Submitted to *Management Science*
manuscript MS-FIN-20-03706.R1

A Casino Gambling Model under Cumulative Prospect Theory: Analysis and Algorithm

Sang Hu
School of Data Science, The Chinese University of Hong Kong, Shenzhen, China 518172, husang@cuhk.edu.cn

Jan Obłój
Mathematical Institute, Oxford-Man Institute of Quantitative Finance and St John's College, University of Oxford, Oxford, UK, Jan.Obloj@maths.ox.ac.uk

Xun Yu Zhou
Department of Industrial Engineering and Operations Research, Columbia University, New York, New York 10027, xz2574@columbia.edu

We develop an approach to solve Barberis (2012)'s casino gambling model in which a gambler whose preferences are specified by the cumulative prospect theory (CPT) must decide when to stop gambling by a prescribed deadline. We assume that the gambler can assist their decision using an independent randomization. The problem is inherently time-inconsistent due to the probability weighting in CPT, and we study both precommitted and naïve stopping strategies. We turn the original problem into a computationally tractable mathematical program, based on which we devise an algorithm to compute optimal precommitted rules which are randomized and Markovian. The analytical treatment enables us to confirm the economic insights of Barberis (2012) for much longer time horizons, and to make additional predictions regarding a gambler's behavior, including that with randomization they may enter the casino even when allowed to play only once and that it is prevalent that a naïf never stops loss.

Key words: casino gambling; cumulative prospect theory; optimal stopping; probability weighting; time inconsistency; randomization; finite time horizon; Skorokhod embedding; potential function.

1. Introduction

Tversky and Kahneman (1992)'s cumulative prospect theory (CPT) is the leading descriptive theory of how people make decisions under risk. Its predictions are thus of paramount interest. Barberis (2012) introduces a simple multi-period setting for studying the dynamic implications of CPT. Specifically, the author proposes a casino gambling model to study the optimal timing to quit gambling and leave the casino, and derives two key economic insights: (1) a CPT gambler may be willing to enter the casino even though its bets offer neither positive expected values nor skewness because, by implementing an appropriate stopping strategy, he would be able to build a positively skewed final winning amount that would be favored by the underlying probability weighting in CPT; (2) there is an inherent time-inconsistency due to the dynamically changing strength of probability weighting on a same event: the gambler may deviate completely from his initial stopping strategy as he gambles along, and his eventual stopping behavior depends on whether he is aware of this time-inconsistency and whether he is able to commit to his original plan.[1]

Barberis (2012) considers a finite horizon model, assuming the gambler can take a sequence of up to T gambles. The model is solved using exhaustive search across Markovian stopping strategies. Barberis (2012) focuses on the case of $T=5$ with which the paper discusses the economic implications of CPT in detail. However, it is not clear whether these implications continue to hold for higher values of T. The exhaustive search method can not[2] yield the answer as the number of admissible strategies is *exponential* in T^2.[3] In general, in order to better understand the implications of a model, it is important to have a *systematic approach* to solve it, not necessarily in an analytically closed form, but in a computationally efficient way.[4] Not only can we then obtain optimal solutions for arbitrary values of parameters, but we may gain (likely more profound) economic insights from the model by post-optimality analyses.

The current paper achieves this aim and develops an analytical treatment of the Barberis (2012) casino gambling model. Our main contribution is to derive a new algorithm for solving the problem and determining the CPT individual's behavior. Our algorithm works in a relaxed setting where the gambler is allowed to use an independent randomization when deciding whether to stop or to continue. The algorithm can be implemented for significantly higher values of T than those attainable with Barberis' exhaustive search method. We show that the behavior documented in Barberis (2012) for $T=5$ continues to hold for higher values of T. Further, we undertake comparative statics analysis for the parameters in the classical CPT preferences specification, and explore the importance of randomization for entering the casino and for playing on.

[1] Barberis (2012) discusses three types of gamblers, following the original classification of Strotz (1955): a *naïve* gambler who is unaware of the time-inconsistency and changes his strategy all the time; a *precommitted* gambler who is aware of time-inconsistency and can commit to his initial plan; and a *sophisticated* gambler who is aware of time-inconsistency yet unable to commit, and at each time takes the future selves' disobedience into account when devising an optimal strategy.

[2] We ran exhaustive search on a desktop with Intel Core i5-4590/CPU 3.30GHz/RAM 8.00GB for different T's while keeping the other parameters same as Barberis (2012)'s. The running times for $T=5,6,7$ were 39 seconds, 771 seconds and 27 hours, respectively. We were unable to obtain the solution for $T=8$ due to lack of memory, with the running time estimated to be 300 days.

[3] The number of nodes is $\frac{T(T+1)}{2}$ in a binomial tree of horizon T, and at each node there is a binary choice of {stop, continue}. Hence the total number of strategies is $2^{\frac{T(T+1)}{2}}$.

[4] Consider, e.g., the simplex method for linear programs or the dynamic programming formulation for optimal control problems.

The main technical hurdle to solve the casino model is probability weighting. Barberis (2012) acknowledges that the nonlinear probability weighting involved in CPT makes it "very difficult" to solve the problem analytically, and "the problem has no known analytical solution for general T" (p. 42). The two main approaches in the classical optimal stopping theory – dynamic programming (variational inequalities) and martingale method – both fail under probability weighting: the former does because of the time-inconsistency, and the latter does because of the absence of a "tower property" with respect to the *weighted* probability.

He et al. (2017, 2019a,b) are probably the first series of papers that aim at an *analytical treatment* of the casino model, albeit in the *infinite* time horizon.[5] The main idea of these papers consists of two deeply intertwined steps: (1) search the optimal *probability distribution* of the final winning/losing amount upon leaving the casino instead of the optimal *time* to leave; (2) once the optimal distribution is found, recover the optimal time that generates it.[6] Both steps call for a complete characterization of the set of all the admissible distributions, and the second step is the discrete-time version of the eminent *Skorokhod embedding theorem* which in the casino setting is solved in He et al. (2019b). The main thrust to make this idea work is to permit *randomization*, namely the gambler can flip an independent, possibly biased, coin to assist his decision each step of the way. The probabilities of the head of the coin are *endogenous* and dynamically changing; thus they are *part* of the final solution. Mathematically, randomization *convexifies* the aforementioned set of admissible distributions, which in turn makes the Skorokhod embedding work. The randomization of decisions is a key feature when studying agents with CPT preferences, as discussed independently by Henderson et al. (2017). He et al. (2017, 2019a) also allow path-dependent strategies, that is, the stopping decision is made based on the whole betting history instead of just the current winning/losing amount. They further show that allowing path-dependent strategies or randomized ones *strictly* improve the optimal CPT values. Based on these analyses, He et al. (2019a) turn the casino model into an infinite dimensional mathematical program that can be solved fairly efficiently. Most of the gambler's behaviors – those of a precommitter and of a naïf – implied from the solutions reconcile qualitatively with Barberis (2012)'s results; but there are also new findings. For example, it is revealed that, for most empirically relevant CPT parameter estimates, a precommitted gambler lets gain run while stops loss, but a naïve one almost surely does not stop at *any* loss level.

As noted, He et al. (2017, 2019a) deal with the infinite horizon gambling model. There are important reasons to study the *finite* horizon model under CPT preferences, despite the existing

[5] Here by "analytical treatment" we mean an optimization analysis not based on heuristics or on brute force such as an exhaustive search.

[6] This idea was first put forth by Xu and Zhou (2012) for a continuous-time optimal stopping model featuring probability weighting. There is considerable difficulty to adapt this idea to the discrete-time setting.

results for the infinite horizon counterpart. Conceptually, the finite horizon problem approximates the reality much better, as a gambler clearly will not be able to play arbitrarily and indefinitely long. Also, the original work of Barberis (2012) considers $T = 5$ and hence we need to solve the finite horizon model in order to be able to make a direct comparison. It is worth noting that solutions to the finite horizon case can *not* be recovered from those of the infinite horizon case by a simple truncation: if τ is optimal for the latter, then, typically, $\tau \wedge T$ will not be optimal for the former.

Methodologically, the finite horizon case is significantly more complex. It is well acknowledged that optimal stopping in a finite horizon is *fundamentally* more difficult than its infinite horizon counterpart, mainly because value function of the former has both time and spatial variables while the latter has only spatial variables. In the infinite time horizon setting in which the accumulated winning/losing amount is modelled by a symmetric random walk S, He et al. (2019b) show that for any centered probability measure μ on the set of integers $\mathbb{Z}$, there exists a *randomized* stopping time τ such that S_τ's distribution is μ.[7] As discussed previously, this is the key theoretical underpinning for the new approach. Unfortunately, this result is no longer true if the stopping time is constrained by a pre-specified deadline. Indeed, *additional* conditions are required for measures that can be embedded by uniformly bounded stopping times. One of the technical contributions of this paper is to identify explicitly these conditions, which in turn enables us to reformulate the original casino model into a mathematical program whose number of constraints is of the order of T and, hence, can be efficiently solved.

Once we have an algorithm to solve the gambling model for *any* parameter values, we are then able to compare our results with those of Barberis (2012)'s. In particular, we consider the case that is solved and discussed in Barberis (2012) with $T = 5$ and find that the respective stopping strategies for a precommitter are identical except in two time–state instances in which our decisions are to stop with very small probabilities (0.00864 and 0.0368 respectively) whereas Barberis' are just to continue. Qualitatively, both strategies are of the so-called *loss-exit* type, namely, they continue in gains but stop after having accumulated sufficient amounts of losses. With randomization, our optimal CPT value improves, if slightly, over Barberis'. Likewise, the respective naïve strategies are the same save for one time–state instance in which ours is to stop with a probability of 0.179 while Barberis' is to continue. Our solution, however, enables us to look beyond the relatively short horizon of $T = 5$. Indeed, we carry out numerical experiments for different values of T up to $T = 20$, and confirm that the earlier insights of Barberis (2012) hold more generally, including that the interplay between the utility function, probability weighting and loss aversion dictates various gambling behaviors.

[7] In the terminology of Skorokhod embedding theorem, we say τ *embeds* μ in S.

Note that our analytical treatment relies on the introduction of randomization in our model, as randomization convexifies the optimization problem. Barberis (2012) does not allow randomization, for which our approach would fail. However, our solution would provide a well-founded relaxation *heuristic* for solving a casino model without randomization: we first relax the problem by introducing randomization, and then, for each time-state pair, round up or round down the probability of stopping to 1 (which means stop) or to 0 (which means continue).

Our approach makes it possible to analyze and understand the impacts of some key attributes of the model in general terms. For example, Barberis (2012) argues that a gambler may be willing to enter a casino because, by implementing a loss-exit strategy, he may be able to generate a positively skewed probability distribution of the final accumulated gain/loss which has a positive CPT preference value. However, he will need to spend time building such a skewed distribution, which requires a sufficiently large T. We show, however, that the same gambler who would have demanded a long horizon for agreeing to enter the casino, will enter even if he is allowed to play only *once* (i.e., $T = 1$), provided that he can flip a coin to decide if he plays or not. The reason for this is that, with randomization, the gambler can *design* a coin *right away* with a very small probability for both a gain and a loss. If the overweighing on gains sufficiently outweighs that on losses then the gambler will enter.

We also study the behaviors of a naïve gambler with various parameter specifications and a longer time horizon ($T = 10$). We find that, unless he does not enter the casino, his behavior is consistently of gain-exit type, i.e., he stops gain but lets loss run, reminiscent of the *disposition effect* in security trading (Odean 1998). In particular, he never stops loss and gambles "until the bitter end". This gamble-until-the-bitter-end behavior is derived by Ebert and Strack (2015) in a model in which a naïve gambler can construct *arbitrarily* small random payoffs. Because he prefers "skewness in the small", he never stops gambling. Henderson et al. (2017), employing the approach developed in Xu and Zhou (2012), investigate a stylized continuous-time model and show that a naïve gambler may stop with a positive probability if she is allowed to randomize, which complements and counters the findings in Ebert and Strack (2015). Both Ebert and Strack (2015) and Henderson et al. (2017) rely on the crucial feature of their models that allows the gambler to construct arbitrarily small random payoffs. This feature is absent in our discrete-time model, in which the gambler cannot construct strategies with arbitrarily small random payoffs due to the minimal stake size fixed to be \$1. Hence, their results are not applicable to our setting. Our finding therefore suggests that the gamble-until-the-bitter-end phenomenon is probably more prevalent of a naïf's behavior.

The paper proceeds as follows. In Section 2, we formulate a casino gambling model under CPT as an optimal stopping problem and discuss why we allow randomization in our model. In Section 3, we develop the key step in our approach to solve the gambling model: characterizing the set of probability distributions of all possible accumulated winning/losing amounts upon leaving the casino. In Section 4, we present a mathematical program that is equivalent to the casino model, and then report the results of a numerical example which is studied in Barberis (2012). We discuss various implications and predictions of our model in Section 5. Finally, we conclude the paper by Section 6. Proofs are placed in Appendices.

2. The Model

In this section we first highlight the key ingredients of Tversky and Kahneman (1992)'s CPT, then formulate the casino gambling model in a finite time horizon as an optimal stopping problem, and finally discuss the reasons why we make randomization available in our model.

2.1. Cumulative prospect theory

There are four key features of CPT that differentiate it from the classical expected utility theory: 1) individuals derive utility from gains and losses with respect to a *reference point* in wealth, rather than from the absolute amount of wealth itself; 2) individuals are *loss-averse*, namely they are more sensitive to losses than to gains; 3) individuals are risk-averse in gains and risk-loving in losses; and 4) individuals exaggerate extremely small probabilities of extremely large gains and losses.

Accordingly, in CPT the utility function is

$$u(x) = \begin{cases} u_+(x-k), & x \geq k, \\ -\lambda u_-(k-x), & x < k, \end{cases}$$

where k is the reference point, $u_+(\cdot)$ and $u_-(\cdot)$ are both *concave* functions and $\lambda > 1$. This renders an overall *S-shaped* utility function $u(\cdot)$ that is concave (risk-averse) in the gain region $x \geq k$ and convex (risk-loving) in the loss region $x < k$. Moreover, $\lambda > 1$ yields that, for the same magnitude of a gain and a loss, the agent is more sensitive to the latter, reflecting loss aversion. Tversky and Kahneman (1992) propose the following parametric form of $u(\cdot)$:

$$u(x) = \begin{cases} (x-k)^{\alpha_+}, & x \geq k, \\ -\lambda(k-x)^{\alpha_-}, & x < k, \end{cases} \tag{1}$$

where $0 < \alpha_\pm \leq 1$ and $\lambda > 1$; see the left panel of Figure 1 for an illustration of this type of function.

CPT uses *probability weighting* (or *distortion*) *functions* to capture the exaggeration of the tail distributions. An *inverse S-shaped* weighting function is first concave and then convex in the domain of probabilities. Such a weighting function overweights both tails of a probability distribution, reflecting the exaggeration of extremely small probabilities of extremely large gains and losses. Tversky and Kahneman (1992) suggest a parametric form of a weighting function $w(\cdot)$:

$$w(p) = \frac{p^\delta}{(p^\delta + (1-p)^\delta)^{\frac{1}{\delta}}}, \tag{2}$$

where $0 < \delta \leq 1$; see the right panel of Figure 1 for an illustration. Note that $\delta = 1$ means that no weighting is applied.

