Yale ICF Working Paper No. 12-33

A Model of Casino Gambling

Nicholas Barberis
Yale University

March 2009

A Model of Casino Gambling

Nicholas Barberis
Yale University

March 2009*

Abstract

Casino gambling is a hugely popular activity around the world, but there are still very few models of why people go to casinos or of how they behave when they get there. In this paper, we show that prospect theory can offer a surprisingly rich theory of gambling, one that captures many features of actual gambling behavior. First, we demonstrate that, for a wide range of parameter values, a prospect theory agent would be willing to gamble in a casino, even if the casino only offers bets with zero or negative expected value. Second, we show that prospect theory predicts a plausible time inconsistency: at the moment he enters a casino, a prospect theory agent plans to follow one particular gambling strategy; but *after* he enters, he wants to switch to a different strategy. The model therefore predicts heterogeneity in gambling behavior: how a gambler behaves depends on whether he is aware of the time-inconsistency; and, if he *is* aware of it, on whether he is able to commit, in advance, to his initial plan of action.

JEL classification: D03, D81

Keywords: gambling, prospect theory, time inconsistency, probability weighting

*I thank Isaiah Andrews, Yaxin Duan, Keith Gamble, and Bjorn Johnson for excellent research assistance and Daniel Benjamin, Keith Chen, James Choi, Stefano Della Vigna, Ed Glaeser, Daniel Gottlieb, Botond Koszegi, David Laibson, Ulrike Malmendier, Ted O'Donoghue, Matthew Rabin, Andrei Shleifer, Richard Thaler, and seminar participants at Cornell University, Harvard University, Princeton University, Stanford University, the Stockholm School of Economics, the University of California at Berkeley, Yale University, and the NBER for very helpful feedback. Comments are welcome at nick.barberis@yale.edu.

1 Introduction

Casino gambling is a hugely popular activity. The American Gaming Association reports that, in 2007, 54 million people made 376 million trips to casinos in the United States alone. U.S. casino revenues that year totalled almost $60 billion.

In order to fully understand how people think about risk, we need to make sense of the existence and popularity of casino gambling. Unfortunately, there are still very few models of why people go to casinos or of how they behave when they get there. The challenge is clear. The standard economic model of risk attitudes couples the expected utility framework with a concave utility function. This model is helpful for understanding a range of phenomena. It cannot, however, explain casino gambling: an agent with a concave utility function will always turn down a wealth bet with a negative expected value.

While casino gambling is hard to reconcile with the standard model of risk attitudes, researchers have made some progress in understanding it better. One approach is to introduce non-concave segments into the utility function. A second approach argues that people derive a separate component of utility from gambling. This utility may be only indirectly related to the bets themselves – for example, it may stem from the social pleasure of going to a casino with friends; or it may be directly related to the bets, in that the gambler enjoys the feeling of suspense as he waits for the bets to play out (see Conlisk (1993) for a model of this last idea). A third approach suggests that gamblers simply overestimate their ability to predict the outcome of a bet; in short, they think that the odds are more favorable than they actually are.

In this paper, we present a new model of casino gambling based on Tversky and Kahneman's (1992) cumulative prospect theory. Cumulative prospect theory, one of the most prominent theories of decision-making under risk, is a modified version of Kahneman and Tversky's (1979) prospect theory. It posits that people evaluate risk using a value function that is defined over gains and losses, that is concave over gains and convex over losses, and that is kinked at the origin, so that people are more sensitive to losses than to gains, a feature known as loss aversion. It also posits that people use *transformed* rather than objective probabilities, where the transformed probabilities are obtained from objective probabilities by applying a weighting function. The main effect of the weighting function is to overweight the tails of the distribution it is applied to. The overweighting of tails does not represent a bias in beliefs; it is simply a modeling device for capturing the common preference for a lottery-like, or positively skewed, wealth distribution.

We choose prospect theory as the basis for a possible explanation of casino gambling because we would like to understand gambling in a framework that also explains *other* evidence on risk attitudes. Prospect theory can explain a wide range of experimental evidence on attitudes to risk – indeed, it was designed to – and it can also shed light on much *field* evidence on risk-taking: for example, it can address a number of facts about risk premia in asset markets (Benartzi and Thaler, 1995; Barberis and Huang, 2008). By offering a prospect theory model of casino gambling, our paper suggests that gambling is not necessarily an isolated phenomenon requiring its own unique explanation, but rather one of a family of facts that can be understood using a single model of risk attitudes.

The idea that prospect theory might explain casino gambling is initially surprising. Through the overweighting of the tails of distributions, prospect theory can easily explain why people buy lottery tickets. Casinos, however, offer gambles that, aside from their low expected values, are also much less skewed than a lottery ticket. Since prospect theory agents are more sensitive to losses than to gains, one would think that they would find these gambles very unappealing. Initially, then, prospect theory does not seem to be a promising starting point for a model of casino gambling. Indeed, it has long been thought that gambling is the one major risk-taking phenomenon that prospect theory is *not* well-suited to explain.

In this paper, we show that, in fact, prospect theory can offer a rich theory of casino gambling, one that captures many features of actual gambling behavior. First, we demonstrate that, for a wide range of preference parameter values, a prospect theory agent *would* be willing to gamble in a casino, even if the casino only offers bets with zero or negative expected value. Second, we show that prospect theory – in particular, its probability weighting feature – predicts a plausible *time inconsistency*: at the moment he enters a casino, a prospect theory agent plans to follow one particular gambling strategy; but *after* he enters, he wants to switch to a different strategy. How a gambler behaves therefore depends on whether he is aware of this time inconsistency; and, if he *is* aware of it, on whether he is able to commit in advance to his initial plan of action.

What is the intuition for why, in spite of loss aversion, a prospect theory agent might still be willing to enter a casino? Consider a casino that offers only zero expected value bets – specifically, 50:50 bets to win or lose some fixed amount $\$h$ – and suppose that the agent makes decisions by maximizing the cumulative prospect theory utility of his accumulated winnings or losses at the moment he leaves the casino. We show that, if the agent enters the casino, his preferred plan is to gamble as long as possible if he is winning, but to stop gambling and leave the casino if he starts accumulating losses. An important property of this plan is that, even though the casino offers only 50:50 bets, the distribution of the agent's perceived *overall* casino winnings becomes positively skewed: by stopping once he starts accumulating losses, the agent limits his downside; and by continuing to gamble when he is winning, he retains substantial upside.

At this point, the probability weighting feature of prospect theory plays an important role. Under probability weighting, the agent overweights the tails of probability distributions. With sufficient probability weighting, then, the agent may *like* the positively skewed distribution generated by his planned gambling strategy. We show that, for a wide range of parameter values, the probability weighting effect indeed outweighs the loss-aversion effect and the agent *is* willing to enter the casino. In other words, while the prospect theory agent would always turn down the basic 50:50 bet if it were offered *in isolation*, he is nonetheless willing to enter the casino because, through a clever choice of exit strategy, he gives his overall casino experience a positively skewed distribution, one which, with sufficient probability weighting, he finds attractive.

