf\{s \geq t : \theta_s = 1\}$.
(ii) Any Snell pair consists of $S \equiv 1$ and the classical Snell envelope $V_t = \operatorname{ess\,sup}_{\tau \geq t} E[G_\tau | \mathcal{F}_t]$.

Proof. Note that $\sigma = \infty$. Let θ be an equilibrium and (V, S) the associated Snell pair. Then

$$V_t = 1_{\{\theta_t=1\}}G_t + 1_{\{\theta_t=0\}}J_t(\theta) = 1_{\{\theta_t=1\}}G_t + 1_{\{\theta_t=0\}}E[G_{\mathcal{L}_t\theta}|\mathcal{F}_t]$$
$$= E[1_{\{\theta_t=1\}}G_t + 1_{\{\theta_t=0\}}G_{\mathcal{L}_t\theta}|\mathcal{F}_t] = E[G_{\tau_t}|\mathcal{F}_t].$$

We have $S_T = 1$ by (6.4)–(6.5). Since S is a supermartingale dominated by 1, we must have $S_t = 1$ for all $t \in \mathbb{T}$. It follows that $SV = V$ is the Snell envelope of $SG = G$; that is, $V_t = \operatorname{ess\,sup}_{\tau \geq t} E[G_\tau|\mathcal{F}_t]$. $\qquad\square$

In the finite horizon-case, Snell pairs correspond to the processes constructed in Section 3.

Corollary 6.7. *Let $T < \infty$. Then there exists a unique Snell pair (V, S) and it is determined by the backward recursion of Theorem 3.1.*

Proof. Let (V', S') and θ be as in Theorem 3.1. By Theorem 6.5, there exists a unique Snell pair (V, S) with $\theta = 1_{\{G \geq V\}}$, and it is completely determined by (6.4)–(6.5). In view of Lemma 3.2 and the definition in Theorem 3.1, (V', S') also satisfies (6.4)–(6.5), thus $(V', S') = (V, S)$. $\qquad\square$

We note that in the infinite-horizon case, the examples in Section 5 show that Snell pairs are not unique in general.

Proof of Theorem 6.5. We focus on the case $T = \infty$; the finite-horizon case is similar but simpler.

(a) Let (V, S) be a Snell pair and $\theta = 1_{\{G \geq V\}}$; we show that $\theta \in \Theta$ and $\Phi(\theta) = \theta$. If $t \geq T_e$, then (i) implies $V_t = G_t$ and hence $\theta_t = 1$ and $S_t = 1_{D_t}$; see (iii). Let $t < T_e$. Note that we are in D_t and $t < \mathcal{L}_t\theta \leq T_e \leq \sigma$. If $\mathcal{L}_t\theta = \infty$, we have $S_{\mathcal{L}_t\theta} = S_\infty = 1_{\{\sigma=\infty\}} = 1$, whereas if $\mathcal{L}_t\theta < \infty$, we have $S_{\mathcal{L}_t\theta} = 1_{D_{\mathcal{L}_t\theta}}$. In summary,

$$S_{\mathcal{L}_t\theta} = 1_{\{\mathcal{L}_t\theta<\infty\}}1_{D_{\mathcal{L}_t\theta}} + 1_{\{\mathcal{L}_t\theta=\infty\}} = 1_{\{\mathcal{L}_t\theta\lhd\sigma\}}. \tag{6.6}$$

As a consequence, recalling Lemma 6.4 for the case $\mathcal{L}_t\theta = \infty$,

$$S_{\mathcal{L}_t\theta}V_{\mathcal{L}_t\theta} = 1_{\{\mathcal{L}_t\theta\lhd\sigma\}}G_{\mathcal{L}_t\theta}. \tag{6.7}$$

Next, we consider separately two cases.