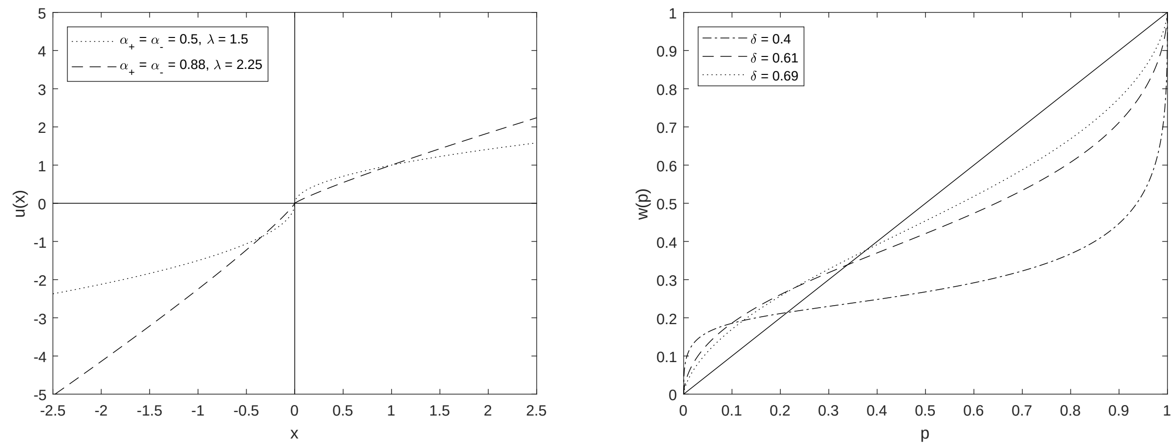

Figure 1 The left panel graphs two S-shaped utility functions (1) with $\alpha_+ = \alpha_- = 0.5, \lambda = 1.5$ and $\alpha_+ = \alpha_- = 0.88, \lambda = 2.25$, respectively. The right panel depicts three inverse S-shaped probability weighting functions (2) with $\delta = 0.4$, $\delta = 0.61$, and $\delta = 0.69$, respectively.

2.2. Formulation of a casino gambling model

We now reformulate Barberis (2012)'s model of casino gambling in a finite time horizon $[0,T]$, where $T \in \mathbb{Z}^+ := \{1,2,3,...\}$ is given. The gambling process proceeds as follows. At time 0, the gambler is offered a fair bet, e.g., one with a roulette wheel: win or lose \$1 with equal probability.[8] If the gambler decides not to play the bet, then he will not even enter the casino. If the gambler enters and takes the bet, then the bet outcome is played out at time 1, leading to either a win or a loss of \$1 at time 1. At that time the gambler is offered the same bet again and he decides whether to play. If he declines the bet, then the game is over and the gambler leaves the casino with \$1 gain or loss. This process continues in the same fashion until time T: the bet is offered and played out repeatedly until either the first time the gambler declines the bet, or at time T when the gambler *must* quit gambling and leave. The *accumulated* gain/loss process can be represented as a binomial tree; see Figure 2. Therein, each node is marked by a pair (t,x), where $t \in \mathbb{N} := \{0,1,2,\dots\}$ stands for the time and $x \in \mathbb{Z} := \{0,\pm1,\pm2,\dots\}$ the amount of cumulative gains or losses. For example, the node $(2,-2)$ signifies a cumulative loss of \$2 at time 2. The process has a terminal time T but the gambler may quit at some earlier time $\tau \le T$.

The gain/loss binomial tree $S = (S_t : t \in \mathbb{N})$ is a *standard symmetric random walk* (SSRW) defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\in\mathbb{N}})$. We assume the probability space is rich enough to support an $\mathcal{F}_0$-measurable random variable Z that is uniformly distributed on $[0,1]$ and *independent* of S.

[8] As in Barberis (2012), we assume in this paper that the gamble is fair. It will not affect the main economic findings and implications of our results. A model of unfair games is more technical, and is left for a future study.

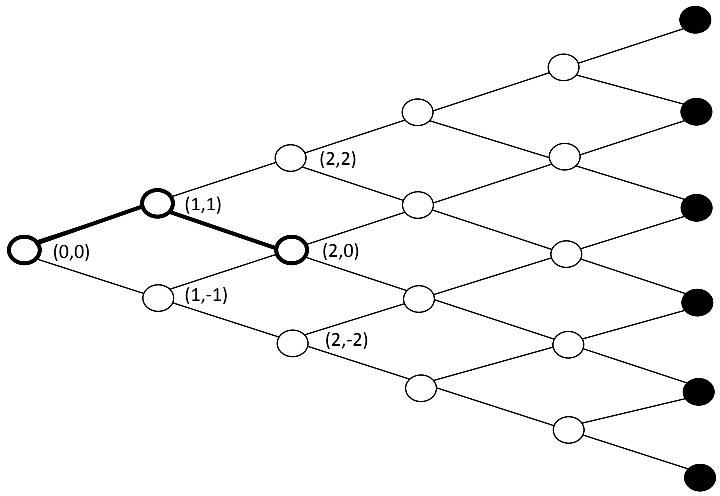

Figure 2 **The gain/loss binomial tree with $T=5$. The gambler must leave the casino by time 5, which is represented by the black nodes.**

Suppose the gambler quits gambling at a random time $\tau \in [0,T]$. Then, with the reference point being his *initial* wealth before he enters the casino, the CPT value of his wealth upon leaving is

$$
\begin{aligned}
V(S_\tau) := & \sum_{n=1}^{T} u_+(n)\left[w_+\left(\mathbb{P}(S_\tau \geq n)\right) - w_+\left(\mathbb{P}(S_\tau \geq n+1)\right)\right] \\
& - \lambda \sum_{n=1}^{T} u_-(n)\left[w_-\left(\mathbb{P}(S_\tau \leq -n)\right) - w_-\left(\mathbb{P}(S_\tau \leq -n-1)\right)\right].
\end{aligned}
\tag{3}
$$

This value reduces to the expected utility of accumulated gains/losses when $w_+(x) = w_-(x) = x$ and $\lambda = 1$. So the formula is a generalization of the expected utility when the decumulative probabilities are weighted and loss aversion is incorporated.

Throughout this paper we assume that both $u_+(\cdot)$ and $u_-(\cdot)$ are concave and both $w_+(\cdot)$ and $w_-(\cdot)$ are inverse S-shaped. The gambler needs to determine the optimal time to quit and leave the casino: such a stopping (exit) strategy τ is made at $t=0$ to maximize $V(S_\tau)$ among all admissible strategies. Note that, due to probability weighting, the problem is inherently *time-inconsistent*; so τ is optimal only at $t=0$ in the sense of a *precommitted* strategy; it may no longer be optimal from the vantage point of any later time $t>0$.

We now define precisely the set of *admissible* stopping strategies

$$
\mathcal{T}_T := \{\tau \in [0,T] : \tau \text{ is an } (\mathcal{F}_t)_{t\in\mathbb{N}}\text{-stopping time}\}.
$$

So a decision whether or not to quit at time $t \in [0,T]$ depends on all the information up to t. In particular, path-dependent strategies are admissible. Moreover, $\mathcal{F}_0$ – and hence all $\mathcal{F}_t$ – contains the information about $Z \sim U(0,1)$, a uniform random variable on $(0,1)$ independent of S. Using Z, we can generate countably many Bernoulli (binary) random variables which are mutually independent and also independent of S.[9] In consequence, an admissible strategy may involve randomization by tossing a (generally biased) coin. In the next subsection we will outline the rationale behind allowing randomized strategies.

The gambler's problem is

$$\max_{\tau \in \mathcal{T}_T} V(S_\tau)\,. \tag{4}$$

2.3. Randomization

In our model (4), the filtration $(\mathcal{F}_t)_{t\in\mathbb{N}}$ includes the information based on a uniform random variable that is *independent* of the underlying random walk. This means that we allow the gambler to *assist* his decision by flipping an independent, most likely biased, coin at each node. We now provide some discussions about the assumption of making this randomization available in our model.

First and foremost, randomization is introduced for a technical reason – it is a relaxation (convexification) of the original Barberis (2012) model that enables our analytical treatment. Moreover, one naturally has a well-founded heuristic strategy for the non-randomized model based on the optimal strategy for the randomized counterpart by rounding up or down the probabilities to 1 or 0. Specifically, suppose r is the probability of stopping at a node produced by an optimal strategy for the randomized model. Then, the following is the "rounded" strategy for the non-randomized model: we stop at the node if $r \geq 0.5$ and continue otherwise. For long time horizons, such rounded strategies are nearly optimal for the non-randomized model; see Subsection 5.3 for details.

Second, randomization is related to the accommodation of path-dependence. It is practically more reasonable, and indeed necessary, to consider path-dependent strategies than mere Markovian ones. At least two path-dependent gambling behaviors or strategies – the gambler's fallacy and the hot hand – have been observed both in a laboratory setting and in actual casinos; see Croson and Sundali (2005) and the rich literature reviewed therein. In both strategies, the gamblers act as if the sequence of i.i.d. events were actually autocorrelated, e.g., seeing a sequence of five red wins in roulette might trigger a specific bet. More broadly, it makes intuitive sense that how the gambler has arrived at his current wealth – whether he has won big first and then lost most of it, or he has gradually accumulated this amount by many small wins – may well affect his decisions. As a matter of fact, almost all our decisions in life are made based on all the information, past and present, rather than just on the current state of affairs. Accordingly, path-dependence is the standard formulation in optimal stopping theory (see, e.g., Shiryaev 1978) and indeed in general stochastic

[9] A Bernoulli random variable ξ_1 with $\mathbb{P}(\xi_1 = 0) = r_1 = 1 - \mathbb{P}(\xi_1 = 1)$ is obtained setting $\xi_1 = \mathbf{1}_{Z>r_1}$, i.e., $\xi_1 = 0$ if $Z \leq r_1$ and $\xi_1 = 1$ if $Z > r_1$. An analogous, and independent, Bernoulli variable ξ_2 can then be defined as $\xi_2 = \mathbf{1}_{r_1 r_2 < Z \leq r_1} + \mathbf{1}_{Z > r_1 + (1-r_1) r_2}$. Likewise, subsequent independent Bernoulli random variables $\xi_3, \xi_4, \ldots$ are obtained via further nested definitions.

control theory (see, e.g., Yong and Zhou 1999). Being Markovian is just a mathematical assumption and convenience that aims to dramatically reduce the dimension of the underlying problem or, in a continuous-time setting, turns an infinite dimensional problem into a finite dimensional one.

He et al. (2017) shows that any non-randomized, path-dependent stopping time can be replaced by a randomized, *Markovian* stopping time in the sense that the latter achieves the same CPT value as the former.[10] As a result, we can consider randomization *in lieu of* considering the past information and path-dependent strategies. One specific example is given in Section 5.1 below. To understand this relationship, suppose that the casino had been going on forever and the gambler has access to the results of *all* the past spins of the roulette wheel: $X_t \in \{0,1\}$, $t=-1,-2,\ldots$. This historical information could be perfectly summarized with a uniform random variable $Z=\sum_{t\le -1} X_t 2^t$ which would be independent of all the future spins. The setting would thus be fully equivalent to our randomized setting. However, in a binomial model as well as in practice, the gambler has access to only a *finite* history of N spins at any given time. He can use it to construct Bernoulli random variables ξ but only with $\mathbb{P}(\xi=1)$ being a multiple of 2^{-N} and can only construct finitely many such independent variables. For this reason randomization may allow for richer strategies than path-dependence. This is precisely the reason why randomized Markovian strategy may not be replicated by a non-randomized, path-dependent strategy, and the optimal CPT value among the former type may be *strictly* greater than that among the latter type, as He et al. (2017) show. The authors of that paper attribute this mathematically to the lack of quasi-convexity of CPT preference (in contrast to the classical expected utility theory preference). This property was also exploited by Henderson et al. (2017) to complement and counter the findings in Ebert and Strack (2015).

While there are instances of randomization in real life, this behavior is not commonly observed.[11] It is interesting to reflect why this is, given that our model says we should observe it. Some possible answers are: (i) while CPT is a descriptive theory, randomization appears as a normative ingredient: its use is advised to strictly improve the optimal CPT value. However, people may not want to leave their decisions to chance; (ii) in many cases especially for longer time periods, randomization does not raise CPT value very much (see Subsection 5.3); (iii) it may be difficult for people to

[10] While this result has been obtained for the infinite horizon model in He et al. (2017), the underlying mathematical argument would be exactly the same for the finite horizon case. The intuition is that, due to the independent increments of a random walk, considering all the past information can be achieved by randomizing at the current state.

[11] These instances include last-minute deals by flight booking apps, "sushi omakase" (you entrust yourself to a sushi chef to choose the ingredients and presentations of your sushi plate), and "fukubukuro" (grab bags filled with unknown and random contents). There is research in experimental psychology and economics that documents experiments involving individuals *deliberately* randomizing when making decisions; see, e.g., Agranov and Ortoleva (2017), Dwenger et al. (2013).

figure out the specific probabilities needed to implement randomization; and (iv) non-randomized strategy may be the best easily available approximation: in our numerical examples, randomization happens rarely and often with very uneven odds.

3. Characterization of Stopped State Distributions

As explained earlier, the main thrust of our approach to solving Problem (4) is to change its decision variable from the stopping time τ to the *distribution* of the stopped state S_τ. A key step is therefore to characterize the admissible set of these distributions. Moreover, once an optimal distribution is obtained there needs to be a way to recover the stopping time that generates this distribution. These two questions are intertwined and will actually be solved together. This section addresses them.

Denote by $\mathcal{P}(\mathbb{R})$ the set of probability measures μ on $\mathbb{R}$ and by $\mathcal{P}_0(\mathbb{R})$ the subset of $\mathcal{P}(\mathbb{R})$ whose elements have finite first moments and are centered: $\int |x|\mu(dx) < \infty$ and $\int x\mu(dx) = 0$. Denote by $\mathcal{P}_0(\mathbb{Z}) = \{\mu \in \mathcal{P}_0(\mathbb{R}) : \mu(\mathbb{Z}) = 1\}$ the subset of $\mathcal{P}_0(\mathbb{R})$ supported on integers.

For $\mu \in \mathcal{P}_0(\mathbb{R})$, define a function

$$U_\mu(x) := \int_{\mathbb{R}} |x-y|\mu(dy), \ \ x \in \mathbb{R},$$

which is called the *potential of* μ.[12] For $\mu \in \mathcal{P}_0(\mathbb{Z})$, U_μ is a linear interpolation of the points $\{U_\mu(k) : k \in \mathbb{Z}\}$. The following are evident:

$$\mu(\{x\}) = \frac{U_\mu(x+1) + U_\mu(x-1)}{2} - U_\mu(x), \quad U_\mu(x) = -2x\mu([x,\infty)) + x + 2\sum_{y \ge x} y\mu(\{y\}). \tag{5}$$

Potential function uniquely determines probability measure, namely, two measures are identical if and only if their potential functions are identical; see Obłój (2004). Finally, for any stopping time τ, with a slight abuse of notation we simply write U_{S_τ} for the potential of the distribution of S_τ, when well defined.

We can use a sequence of piecewise linear functions, called *evolutional functions*, to approach a potential function. Indeed, given $\mu \in \mathcal{P}_0(\mathbb{Z})$, we define recursively the following sequence of functions:

$$\begin{aligned} &U_0^\mu(x) := |x|, \\ &U_t^\mu(x) := \frac{U_{t-1}^\mu(x-1) + U_{t-1}^\mu(x+1)}{2} \wedge U_\mu(x), \quad t = 1, 2, ..., \ x \in \mathbb{Z}. \end{aligned} \tag{6}$$

We then extend each U_t^μ to non-integers $x \in \mathbb{R}$ by linear interpolation. When μ is fixed, we may drop the superscript μ and just write U_t for simplicity.

[12] Note that our definition here is the negative of the usual definition of potential.

The optimal stopping time we will derive belongs to a special class of randomized, Markovian stopping times called the *Root stopping times*. The original version of the Root stopping times was developed in Root (1969) to solve the classical Skorokhod embedding problem for a Brownian motion B on an infinite time horizon.[13] We now develop the analogous ideas in discrete time for the SSRW. Intuitively speaking, our stopping times will be defined using a barrier which splits the space-time into two regions: "before" and "after". The SSRW is allowed to evolve freely in the "before" (continuation), region but is stopped upon the entry to the "after" (stopping) region. The decision whether it is stopped at the barrier between the two regions will be randomized.