Prospect theory offers more than just an explanation of why people go to casinos. Through the probability weighting function, it also predicts a time inconsistency. At the moment he enters a casino, the agent's preferred plan is to keep gambling if he is winning but to stop gambling if he starts accumulating losses. We show, however, that once he starts gambling, he wants to do the opposite: to keep gambling if he is losing and to stop gambling if he accumulates a significant gain.

As a result of this time inconsistency, our model predicts significant heterogeneity in gambling behavior. How a gambler behaves depends on whether he is aware of the time inconsistency. A gambler who *is* aware of the time inconsistency has an incentive to try to commit to his initial plan of action. For gamblers who are aware of the time inconsistency, then, their behavior further depends on whether they are indeed able to find a commitment device.

To study these distinctions, we consider three types of agents. The first type is "naive": he is unaware that he will exhibit a time inconsistency. This gambler *plans* to keep gambling as long as possible if he is winning and to exit only if he starts accumulating losses. After entering the casino, however, he deviates from this plan and instead gambles as long as possible when he is losing and stops only after making some gains.

The second type of agent is "sophisticated" but unable to commit: he recognizes that, if he enters the casino, he will deviate from his initial plan; but he is unable to find a way of committing to his initial plan. He therefore knows that, if he enters the casino, he will keep gambling when he is losing and will stop gambling after making some gains, a strategy that will give his overall casino experience a *negatively* skewed distribution. Since he overweights the tails of probability distributions, he finds this unattractive and therefore refuses to enter the casino in the first place.

The third type of agent is sophisticated and able to commit: he also recognizes that, if he enters the casino, he will want to deviate from his initial plan; but he is able to find a way of committing to his initial plan. Just like the naive agent then, this agent plans, on entering the casino, to keep gambling as long as possible when winning and to exit only if he starts accumulating losses. Unlike the naive agent, however, he is able, through the use of a commitment device, to stick to this plan. For example, he may bring only a small amount of cash to the casino while also leaving his ATM card at home; this guarantees that he will indeed leave the casino if he starts accumulating losses. According to our model, we should observe some actual gamblers behaving in this way. Anecdotally, at least, some gamblers do use techniques of this kind.

In summary, under the view proposed in this paper, casinos are popular because they cater to two aspects of our psychological make-up. First, they cater to the tendency to overweight the tails of distributions, which makes even the small chance of a large win at the casino seem very alluring. And second, they cater to what we could call "naivete," namely the failure to recognize that, after entering a casino, we may deviate from our initial plan of action.

Our model is a complement to existing theories of gambling, not a replacement. The popularity of casinos is probably driven by many factors, and we suspect that some of the factors that have already been mentioned in the literature – the utility of gambling, for example, and the misperception of casino odds – play at least as large a role as prospect theory.

At the same time, we think that prospect theory can add significantly to our understanding of casino gambling. As noted above, one attractive feature of the prospect theory approach is that it not only explains why people go to casinos, but also offers a rich description of what they do once they get there. Moreover, it explains a number of features of casino gambling that have not emerged from earlier models: for example, the tendency to gamble longer than planned in the region of losses, the strategy of leaving one's ATM card at home, and casinos' practice of issuing free vouchers to people who are winning. Finally, our approach shows that we can understand casino gambling in the context of a model – cumulative prospect theory – that already explains a range of *other* evidence on attitudes to risk.

In recent years, there has been a surge of interest in the time inconsistency that stems from hyperbolic discounting.[1] While it has long been understood that probability weighting can also lead to a time inconsistency, there has been very little analysis of this idea. In this paper, we make the case that this second type of inconsistency may also be important in practice. While casino gambling is its most obvious application, it may also play a significant role in other contexts. For example, in Section 5, we briefly outline an application to stock market trading.

In Section 2, we review both prospect theory and cumulative prospect theory. In Section 3, we present a model of casino gambling. Section 4 discusses the model further and Section 5 presents an application to stock market trading. Section 6 concludes.

[1]See, for example, Laibson (1997), O'Donoghue and Rabin (1999), Della Vigna and Malmendier (2006), and the references therein.

2 Cumulative Prospect Theory

In Section 3, we present a model of casino gambling in which agents evaluate risk in the way described by cumulative prospect theory. In this section, we introduce cumulative prospect theory after first reviewing the original version of prospect theory, due to Kahneman and Tversky (1979). Readers who are already familiar with cumulative prospect theory may prefer to jump directly to Section 3.

Consider the gamble

$$(x, p; y, q), \tag{1}$$

to be read as "gain x with probability p and y with probability q, independent of other risks," where $x \leq 0 \leq y$ or $y \leq 0 \leq x$, and where $p + q = 1$. In the expected utility framework, an agent with utility function $U(\cdot)$ evaluates this gamble by computing

$$pU(W + x) + qU(W + y), \tag{2}$$

where W is his current wealth. In the original version of prospect theory, the agent assigns the gamble the value

$$\pi(p)v(x) + \pi(q)v(y), \tag{3}$$

where $v(\cdot)$ and $\pi(\cdot)$ are known as the value function and the probability weighting function, respectively. Figure 1 shows the forms of $v(\cdot)$ and $\pi(\cdot)$ suggested by Kahneman and Tversky (1979). The functions satisfy $v(0) = 0$, $\pi(0) = 0$, and $\pi(1) = 1$.

There are four important differences between (2) and (3). First, the carriers of value in prospect theory are gains and losses, not final wealth levels: the argument of $v(\cdot)$ in (3) is x, not $W + x$. Second, while $U(\cdot)$ is typically concave everywhere, $v(\cdot)$ is concave only over gains; over losses, it is convex. This captures the experimental finding that people tend to be risk averse over moderate-probability gains – they prefer a certain gain of \$500 to (\$1000, $\frac{1}{2}$) – but risk-*seeking* over moderate-probability losses, in that they prefer ($-$\$1000, $\frac{1}{2}$) to a certain loss of \$500.[2]

Third, while $U(\cdot)$ is typically differentiable everywhere, the value function $v(\cdot)$ is kinked at the origin so that the agent is more sensitive to losses – even small losses – than to gains of the same magnitude. As noted in the Introduction, this element of prospect theory is known as loss aversion. Kahneman and Tversky (1979) infer it from the widespread aversion to bets such as (\$110, $\frac{1}{2}$; $-$\$100, $\frac{1}{2}$).

Finally, under prospect theory, the agent does not use objective probabilities when evaluating the gamble, but rather, transformed probabilities obtained from objective probabilities via the probability weighting function $\pi(\cdot)$. The most important feature of this function is

[2] We abbreviate $(x, p; 0, q)$ to (x, p).

that low probabilities are overweighted: in the right panel of Figure 1, the solid line lies above the 45-degree dotted line for low p. This is inferred from subjects' preference for $(\$5000, 0.001)$ over a certain \$5, and from their preference for a certain loss of \$5 over $(-\$5000, 0.001)$; in other words, it is inferred from their simultaneous demand for both lotteries and insurance. Spelling this out in more detail,

$$\begin{aligned} (\$5, 1) &\prec (\$5000, 0.001) \\ &\Rightarrow v(5)\pi(1) < v(5000)\pi(0.001) < 1000\, v(5)\pi(0.001) \\ &\Rightarrow \pi(0.001) > 0.001, \end{aligned} \tag{4}$$

so that low probabilities are overweighted. A similar calculation in the case of the $(-\$5000, 0.001)$ gamble, using the fact that $v(\cdot)$ is convex over losses, produces the same result.