Case $\theta_t = 0$: Using (v), the Optional Sampling Theorem (with the boundedness of S and the uniform integrability from Lemma 6.4) as well as (6.6) and (6.7), we see that

$$S_t = E[S_{\mathcal{L}_t\theta}|\mathcal{F}_t] = P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t) \tag{6.8}$$

29

and
$$S_t V_t = E[S_{\mathcal{L}_t\theta} V_{\mathcal{L}_t\theta} | \mathcal{F}_t] = E[1_{\{\mathcal{L}_t\theta \lhd \sigma\}} G_{\mathcal{L}_t\theta} | \mathcal{F}_t]. \qquad (6.9)$$

In view of (i), Equation (6.8) yields in particular that $P(\mathcal{L}_t\theta \lhd \sigma | \mathcal{F}_t) = S_t > 0$ since we are in D_t, as required for the admissibility of θ. Moreover, as $\theta_t = 0$, (6.8) and (6.9) together imply that

$$G_t < V_t = \frac{S_t V_t}{S_t} = \frac{E[1_{\{\mathcal{L}_t\theta \lhd \sigma\}} G_{\mathcal{L}_t\theta} | \mathcal{F}_t]}{P(\mathcal{L}_t\theta \lhd \sigma | \mathcal{F}_t)} = J_t(\theta).$$

Case $\theta_t = 1$: In this case, S is a martingale from time $t+1$ to time $\mathcal{L}_t\theta$ and hence, similarly to the previous case,

$$S_{t+1} = E[S_{\mathcal{L}_t\theta} | \mathcal{F}_{t+1}] = E[1_{\{\mathcal{L}_t\theta \lhd \sigma\}} | \mathcal{F}_{t+1}].$$

Taking conditional expectations on both sides, we deduce that

$$S_t^t = E[S_{t+1} | \mathcal{F}_t] = P(\mathcal{L}_t\theta \lhd \sigma | \mathcal{F}_t).$$

In view of $t < T_e$ and (i), we have $P(S_{t+1} > 0 | \mathcal{F}_t) = P(D_{t+1} | \mathcal{F}_t) > 0$ which then implies $P(\mathcal{L}_t\theta \lhd \sigma | \mathcal{F}_t) = E[S_{t+1} | \mathcal{F}_t] > 0$ and finishes the proof of admissibility. Moreover, by the supermartingale property of $S^t V$, the Optional Sampling Theorem with the uniform integrability from Lemma 6.4, and (6.7), we have

$$\begin{aligned}
S_t^t V_t &\geq E[S_{\mathcal{L}_t\theta}^t V_{\mathcal{L}_t\theta} | \mathcal{F}_t] = E[S_{\mathcal{L}_t\theta} V_{\mathcal{L}_t\theta} | \mathcal{F}_t] \\
&= E[1_{\{\mathcal{L}_t\theta \lhd \sigma\}} G_{\mathcal{L}_t\theta} | \mathcal{F}_t] = P(\mathcal{L}_t\theta \lhd \sigma | \mathcal{F}_t) J_t(\theta) \\
&= E[S_{t+1} | \mathcal{F}_t] J_t(\theta) = S_t^t J_t(\theta).
\end{aligned}$$

As $S_t^t > 0$ and $\theta_t = 0$, we conclude that $G_t = V_t \geq J_t(\theta)$.

Putting the two cases together and noting $\{\theta_t = 0\} \subseteq \{t < T_e\} \subseteq D_t$, we conclude that (6.4) and (6.5) hold. We also recall that the condition on S_∞ was already established in (6.2). Finally, (6.4) shows that

$$\Phi(\theta)_t = 1_{\{t < T_e\} \cap \{G_t \geq J_t(\theta)\}} + 1_{\{t \geq T_e\}} = \theta_t$$

and the proof of (a) is complete.

(b) Let θ be an equilibrium stopping policy with early stopping preference; we show that the pair (V, S) defined by (6.4) and (6.5) is a Snell pair.

First, we check that (6.5) implies $S_\infty := \lim_t S_t = 1_{D_\infty}$. Indeed, we have $P(\mathcal{L}_t\theta \lhd \sigma | \mathcal{F}_t) \geq P(\sigma = \infty | \mathcal{F}_t) = P(D_\infty | \mathcal{F}_t) \to 1_{D_\infty}$. Thus, (6.5) implies $\lim_t S_t = 1_{D_\infty}$ as desired.