Consider an integer-valued vector $\mathbf{b} := (\ldots, b(-1), b(0), b(1), \ldots)$ where, for any $x \in \mathbb{Z}$, $b(x) \geq |x|$ and $b(x) - x$ is even, and another vector $\mathbf{r} := (\ldots, r(-1), r(0), r(1), \ldots)$, where $r(x) \in [0,1]$. Given $\mathbf{b}$ and $\mathbf{r}$, define the probability distributions of a family of independent Bernoulli random variables $\{\xi_{t,x} : t \in \mathbb{N}, x \in \mathbb{Z}\}$ as follows:

$$\begin{cases} \mathbb{P}(\xi_{t,x} = 0) = 1 - \mathbb{P}(\xi_{t,x} = 1) = 0, & t < b(x), \\ \mathbb{P}(\xi_{t,x} = 0) = 1 - \mathbb{P}(\xi_{t,x} = 1) = r(x), & t = b(x), \\ \mathbb{P}(\xi_{t,x} = 0) = 1 - \mathbb{P}(\xi_{t,x} = 1) = 1, & t > b(x). \end{cases}$$

Graphically, $\mathbf{b}$ is a barrier that defines a time-space stopping region

$$\mathcal{R}_b := \big\{(t,x) : t \in \mathbb{N},\ x \in \mathbb{Z},\ t \geq b(x)\big\},$$

and the components of $\mathbf{r}$ are the probabilities to stop exactly on the boundary of this stopping region. The randomized Root stopping time is defined as

$$\tau_R(\mathbf{b}, \mathbf{r}) := \inf\{t \in \mathbb{N} : (t, S_t) \in \mathcal{R}_b \text{ and } \xi_{t,S_t} = 0\}. \tag{7}$$

This stopping time is Markovian, because it depends only on the current state of the random walk S. It is randomized because it depends on the outcome of the Bernoulli random variables $\xi_{t,x}$'s.

Figure 3 illustrates such a stopping time. The grey boundary divides the area into two subareas: the one on the left hand side has white nodes representing "continue", and that on the right hand side consists of black nodes indicating "stop". Stopping at a grey node (t,x) is randomized with $r(x)$ being the probability of stopping.

The following theorem is one of the main results of the paper. It provides a theoretical foundation for the numerical algorithm we are going to present to solve our casino gambling model. It characterizes the admissible set of stopped distributions under stopping times in $\mathcal{T}_T$, and reveals that

[13] The (original) Root stopping time τ is the first entry time of B to a region $\mathcal{R}$ in the time-space $\mathbb{R}_+ \times \mathbb{R}$, which has the property that if $(t,x) \in \mathcal{R}$ than also $(s,x) \in \mathcal{R}$ for all $s \geq t$. For any centred μ on $\mathbb{R}$ with a finite second moment v, a region $\mathcal{R}$ exists such that the time τ embeds μ in B, i.e., $B_\tau \sim \mu$ and $\mathbb{E}[\tau] = v$; see Root (1969), Rost (1976) and Obłój (2004).

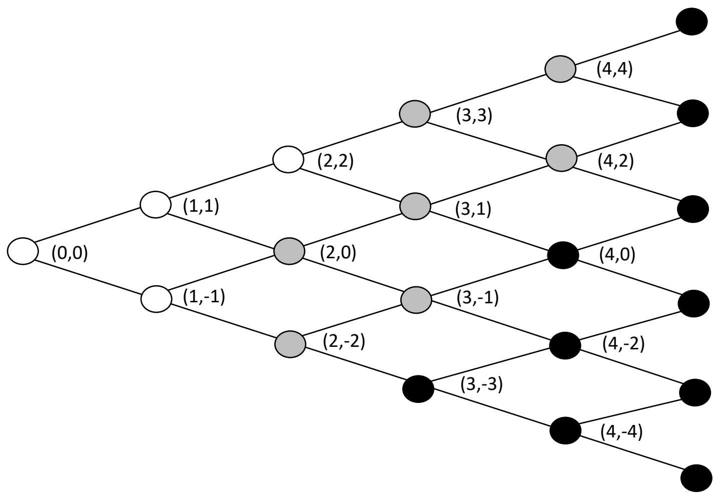

Figure 3 An example of the Root stopping time with $T=5$. Black nodes mean "stop", white nodes mean "continue", and grey nodes mean "randomize". The boundary b is given as follows: $b(4)=4$, $b(3)=3$, $b(2)=4$, $b(1)=3$, $b(0)=2$, $b(-1)=3$, $b(-2)=2$, $b(-3)=3$, $b(-4)=4$.

the set is the same as that of stopped distributions using only randomized Root stopping times. As a consequence, any admissible stopping strategy may be replaced with a suitable randomized Root stopping time without decreasing the value function.

THEOREM 3.1. *Let $T \geq 1$, $\mu \in \mathcal{P}_0(\mathbb{Z})$ such that $\mu([-T,T])=1$. Then there exists a stopping time $\tau \in \mathcal{T}_T$ such that $S_\tau \sim \mu$ if and only if*

$$U_\mu(x) \leq \frac{U^\mu_{T-1}(x+1)+U^\mu_{T-1}(x-1)}{2}, \quad x=-(T-2),-(T-4),\ldots,T-4,T-2. \tag{8}$$

Moreover, in this case there exists a randomized Root stopping time $\tau_R(\mathbf{b},\mathbf{r}) \in \mathcal{T}_T$ such that $S_{\tau_R(\mathbf{b},\mathbf{r})} \sim \mu$.

4. A Mathematical Program

Theorem 3.1 hints that we can, instead of endeavoring to find the stopping time τ in Problem (4), try to find the probability distribution μ of the stopped state S_τ. Namely we change decision variable from τ to μ for Problem (4). The resulting problem is a (nonlinear) *mathematical program* (i.e., a constrained optimization problem) with the condition (8) translating into certain constraints.

Moreover, once we solve this problem and find the optimal distribution μ, then it follows from Theorem 3.1 that there exists a randomized Root stopping time $\tau_R(\mathbf{b},\mathbf{r})$ that achieves the same stopped distribution and, hence, solves (4). Furthermore, based on the proof of Theorem 3.1 (see Appendix A), we can devise an algorithm to find $(\mathbf{b},\mathbf{r})$ and, consequently, $\tau_R(\mathbf{b},\mathbf{r})$. We now formulate the mathematical program and provide its solution algorithm.

4.1. The mathematical program formulation and solution

Given $\tau \in \mathcal{T}_T$, let $\mu \sim S_\tau$. Define two T-dimensional vector variables, $\mathbf{x} := (x_1, x_2, ..., x_T)$ and $\mathbf{y} := (y_1, y_2, ..., y_T)$, where $x_n = \mu([n,T])$, $y_n = \mu([-T,-n])$, $n = 1,2,...,T$. Clearly, $\mathbf{x}$ and $\mathbf{y}$ are gambler's decumulative gain distribution and cumulative loss distribution, respectively. Then the original objective function (3) is equivalent to, as a function of $(\mathbf{x},\mathbf{y})$,

$$\mathbb{U}(\mathbf{x},\mathbf{y}) := \sum_{n=1}^{T} [u_+(n) - u_+(n-1)]\, w_+(x_n) - \lambda \sum_{n=1}^{T} [u_-(n) - u_-(n-1)]\, w_-(y_n). \tag{9}$$

Naturally, we must have $1 \ge x_1 \ge x_2 \ge ... \ge x_T \ge 0$, $1 \ge y_1 \ge y_2 \ge ... \ge y_T \ge 0$, $x_1 + y_1 \le 1$. On the other hand, μ has zero expectation due to optional sampling theorem; so

$$\begin{aligned} 0 &= \sum_{n=-T}^{T} n\mu(\{n\}) = \sum_{n=1}^{T} n\mu(\{n\}) - \sum_{n=1}^{T} n\mu(\{-n\}) = \sum_{n=1}^{T} \mu([n,T]) - \sum_{n=1}^{T} \mu([-T,-n]) \\ &= \sum_{n=1}^{T} x_n - \sum_{n=1}^{T} y_n. \end{aligned}$$

In summary, for $\tau \in \mathcal{T}_T$, $\mu \sim S_\tau$ necessarily satisfies the following constraints:

$$\begin{cases} 1 \ge x_1 \ge x_2 \ge ... \ge x_T \ge 0, \\ 1 \ge y_1 \ge y_2 \ge ... \ge y_T \ge 0, \\ x_1 + y_1 \le 1, \\ \sum_{n=1}^{T} x_n = \sum_{n=1}^{T} y_n. \end{cases} \tag{10}$$

Moreover, Theorem 3.1 necessitates condition (8), which constitutes a family of inequalities on μ's potential function and the corresponding evolutional functions, which will later be translated into constraints on S_τ's distribution functions. Here, let us illustrate (8) for some values of T's. To ease notation, we will suppress the superscript μ on the evolutional functions. For $T = 1$, the condition is satisfied automatically for any $\mu \in \mathcal{P}_0(\mathbb{Z})$ with $\mu([-T,T]) = 1$. For $T = 2$, (8) amounts to $U_\mu(0) \le \frac{U_1(1)+U_1(-1)}{2} = 1$. For $T = 4$, (8) is equivalent to

$$\begin{cases} U_\mu(2) \le \frac{3+U_3(1)}{2}, & U_3(1) = \min\left\{U_\mu(1), \frac{2+\min(U_\mu(0),1)}{2}\right\}, \\ U_\mu(0) \le \frac{U_3(1)+U_3(-1)}{2}, & U_3(-1) = \min\left\{U_\mu(-1), \frac{2+\min(U_\mu(0),1)}{2}\right\}, \\ U_\mu(-2) \le \frac{3+U_3(-1)}{2}. \end{cases}$$

The following lemma, which follows a direct, if somewhat lengthy, computation, expresses $U_\mu(n)$ and, consequently, the constraints (8), in terms of $\mathbf{x}$ and $\mathbf{y}$.

LEMMA 4.1. *For $n \in \mathbb{Z} \cap [-T,T]$,*

$$U_\mu(n) = \begin{cases} 2\sum_{j=n+1}^{T} x_j + n, & n \ge 0, \\ 2\sum_{j=|n|+1}^{T} y_j + |n|, & n < 0. \end{cases}$$

To illustrate, take $T=5$. Then

$$U_\mu(3)=2\sum_{n=4}^{5}x_n+3\,,\;U_\mu(-3)=2\sum_{n=4}^{5}y_n+3\,,$$
$$U_\mu(1)=2\sum_{n=2}^{5}x_n+1\,,\;U_\mu(-1)=2\sum_{n=2}^{5}y_n+1\,.$$

We are now ready to formulate the mathematical program that is equivalent to the original stopping problem (4). Define

$$\mathbf{A}=\begin{bmatrix}1 & 0 & \dots & 0\\ -1 & 1 & \ddots & \vdots\\ 0 & -1 & \ddots & 0\\ \vdots & \ddots & \ddots & 1\\ 0 & \dots & 0 & -1\end{bmatrix}_{(T+1)\times T},\quad \mathbf{c}=\begin{bmatrix}1\\0\\0\\ \vdots\\0\end{bmatrix}_{(T+1)\times 1},\quad \mathbf{1}=\begin{bmatrix}1\\1\\ \vdots\\1\end{bmatrix}_{T\times 1},\quad \mathbf{e_j}=\begin{bmatrix}\vdots\\1\\0\\ \vdots\end{bmatrix}_{T\times 1},$$

along with a set of functions $f_n^m:\mathbb{R}^T\times\mathbb{R}^T\to\mathbb{R}$, $m=1,\dots T$, $n=1,\dots 2T+1$, in the following way. For $\mathbf{x},\mathbf{y}\in\mathbb{R}^T$, let

$$f_1^m(\mathbf{x},\mathbf{y})=f_{2T+1}^m(\mathbf{x},\mathbf{y})\equiv T,\;\; m=1,\dots T,$$

$$f_n^1(\mathbf{x},\mathbf{y})\equiv|n-(T+1)|,\;\; n=2,\dots 2T,$$

and for $m=2,3,\dots T$:

$$f_n^m(\mathbf{x},\mathbf{y})=\begin{cases}\min\left(\frac{f_{n-1}^{m-1}(\mathbf{x},\mathbf{y})+f_{n+1}^{m-1}(\mathbf{x},\mathbf{y})}{2},2\sum_{j=T+2-n}^{T}\mathbf{e_j}'\mathbf{y}+(T+1)-n\right), & n=2,3,\dots T,\\ \min\left(\frac{f_{n-1}^{m-1}(\mathbf{x},\mathbf{y})+f_{n+1}^{m-1}(\mathbf{x},\mathbf{y})}{2},2\sum_{j=n-T}^{T}\mathbf{e_j}'\mathbf{x}+n-(T+1)\right), & n=T+1,\dots 2T-1,2T.\end{cases}$$

Then, the mathematical program is

$$\begin{aligned}\max_{\mathbf{x},\mathbf{y}}\quad & \mathbb{U}(\mathbf{x},\mathbf{y})\,,\\ \text{subject to}\quad & \mathbf{Ax}\le\mathbf{c},\;\mathbf{Ay}\le\mathbf{c},\;\mathbf{e_1}'\mathbf{x}+\mathbf{e_1}'\mathbf{y}\le 1,\;\mathbf{1}'\mathbf{x}-\mathbf{1}'\mathbf{y}=0,\\ & \tfrac{f_{n-1}^T(\mathbf{x},\mathbf{y})+f_{n+1}^T(\mathbf{x},\mathbf{y})}{2}\ge 2\sum_{j=T+2-n}^{T}\mathbf{e_j}'\mathbf{y}+(T+1)-n\\ & \qquad\text{for } n=2k+1,k=1,2,\dots,n\le T,\\ & \tfrac{f_{n-1}^T(\mathbf{x},\mathbf{y})+f_{n+1}^T(\mathbf{x},\mathbf{y})}{2}\ge 2\sum_{j=n-T}^{T}\mathbf{e_j}'\mathbf{x}+n-(T+1)\\ & \qquad\text{for } n=2T-2k+1,k=1,2,\dots,n\ge T+1.\end{aligned}\tag{11}$$

The number of decision variables ($\mathbf{x}$ and $\mathbf{y}$) and the number of constraints in (11) are both *linear* in T; hence the complexity of the problem is manageable. Moreover, there are standard solvers to solve this type of mathematical program.[14]

[14] In the following numerical experiments, we employ nonlinear optimization solver 'fmincon' from **MATLAB** Optimization Toolbox, on a desktop with Intel Core i5-4590/CPU 3.30GHz/RAM 8.00GB. For the Barberis (2012)'s parameters $\alpha_+=\alpha_-=0.95$, $\delta_+=\delta_-=0.5$, $\lambda=1.5$ with $T=5,6,7,8$, MATLAB uses 205 seconds, 220 seconds, 280 seconds, 350 seconds respectively. (Compare with those of the brute force reported in Footnote 3.) The running times for $T=10,20,30,40,50$ are 9.35 minutes, 29 minutes, 89 minutes, 3.5 hours, 7.17 hours, respectively. The codes are available at https://github.com/hucuhksz/optimalgambling.

Once we solve this problem to get optimal $(\mathbf{x}^*, \mathbf{y}^*)$, we then run the following algorithm to find the optimal randomized Root stopping time:

Step 1 Given $\mu^* \equiv (\mathbf{x}^*, \mathbf{y}^*)$, compute the corresponding potential function: $U_{\mu^*}(n) = 2\sum_{j=n+1}^{T} x_j^* + n$ for $n \geq 0$, and $U_{\mu^*}(n) = 2\sum_{j=|n|+1}^{T} y_j^* + |n|$ for $n < 0$. Then, compute its evolutional functions $U_t^{\mu^*}$ by (6), $t = 0, 1, ...T$.