The transformed probabilities $\pi(p)$ and $\pi(q)$ in (3) should not be thought of as beliefs, but as decision weights which help us capture the experimental evidence on risk attitudes. In Kahneman and Tversky's (1979) framework, an agent evaluating the lottery-like $(\$5000, 0.001)$ gamble understands that he will only receive the \$5000 with probability 0.001. The overweighting of 0.001 introduced by prospect theory is simply a modeling device which captures the agent's preference for the lottery over a certain \$5.

In this paper, we do not work with the original prospect theory, but with a modified version, cumulative prospect theory, proposed by Tversky and Kahneman (1992). In this modified version, Tversky and Kahneman (1992) suggest explicit functional forms for $v(\cdot)$ and $\pi(\cdot)$. Moreover, they apply the probability weighting function to the *cumulative* probability distribution, not to the probability density function. This ensures that cumulative prospect theory does not violate first-order stochastic dominance – a weakness of the original prospect theory – and also that it can be applied to gambles with any number of outcomes, not just two. Finally, Tversky and Kahneman (1992) allow the probability weighting functions for gains and losses to differ.

Formally, under cumulative prospect theory, the agent evaluates the gamble

$$(x_{-m}, p_{-m}; \ldots; x_{-1}, p_{-1}; x_0, p_0; x_1, p_1; \ldots; x_n, p_n), \tag{5}$$

where $x_i < x_j$ for $i < j$ and $x_0 = 0$, by assigning it the value

$$\sum_{i=-m}^{n} \pi_i v(x_i), \tag{6}$$

where[3]

$$\pi_i = \begin{cases} w^+(p_i + \ldots + p_n) - w^+(p_{i+1} + \ldots + p_n) & \text{for} \quad 0 \le i \le n \\ w^-(p_{-m} + \ldots + p_i) - w^-(p_{-m} + \ldots + p_{i-1}) & \phantom{\text{for}} \quad -m \le i < 0 \end{cases} \tag{7}$$

[3]When $i = n$ and $i = -m$, equation (7) reduces to $\pi_n = w^+(p_n)$ and $\pi_{-m} = w^-(p_{-m})$, respectively.

and where $w^+(\cdot)$ and $w^-(\cdot)$ are the probability weighting functions for gains and losses, respectively. Tversky and Kahneman (1992) propose the functional forms

$$v(x) = \begin{cases} x^\alpha & \text{for} \quad x \geq 0 \\ -\lambda(-x)^\alpha & \phantom{\text{for}} \quad x < 0 \end{cases} \tag{8}$$

and

$$w^+(P) = \frac{P^\gamma}{(P^\gamma + (1-P)^\gamma)^{1/\gamma}}, \quad w^-(P) = \frac{P^\delta}{(P^\delta + (1-P)^\delta)^{1/\delta}}, \tag{9}$$

where α, γ, $\delta \in (0,1)$ and $\lambda > 1$.

For $\alpha \in (0,1)$ and $\lambda > 1$, the value function $v(\cdot)$ in (8) captures the features highlighted earlier: it is concave over gains, convex over losses, and exhibits a greater sensitivity to losses than to gains. The degree of sensitivity to losses is determined by λ, the coefficient of loss aversion. For $\gamma, \delta \in (0,1)$, the weighting functions $w^+(\cdot)$ and $w^-(\cdot)$ in (9) capture the overweighting of low probabilities described earlier: for low, positive P, $w^-(P) > P$ and $w^+(P) > P$.

Equation (7) shows that, under cumulative prospect theory, the weighting function is applied to the cumulative probability distribution. If it were instead applied to the probability density function, as in the original prospect theory, the probability weight π_i, for $i < 0$ say, would be $w^-(p_i)$. Instead, equation (7) shows that, under cumulative prospect theory, the probability weight π_i is obtained by taking the total probability of all outcomes equal to or worse than x_i, namely $p_{-m} + \ldots + p_i$, the total probability of all outcomes strictly worse than x_i, namely $p_{-m} + \ldots + p_{i-1}$, applying the weighting function to each, and computing the difference.

The effect of applying the weighting function to a *cumulative* probability distribution is to make the agent overweight the *tails* of that distribution. In equations (6)-(7), the most extreme outcomes, x_{-m} and x_n, are assigned the probability weights $w^-(p_{-m})$ and $w^+(p_n)$, respectively. If p_{-m} and p_n are small, we then have $w^-(p_{-m}) > p_{-m}$ and $w^+(p_n) > p_n$. The most extreme outcomes – the outcomes in the tails – are therefore overweighted. Just as in the original prospect theory, then, a cumulative prospect theory agent likes positively skewed, or lottery-like, wealth distributions. This will play an important role in our analysis.

Using experimental data, Tversky and Kahneman (1992) estimate $\alpha = 0.88$, $\lambda = 2.25$, $\gamma = 0.61$, and $\delta = 0.69$ for their median subject. Since the estimates of γ and δ are similar, we set $\gamma = \delta$ for simplicity, so that

$$w^+(P) = w^-(P) \equiv w(P) = \frac{P^\delta}{(P^\delta + (1-P)^\delta)^{1/\delta}}. \tag{10}$$

To ensure the monotonicity of $w(\cdot)$, we require $\delta \in (0.28, 1)$.

Figure 2 plots the weighting function $w(\cdot)$ in (10) for $\delta = 0.65$ (the dashed line), for $\delta = 0.4$ (the dash-dot line), and for $\delta = 1$, which corresponds to no probability weighting at all (the solid line). The overweighting of low probabilities is clearly visible for $\delta < 1$.

3 A Model of Casino Gambling

In the United States, the term "gambling" typically refers to one of four things: (i) casino gambling, of which the most popular forms are slot machines and the card game of blackjack; (ii) the buying of lottery tickets; (iii) pari-mutuel betting on horses at racetracks; and (iv) fixed-odds betting through bookmakers on sports such as football, baseball, basketball, and hockey – a form of gambling that is legal only in Nevada. The American Gaming Association estimates the 2007 revenues from each type of gambling at \$59 billion, \$24 billion, \$4 billion, and \$200 million, respectively.[4]

While the four types of gambling listed above have some common characteristics, they also differ in some ways. Casino gambling differs from playing the lottery in that the payoff of a casino game is typically much less positively skewed than that of a lottery ticket. And it differs from racetrack-betting and sports-betting in that casino games usually require less skill: while some casino games have an element of skill, many are purely games of chance.