We readily see that (i') holds, so it suffices to show (ii'), (iii') and (iv). Note that

$$\{\theta_t = 0\} = \{\Phi(\theta)_t = 0\} = \{t < T_e\} \cap \{J_t(\theta) > G_t\} = \{V_t > G_t\}. \quad (6.10)$$

Let $t < T_e$ (which implies that we are in D_t), then

$$
\begin{aligned}
S_t^t &= E[S_{t+1}|\mathcal{F}_t] \\
&= E[1_{D_{t+1}} \left(1_{\{V_{t+1}=G_{t+1}\}} + 1_{\{V_{t+1}>G_{t+1}\}} P(\mathcal{L}_{t+1}\theta \lhd \sigma|\mathcal{F}_{t+1})\right)|\mathcal{F}_t] \\
&= E[1_{\{V_{t+1}=G_{t+1}\}\cap D_{t+1}} + 1_{\{V_{t+1}>G_{t+1}\}\cap\{\mathcal{L}_{t+1}\theta\lhd\sigma\}}|\mathcal{F}_t] \\
&= E[1_{\{\theta_{t+1}=1\}\cap\{t+1\lhd\sigma\}} + 1_{\{\theta_{t+1}=0\}\cap\{\mathcal{L}_{t+1}\theta\lhd\sigma\}}|\mathcal{F}_t] \\
&= E[1_{\{\theta_{t+1}=1\}\cap\{\mathcal{L}_t\theta\lhd\sigma\}} + 1_{\{\theta_{t+1}=0\}\cap\{\mathcal{L}_t\theta\lhd\sigma\}}|\mathcal{F}_t] \\
&= P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t) \le S_t
\end{aligned}
$$

and

$$
\begin{aligned}
&E[S_{t+1}V_{t+1}|\mathcal{F}_t] \\
&= E[1_{D_{t+1}}(1_{\{V_{t+1}=G_{t+1}\}}G_{t+1} + 1_{\{V_{t+1}>G_{t+1}\}} P(\mathcal{L}_{t+1}\theta \lhd \sigma|\mathcal{F}_{t+1})V_{t+1}|\mathcal{F}_t] \\
&= E[1_{\{\theta_{t+1}=1\}\cap D_{t+1}}G_{t+1} + 1_{\{\theta_{t+1}=0\}} P(\mathcal{L}_{t+1}\theta \lhd \sigma|\mathcal{F}_{t+1})J_{t+1}(\theta)|\mathcal{F}_t] \\
&= E[1_{\{\theta_{t+1}=1\}\cap\{t+1\lhd\sigma\}}G_{t+1} + 1_{\{\theta_{t+1}=0\}} E[G_{\mathcal{L}_{t+1}\theta}1_{\{\mathcal{L}_{t+1}\theta\lhd\sigma\}}|\mathcal{F}_{t+1}]|\mathcal{F}_t] \\
&= E[G_{\mathcal{L}_t\theta}1_{\{\mathcal{L}_t\theta\lhd\sigma\}}|\mathcal{F}_t] = P(\mathcal{L}_t\theta \lhd \sigma|\mathcal{F}_t)J_t(\theta) \le S_t^t V_t \le S_t V_t.
\end{aligned}
$$

This shows that $(S)_{\cdot\wedge T_e}$, $(SV)_{\cdot\wedge T_e}$ and $(S^t V)_{\cdot\wedge T_e}$ are supermartingales on $\mathbb{T}$. In particular, (iv) holds. In fact, S is a supermartingale up to T: for any finite $t \ge T_e$, we have $V_t = G_t$ and $V_{t+1} = G_{t+1}$ and consequently

$$E[S_{t+1}|\mathcal{F}_t] = E[1_{D_{t+1}}|\mathcal{F}_t] \le 1_{D_t} = S_t.$$

As S is bounded and $S_\infty = \lim S_t$, the supermartingale property up to T follows.

Next, let Y be the Snell envelope of $1_{\{t<\infty\}\cap\{V_t=G_t\}\cap D_t} + 1_{\{t=\sigma=\infty\}}$. On the one hand, $S \ge Y$ since Y is the smallest supermartingale dominating $1_{\{t<\infty\}\cap\{V_t=G_t\}\cap D_t} + 1_{\{t=\sigma=\infty\}}$. On the other hand, let $t \in \mathbb{T}$ and define $\hat{\tau} := t1_{\{V_t=G_t\}} + \mathcal{L}_t\theta1_{\{V_t>G_t\}}$ as the stopping time induced by θ at t, then

the stopping representation of the Snell envelope yields

$$Y_t = \operatorname*{ess\,sup}_{\tau \geq t} E[1_{\{\tau < \infty\} \cap \{V_\tau = G_\tau\} \cap D_\tau} + 1_{\{\tau = \sigma = \infty\}} | \mathcal{F}_t]$$