Step 2 Compute the boundary $\mathbf{b}$ that separates the "continue" region from the "stop" region: $b(n) = \inf\{t \geq |n|, t \in \mathbb{Z} : U_{t+1}^{\mu^*}(n) = U_{\mu^*}(n)\}$, $n \in [-T, T] \cap \mathbb{Z}$. Note that the constraints in (11) guarantee that the set involved is non-empty and $b(n) \leq T$ $\forall n \in [-T, T] \cap \mathbb{Z}$.

Step 3 Compute the probability $\mathbf{r}$ to stop at the boundary:

$$r(n) = \frac{U_{b(n)}^{\mu^*}(n-1) + U_{b(n)}^{\mu^*}(n+1) - 2U_{\mu^*}(n)}{U_{b(n)}^{\mu^*}(n-1) + U_{b(n)}^{\mu^*}(n+1) - 2U_{b(n)}^{\mu^*}(n)}, \ n \in [-T, T] \cap \mathbb{Z}.$$

Step 4 Construct $\tau_R(\mathbf{b}, \mathbf{r})$ according to (7).

4.2. A numerical example

We present an example to illustrate the solution procedure, using the same parameters as in Barberis (2012) with $T = 5$, $\alpha_+ = \alpha_- = 0.95$, $\delta_+ = \delta_- = 0.5$, $\lambda = 1.5$.[15] Solving the corresponding mathematical program for the optimal distribution μ^* yields

$$\begin{aligned} &x_1^* = 0.1875,\ x_2^* = 0.1273,\ x_3^* = 0.1227,\ x_4^* = 0.03152,\ x_5^* = 0.03098, \\ &y_1^* = 0.5,\ y_2^* = 0,\ y_3^* = 0,\ y_4^* = 0,\ y_5^* = 0. \end{aligned}$$

The corresponding potential function U_{μ^*} is

$$\begin{aligned} &U_{\mu^*}(0) = 1,\ U_{\mu^*}(1) = 1.625,\ U_{\mu^*}(2) = 2.3704,\ U_{\mu^*}(3) = 3.125,\ U_{\mu^*}(4) = 4.06196, \\ &U_{\mu^*}(n) = |n| \text{ for } n \geq 5 \text{ and } n \leq -1. \end{aligned}$$

We then apply the algorithm previously presented to recover the optimal randomized Root stopping time τ^* from the optimal distribution μ^*, with $S_{\tau^*} \sim \mu^*$. The strategy, which is optimal at $t = 0$ (only) and implemented by the precommitted gambler, is drawn in the left panel of Figure 4. Note that black nodes mean "stop", white ones mean "continue", and grey ones mean "randomization". The number above a grey node is the probability to stop.

The main feature of this precommitted optimal strategy is to continue in the gain domain and to stop in the loss domain until $T = 5$, except at time 4 where there are positive probabilities to stop in gains. In particular, randomization takes place at nodes $(4, 4)$ and $(4, 2)$, with the (very small) probabilities to stop equal to 0.00864 and 0.0368, respectively. The CPT value of this randomized

[15] More examples with much longer time horizons will be presented in the next section.

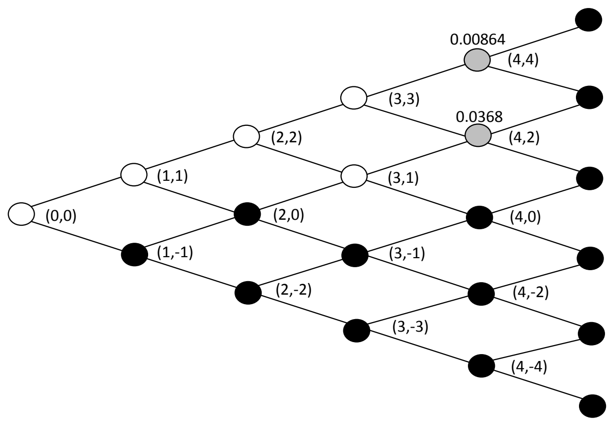

Figure 4 The left panel shows the precommitter's strategy and the right panel shows the naïf's strategy, for $T=5$, $\alpha_+=\alpha_-=0.95$, $\delta_+=\delta_-=0.5$, $\lambda=1.5$. Black nodes mean "stop", white nodes mean "continue", and grey nodes mean "randomize". The numbers above the grey node stand for the probability to stop. While the precommitter's strategy is mainly to continue in gains and stop in losses, the naïf's behavior is almost completely reversed.

strategy is 0.3369592. Compared with Barberis (2012) where the CPT value is 0.3369398, the optimal non-randomized, Markovian strategy has white nodes at $(4,4)$ and $(4,2)$, instead of grey nodes that involve randomization.[16] In summary, allowing path-dependent and randomized strategies does indeed improve optimal CPT values (albeit only slightly) over non-randomized, Markovian ones.[17] Moreover, one can achieve this improved optimal value by implementing a Markovian randomization, with the overall strategy very similar qualitatively to Barberis (2012)'s – both are of the loss-exit type.[18]

While the precommitted gambler follows through the optimal strategy originally determined at time 0, a *naïve* gambler thought he would do the same but in actuality constantly deviates from previously planned strategies. More precisely, at *any* time $t>0$, a naïf *re-considers* the optimal stopping problem starting from t, devises a precommitted strategy but carries it out for only *one* period (because he will re-optimize again at the next time instant). Here, we assume this gambler

[16] These are the only two nodes that are different between Barberis (2012) and the present paper. Note they occur at $T-1$ and when there are sufficient gains. The intuition why the gambler randomizes at these two nodes will be explained in Subsection 5.4 after we have presented more examples and established more general results.

[17] This improvement can be *significant* with other parameter specifications. For example, with $T=2$ and $(\alpha_\pm,\delta_\pm,\lambda)=(0.9,0.5,1.25)$, the optimal CPT value among non-randomized, path-independent strategies is 0.058069135, and that among randomized ones is 0.065696808, representing a 13% increase. When $T=2$ and $(\alpha_\pm,\delta_\pm,\lambda)=(0.5,0.5,1)$, the corresponding values are 0.0253839 and 0.0492624, representing a 94% increase.

[18] We have also revisited the $T=6$ example considered also in He et al. (2017). In that paper, a randomized strategy, found by trial and error, leads to the value function $V=0.250702$ compared with $V=0.250440$ for the best non-randomized strategy. Using our algorithm, we see that the best randomized strategy actually gives $V=0.257483$. It is still a loss-exit type and it stops at node $(0,0)$ with probability 0.201, node $(5,1)$ with 0.436, node $(5,3)$ with 0.0292 and node $(5,5)$ with 0.0113.

keeps his initial wealth at time 0 as the reference point.[19] The naïve gambler's strategy can be computed by deriving all the time-t precommitted strategies, $t = 0, 1, ..., T$, implementing each of them for just one period, and then "pasting" them together. As a result, his *actual* quitting strategy could be drastically different from the precommitted one, the one he originally planned before he enters the casino; see the right panel of Figure 4. There, the only node calling for randomization is now (2,0), with a probability of 0.179 to quit.[20] Comparing the two strategies depicted in Figure 4, we find that the naïve strategy is not only significantly different from the precommitted one, but indeed almost completely *opposite* in character: the latter is mainly to continue in the gain domain and to stop in the loss domain, while the former is reversed. This sharp contrast between the planned and actual behaviors along with experimental evidences was first noted and discussed in Barberis (2012, Section 4.3) for the non-randomized model. Heimer et al. (2020) presents strong evidence from lab and field that supports the dynamic inconsistency that emerges in the framework of Barberis (2012). In the context of stock trading, such a naïve behavior – the tendency of selling winners too soon and keeping losers too long – is widely observed especially for retail investors, and is termed the *disposition effect* by Odean (1998).

5. Discussion

5.1. To enter or not to enter: the power of randomization

One of the main takeaways of Barberis (2012) is that CPT offers an explanation why a gambler would be willing to enter a casino even if the bets there have neither skewness nor positive expected values. By implementing a loss-exit strategy, namely keep gambling when winning but stop gambling when accumulating a sufficient loss, he envisions a positively skewed probability distribution of the accumulated gain/loss at the exit time which is favored by the CPT preference. However, he would need a sufficiently *long* time period to build such a skewed distribution in order to have a *positive* CPT value to justify the entry (recall that the CPT value of not playing at all is zero). For the case of a piece-wise power utility function (1) and an inverse S-shaped weighting function (2), Barberis (2012, Proposition 1) provides a sufficient condition for this to happen.[21] Moreover, for the parameter values $\alpha_+ = \alpha_- = 0.88$, $\delta_+ = \delta_- = 0.65$, $\lambda = 2.25$, this sufficient condition translates into $T \geq 26$; see Barberis (2012, Corollary 1).[22]

[19] This is also the assumption made in Barberis (2012) when analyzing a naïve gambler's behavior. It is both natural and plausible that a gambler remembers the initial amount of cash he brought into the casino and always compares wins and losses against that amount.

[20] This is also the only node that makes our naïve strategy different from Barberis (2012)'s in which the node (2,0) is white meaning "continue"; see the right panel of Figure 4 therein.

[21] Barberis (2012, Proposition 1) is stated for a naïve gambler. However, the result holds for a precommitter as well because both gamblers face the same problem at $t = 0$.

[22] These parameter values are close to those given by Tversky and Kahneman (1992), i.e., $\alpha_+ = \alpha_- = 0.88$, $\delta_+ = 0.61$, $\delta_- = 0.69$, $\lambda = 2.25$. If we apply the exact Tversky and Kahneman (1992) parameter values to Barberis (2012,

However, with randomization allowed, the gambler may be willing to enter the casino even if he is allowed to play only *once* (i.e., $T = 1$).

PROPOSITION 5.1. *Suppose $T = 1$. If $\lim_{p\to 0}[w'_+(p)/w'_-(p)] > \lambda[u_-(1)/u_+(1)]$, then the optimal CPT value is strictly positive.*

Recall that in our model, randomization is available; so the optimal CPT value being strictly positive means that the gambler will enter the casino, possibly tossing a coin to decide whether to actually play (the only) one round of bet. It is straightforward to show that for the weighting function (2), when $\delta_+ < \delta_-$ (which is the case with Tversky and Kahneman (1992)'s estimates), we have $\lim_{p\to 0}[w'_+(p)/w'_-(p)] = +\infty$. Hence, Proposition 5.1 yields that, as long as the loss-aversion degree λ is finite, a randomized gambling strategy is always preferred to non-gamble, even when $T = 1$. As an example, consider a piece-wise power utility function (1) with $\alpha_\pm \in (0,1)$, $\lambda = 1.5$ and an inverse S-shaped weighting function (2) with $\delta_+ = 0.5$, $\delta_- = 0.95$. The optimal CPT value with $T = 1$ is 0.07596, and the gambler will enter the casino with a (small) probability 0.057. Without randomization, however, the gambler will not enter the casino under this group of parameters.[23] In line with our discussion of the relation between path-dependence and randomization in Section 2.3, we recall that this randomization may be achieved endogenously by replacing it with path-dependence. The gambler who enters the casino will typically see a strip which displays the results of the past N roulette spins. Even looking at only the last 4 results, the gambler in this particular example can very closely replicate the randomized strategy: if he decides to bet on red only if the last 4 spins were red, and does not bet otherwise, he will bet with probability 0.0625.

What if $T \geq 2$? Naturally, as T increases, the optimal CPT values increase. Figure 5 graphs the optimal CPT values for $T = 1, 2, \ldots, 20$ with the Tversky and Kahneman (1992) estimates. Therefore, if the gambler will enter the casino for $T = 1$ with a given set of parameters, so will he for $T \geq 2$ with the same parameters. As a consequence, Proposition 5.1 holds for $T \geq 2$ as well.

The intuition of Proposition 5.1 is as follows. Let r be the probability of not entering the casino based on the coin flip at $t = 0$. Then the gambler's gamble becomes a gain of \$0 with probability r, a gain of \$1 with probability $(1-r)/2$, and a loss of \$1 with probability $(1-r)/2$. For sufficiently high r, the decision to enter is driven by the relative overweighting of low probabilities for the gain vs. the loss; and if the overweighting for the gain is high enough, the optimal CPT value is positive and hence the gambler will enter with a positive probability. Note that even though randomization

Proposition 1), then the corresponding $T \geq 20$. Such a shorter period is expected because the probability weighting in gains is stronger than that in losses with Tversky and Kahneman (1992)'s parameters; thus it takes less time to build the desired positively skewed distribution with a positive CPT value.

[23] Such a result also holds under the Tversky and Kahneman (1992) estimates; that is, the gambler will enter the casino with $T = 1$ if randomization is allowed, but will not without randomization.

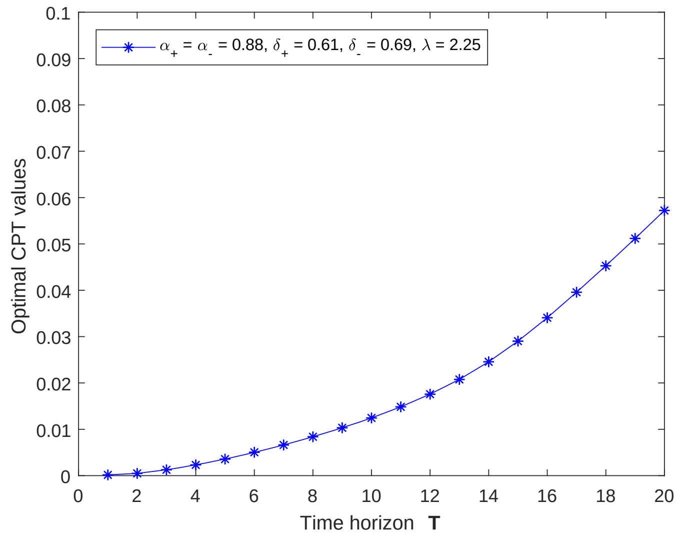

Figure 5 **Optimal CPT values for $T = 1, 2, ...20$ under the parameter values of Tversky and Kahneman (1992), i.e., $\alpha_+ = \alpha_- = 0.88$, $\delta_+ = 0.61$, $\delta_- = 0.69$, $\lambda = 2.25$.**

still gives rise to a *symmetric* distribution of gains and losses and hence the loss aversion seemingly would prevent the gambler from entering, the condition $\lim_{p\to 0}[w'_+(p)/w'_-(p)] > \lambda[u_-(1)/u_+(1)]$ means that exaggeration of the gain outweighs exaggeration of the loss and the loss aversion *combined*, yielding the contrary.[24]

On the other hand, the effectiveness of randomization crucially depends on the chosen parameters. If the degree of probability weighting in gains is equal to or less than that in losses, and the level of loss-aversion is sufficiently large so that $\lambda[u_-(1)/u_+(1)] > 1$, then the above proposition does not apply because $\lim_{p\to 0}[w'_+(p)/w'_-(p)] \le 1 < \lambda[u_-(1)/u_+(1)]$. For example, for utility function (1) and probability weighting function (2), let $\alpha_+ = \alpha_- = 0.88$, $\delta_+ = \delta_- = 0.65$, $\lambda = 2.25$, the values used in Barberis (2012, Corollary 1). In this case, we find a positive CPT value of *randomized* strategies only when the time horizon is at least $T = 25$, slightly shorter than that ($T = 26$) presented in Barberis (2012, Corollary 1) using a special *non-randomized* loss-exit strategy. The corresponding optimal precommitted strategy is also of a loss-exit type: stop once losing \$1, and continue with possible randomization when wining. If we further let the probability weighting in gains be weaker than that in losses, e.g., $\delta_+ = 0.69$ and $\delta_- = 0.61$, then a positive preference value is found only at $T \ge 39$.