In this paper, we focus our attention on casino gambling, largely because, from the perspective of prospect theory, it is the hardest to explain. The buying of lottery tickets is already directly captured by prospect theory through the overweighting of low probabilities. Casino games are much less skewed than a lottery ticket, however. It is therefore not at all clear that we can use the overweighting of low probabilities to explain the popularity of casinos. Meanwhile, many authors have suggested that the popularity of racetrack-betting and sports-betting stems from the bettors' belief that they are informed about the sporting event in question.

We model a casino in the following way. There are $T + 1$ dates, $t = 0, 1, \ldots, T$. At time 0, the casino offers the agent a 50:50 bet to win or lose a fixed amount \$$h$. If the agent turns the gamble down, the game is over: he is offered no more gambles and we say that he has declined to enter the casino. If the agent *accepts* the 50:50 bet, we say that he has agreed to enter the casino. The gamble is then played out and, at time 1, the outcome is announced. At that time, the casino offers the agent another 50:50 bet to win or lose \$$h$. If he turns it down, the game is over: the agent settles his account and leaves the casino. If he *accepts* the gamble, it is played out and, at time 2, the outcome is announced. The game then continues

[4]The \$200 million figure refers to sports-betting through *legal* bookmakers. It it widely believed that this figure is dwarfed by the revenues from illegal sports-betting. Also excluded from these figures are the revenues from online gambling.

in the same way. If, at time $t \in [0, T-2]$, the agent agrees to play a 50:50 bet to win or lose \$$h$, then, at time $t+1$, he is offered another such bet and must either accept it or decline it. If he declines it, the game is over: he settles his account and leaves the casino. At time T, the agent *must* leave the casino if he has not already done so. We think of the interval from 0 to T as an evening of play at a casino.

By assuming an exogeneous date, date T, at which the agent must leave the casino if he has not already done so, we make our model somewhat easier to solve. This is not, however, the reason we impose the assumption. Rather, we impose it because we think that it makes the model more realistic: whether because of fatigue or because of work and family commitments, most people simply cannot stay in a casino indefinitely.

Of the major casino games, our model most closely resembles blackjack: under optimal play, the odds of winning a round of blackjack are close to 0.5, which matches the 50:50 bet offered by our casino. Slot machines offer a positively skewed payoff and therefore, at first sight, do not appear to fit the model as neatly. In Section 4, however, we argue that the model may be able to shed as much light on slot machines as it does on blackjack.

In the discussion that follows, it will be helpful to think of the casino as a binomial tree. Figure 3 illustrates this for $T = 5$. Each column of nodes corresponds to a particular time: the left-most node corresponds to time 0 and the right-most column to time T. The various nodes within a column represent the different possible accumulated winnings or losses at that time. At time 0, then, the agent starts in the left-most node. If he takes the time 0 bet and wins, he moves one step *up* and to the right; if he takes the time 0 bet and loses, he moves one step *down* and to the right, and so on. Whenever the agent wins a bet, he moves up a step in the tree, and whenever he loses, he moves down a step.

We refer to the nodes in the tree by a pair of numbers (t, j). The first number, t, which ranges from 0 to T, indicates the time that the node corresponds to. The second number, j, which, for given t, can range from 1 to $t+1$, indicates how far down the node is within the column of $t+1$ nodes for that time: the highest node in the column corresponds to $j = 1$ and the lowest node to $j = t+1$. The left-most node in the tree is therefore node $(0, 1)$. The two nodes in the column immediately to the right, starting from the top, are nodes $(1, 1)$ and $(1, 2)$; and so on.

Throughout the paper, we use a simple color scheme to represent the agent's behavior. If a node is colored white, this means that, at that node, the agent agrees to play a 50:50 bet. If the node is black, this means that the agent does *not* play a 50:50 bet at that node, either because he leaves the casino when he arrives at that node, or because he has already left the casino in an earlier round and therefore never even reaches the node. For example, the interpretation of Figure 3 is that the agent agrees to enter the casino at time 0 and then keeps gambling until time $T = 5$ or until he hits node $(3, 1)$, whichever comes first. Clearly, a node that can only be reached by passing through a black node must itself be black. In Figure 3, the fact that node $(3, 1)$ has a black color immediately implies that node $(4, 1)$ must also have a black color.

As noted above, the basic gamble offered by the casino in our model is a 50:50 bet to win or lose $\$h$. We assume that the gain and the loss are equally likely only because this simplifies the exposition, not because it is necessary for our analysis. In fact, our analysis can easily be extended to the case in which the probability of winning $\$h$ is different from 0.5. Indeed, we find that the results we obtain below continue to hold even if, as in actual casinos, the basic gamble has a slightly *negative* expected value: even if it entails a 0.48 chance of winning $\$h$, say, and a 0.52 chance of losing $\$h$. We discuss this issue again in Section 4.1.

Now that we have described the structure of the casino, we are ready to present the behavioral assumption that drives our analysis. Specifically, we assume that the agent in our model *maximizes the cumulative prospect theory utility of his accumulated winnings or losses at the moment he leaves the casino*, where the cumulative prospect theory value of a distribution is given by (6)-(8) and (10). In making this assumption, we recognize that we are almost certainly leaving out other factors that also affect the agent's decision-making. Nonetheless, we hope to show in this and subsequent sections that our assumption is not only parsimonious but also leads to a rich theory of gambling.

Our behavioral assumption immediately raises an important issue, one that plays a central role in our analysis. This is the fact that cumulative prospect theory – in particular, its probability weighting feature – introduces a time inconsistency: the agent's *plan*, at time t, as to what he would do if he reached some later node is not necessarily what he actually does when he reaches that node.

To see this, consider the following example with $T = 5$ and $h = \$10$. Suppose that, at time 0, the agent is trying to decide between two exit strategies. Under exit strategy A, shown in the left panel of Figure 4, he would leave the casino only at the last date, $T = 5$. Under exit strategy B, shown in the right panel of Figure 4, he would leave the casino only at the last date *or* in node $(4, 1)$, whichever comes first. We now show that, from the perspective of time 0, the agent prefers strategy A, while, from the perspective of time 4, he prefers strategy B. In other words, there is a time inconsistency: at time 0, the agent *plans* to gamble in node $(4, 1)$; but if he actually reaches that node, he instead leaves the casino.

Which of the two exit strategies offers higher utility from the perspective of time 0? Under exit strategy A, the accumulated win or loss at the moment the agent leaves the casino has the distribution

$$(\$50, \frac{1}{32}; \$30, \frac{5}{32}; \$10, \frac{10}{32}; -\$10, \frac{10}{32}; -\$30, \frac{5}{32}; -\$50, \frac{1}{32}).$$

Under exit strategy B, the accumulated win or loss at the moment the agent leaves the casino has the distribution

$$(\$40, \frac{1}{16}; \$30, \frac{4}{32}; \$10, \frac{10}{32}; -\$10, \frac{10}{32}; -\$30, \frac{5}{32}; -\$50, \frac{1}{32}).$$

The two strategies differ only in the size and the probability of the two highest winnings that they offer. This means that, to see which of the two strategies the agent prefers at time 0, we need only look at the contribution of the two highest winnings to total utility.