$$\geq E[1_{\{\hat{\tau} < \infty\} \cap \{V_{\hat{\tau}} = G_{\hat{\tau}}\} \cap D_{\hat{\tau}}} + 1_{\{\hat{\tau} = \sigma = \infty\}} | \mathcal{F}_t]$$

$$= 1_{\{V_t = G_t\} \cap D_t} + 1_{\{V_t > G_t\}} E[1_{\{\mathcal{L}_t \theta < \infty\} \cap D_{\mathcal{L}_t \theta}} + 1_{\{\mathcal{L}_t \theta = \sigma = \infty\}} | \mathcal{F}_t]$$

$$= 1_{\{V_t = G_t\} \cap D_t} + 1_{\{V_t > G_t\}} E[1_{\{\mathcal{L}_t \theta < \sigma\}} + 1_{\{\mathcal{L}_t \theta = \sigma = \infty\}} | \mathcal{F}_t]$$

$$= 1_{\{V_t = G_t\} \cap D_t} + 1_{\{V_t > G_t\}} E[1_{\{\mathcal{L}_t \theta \lhd \sigma\}} | \mathcal{F}_t] = S_t.$$

Thus, we have shown $S = Y$ and (iii') is proved.

Similarly, let Z be the Snell envelope of $(SG)_{\cdot \wedge T_e}$. We have $(SV)_{\cdot \wedge T_e} \geq Z$ since Z is the smallest supermartingale dominating $(SG)_{\cdot \wedge T_e}$. Let $t \in \mathbb{T}$. On the set $\{V_t = G_t\}$, we trivially have $Z_t \geq (SG)_{t \wedge T_e} = (SV)_{t \wedge T_e}$ by the definition of V. Whereas on the set $\{V_t > G_t\} \subseteq \{t < T_e\}$,

$$Z_t = \operatorname*{ess\,sup}_{\tau \geq t} E[(SG)_{\tau \wedge T_e} | \mathcal{F}_t] \geq \limsup_{N \to \infty} E[(SG)_{\mathcal{L}_t \theta \wedge N} | \mathcal{F}_t]$$

$$= E[(SG)_{\mathcal{L}_t \theta} | \mathcal{F}_t] = E[G_{\mathcal{L}_t \theta} 1_{\{\mathcal{L}_t \theta \lhd \sigma\}} | \mathcal{F}_t]$$

$$= S_t V_t = (SV)_{t \wedge T_e},$$

where we have used the Dominated Convergence Theorem, (6.5), (6.10) and the definitions of V_t, S_t and $J_t(\theta)$. We conclude that $(SV)_{\cdot \wedge T_e} = Z$; that is, (ii') holds.

It remains to observe the uniqueness. Indeed, if (V', S') is another Snell pair such that $\theta = 1_{\{G \geq V'\}}$, then (V', S') satisfies (6.4) and (6.5) by (a). But (6.4) and (6.5) uniquely define the two processes, so we must have $(V', S') = (V, S)$. $\qquad\square$

References

[1] N. Barberis. A model of casino gambling. *Manag. Sci.*, 58(1):35–51, 2012.

[2] S. Basak and G. Chabakauri. Dynamic mean-variance asset allocation. *Rev. Financ. Stud.*, 23(8):2970–3016, 2010.

[3] T. Björk, M. Khapko, and A. Murgoci. On time-inconsistent stochastic control in continuous time. *Finance Stoch.*, 21(2):331–360, 2017.

[4] T. Björk and A. Murgoci. A theory of Markovian time-inconsistent stochastic control in discrete time. *Finance Stoch.*, 18(3):545–592, 2014.

[5] T. Björk, A. Murgoci, and X. Y. Zhou. Mean-variance portfolio optimization with state-dependent risk aversion. *Math. Finance*, 24(1):1–24, 2014.

[6] S. Christensen and K. Lindensjö. On finding equilibrium stopping times for time-inconsistent Markovian problems. *SIAM J. Control Optim.*, 56(6):4228–4255, 2018.

[7] S. Christensen and K. Lindensjö. On time-inconsistent stopping problems and mixed strategy stopping times. *To appear in Stochastic Process. Appl.*, 2018.