[24] Here we must put in a caveat that, as discussed in Subsection 2.3, it is not widely observed that people toss coins to decide whether to enter a casino. So, the discussion here is more hypothetical – whether a CPT gambler would enter a casino *if* he knew randomization might increase his preference value.

5.2. Prospect theory parameter values

There are three components in the risk/loss preferences under CPT: the utility function, the probability weighting and the loss aversion. They are intertwined and compete with each other in determining the overall preference and dictating the final behavior. In this subsection, we study the roles they play in the case when the utility function is (1) and the weighting function is (2), with parameters $\alpha_\pm$, $\delta_\pm$ and λ. Note that Barberis (2012, Section 3.1) also discusses them broadly (see especially Figure 3 in that paper), albeit for $T=5$. We are able to consider larger values of T thanks to the algorithm derived in the current paper.

Other parameters being kept unchanged, the effect of each of these parameters is as follows: a smaller α_+ implies a higher degree of risk-aversion in gains and a smaller α_- implies a higher degree of risk-seeking in losses, a smaller $\delta_\pm$ yields a higher level of probability weighting in gains/losses, and a smaller λ indicates a smaller extent of loss aversion. To understand the overall impact of these parameters on exit decisions, we first fix λ and consider four sets of scenarios: large $\alpha_\pm$ and small $\delta_\pm$; small $\alpha_\pm$ and large $\delta_\pm$; small $\alpha_\pm$ and small $\delta_\pm$; and large $\alpha_\pm$ and large $\delta_\pm$. Then we examine the effect of λ. In the following discussions we fix $T=10$.

The left panel of Figure 6 draws the optimal precommitted strategy when $\alpha_\pm=0.95$, $\delta_\pm=0.5$ and $\lambda=1.5$. These are the same parameters as in the numerical example presented in Subsection 4.2, except now we have a much longer horizon. Again, black nodes mean "stop", white nodes "continue", grey nodes "randomize", and the number above the grey node is the probability to stop. This strategy is mainly to continue or toss a coin in gains until the final time and to stop in losses, which is thus a loss-exit one. The intuition is as follows. This is the case where $\alpha_\pm$ are relatively large (lower risk-aversion/-seeking) and $\delta_\pm$ relatively small (heavier probability weighting). In the gain region, the stronger exaggeration of the small probability of winning a large amount outweighs the weaker risk aversion; hence the gambler is willing to take more risk and stay longer. In the loss region, the stronger exaggeration of the small probability of losing a large amount, together with the loss aversion, outweighs the weaker risk-seeking appetite and prompts the gambler to play safe and quit earlier.

The above argument is reversed, leading to a gain-exit type of strategy, when $\alpha_\pm$ are relatively small and $\delta_\pm$ relatively large, such as the one depicted in the left panel of Figure 7 where $\alpha_\pm=0.5$, $\delta_\pm=0.95$, $\lambda=1.5$. An interesting small variation of this case is when probability weighting is absent, i.e., $\alpha_\pm=0.5$, $\delta_\pm=1$, $\lambda=1.5$, in which the optimal CPT value is positive and the precommitted strategy is still a gain-exit one. Indeed, a positive preference value is found at a much shorter horizon $T=4$ under this group of parameters, and the optimal distribution of S_τ is left-skewed (which is favored by a strong risk-seeking preference in losses represented by α_-).

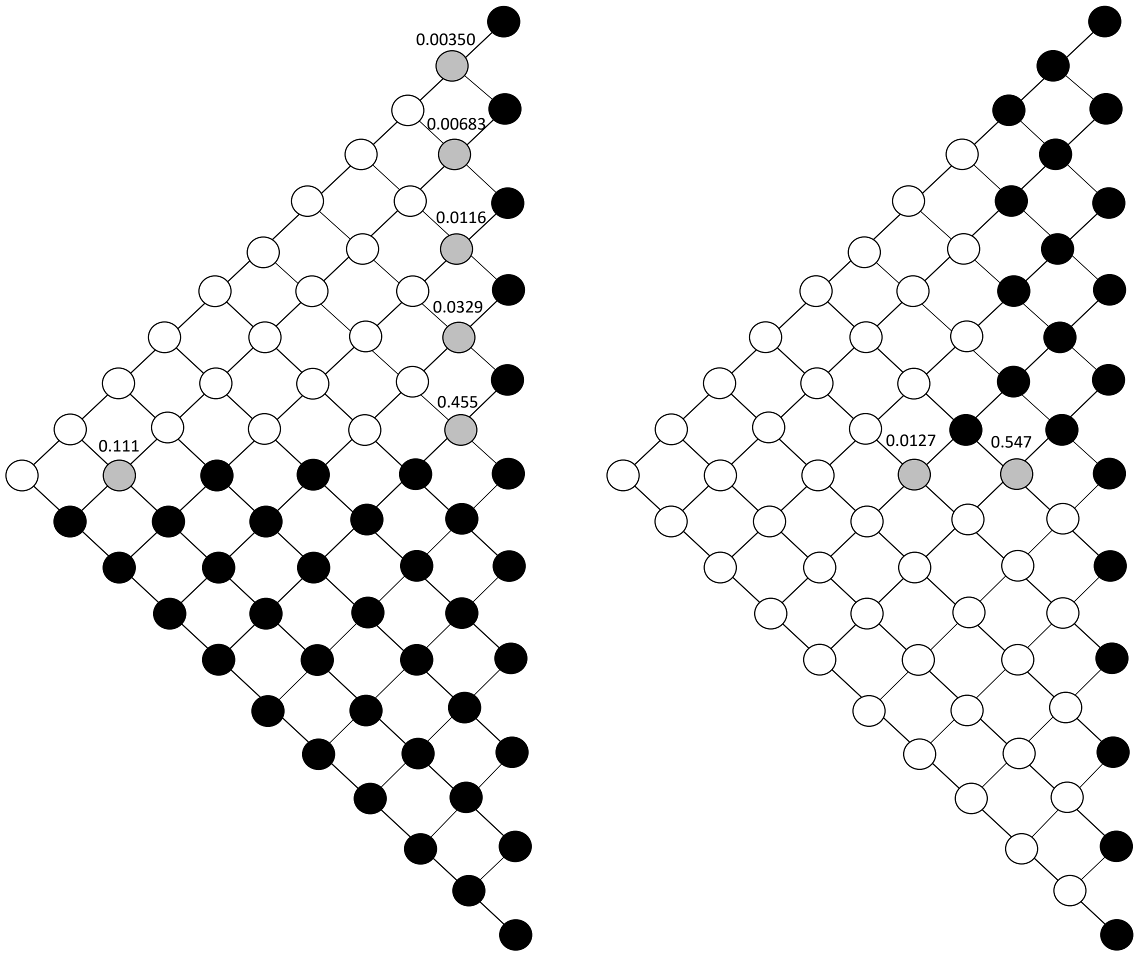

Figure 6 **The precommitted (left panel) and naïve (right panel) strategies for $T = 10$, $\alpha_{\pm} = 0.95$, $\delta_{\pm} = 0.5$, $\lambda = 1.5$. Black nodes are "stop", white nodes are "continue", and grey nodes are "randomize". The numbers above the grey nodes are the probabilities to stop.**

The left panel of Figure 8 shows the precommitted strategy for the parameter values $\alpha_{\pm} = \delta_{\pm} = 0.5$ and $\lambda = 1.5$, which is the case of small $\alpha_{\pm}$ and small $\delta_{\pm}$. This is still a loss-exit strategy, but the main differences from that visualized by the left panel of Figure 6 are that, in the gain region, there are now more black nodes and the numbers above the grey nodes are larger, implying a higher likelihood of stop even when the gambler has accumulated a gain. The reason is that with a smaller α_+, the exaggeration of the small probability of winning a large amount still outweighs the risk aversion in gains, but with a lesser degree than the previous case.

The last set of parameters are $\alpha_{\pm} = \delta_{\pm} = 0.95$, $\lambda = 1.5$, with which the optimal CPT value is zero and the gambler will simply not enter the casino. This is because these parameter values render a risk preference close to both risk-neutral and probability–weighting–free, while a zero-mean bet and a loss-aversion degree $\lambda > 1$ prevent the gambler from playing the game at all.

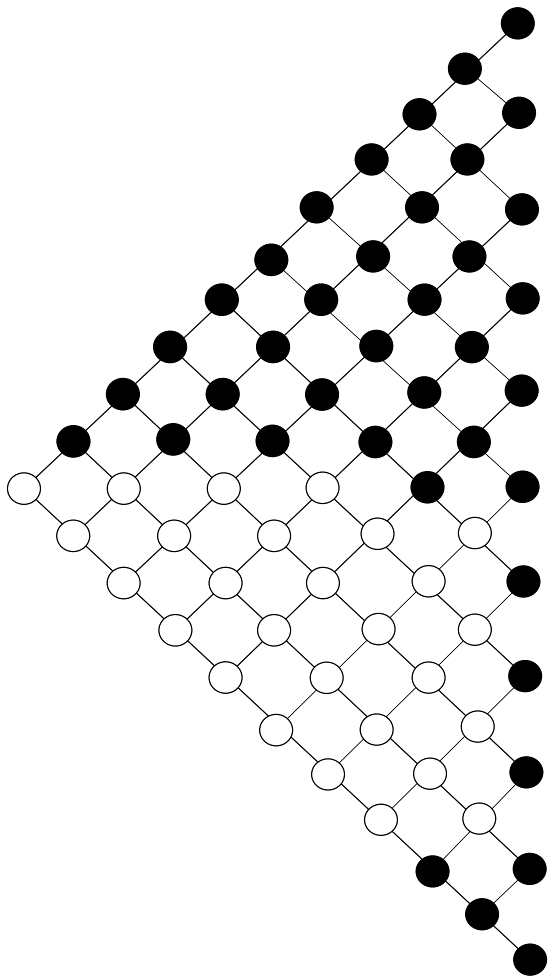

Figure 7 The precommitted (left panel) and naïve (right panel) strategies for $T = 10$, $\alpha_\pm = 0.5$, $\delta_\pm = 0.95$, $\lambda = 1.5$. Black nodes are "stop" and white nodes are "continue". There is no grey node.

The impact of λ is more straightforward, which we now examine. For each group of $\alpha_\pm$ and $\delta_\pm$ considered above, we obtain the optimal CPT value by varying λ from 1 to 3; see Figure 9, the left panel. Quite naturally, each of the optimal CPT values decreases as λ increases, and three of them hit zero before λ reaches 3. As a result, the gambler will be increasingly reluctant to stay in or even enter the casino as his level of loss aversion increases.

The analysis in this subsection shows that the CPT casino modeling with various constellations of parameter specifications can predict and explain a rich array of gambler behaviors. In particular, whether the strategy is loss-exit or otherwise depends on the interplay between the three intertwining and competing forces represented by $\alpha_\pm$, $\delta_\pm$, and λ.

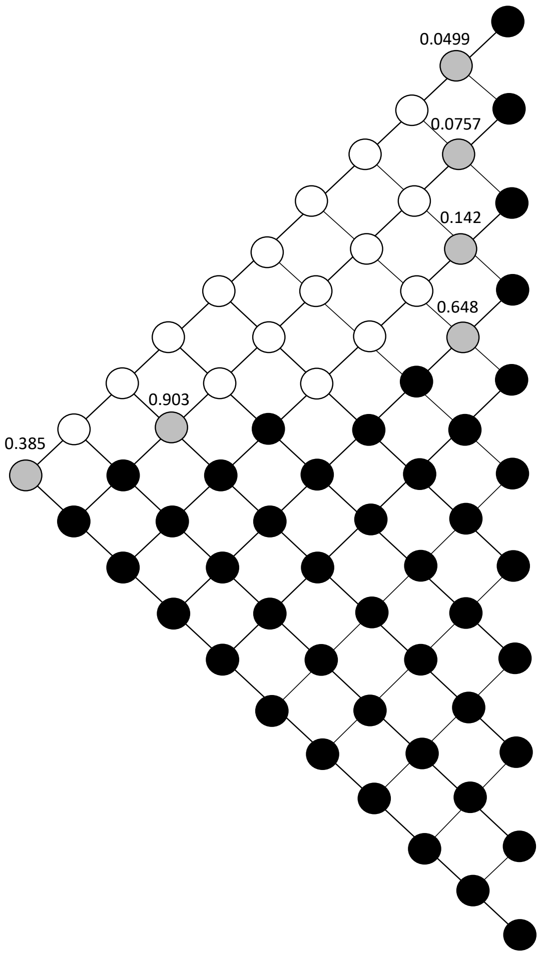

Figure 8 **The precommitted (left panel) and naïve (right panel) strategy for $T = 10$, $\alpha_{\pm} = \delta_{\pm} = 0.5$, $\lambda = 1.5$. Black nodes are "stop", white nodes are "continue", and grey nodes are "randomize". The numbers above the grey nodes are the probabilities to stop.**

5.3. Non-randomized strategy

As discussed earlier randomization is a relaxation approach that turns an analytically intractable problem into a tractable one. Moreover, it gives rise to a natural, rounded strategy for the non-randomized model.

The "luckiest" situation with this relaxation approach is when an optimal strategy happens to be non-randomized, as exemplified by the case when $T = 10$, $\alpha_{\pm} = 0.5$, $\delta_{\pm} = 0.95$, $\lambda = 1.5$. Figure 7 left panel shows the precommitted strategy in this case turns out to be a non-randomized one. A less trivial example is seen in Figure 4 left panel, which shows the precommitted strategy with randomization for $T = 5$, $\alpha_{\pm} = 0.95$, $\delta_{\pm} = 0.5$, $\lambda = 1.5$. If we apply the rounding method to the probabilities associated with the grey nodes, then we obtain a non-randomized strategy which is *exactly* the optimal one given by Barberis (2012).

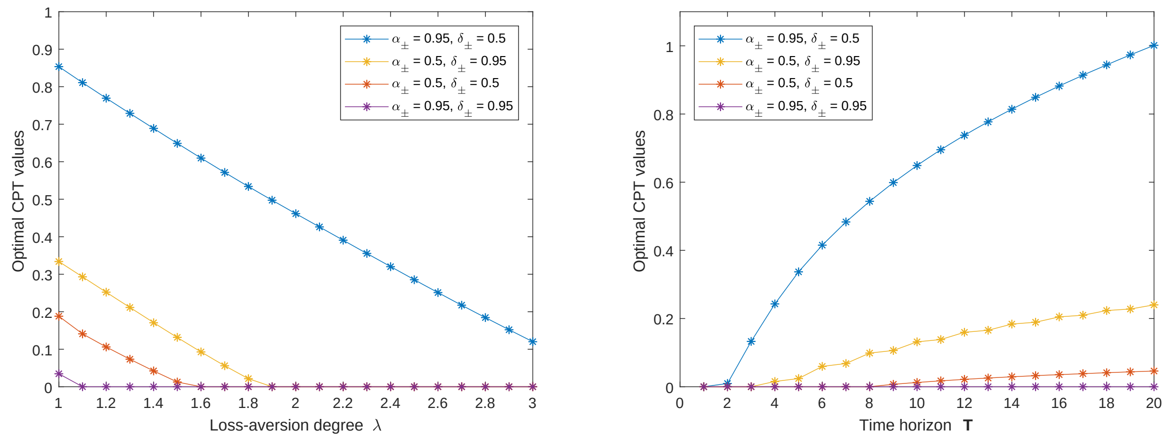

Figure 9 Optimal CPT values for λ from 1 to 3, while $T = 10$, are shown in the left panel. Optimal CPT values for $T = 1, \ldots 20$, while $\lambda = 1.5$, are shown in the right panel. In both panels, $(\alpha_\pm, \delta_\pm) \in \{(0.95, 0.5), (0.5, 0.95), (0.5, 0.5), (0.95, 0.95)\}$.