Applying (6)-(8) and (10), the contribution of the two highest potential winnings to the cumulative prospect theory utility of strategy A, as evaluated at time 0, is

$$v(50)w(\frac{1}{32}) + v(30)\left[w(\frac{5}{32} + \frac{1}{32}) - w(\frac{1}{32})\right]. \tag{11}$$

The contribution of the two highest potential winnings to the cumulative prospect theory utility of strategy B, as evaluated at time 0, is

$$v(40)w(\frac{1}{16}) + v(30)\left[w(\frac{4}{32} + \frac{1}{16}) - w(\frac{1}{16})\right]. \tag{12}$$

The agent therefore prefers strategy A – in other words, from the perspective of time 0, he would prefer to keep gambling after arriving at node $(4, 1)$ – if

$$(v(50) - v(30))w(\frac{1}{32}) > (v(40) - v(30))w(\frac{1}{16}). \tag{13}$$

Since the argument of $v(\cdot)$ in this condition is always a gain, the condition depends only on α, which governs the concavity of the value function in the region of gains, and on δ, which controls the degree of probability weighting. It does not depend on the degree of loss aversion λ.

The shaded area in Figure 5 shows the range of values of α and δ for which condition (13) holds. The figure shows that, for the vast majority of values of α and δ, the agent, *from the perspective of time 0*, prefers exit strategy A: in other words, he would prefer to keep gambling in node $(4, 1)$. The intuition is that, from the perspective of time 0, the agent is keen to give himself the chance of reaching node $(5, 1)$: although the \$50 prize in that node has low probability, namely $\frac{1}{32}$, this low probability is overweighted under cumulative prospect theory, making the node very appealing to the agent. But in order to reach node $(5, 1)$, he must, of course, keep gambling at node $(4, 1)$.[5]

Now consider what the agent *actually* does if he arrives at node $(4, 1)$. If he leaves the casino at this node, he earns utility of

$$v(40). \tag{14}$$

[5] The concavity of $v(\cdot)$ in the region of gains makes it harder for condition (13) to hold. What Figure 5 shows, however, is that, for the vast majority of values of α and δ, the probability weighting function overcomes the concavity of $v(\cdot)$ and ensures that condition (13) *does* hold.

Alternatively, he can keep gambling, which, from the perspective of time 4, offers cumulative prospect theory utility of

$$v(50)w(\frac{1}{2}) + v(30)\left[1 - w(\frac{1}{2})\right]. \tag{15}$$

The agent therefore wants to keep gambling at node $(4, 1)$ if

$$(v(50) - v(30))w(\frac{1}{2}) > v(40) - v(30). \tag{16}$$

It is straightforward to check that condition (16) *never* holds for $\alpha, \delta \in (0, 1)$. In other words, *from the perspective of time 4*, the agent prefers exit strategy B: if he reaches node $(4, 1)$, he always wants to leave the casino. This means that, for the vast majority of values of α and δ – specifically, for all the values indicated by the shaded area in Figure 5 – the agent is time inconsistent: if he arrives at node $(4, 1)$, he no longer wants to keep gambling as his original time 0 plan stipulated that he should. What is the intuition? From the time 0 perspective, node $(5, 1)$ was unlikely, overweighted, and hence appealing. From the time 4 perspective, however, it is no longer unlikely: once the agent is at node $(4, 1)$, node $(5, 1)$ can be reached with probability 0.5. The probability weighting function *underweights* moderate probabilities like 0.5. From the perspective of time 4, then, the $50 win in node $(5, 1)$ is no longer as appealing.[6]

The time inconsistency we have just described stems entirely from probability weighting. In the absence of probability weighting, conditions (13) and (16) are identical and the agent always prefers to leave the casino at node $(4, 1)$ rather than to continue gambling, whether this is judged from the perspective of time 0 or from the perspective of time 4.

Our example illustrates a time inconsistency in the upper part of the binomial tree. For the vast majority of values of α and δ, there is an analogous time inconsistency in the *bottom* part of the tree. From the perspective of time 0, the agent would like to stop gambling if he were to arrive at node $(4, 5)$, the bottom node in the second column from the right. However, if he actually arrives in node $(4, 5)$, he wants to keep gambling, contrary to his initial plan. The intuition for this inconsistency parallels the intuition for the inconsistency in the upper part of the tree.

Given the time inconsistency, the agent's behavior depends on two things. First, it depends on whether he is aware of the time inconsistency. An agent who *is* aware of the time inconsistency has an incentive to try to commit to his initial plan of action. For this agent, then, his behavior further depends on whether he is indeed able to commit. To explore these distinctions, we consider three types of agents. Our classification parallels the one used in the related literature on hyperbolic discounting.

[6] The concavity of $v(\cdot)$ in the region of gains only strengthens the agent's desire to leave the casino if he arrives at node $(4, 1)$.

The first type of agent is "naive". An agent of this type does not realize that, at time $t > 0$, he will deviate from his initial plan. We analyze his behavior in Section 3.1.

The second type of agent is "sophisticated" but unable to commit. An agent of this type recognizes that, at time $t > 0$, he will deviate from his initial plan. He would therefore like to commit to his initial plan – but is unable to find a way to do so. We analyze his behavior in Section 3.2.

The third and final type of agent is sophisticated and able to commit. An agent of this type also recognizes that, at time $t > 0$, he will want to deviate from his initial plan. However, he is able to find a way of committing to this initial plan. We analyze his behavior in Section 3.3.[7]

3.1 Case I: The naive agent

The naive agent is unaware of his time inconsistency: at time t, he does not realize that, at time $t' > t$, he will deviate from the plan of action he crafts at time t. We analyze his behavior in two steps. First, we study his behavior at time 0 as he decides whether to enter the casino. If we find that, for some parameter values, he is willing to enter the casino, we then look, for those parameter values, at his behavior *after* entering the casino, in other words, at his behavior for $t > 0$.

The initial decision

At time 0, the naive agent chooses a plan of action. A "plan" is a mapping from every node in the binomial tree between $t = 1$ and $t = T - 1$ to one of two possible actions: "exit," which indicates that the agent plans to leave the casino if he arrives at that node; and "continue," which indicates that he plans to keep gambling if he arrives at that node. We denote the set of all possible plans as $S_{(0,1)}$, with the subscript $(0, 1)$ indicating that this is the set of plans that is available at node $(0, 1)$, the left-most node in the tree. Even for low values of T, the number of possible plans is very large.[8]

For each plan $s \in S_{(0,1)}$, there is a random variable $\widetilde{G}_s$ which represents the accumulated

[7] An implicit assumption here is that, at time 0, the agent disapproves of his future preferences, in other words, of the preferences that will lead him to deviate from his initial plan. It is this disapproval that makes the sophisticated agent want to commit. In Section 4.4, we discuss this idea in more detail.