[8] C. Czichowsky. Time-consistent mean-variance portfolio selection in discrete and continuous time. *Finance Stoch.*, 17(2):227–271, 2013.

[9] S. Ebert and P. Strack. Never, ever getting started: On prospect theory without commitment. *Preprint SSRN:2765550*, 2018.

[10] I. Ekeland and A. Lazrak. Being serious about non-commitment: subgame perfect equilibrium in continuous time. *Preprint arXiv:0604264v1*, 2006.

[11] I. Ekeland and A. Lazrak. The golden rule when preferences are time inconsistent. *Math. Financ. Econ.*, 4(1):29–55, 2010.

[12] I. Ekeland, O. Mbodji, and T. A. Pirvu. Time-consistent portfolio management. *SIAM J. Financial Math.*, 3(1):1–32, 2012.

[13] I. Ekeland and T. A. Pirvu. Investment and consumption without commitment. *Math. Financ. Econ.*, 2(1):57–86, 2008.

[14] S. Frederick, G. Loewenstein, and T. O'Donoghue. Time discounting and time preference: A critical review. *J. Econ. Lit.*, 40(2):351–401, 2002.

[15] X. D. He and X. Y. Zhou. Portfolio choice under cumulative prospect theory: An analytical treatment. *Manag. Sci.*, 57(2):315–331, 2011.

[16] Y. Hu, H. Jin, and X. Y. Zhou. Time-inconsistent stochastic linear-quadratic control. *SIAM J. Control Optim.*, 50(3):1548–1572, 2012.

[17] Y. Hu, H. Jin, and X. Y. Zhou. Time-inconsistent stochastic linear-quadratic control: characterization and uniqueness of equilibrium. *SIAM J. Control Optim.*, 55(2):1261–1279, 2017.

[18] Y.-J. Huang and A. Nguyen-Huu. Time-consistent stopping under decreasing impatience. *Finance Stoch.*, 22(1):69–95, 2018.

[19] Y.-J. Huang, A. Nguyen-Huu, and X. Y. Zhou. General stopping behaviors of naive and non-committed sophisticated agents, with application to probability distortion. *To appear in Math. Finance*, 2017.

[20] Y.-J. Huang and Z. Zhou. Optimal equilibria for time-inconsistent stopping problems in continuous time. *To appear in Math. Finance*, 2017.

[21] Y.-J. Huang and Z. Zhou. Strong and weak equilibria for time-inconsistent stochastic control in continuous time. *Preprint arXiv:1809.09243v1*, 2018.

[22] Y.-J. Huang and Z. Zhou. The optimal equilibrium for time-inconsistent stopping problems – the discrete-time case. *SIAM J. Control Optim.*, 57(1):590–609, 2019.

[23] H. Jin and X. Y. Zhou. Behavioral portfolio selection in continuous time. *Math. Finance*, 18(3):385–426, 2008.

[24] C. Karnam, J. Ma, and J. Zhang. Dynamic approaches for some time-inconsistent optimization problems. *Ann. Appl. Probab.*, 27(6):3435–3477, 2017.

[25] P.-L. Lions. HJB equations and extensions of classical stochastic control theory. *Lectures at Collège de France*, 2017. Video available at `https://www.college-de-france.fr/site/en-pierre-louis-lions/course-2016-10-21-09h00.htm`. Transcript by Charafeddine Mouzouni available at `https://hal.archives-ouvertes.fr/cel-01568969`.

[26] B. Peleg and M. E. Yaari. On the existence of a consistent course of action when tastes are changing. *Rev. Econ. Stud.*, 40(3):391–401, 1973.

[27] R. A. Pollak. Consistent planning. *Rev. Econ. Stud.*, 35(2):201–208, 1968.

[28] P. A. Samuelson. An exact consumption-loan model of interest with or without the social contrivance of money. *Ann. Math. Statist.*, 66(6):467–482, 1958.

[29] R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud*, 23(3):165–180, 1955.

[30] K. S. Tan, W. Wei, and X. Y. Zhou. Failure of smooth pasting principle and nonexistence of equilibrium stopping rules under time-inconsistency. *Preprint arXiv:1807.01785v1*, 2018.

[31] J. Yong. Time-inconsistent optimal control problems and the equilibrium HJB equation. *Math. Control Relat. Fields*, 2(3):271–329, 2012.