These two examples are, naturally, more the exception than the rule; but in general we may still be able to obtain some error bound of a rounded sub-optimal strategy for the non-randomized model. For example, with $T = 10$, $\alpha_\pm = 0.95$, $\delta_\pm = 0.5$, $\lambda = 1.5$, Figure 6 left panel shows the precommitted randomized strategy, with the corresponding optimal CPT value being 0.64886. A straightforward calculation yields that the non-randomized strategy using round-up/down delivers a CPT value of 0.64853. Since the optimal CPT value of the non-randomized model must lie between these two values, we conclude that the rounded strategy is sub-optimal for the non-randomized model with a relative error less than 0.05%. However, we recall from Footnote 18, which dealt with $T = 2$, that this error may become larger for shorter horizons. The reason for that, as discussed in Section 5.1, is that without randomization a distribution perceived (after probabilities are distorted) to be positively skewed cannot be constructed in a relatively short time. Randomization greatly enlarges the set of achievable distributions for short horizons. As a result, the effect of randomization in terms of improving the optimal CPT value dissipates in time. This nonetheless is a good news in terms of solving the original Barberis (2012)'s model without randomization: when the horizon is short we use the brute force, and when it is long we use the rounding method.

5.4. Grey nodes at time $T-1$

When examining the $t = T-1$ nodes (i.e., those just before the terminal date) in the examples of loss-exit precommitted strategies shown in Figure 4 with $T = 5$ and in Figures 6 and 8 with $T = 10$, we observe that the upper nodes are grey (where randomization takes place) and the lower ones are black (where the gamble is stopped). These colors are actually prevalent and not coincidences.

Take the node $(T-1,T-1)$ for example, which is in the gain region. Let $\tau^* \in \mathcal{T}_T$ be the optimal precommitted strategy, and p_{T-1} be the probability (which is assumed to be positive) that the node $(T-1,T-1)$ will be reached under τ^*. Now, *given* the gambler is at $(T-1,T-1)$, denote by $r_{T-1} \in [0,1]$ the probability to stop, and $q_{T-1} = \frac{1-r_{T-1}}{2} \in [0,\frac{1}{2}]$. Because the value of q_{T-1} only affects the decumulative probabilities of $\mathbb{P}(S^*_\tau \geq T-1)$ and $\mathbb{P}(S^*_\tau \geq T)$, the optimal value of q_{T-1} (corresponding to τ^*), denoted by q^*_{T-1}, must maximize the following function in q_{T-1}:

$$w_+(q_{T-1}p_{T-1})[u_+(T)-u_+(T-1)] + w_+\big((1-q_{T-1})p_{T-1}\big)[u_+(T-1)-u_+(T-2)].$$

For T large enough, both $q_{T-1}p_{T-1}$ and $(1-q_{T-1})p_{T-1}$ are small enough to fall into the *concave* region of the probability weighting function $w_+(\cdot)$. Hence the above is a concave maximization and the following first-order condition is necessary and sufficient for a maximum q^*_{T-1}:

$$0 = \left\{w'_+(q^*_{T-1}p_{T-1})[u_+(T)-u_+(T-1)] - w'_+\big((1-q^*_{T-1})p_{T-1}\big)[u_+(T-1)-u_+(T-2)]\right\} p_{T-1},$$

or equivalently,

$$\frac{w'_+(q^*_{T-1}p_{T-1})}{w'_+\big((1-q^*_{T-1})p_{T-1}\big)} = \frac{u_+(T-1)-u_+(T-2)}{u_+(T)-u_+(T-1)}\,. \tag{12}$$

Assuming $u_+(\cdot)$ is *strictly* concave (e.g., that given by (1)), the right hand side of (12) is strictly greater than one. Hence, the equation is satisfied by some $q^*_{T-1} \in (0,\frac{1}{2})$, but *not* $q^*_{T-1}=0$ (noting $w'_+(0)=+\infty$) or $q^*_{T-1}=\frac{1}{2}$. Recall that $q^*_{T-1}=0$ and $q^*_{T-1}=\frac{1}{2}$ correspond to $r^*_{T-1}=1$ and $r^*_{T-1}=0$ respectively. So, at node $(T-1,T-1)$, he will *not* have a black-and-white decision of either "continue" or "stop"; rather he will *always* engage in randomization at that node under the optimal strategy τ^*.

What is the intuition behind this result? Due to the overweighting on gains, stopping completely at the gain nodes at $t=T-1$ is not optimal when randomization is disallowed, and hence more so when randomization is allowed. On the other hand, playing the last bet without tossing a coin (i.e., *definitely* continuing) is not optimal because of the *strict* risk aversion – randomization with a small probability of stopping triggers probability weighting in gains which in turn offsets the risk aversion level.

In contrast, the time $T-1$ nodes in the *loss* region of the left panels in Figures 4, 6 and 8 are all black. This is because they are all "blocked" by other black nodes before them, and hence simply not reachable by the optimal strategy τ^*. However, in general (under certain mild conditions) these nodes will still be black even if they *are* reachable. Take the node $(T-1,-(T-1))$ for example. Suppose $p_{-(T-1)}>0$ is the probability of reaching $(T-1,-(T-1))$ under τ^*, and $r_{-(T-1)} \in [0,1]$ is the probability to stop at $(T-1,-(T-1))$ and $q_{-(T-1)} = \frac{1-r_{-(T-1)}}{2} \in [0,\frac{1}{2}]$. Then a similar analysis to the gain case shows that the optimal $q^*_{-(T-1)}$ minimizes the following function in $q_{-(T-1)}$:

$$w_-(q_{-(T-1)}p_{-(T-1)})[u_-(T)-u_-(T-1)] + w_-\big((1-q_{-(T-1)})p_{-(T-1)}\big)[u_-(T-1)-u_-(T-2)]. \tag{13}$$

Different from the gain region, in the loss region the optimality is achieved by *minimizing* a *concave* function when $p_{-(T-1)}$ is sufficiently small. Hence, the optimal $q^*_{-(T-1)}$ must be either 0 or $\frac{1}{2}$, corresponding to "stop" or "continue" respectively. This means that the gambler will *not* flip a coin this time. To investigate which is better between "stop" and "continue", we calculate the difference between the objective values (13) at $q^*_{-(T-1)} = 0$ and at $q^*_{-(T-1)} = \frac{1}{2}$:

$$\begin{aligned} & w_-(p_{-(T-1)})[u_-(T-1) - u_-(T-2)] - w_-(p_{-(T-1)}/2)[u_-(T) - u_-(T-2)] \\ = & [u_-(T-1) - u_-(T-2)] w_-(p_{-(T-1)}/2) \left[\frac{w_-(p_{-(T-1)})}{w_-(p_{-(T-1)}/2)} - \frac{u_-(T) - u_-(T-2)}{u_-(T-1) - u_-(T-2)} \right]. \end{aligned}$$

As $T \to \infty$, we have $p_{-(T-1)} \to 0$ and, hence,

$$\frac{w_-(p_{-(T-1)})}{w_-(p_{-(T-1)}/2)} \to 2^{\delta_-}, \quad \frac{u_-(T) - u_-(T-2)}{u_-(T-1) - u_-(T-2)} = 1 + \frac{u_-(T) - u_-(T-1)}{u_-(T-1) - u_-(T-2)} \to 2,$$

assuming $w_-(\cdot)$ is given by (2) with $0 < \delta_- < 1$ and $u_-(\cdot)$ has diminishing marginal (dis)utility, namely, $u_-'(x) \to 0$ as $x \to \infty$ (which holds for (1)). This implies that the value (13) at $q^*_{-(T-1)} = 0$ is smaller than that at $q^*_{-(T-1)} = \frac{1}{2}$, when T is sufficiently large. Consequently, the gambler will optimally choose to stop at node $(T-1, -(T-1))$.

For general utility and weighting functions, the above analysis requires sufficiently large T; but for the utility function (1) and probability weighting function (2) with Tversky and Kahneman (1992)'s estimates, T does not need to be excessively large. For example, it follows from the analysis that all we need is to ensure p_{T-1} falls into the concave domain of $w_+(\cdot)$. For $\delta_+ = 0.61$, this requires $p_{T-1} < 0.3$ (refer to the right panel of Figure 1) which is satisfied when $T = 3$. Similarly, for $\alpha_- = 0.88$ and $\delta_- = 0.69$, a straightforward calculation yields that when $T = 3$, $p_{-(T-1)}$ falls into the concave domain of $w_-(\cdot)$ and consequently $r_{-(T-1)} = 1$ dominates the other choices.

5.5. Naïve gamblers

While a precommitted gambler follows the optimal strategy determined at time 0, a naïve gambler constantly deviates from it. We have shown in Subsection 4.2 that, under the parameter specification therein, the naïf's actual behavior changes from the originally planned loss-exit strategy to an eventual gain-exit one.

Numerically, the naïf's strategy can be obtained by computing each time-t precommitted strategy, carrying it out for just one period, and then pasting them together; see Subsection 4.2 for details. We apply this scheme to the first three groups of parameters studied in Subsection 5.2, and draw the naïve strategies in the right panels of Figure 6 – 8.

The problem in Figure 6 has the same parameter values as that in Figure 4 but a longer horizon. The changes from the left panels to the right ones in the two figures are qualitatively the same, namely the naïf turns a loss-exit strategy to a gain-exit one eventually. The same happens in Figure

8.[25] In Figure 7, the two panels are almost identical – both are gain-exit – except the two lowest nodes at $t=8,9$. This is because the difference in behaviors of the precommitter and the naïf emanates from time-inconsistency, which in turn stems from probability weighting. In this case, the strength of probability weighting is very low with $\delta_{\pm}=0.95$, leading to a low level of time-inconsistency than the other two cases and hence the high similarity between the precommitted and naïve strategies.

It is very interesting to note that, in *all* the cases, the naïve gambler's behavior is *consistent*, irrespective of the underlying parameter specifications: once he enters the casino he *always* takes gain-exit strategies, reminiscent of the disposition effect in security trading. In particular, he never stops losses and gambles "until the bitter end" (Ebert and Strack 2015).[26]

We now give a CPT-based explanation for this phenomenon. Suppose a naïve gambler has accumulated a gain equal to $x>0$ at time $T-1$, the date just before the terminal one. Then his decision problem regarding whether he should quit at $T-1$ can be formulated as

$$\max_{q\in[0,1/2]} g(q), \quad \text{where } g(q):=\big(u_+(x+1)-u_+(x)\big)w_+(q)-\big(u_+(x)-u_+(x-1)\big)(1-w_+(1-q)),$$

and, as before, $q=\frac{1-r}{2}$ and $r\in[0,1]$ is the probability to stop. Suppose w_+ satisfies the so-called subcertainty, i.e., $1-w_+(1-p)\geq w_+(p)$ for $p\in[0,1/2]$, a property that is proposed by Kahneman and Tversky (1979) and shared by many probability weighting functions including (2). Then

$$g(q)\leq\Big(\big(u_+(x+1)-u_+(x)\big)-\big(u_+(x)-u_+(x-1)\big)\Big)w_+(q)\leq 0,$$

where the second inequality follows from the concavity of u_+, while the equality is achieved when $q=0$, corresponding to the decision of "stop". We have established the following result.

PROPOSITION 5.2. *Assume that w_+ satisfies subcertainty. Then it is optimal for a naïve gambler to stop in gain at $T-1$.*

Next, suppose the naïf has accumulated a loss $-x<0$ at $T-1$. His decision problem to continue or stop at $T-1$ is

$$\min_{q\in[0,1/2]} l(q), \quad \text{where } l(q):=\big(u_-(x+1)-u_-(x)\big)w_-(q)-\big(u_-(x)-u_-(x-1)\big)(1-w_-(1-q)).$$

[25] In the right panel of Figure 8, all the nodes with state $x=1$ are black, which "block" the gambler from accessing the nodes beyond state 1. This is why the nodes above state 1 are also all black.

[26] We reiterate that the result of Ebert and Strack (2015) depends critically on the assumption that the gambler can construct arbitrarily small random payoffs, which is possibly valid only in a continuous-time model. The finding that "gamble-until-bitter-end" is also present in the discrete-time casino model suggests that the behavior is probably more prevalent characterizing broadly a naïf (be it a gambler or an investor).

Suppose probability weighting function w_- is differentiable and $w_-'(1-p)/w_-'(p) \geq 1$ for $p \in [0,1/2]$, with the left hand side in the sense of limit for $p=0$. A straightforward calculation verifies that this condition is satisfied by the Tversky–Kahneman weighting function (2). Then

$$
\begin{aligned}
l'(q) &= \big(u_-(x+1)-u_-(x)\big)w_-'(q) - \big(u_-(x)-u_-(x-1)\big)w_-'(1-q) \\
&= \big(u_-(x)-u_-(x-1)\big)w_-'(q)\left(\frac{u_-(x+1)-u_-(x)}{u_-(x)-u_-(x-1)} - \frac{w_-'(1-q)}{w_-'(q)}\right) \\
&\leq \big(u_-(x)-u_-(x-1)\big)w_-'(q)\left(\frac{u_-(x+1)-u_-(x)}{u_-(x)-u_-(x-1)} - 1\right) \leq 0,
\end{aligned}
$$

where the last inequality comes from the concavity of u_-. As a result, $l(q)$ is non-increasing in $q \in [0,1/2]$ and the minimum is achieved when $q=1/2$, corresponding to the "continue" decision.

PROPOSITION 5.3. *Assume that w_- is differentiable and $w_-'(1-p)/w_-'(p) \geq 1$ for $p \in [0,1/2]$. Then it is optimal for a naïve gambler to continue in loss at $T-1$.*

A corollary of Proposition 5.3 is that the naïf will definitely continue even if there is only one round of play left as long as he is in loss, let alone when a longer horizon is allowed. As a consequence, he will not stop loss in any case, until the bitter end.

Note that the above analysis is based on the assumption that the reference point at any time $t>0$ is the same as the one at time 0, which is also assumed in Barberis (2012). One may wonder what if the reference point dynamically changes. The general problem of reference point updating and the resulting stopping behaviors for naïfs calls for a separate study and is beyond the scope of this paper. Here, let us consider as an example a simple yet natural rule of updating the reference point: At each node a naïve gambler uses the current wealth as the reference point. It follows that the problem he faces at node (t, S_t) is exactly the entry problem a precommitter faces in the shortened horizon $T-t$. As a result, the naïve gambler's strategy is state *independent*, and dependent of only the number of the *remaining* periods. In particular, the decision no longer depends on whether in the gain or loss regions.

We use the previous three groups of parameters to illustrate the resulting strategies. Let $T=10$. For $\alpha_\pm = 0.95$, $\delta_\pm = 0.5$, $\lambda = 1.5$, the naïve gambler continues (perhaps with randomization) up to time 8 and stops completely at time 9, whether in the gain region or the loss one; see Figure 10 left panel. For $\alpha_\pm = 0.5$, $\delta_\pm = 0.95$, $\lambda = 1.5$, the naïve gambler continues until up to time 6; see Figure 10 middle panel. For $\alpha_\pm = 0.5$, $\delta_\pm = 0.5$, $\lambda = 1.5$, the naïve gambler continues with randomization at $t=1$ and stops for sure at $t=2$; see Figure 10 right panel.

As predicted, all these strategies differ fundamentally from their counterparts for the case when the reference point does not change. In particular, they are neither gain-exit nor loss-exit, and the stopping decisions are identical across different states at the same time.