[8] Since, for each of the $T(T+1)/2 - 1$ nodes between time 1 and time $T - 1$, the agent can either exit or continue, an upper bound on the number of elements of $S_{(0,1)}$ is 2 to the power of $T(T+1)/2 - 1$. For $T = 5$, this equals $16,384$; for $T = 6$, it equals $1,048,576$. The number of *distinct* plans is lower than 2 to the power of $T(T+1)/2 - 1$, however. For example, for any $T \geq 2$, all plans that assign the action "exit" to nodes $(1, 1)$ and $(1, 2)$ are effectively the same.

winnings or losses the agent will experience if he exits the casino at the nodes specified by plan s. For example, if s is the exit strategy shown in Figure 3, then

$$\widetilde{G}_s \sim (\$30, \frac{7}{32}; \$10, \frac{9}{32}; -\$10, \frac{10}{32}; -\$30, \frac{5}{32}; -\$50, \frac{1}{32}).$$

With this notation in hand, we can write down the problem that the naive agent solves at time 0. It is:

$$\max_{s \in S_{(0,1)}} V(\widetilde{G}_s), \tag{17}$$

where $V(\cdot)$ computes the cumulative prospect theory value of the gamble that is its argument. We emphasize that the naive agent chooses a plan at time 0 without regard for the possibility that he might stray from the plan in future periods. After all, he is naive: he does not realize that he might later depart from the plan.

The non-concavity and nonlinear probability weighting embedded in $V(\cdot)$ make it very difficult to solve problem (17) analytically. However, we can solve it numerically and find that this approach allows us to draw out the economic intuition in full. Throughout the paper, we check the robustness of our conclusions by solving (17) for a wide range of preference parameter values.

The time inconsistency introduced by probability weighting means that we cannot use dynamic programming to solve the above problem. Instead, we use the following procedure. For each plan $s \in S_{(0,1)}$ in turn, we compute the gamble $\widetilde{G}_s$ and calculate its cumulative prospect theory value $V(\widetilde{G}_s)$. We then look for the plan s^* with the highest cumulative prospect theory value $V^* = V(\widetilde{G}_{s^*})$. The naive agent enters the casino – in other words, he plays a gamble at time 0 – if and only if $V^* \geq 0$.[9]

We now present some results from our numerical analysis. We set $T = 5$ and $h = \$10$. The shaded areas in Figure 6 show the range of values of the preference parameters α, δ, and λ for which the naive agent is willing to enter the casino, in other words, the range for which $V^* \geq 0$. To understand the figure, recall that, based on experimental data, Tversky and Kahneman's (1992) median estimates of the preference parameters are

$$(\alpha, \delta, \lambda) = (0.88, 0.65, 2.25). \tag{18}$$

Each of the three panels in the figure fixes one of the three parameters at its median estimate and shows the range of the other two parameters for which the agent enters the casino. The small circles correspond to the median estimates in (18).

The key result in Figure 6 is that, even though the agent is loss averse and even though the casino offers only 50:50 bets with zero expected value, there is still a wide range of

[9] Recall that the set $S_{(0,1)}$ consists only of plans that involve gambling at node $(0, 1)$. The agent is therefore willing to gamble at this node if the best plan that involves gambling, plan s^*, offers higher utility than not gambling; in other words, higher utility than zero.

parameter values for which the agent *is* willing to enter the casino. Note that, for Tversky and Kahneman's median estimates in (18), the agent is not willing to enter the casino. Nonetheless, for parameter values that are not far from those in (18), he *is* willing to gamble.

To understand why, for some parameter values, the agent is willing to gamble, we examine his optimal exit plan s^*. Consider the case of $(\alpha, \delta, \lambda) = (0.95, 0.5, 1.5)$; we find that, for these parameter values, the agent is willing to enter the casino. The left panel in Figure 7 shows the agent's optimal exit plan in this case. Recall that, if the agent arrives at a solid black node, he leaves the casino at that node; otherwise, he continues gambling. The figure shows that, roughly speaking, the agent's optimal plan is to keep gambling until time T *or* until he starts accumulating losses, whichever comes first. Through extensive checks, we find that, for almost all the parameter values for which the naive agent is willing to enter the casino, the optimal exit strategy is similar to the one in Figure 7.

The exit plan in Figure 7 helps us understand why it is that, even though the agent is loss averse and even though the casino offers only zero expected value bets, the agent is still willing to enter the casino. The reason is that, even though the basic 50:50 bet offered by the casino is unappealing when considered on its own, the agent is able, through his exit plan, to give his *overall* casino experience a positively skewed distribution: by exiting once he starts accumulating losses, he limits his downside; and by continuing to gamble when he is winning, he retains substantial upside.

At this point, probability weighting plays an important role. Since the agent overweights the tails of probability distributions, he may *like* the positively skewed distribution offered by the overall casino experience. In particular, under probability weighting, the chance, albeit small, of winning the large jackpot $\$Th$ in the top-right node $(T, 1)$ becomes particularly enticing. In summary, then, while the agent would always turn down the basic 50:50 bet offered by the casino if that bet were offered *in isolation*, he is nonetheless able, through a clever choice of exit strategy, to give his overall casino experience a positively skewed distribution, one which, with sufficient probability weighting, he finds attractive.[10][11]

We suspect that when actual gamblers enter a casino, they often have in mind a plan that

[10]For a very small range of parameter values – a range in which α and λ are much lower than Tversky and Kahneman's (1992) estimates and δ much higher – the naive agent enters the casino with a different plan in mind, namely one in which he keeps gambling if he is losing and stops if he accumulates some gains. This strategy gives his perceived overall casino experience a negatively skewed distribution; but since α is so low and δ is so high, he does not find this unappealing.

[11]The formulation in (17) assumes that the agent's "reference point" for computing gains and losses is always fixed at his initial wealth at the moment he enters the casino. We know little about how reference points move over time. Our strategy is therefore to pick one simple assumption – that the reference point remains fixed – and to show that this leads to a rich model of gambling. Intuitively, a model in which the agent updates his reference point over time would have a harder time explaining casino gambling: in such a model, the agent would often be at the most risk averse point of the value function, the kink.

is *broadly* similar to the one in the left panel of Figure 7 – specifically, a plan under which they continue to gamble when they are winning but stop gambling once their accumulated losses reach *some* cutoff level. However, we also suspect that they may not have in mind the *exact* plan in Figure 7. In particular, they may be uncomfortable with a plan under which they might have to leave the casino after just one bet: it might feel silly to leave the casino so early if they have just traveled a long time to get there. While our thesis in this paper is that prospect theory can shed much light on casino gambling, one thing it does not capture, at least in the basic model we have outlined so far, is an aversion to leaving the casino soon after arriving.

It is straightforward to incorporate an aversion to an early exit into our model. Specifically, at time 0, instead of solving (17), the naive agent can maximize $V(\widetilde{G}_s)$ over a *subset* of the plans in $S_{(0,1)}$, namely that subset for which the probability of leaving the casino in the first few rounds is lower than some given number. We find that there are several plans that have a positive cumulative prospect theory value – so that, under these plans, the agent would be willing to enter the casino – but that nonetheless entail a low probability of exit in the early rounds. Figure 8 illustrates one such plan for the same preference parameter values as in Figure 7. This plan, which we suspect is more typical of the plans that many actual gamblers have in mind, is not optimal in our basic model; but it may be optimal in a slightly extended model that combines cumulative prospect theory with an aversion to leaving the casino in the very early rounds.[12]

Figure 6 shows that the agent is more likely to enter the casino for *low* values of δ, for *low* values of λ, and for *high* values of α. The intuition is straightforward. By adopting an exit plan under which he rides gains as long as possible but stops gambling once he starts accumulating losses, the agent gives his overall casino experience a positively skewed distribution. As δ falls, the agent overweights the tails of probability distributions all the more heavily. He is therefore all the more likely to find a positively skewed distribution attractive and hence all the more likely to enter the casino. As λ falls, the agent becomes less loss averse. He is therefore less scared by the potential losses he could incur at the casino and therefore more willing to enter. Finally, as α falls, the marginal utility of additional gains diminishes more rapidly. The agent is therefore less excited about the possibility of a large win and hence less likely to enter the casino.

We noted above that, due to the convexity of the value function in the region of losses and the use of transformed probabilities, it is difficult to solve problem (17) analytically. We

[12]In particular, we conjecture that a plan similar to the one in Figure 8 is optimal in a model in which the agent, in addition to deriving cumulative prospect theory utility from his gain or loss at the moment of exit, also derives a per-period utility of gambling. As such, this would be an example of the way in which our model complements earlier models based on utility of gambling. In Section 4.4, we discuss an alternative extension of our model in which the agent's optimal plan is less extreme, and therefore potentially more realistic, than the one in Figure 7.

have, however, been able to derive the follow result, which states a sufficient condition for the naive agent to be willing to enter the casino. The proof is in the Appendix.

Proposition 1: For given preference parameters $(\alpha, \delta, \lambda)$ and a given number of rounds of gambling T, the naive agent is willing to enter the casino at time 0 if[13]

$$\sum_{j=1}^{T-\left[\frac{T}{2}\right]} (T+2-2j)^{\alpha} \left(w(2^{-T}\binom{T-1}{j-1}) - w(2^{-T}\binom{T-1}{j-2}) \right) \geq \lambda w(\frac{1}{2}). \tag{19}$$

To derive condition (19), we take one particular exit strategy which, from extensive numerical analysis, we know to be either optimal or close to optimal for a wide range of parameter values – roughly speaking, a strategy in which the agent keeps gambling when he is winning but stops gambling once he starts accumulating losses – and compute its cumulative prospect theory value explicitly. Condition (19) checks whether this value is positive; if it is, we know that the naive agent enters the casino. The condition is useful because it can shed light on the agent's behavior when T is high without requiring us to solve problem (17) explicitly, something which, for high values of T, is computationally very taxing.

For four different values of T, Figure 9 sets $\alpha = 0.88$ and plots the range of values of δ and λ for which condition (19) holds. We emphasize that the condition is sufficient but not necessary. If it holds, the naive agent enters the casino; but he may enter the casino even if it does not hold. Nonetheless, by comparing the top-left panels in Figures 6 and 9, both of which correspond to $T = 5$, we see that the parameter values for which condition (19) holds and the parameter values for which the naive agent actually enters the casino are very similar. In this sense, condition (19) is not only sufficient but almost necessary as well.

The top-right and bottom panels in Figure 9 suggest that, as the number of rounds of gambling T goes up, the naive agent is willing to enter the casino for a wider range of preference parameter values. Intuitively, as T goes up, the agent, through a careful choice of exit strategy, can create an overall casino experience that is all the more positively skewed and therefore, for someone who overweights tails, all the more attractive.[14]

Figures 6 and 9 show that, for Tversky and Kahneman's (1992) *median* estimates of α, δ, and λ, the prospect theory agent is only willing to enter the casino for high values of T; and Figure 9 suggests that even for high values of T, he is just barely willing to enter. There is

[13]In this expression, $\binom{T-1}{-1}$ is assumed to be equal to 0.

[14]It is easy to prove that the range of preference parameter values for which the naive agent enters the casino when $T = \tau$ is at least as large as the range for which he enters when $T = \tau + 1$. In particular, this follows from the fact that any plan that can be implemented in τ rounds of gambling can also be implemented in $\tau + 1$ rounds of gambling. Figure 9 gives us a sense of *how much* the range expands as T goes up.

a sense in which this fits with the evidence. Although 54 million people visited U.S. casinos in 2007, this still represents a minority of the U.S. population. The fact that the median U.S. resident does not gamble is consistent with the fact that, for the median values of the preference parameters, the prospect theory agent in our model often refuses to gamble. From the perspective of our model, the people who visit casinos are those with *lower* values of δ or λ than the median U.S. resident.

We noted earlier that we are dividing our analysis of the naive agent into two parts. We have just completed the first part: the analysis of the agent's time 0 decision as to whether or not to enter the casino. We now turn to the second part: the analysis of what the agent does at time $t > 0$. We know that, at time $t > 0$, the agent will depart from his initial plan. Our goal is to understand exactly how he departs from it.

Subsequent behavior

Suppose that, at time 0, the naive agent decides to enter the casino. In node j at some later time $t \geq 1$, he solves

$$\max_{s \in S_{(t,j)}} V(\widetilde{G}_s). \tag{20}$$

Here, $S_{(t,j)}$ is the set of plans the agent could follow subsequent to time t, where, in a similar way to before, a "plan" is a mapping from every node between time $t+1$ and time $T-1$ to one of two actions: "exit," indicating that the agent plans to leave the casino if he reaches that node, and "continue," indicating that the agent plans to keep gambling if he reaches that node. As before, $\widetilde{G}_s$ is a random variable which represents the accumulated winnings or losses the agent will experience if he exits the casino at the nodes specified by plan s, and $V(\widetilde{G}_s)$ is its cumulative prospect theory value. If s^* is the plan that solves problem (20), the agent gambles in node j at time t if

$$V(\widetilde{G}_{s^*}) \geq v(h(t+2-2j)), \tag{21}$$

where the right-hand side of condition (21) is the utility of leaving the casino at this node.

To see how the naive agent actually behaves for $t \geq 1$, we return to the example from earlier in this section in which $T = 5$, $h = \$10$, and $(\alpha, \delta, \lambda) = (0.95, 0.5, 1.5)$. Recall that, for these parameter values, the naive agent is willing to enter the casino at time 0. The right panel of Figure 7 shows what the naive agent does subsequently, at time $t \geq 1$. By way of reminder, the left panel in the figure shows the initial plan of action he constructs at time 0.

Figure 7 shows that, while the naive agent's *initial* plan was to keep gambling as long as possible when winning but to stop gambling once he started accumulating losses, he *actually*, roughly speaking, does the opposite: he stops gambling once he accumulates some gains and instead continues gambling as long as possible when he is losing. We find a similar pattern of behavior across all parameter values for which the naive agent is willing to enter the casino at time 0. Our model therefore captures a commonly heard intuition, namely that people often gamble more than they planned to in the region of losses.