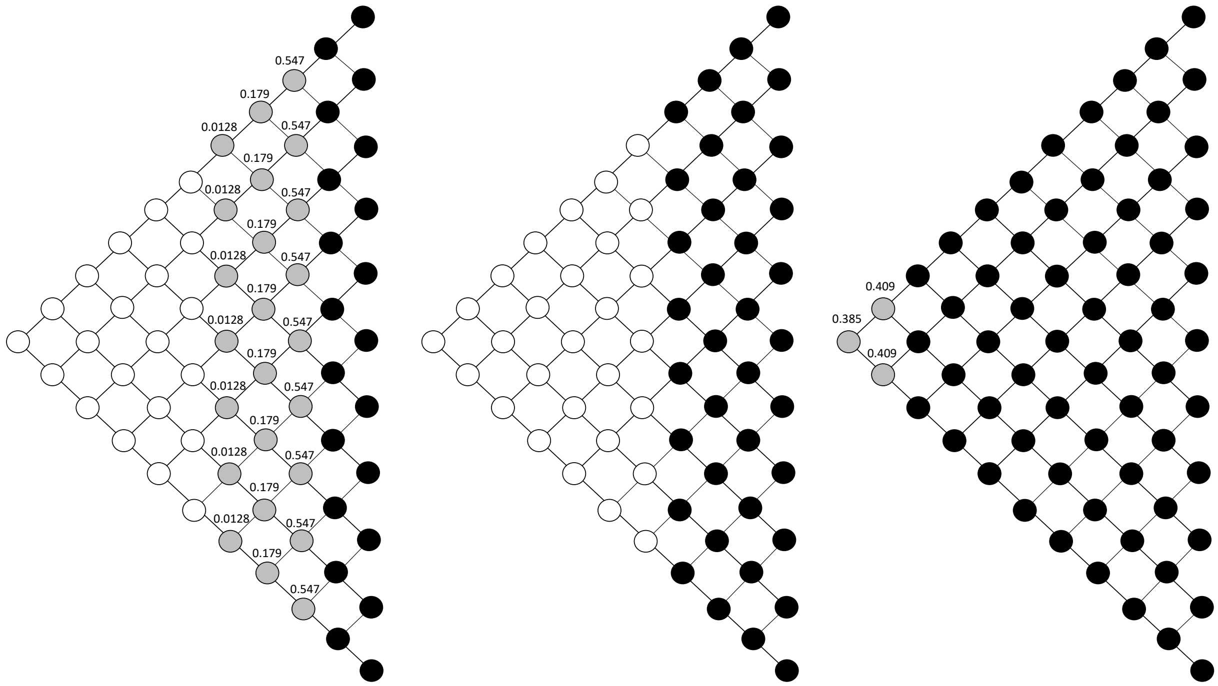

Figure 10 Naïve strategies with reference point updated to be the current wealth, for $T = 10$, $\alpha_\pm = 0.95$, $\delta_\pm = 0.5$, $\lambda = 1.5$ (left panel), $\alpha_\pm = 0.5$, $\delta_\pm = 0.95$, $\lambda = 1.5$ (middle panel), and $\alpha_\pm = 0.5$, $\delta_\pm = 0.5$, $\lambda = 1.5$ (right panel).

5.6. Sophisticated gamblers

A sophisticated gambler is unable to precommit and realizes that her future selves will deviate from whatever plans she makes now. Her resolution is to compromise and choose *consistent planning* in the sense that she optimizes taking the future disobedience as a constraint. Consequently, strategies of sophisticated gamblers can be obtained using backward induction as in dynamic programming. Same as before, randomization is allowed.

To start, we note that at $T-1$, a sophisticated gambler and a naïve one face the same problem; hence we have the following immediate result.

Proposition 5.4. *Propositions 5.2 and 5.3 hold true for a sophisticated gambler.*

Next, we derive a sophisticated gambler's stopping strategies for the four cases studied in Subsection 5.2, where $T = 10$. It turns out that, of the four cases, she will enter the casino *only* in the case when $(\alpha_\pm, \delta_\pm, \lambda) = (0.5, 0.95, 1.5)$, corresponding to Figure 7.[27] Moreover, her strategy is identical to the one depicted in the right panel of Figure 7, which is the actual strategy of the naïve gambler and close to the precommitted strategy. This is because when $\delta_\pm$ is close to 1, the level of

[27] The result about when the sophisticated gambler enters is consistent with Barberis (2012)'s results regarding the parameter range for which the sophisticate enters.

probability weighting is low, hence so is that of time-inconsistency, leading to similar strategies of all the three types of gamblers.

Note that in the case above, the sophisticated gambler takes the gain-exit type of strategy. Indeed, so long as she enters the casino, she essentially stops in gain under some mild conditions. This follows from the following argument: by Proposition 5.4, the sophisticated gambler will stop in gain at $T-1$. Knowing this, she will also stop in gain at $T-2$ by virtue of exactly the same reason. Inductively, this leads to an overall gain-exit type of strategy.

On the other hand, the sophisticated gambler always stops no later than her naïve counterpart does. This is because while the latter solves an optimal stopping problem at every node, the former solves the *same* problem but with constraints from her future selves' decisions. Hence, if the latter finds that stopping immediately is optimal at a current node, so will the former because the strategy of an immediate stop automatically satisfies the aforementioned constraints.

PROPOSITION 5.5. *Under any specification of parameters, a sophisticated gambler stops no later than a naïve gambler does.*

An implication of this result is that the naïf is at least as risk-taking as the sophisticated, if not more.

5.7. Finite horizon versus infinite horizon

This section explores connection between the finite horizon and infinite horizon casino models.

Define

$$\mathcal{T}_\infty := \{\tau \in [0,\infty) : \tau \text{ is an } (\mathcal{F}_t)_{t\in\mathbb{N}}\text{-stopping time}\},$$

which is the set of admissible stopping strategies (allowing randomization) in the infinite time horizon. Suppose $\tau \in \mathcal{T}_\infty$ is optimal for the infinite horizon model and achieves a finite CPT value. Then we have

$$V(S_{\tau\wedge T}) \le \sup_{\sigma\in\mathcal{T}_T} V(S_\sigma) \le V(S_\tau).$$

We see immediately that the value of the finite horizon model converges to that of the infinite horizon one as the horizon approaches infinity. The following makes this formal.

THEOREM 5.1. *Assume τ^* achieves the optimal value of the gambling model in the infinite time horizon with $\tau^* < \infty$ a.s., $V(S_{\tau^*}) = v^* < \infty$, and S_{τ^*} is lower-bounded a.s. Then*

$$\lim_{T\to\infty} \sup_{\tau\in\mathcal{T}_T} V(S_\tau) = \lim_{T\to\infty} V(S_{\tau^*\wedge T}) = v^*.$$

We stress that this result only reveals the relationship between the two models in terms of the optimal values. It does not offer a solution to the finite horizon problem (which is harder) from a solution to the infinite one (which is comparatively easier), nor does it tell the error in the optimal values when T is given and fixed. That said, the result suggests that the optimal value of the infinite horizon model is an upper bound of that of the finite horizon one, and it is a *tight* upper bound if T is sufficiently large. Moreover, while the truncation method mentioned earlier does not provide an *exact* optimal solution to the finite horizon model, it does nevertheless offer a *good* solution when T is large enough.

6. Conclusion

In this paper we develop a systematic approach to studying the stopping behaviors of CPT gamblers in a finite time horizon. We hope that this work opens an avenue of thoroughly understanding Barberis (2012)'s model and beyond. Indeed, as Barberis (2012) points out, casino gambling is not an isolated behavior requiring a unique treatment; rather, it is one of many phenomena, including ones in financial markets, that share a common driver of probability weighting.

References

Agranov M, Ortoleva P (2017) Stochastic choice and preferences for randomization. *Journal of Political Economy* 125(1):40–68.

Barberis N (2012) A model of casino gambling. *Management Science* 58(1):35–51.

Croson R, Sundali J (2005) The gambler's fallacy and the hot hand: Empirical data from casinos. *Journal of Risk and Uncertainty* 30(3):195–209.

Dwenger N, Kübler D, Weizsacker G (2013) Flipping a coin: Theory and evidence, URL http://ssrn.com/abstract=2353282., working Paper.

Ebert S, Strack P (2015) Until the bitter end: on prospect theory in a dynamic context. *American Economic Review* 105(4):1618 – 1633.

He XD, Hu S, Obłój J, Zhou XY (2017) Path-dependent and randomized strategies in barberis' casino gambling model. *Operations Research* 65(1):97–103.

He XD, Hu S, Obłój J, Zhou XY (2019a) Optimal exit time from casino gambling: Strategies of pre-committed and naive gamblers. *SIAM Journal on Control and Optimization* 57(3):1845–1868.

He XD, Hu S, Obłój J, Zhou XY (2019b) Two explicit skorokhod embeddings for simple symmetric random walk. *Stochastic Processes and their Applications* 129(9):3431–3435.

Heimer R, Iliewa Z, Imas A, Weber M (2020) Dynamic inconsistency in risky choice: Evidence from the lab and field, URL https://ssrn.com/abstract=3600583., working Paper.

Henderson V, Hobson D, Tse A (2017) Randomized strategies and prospect theory in a dynamic context. *Journal of Economic Theory* 168(3):287–300.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–291.

Obłój J (2004) The skorokhod embedding problem and its offspring. *Probability Surveys* 1:321–392.

Odean T (1998) Are investors reluctant to realize their losses. *Journal of Finance* 53(5):1775–1798.

Root DH (1969) The exitstence of certain stopping times on brownian motion. *The Annuals of Mathematical Statistics* 40(2):715–718.

Rost H (1976) Skorokhod stopping times of minimal variance. *Sḿinaire de Probabilitś X*, volume 511 of *Lecture Notes in Mathematics*, 194–208 (Springer).

Shiryaev A (1978) *Optimal Stopping Rules* (New York: Springer–Verlag).

Strotz R (1955) Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies* 23:165–180.

Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4):297–323.

Xu ZQ, Zhou XY (2012) Optimal stopping under probability distortion. *Annals of Applied Probability* 23(1):251–282.

Yong J, Zhou XY (1999) *Stochastic Controls: Hamiltonian Systems and HJB Equations* (New York: Springer).

Appendix

Appendix A: Proof of Theorem 3.1

We prove this theorem through a series of results. We start by recalling some properties of the potential and its link to the first exit times.

Proposition A.1. *Let τ be an $(\mathcal{F}_t)_{t\in\mathbb{N}}$-stopping time such that $\{S_{\tau\wedge t} : t \in \mathbb{N}\}$ is uniformly integrable. Then*

(i) *For any $t\in\mathbb{N}$, $U_{S_{\tau\wedge t}}$ is a convex function, $U_{S_\tau}(x) \geq U_{S_{\tau\wedge t}}(x) \geq |x|$ $\forall x\in\mathbb{R}$, with $U_{S_{\tau\wedge t}}(x) = |x|$ $\forall x \notin (-t,t)$.*

(ii) *For any two integers $a < b$ and $\rho := \inf\{u \geq \tau : S_u \notin (a,b)\}$, $U_{S_\rho}(x) = U_{S_\tau}(x)$ $\forall x \notin (a,b)$, and U_{S_ρ} is linear on $[a,b]$.*

(iii) *Fix $t\geq 1$ and let $\mathcal{K} := \{k\in\mathbb{Z} | k = t-1+2j,\ j\in\mathbb{Z}\}$. Then*

$$U_{S_{\tau\wedge t}}(x) = U_{S_{\tau\wedge(t-1)}}(x) + \mathbb{P}(S_{t-1} = x, \tau\geq t)\mathbf{1}_{x\in\mathcal{K}}\ \ \forall x\in\mathbb{Z}. \tag{14}$$

In particular, if t is odd, then $U_{S_{\tau\wedge t}}(x) = U_{S_{\tau\wedge(t-1)}}(x)$ for any odd x; and if t is even, then $U_{S_{\tau\wedge t}}(x) = U_{S_{\tau\wedge(t-1)}}(x)$ for any even x.

Proof The first two properties are standard; see Obłój (2004, Section 2). So we only establish (iii). Note that S_{t-1} is supported on $\mathcal{K}$. We have $|S_{\tau\wedge t}-S_{\tau\wedge(t-1)}|\leq 1$ so that $\{S_{\tau\wedge t}\geq x\}=\{S_{\tau\wedge(t-1)}\geq x\}$ $\forall x\notin\mathcal{K}$. In particular, since S is a martingale, we have $U_{S_{\tau\wedge t}}(x)=U_{S_{\tau\wedge(t-1)}}(x)$ $\forall x\notin\mathcal{K}$. Now take $x\in\mathcal{K}$. Since $x+1,x-1\notin\mathcal{K}$, using (5), we have

$$\begin{aligned}\mathbb{P}(S_{\tau\wedge t}=x)&=\frac{U_{S_{\tau\wedge t}}(x+1)+U_{S_{\tau\wedge t}}(x-1)}{2}-U_{S_{\tau\wedge t}}(x)\\&=\frac{U_{S_{\tau\wedge(t-1)}}(x+1)+U_{S_{\tau\wedge(t-1)}}(x-1)}{2}-U_{S_{\tau\wedge t}}(x)\\&=\mathbb{P}(S_{\tau\wedge(t-1)}=x)+U_{S_{\tau\wedge(t-1)}}(x)-U_{S_{\tau\wedge t}}(x).\end{aligned}\tag{15}$$

Rearranging and observing that $\mathbb{P}(S_t=x,\tau\geq t)=0$ the thesis follows. $\square$

The following proposition provides some useful properties of U_t.

Proposition A.2. *Let $\mu\in\mathcal{P}_0(\mathbb{Z})$ and $U_t=U_t^\mu$ be defined in (6). Then*

(i) $U_0(x)\leq U_t(x)\leq U_\mu(x)\wedge U_{S_t}(x)$ $\forall x\in\mathbb{Z}$, $t\in\mathbb{N}$.

(ii) *$U_t(x)=U_{t+1}(x)$ when t is odd and x is even, or when t is even and x is odd.*

(iii) *$U_t(x)$ is convex in $x\in\mathbb{R}$ and non-decreasing in $t\in\mathbb{N}$.*

Proof (i) By the construction of U_t we have $U_0(x)\leq U_t(x)\leq U_\mu(x)$. On other hand, by (5) and the structure of SSRW that $\mathbb{P}(S_t=x)=\big(\mathbb{P}(S_{t-1}=x-1)+\mathbb{P}(S_{t-1}=x+1)\big)/2$, one can show easily that $U_{S_t}(x)=\big(U_{S_{t-1}}(x-1)+U_{S_{t+1}}(x+1)\big)/2$, $x\in\mathbb{Z}$. Then by induction, we have $U_t(x)\leq U_{S_t}(x)$.

(ii) Again, by construction we have $U_0(x)=U_1(x)$ for all odd x and $U_1(x)=U_2(x)$ for all even x. The conclusions follow immediately from induction.

(iii) Clearly U_0 is convex. Suppose U_t is convex and fix $m\in\mathbb{Z}$. If we put $\tilde{U}(x)=U_t(x)$ for $x\in\mathbb{Z}\setminus\{m\}$, pick any

$$\tilde{U}(m)\in\left[U_t(m),\frac{1}{2}(U_t(m-1)+U_t(m+1))\right],$$

and finally define $\tilde{U}$ by a linear interpolation for $x\in\mathbb{R}$, then $\tilde{U}$ is convex. Observe that U_{t+1} is obtained exactly by repeating this procedure for all $m\in\mathbb{Z}$ and, hence, is also convex. Moreover, it now follows, by its definition, that $U_t(x)$ is non-decreasing in t. $\square$

Proposition A.3. *Let $T\geq 1$, $\mu\in\mathcal{P}_0(\mathbb{Z})$ such that $\mu([-T,T])=1$ and $U_t=U_t^\mu$ be defined in (6). Then, the following are equivalent:*

(i) $U_T(x)=U_\mu(x)$ $\forall x\in\mathbb{Z}$.

(ii) *There exists a randomized Root stopping time $\tau_R(\mathbf{b},\mathbf{r})$ such that $\tau_R(\mathbf{b},\mathbf{r})\leq T$ and $U_{S_{\tau_R(\mathbf{b},\mathbf{r})\wedge t}}=U_t$ $\forall t\leq T$; in particular $S_{\tau_R(\mathbf{b},\mathbf{r})}\sim\mu$.*

(iii) *There exists $\tau\in\mathcal{T}_T$ such that $S_\tau\sim\mu$.*

Furthermore, for any $\tau\in\mathcal{T}_T$ such that $S_\tau\sim\mu$ we have $U_{S_{\tau\wedge t}}(x)\leq U_t(x)$ $\forall x\in\mathbb{R}$, $t\leq T$.

Proof Proof of (i) $\Rightarrow$ (ii). To show the existence of a randomized Root stopping time embedding μ we first construct its stopping barrier $\mathbf{b}$. For $x\in\mathbb{Z}$, define

$$b(x):=\inf\{t\geq|x|:U_{t+1}(x)=U_\mu(x)\}.\tag{16}$$

It follows from Proposition A.2 that $b(x) = x + 2k$ for some $k \in \mathbb{Z}$. Next define the probabilities of the binary random variables $\{\xi_{t,x}\}$, $\mathbb{P}(\xi_{t,x} = 0) = 1 - \mathbb{P}(\xi_{t,x} = 1)$. For each $x \in \mathbb{Z}$,

$$\begin{cases} \mathbb{P}(\xi_{t,x} = 0) = 0 & \text{for } t < b(x)\,, \\ \mathbb{P}(\xi_{t,x} = 0) = r(x) := \frac{U_t(x-1)+U_t(x+1)-2U_\mu(x)}{U_t(x-1)+U_t(x+1)-2U_t(x)} & \text{for } t = b(x)\,, \\ \mathbb{P}(\xi_{t,x} = 0) = 1 & \text{for } t > b(x)\,. \end{cases} \tag{17}$$

Note that $r(x) = 1$ is only possible if $U_\mu(x) = |x|$ which happens for x outside of the support of μ. For other x we have $U_\mu(x) > |x|$ and a randomization, i.e., $0 < r(x) < 1$, happens at a node (t, x) when $t = b(x)$ and

$$\frac{U_t(x-1) + U_t(x+1)}{2} > U_\mu(x) > U_t(x)\,.$$

Let $\tau = \tau_R(\mathbf{b}, \mathbf{r})$ be the randomized Root stopping time in (7). By (i), $U_T \geq U_\mu$ and hence $b(x) \leq T - 1$ $\forall x \in \mathbb{Z} \cap (-T, T)$. It follows that $\tau \leq T$ as required.

To show $S_\tau \sim \mu$, we need only to establish $U_{S_\tau}(x) = U_\mu(x)$ $\forall x \in \mathbb{Z}$. Note $U_{S_0}(x) = U_{S_{\tau\wedge 0}}(x) = U_0(x)$. Suppose we have $U_{S_{\tau\wedge t}}(x) = U_t(x)$ for $t \leq n-1$. It follows from (5) that

$$\mathbb{P}(S_{\tau\wedge t} = x) = \frac{U_t(x+1) + U_t(x-1)}{2} - U_t(x), \;\; t \leq n-1.$$

On the other hand, by Proposition A.1, we have

$$U_{S_{\tau\wedge n}}(x) = U_{S_{\tau\wedge(n-1)}}(x) + \mathbb{P}(S_{n-1} = x, \tau \geq n)\mathbf{1}_{x\in\mathcal{K}} \;\forall x \in \mathbb{Z},$$

where $\mathcal{K} = \{k \in \mathbb{Z} | k = n - 1 + 2j, j \in \mathbb{Z}\}$.

If $U_{n-1}(x) = U_\mu(x) = U_n(x)$, then $b(x) < n-1$ and $\mathbb{P}(S_{n-1} = x, \tau \geq n) = 0$; hence

$$U_{S_{\tau\wedge n}}(x) = U_{S_{\tau\wedge(n-1)}}(x) = U_{n-1}(x) = U_\mu(x) = U_n(x)\,.$$

If $U_{n-1}(x) < U_\mu(x) = U_n(x)$, then $b(x) = n-1$ and necessarily $x \in \mathcal{K}$. We have, by definition,

$$\begin{aligned} \mathbb{P}(S_{n-1} = x, \tau \geq n) &= \mathbb{P}(S_{\tau\wedge(n-1)} = x)\mathbb{P}(\xi_{n-1,x} = 1) \\ &= \left(\frac{U_{n-1}(x+1) + U_{n-1}(x-1)}{2} - U_{n-1}(x)\right)\mathbb{P}(\xi_{n-1,x} = 1). \end{aligned}$$

It then follows that

$$\begin{aligned} U_{S_{\tau\wedge n}}(x) &= U_{S_{\tau\wedge(n-1)}}(x) + \mathbb{P}(S_{n-1} = x, \tau \geq n)\mathbf{1}_{x\in\mathcal{K}} \\ &= U_{n-1}(x)\mathbb{P}(\xi_{n-1,x} = 0) + \frac{U_{n-1}(x+1) + U_{n-1}(x-1)}{2}\mathbb{P}(\xi_{n-1,x} = 1) \\ &= U_{n-1}(x)\frac{U_{n-1}(x-1) + U_{n-1}(x+1) - 2U_\mu(x)}{U_{n-1}(x-1) + U_{n-1}(x+1) - 2U_{n-1}(x)} \\ &\quad + \frac{U_{n-1}(x+1) + U_{n-1}(x-1)}{2}\frac{2U_\mu(x) - 2U_{n-1}(x)}{U_{n-1}(x-1) + U_{n-1}(x+1) - 2U_{n-1}(x)} \\ &= U_\mu(x) = U_n(x)\,. \end{aligned}$$

Finally, if $U_n(x) < U_\mu(x)$, then $b(x) > n-1$. By definition, we have $U_n(x) = \frac{U_{n-1}(x+1)+U_{n-1}(x-1)}{2}$ and $\mathbb{P}(S_\tau = x) = \mathbb{P}(S_\tau = x, \tau \geq n)$. Consequently,

$$\mathbb{P}(S_{n-1} = x, \tau \geq n) = \mathbb{P}(S_{\tau\wedge(n-1)} = x) = \frac{U_{n-1}(x+1) + U_{n-1}(x-1)}{2} - U_{n-1}(x)\,.$$

Thus, if $x\in\mathcal{K}$, then

$$\begin{aligned}U_{S_{\tau\wedge n}}(x) &= U_{S_{\tau\wedge(n-1)}}(x)+\mathbb{P}(S_{n-1}=x,\tau\geq n)\mathbf{1}_{x\in\mathcal{K}}\\ &= U_{n-1}(x)+\frac{U_{n-1}(x+1)+U_{n-1}(x-1)}{2}-U_{n-1}(x)\\ &= \frac{U_{n-1}(x+1)+U_{n-1}(x-1)}{2}=U_n(x).\end{aligned}$$

If $x\notin\mathcal{K}$, then, noting that $\mathbb{P}(S_\tau=x,\tau<n)=0$, we have $\mathbb{P}(S_{\tau\wedge(n-1)}=x)=0$. As a result, $\frac{U_{n-1}(x+1)+U_{n-1}(x-1)}{2}=U_{n-1}(x)$ and

$$\begin{aligned}U_{S_{\tau\wedge n}}(x) &= U_{S_{\tau\wedge(n-1)}}(x)+\mathbb{P}(S_{n-1}=x,\tau\geq n)\mathbf{1}_{x\in\mathcal{K}}\\ &= U_{n-1}(x)=\frac{U_{n-1}(x+1)+U_{n-1}(x-1)}{2}=U_n(x)\,.\end{aligned}$$

In summary, $U_{S_{\tau\wedge n}}(x)=U_n(x)$ $\forall n\in\mathbb{Z}^+$. As a result, $U_{S_\tau}(x)=U_{S_{\tau\wedge T}}(x)=U_T(x)=U_\mu(x)$ $\forall x\in\mathbb{Z}$, namely, $S_\tau\sim\mu$.

Proof of (ii) $\Rightarrow$ (iii). This is trivial.

Proofs of (iii) $\Rightarrow$ (i) and the last assertion of the theorem. We start with the latter assuming (iii) holds.

Let $\tau\in\mathcal{T}_T$ such that $S_\tau\sim\mu$. Note that $U_{S_{\tau\wedge 0}}\equiv U_0$. Suppose $U_{S_{\tau\wedge t-1}}(x)\leq U_{t-1}(x)$ $\forall x$, for some $1\leq t\leq T$. Let $\tilde{S}_t=|S_{\tau\wedge t}-x|$, then $(\tilde{S}_t:t\geq 0)$ is a submartingale. Hence, $U_{S_{\tau\wedge 0}}(x)\leq\ldots\leq U_{S_{\tau\wedge(t-1)}}(x)\leq U_{S_{\tau\wedge t}}(x)\leq\ldots\leq U_{S_{\tau\wedge T}}(x)=U_\mu(x)$ $\forall x$. By (14), if $x\notin\mathcal{K}$, then $U_{S_{\tau\wedge t}}(x)=U_{S_{\tau\wedge(t-1)}}(x)\leq U_{t-1}(x)\leq U_t(x)$; if $x\in\mathcal{K}$ and $U_t(x)=U_\mu(x)$, then $U_{S_{\tau\wedge t}}(x)\leq U_\mu(x)=U_t(x)$; and if $x\in\mathcal{K}$ and $U_t(x)<U_\mu(x)$, then

$$\begin{aligned}U_{S_{\tau\wedge t}}(x) &\leq \frac{U_{S_{\tau\wedge t}}(x-1)+U_{S_{\tau\wedge t}}(x+1)}{2}\\ &= \frac{U_{S_{\tau\wedge(t-1)}}(x-1)+U_{S_{\tau\wedge(t-1)}}(x+1)}{2}\leq\frac{U_{t-1}(x-1)+U_{t-1}(x+1)}{2}=U_t(x)\,,\end{aligned}$$

where the first inequality is due to the convexity of $U_{S_{\tau\wedge t}}(\cdot)$, and the second equality is due to $x-1,x+1\notin\mathcal{K}$. This proves the last assertion of the theorem via an inductive argument in t. Next, taking $t=T$ and noting that $\tau\leq T$ we have $U_\mu=U_{S_{\tau\wedge T}}\leq U_T$ which shows (iii) $\Rightarrow$ (i). $\square$

We are now ready to prove Theorem 3.1. The "only if" part follows immediately from Proposition A.3-(i) and the construction of $U_T(x)$. To prove the "if" part, supposed (8) holds. First, we have $U_T(x)=U_\mu(x)=|x|$ for $|x|\geq T$. For $x=-(T-2),-(T-4),\ldots,T-4,T-2$, it follows from (8) that $U_T(x)=\frac{U_{T-1}(x+1)+U_{T-1}(x-1)}{2}\wedge U_\mu(x)=U_\mu(x)$. Next, by Proposition A.2, $U_T(x)=U_{T-1}(x)$ for all x with $x=T+2j$ for some $j\in\mathbb{Z}$. As a result, for $x=-(T-1),-(T-3),\ldots,T-3,T-1$, we have $U_\mu(x+1)=U_T(x+1)=U_{T-1}(x+1)$, $U_\mu(x-1)=U_T(x-1)=U_{T-1}(x-1)$, and, hence, $U_\mu(x)\leq\frac{U_\mu(x+1)+U_\mu(x-1)}{2}=\frac{U_{T-1}(x+1)+U_{T-1}(x-1)}{2}$, where the first inequality is due to the convexity of U_μ, and it follows that there exists the randomized Root stopping time that embeds μ in the random walk with finite time T. We conclude that $U_T(x)\geq U_\mu(x)$ $\forall x\in\mathbb{Z}$ and, hence, Proposition A.3 yields the desired result.

Appendix B: Proof of Proposition 5.1

Suppose at time 0, the gambler takes a randomized strategy with probability r of "stop" and probability $1-r$ of "continue", where $r \in [0,1]$. Let $q = (1-r)/2 \in [0,1/2]$. With utility function $u(x) = u_+(x)\mathbf{1}_{x\geq 0} - \lambda u_-(-x)\mathbf{1}_{x<0}$, the CPT value of this strategy is given by $u_+(1)w_+(q) - \lambda u_-(1)w_-(q)$, whose derivative in q is $u_+(1)w'_+(q) - \lambda u_-(1)w'_-(q)$. If follows from the assumption $\lim_{p\to 0}[w'_+(p)/w'_-(p)] > \lambda[u_-(1)/u_+(1)]$ that $u_+(1)w_+(q) - \lambda u_-(1)w_-(q)$ is *strictly* increasing in $q \in [0,\tilde{q}]$ for some $\tilde{q} \in (0,1/2]$. Hence, there exists $\bar{q} > 0$ such that $u_+(1)w_+(\bar{q}) - \lambda u_-(1)w_-(\bar{q}) > 0$.

Appendix C: Proof of Theorem 5.1

For any T, we have

$$V(S_{\tau^*\wedge T}) \leq \sup_{\tau\in\mathcal{T}_T} V(S_\tau) \leq V(S_{\tau^*}) = v^*.$$

Since S_{τ^*} is lower-bounded a.s., there exists $N > 0$ such that $S_{\tau^*} > -N$ a.s. For any $\epsilon > 0$, we can choose M large enough such that

$$\begin{aligned}&\sum_{n=1}^{M} u_+(n)\left(w_+\left(\mathbb{P}(S_{\tau^*} \geq n)\right) - w_+\left(\mathbb{P}(S_{\tau^*} \geq n+1)\right)\right)\\ &\quad - \lambda\sum_{n=1}^{N} u_-(n)\left(w_-\left(\mathbb{P}(S_{\tau^*} \leq -n)\right) - w_-\left(\mathbb{P}(S_{\tau^*} \leq -n-1)\right)\right) =: \tilde{v} > v^* - \epsilon/2.\end{aligned}$$

On the other hand, since τ^* is finite a.s., the distribution of $S_{\tau^*\wedge T}$ converges to that of S_{τ^*}. Then there is sufficiently large T such that

$$\begin{aligned}V(S_{\tau^*\wedge T}) \geq &\sum_{n=1}^{M} u_+(n)\left(w_+\left(\mathbb{P}(S_{\tau^*\wedge T} \geq n)\right) - w_+\left(\mathbb{P}(S_{\tau^*\wedge T} \geq n+1)\right)\right)\\ &\quad - \lambda\sum_{n=1}^{N} u_-(n)\left(w_-\left(\mathbb{P}(S_{\tau^*\wedge T} \leq -n)\right) - w_-\left(\mathbb{P}(S_{\tau^*\wedge T} \leq -n-1)\right)\right)\\ > &\,\tilde{v} - \epsilon/2 > v^* - \epsilon.\end{aligned}$$

This establishes the desired result.

Acknowledgments

The first author would like to acknowledge the funding of National Natural Science of China (Grant No. 11901494) and Natural Science of Guangdong Province (Grant No. 2019A1515011396). Part of this research was completed whilst the second author was visiting CUHK and he is grateful for the support from the host. He also gratefully acknowledges support from ERC Starting Grant RobustFinMath 335421. The third author gratefully acknowledges financial supports through start-up grants at both University of Oxford and Columbia University, and through Oxford–Nie Lab for Financial Big Data and the Nie Center for Intelligent Asset Management. The authors would like to thank the editors and three anonymous reviewers for their most helpful comments and suggestions on the earlier versions of this manuscript. We also thank for comments and questions from participants at the 2020 INFORMS annual meeting, the Financial Engineering and Risk Management Seminar of Society of Operations Research of China, and the Financial Engineering workshop at SUST.