Why does the naive agent behave in this way? Suppose that he has accumulated some gains. Whether or not he continues to gamble depends on two opposing forces. On the one hand, since he has accumulated gains, he is in the concave section of the value function. This induces risk aversion which, in turn, encourages him to stop gambling and to leave the casino. On the other hand, the probability weighting function encourages him to keep gambling: by continuing to gamble, he keeps alive the chance of winning a much larger amount of money; while this is a low probability event, the low probability is overweighted, making it attractive to keep gambling. As the agent approaches the end of the tree, however, the possibility of winning a large prize becomes less unlikely; it is therefore overweighted less, and continuing to gamble becomes less attractive. In other words, as the agent approaches the end of the tree, the concavity effect overwhelms the probability weighting effect and the agent stops gambling.

A similar set of opposing forces is at work in the bottom part of the binomial tree. Since, here, the agent has accumulated losses, he is in the convex part of the value function. This induces risk-seeking which encourages him to keep gambling. On the other hand, the probability weighting function encourages him to stop gambling: if he keeps gambling, he runs the risk of a large loss; while this is a low probability event, the low probability is overweighted, making gambling a less attractive option. The right panel in Figure 7 shows that, at *all* points in the lower part of the tree, the convexity effect overwhelms the probability weighting effect and the agent continues to gamble.[15]

3.2 Case II: The sophisticated agent, without commitment

In section 3.1, we considered the case of a naive agent – an agent who, at time t, does not realize that, at time $t' > t$, he will deviate from his time t plan. In Sections 3.2 and 3.3, we study sophisticated agents, in other words, agents who do recognize that they will deviate from their initial plan. A sophisticated agent has an incentive to find a commitment device that will enable him to stick to his time 0 plan. In this section, we consider the case of a sophisticated agent who is *unable* to find a way of committing to his time 0 plan; we label this agent a "no-commitment sophisticate" for short. In Section 3.3, we study the case of a

[15]The naive agent's "naivete" can be interpreted in two ways. The agent may fail to realize that, after he starts gambling, he will be tempted to depart from his initial plan. Alternatively, he may recognize that he will be tempted to depart from his initial plan, but he may erroneously think that he will be able to resist the temptation. Over many repeated casino visits, the agent may learn his way out of the first kind of naivete. It may take much longer, however, for him to learn his way out of the second kind. People often continue to believe that they will be able to exert self-control in the future even when they have repeatedly failed to do so in the past.

sophisticated agent who *is* able to commit to his initial plan.

To determine a course of action, the no-commitment sophisticate uses dynamic programming, working leftward from the right-most column of the binomial tree. If he has not yet left the casino at time T, he must necessarily exit at that time. His value function in node j at time T – here, we mean "value function" in the dynamic programming sense rather than in the prospect theory sense – is therefore

$$J_{T,j} = v(h(T + 2 - 2j)). \tag{22}$$

The agent then continues the backward iteration from $t = T - 1$ to $t = 0$ using

$$J_{t,j} = \max\{v(h(t + 2 - 2j)), V(\widetilde{G}_{t,j})\}, \tag{23}$$

where $J_{t,j}$ is the value function in node j at time t. The term before the comma on the right-hand side is the agent's utility if he leaves the casino in node j at time t. The term after the comma is the utility of continuing to gamble: specifically, it is the cumulative prospect theory value of the random variable $\widetilde{G}_{t,j}$ which measures the winnings or losses the agent will exit the casino with if he continues gambling at time t. The gamble $\widetilde{G}_{t,j}$ is determined by the exit strategy computed in earlier steps of the backward iteration. Continuing this iteration back to $t = 0$, the agent can see whether or not it is a good idea to enter the casino in the first place.

We now return to the example of Section 3.1 in which $T = 5$, $h = \$10$, and $(\alpha, \delta, \lambda) = (0.95, 0.5, 1.5)$. Figure 10 illustrates the outcome of the backward iteration in this case. A solid black node indicates that the agent will leave the casino if he reaches that node; at the other nodes, the agent keeps gambling. The fact that the left-most node is black means that, for these parameter values, the agent chooses not to enter the casino at all.

The intuition for why the no-commitment sophisticate chooses not to enter the casino is straightforward. He realizes that, if he does enter the casino, he will leave as soon as he accumulates some gains but will keep gambling as long as possible if he is losing. This exit policy gives his overall casino experience a *negatively* skewed distribution. Recognizing this in advance, he decides not to enter the casino: since he overweights the tails of distributions, the negative skewness is unattractive.

The result in Figure 10 – that the no-commitment sophisticate refuses to enter the casino – holds for a wide range of preference parameter values. Indeed, after extensive checks, we have been unable to find *any* $(\alpha, \delta, \lambda) \in (0.5, 1) \times (0.28, 0.8) \times (1.3, \infty)$ for which the no-commitment sophisticate is willing to enter the casino at time 0.[16]

[16] For a very small range of parameter values – a range in which α and λ are much lower than Tversky and Kahneman's (1992) estimates and δ much higher – the no-commitment sophisticate *is* willing to enter the casino. While he recognizes that his overall casino experience has a negatively skewed distribution, the fact that α is so low and δ so high means that he does not find this unappealing.

3.3 Case III: The sophisticated agent, with commitment

A sophisticated agent – an agent who recognizes that, at time $t > 0$, he will want to deviate from his initial plan – has an incentive to find a commitment device that will enable him to stick to his initial plan. In this section, we study the behavior of a sophisticated agent who is able to commit. We call this agent a "commitment-aided sophisticate."

We proceed in the following way. We assume that, at time 0, the agent can find a way of committing to *any* exit strategy $s \in S_{(0,1)}$. Once we identify the strategy that he would choose, we then discuss how he might actually commit to this strategy in practice.

At time 0, then, the commitment-aided sophisticate solves exactly the same problem as the naive agent, namely:

$$\max_{s \in S_{(0,1)}} V(\widetilde{G}_s). \tag{24}$$

In particular, since the agent can commit to any exit strategy, we do not need to restrict the set of strategies he considers. He searches across *all* elements of $S_{(0,1)}$ until he finds the strategy s^* with the highest cumulative prospect theory value $V^* = V(\widetilde{G}_{s^*})$. He enters the casino if and only if $V^* \geq 0$.

Since the commitment-aided sophisticate and the naive agent solve exactly the same problem at time 0, they will, for given preference parameter values, choose exactly the same optimal strategy. Moreover, they will enter the casino for exactly the same range of preference parameter values. For $T = 5$ and $h = \$10$, for example, the commitment-aided sophisticate enters the casino for the parameter values indicated by the shaded areas in Figure 6. And for $(\alpha, \delta, \lambda) = (0.95, 0.5, 1.5)$, his optimal plan is the one in the left panel of Figure 7, a plan under which he continues to gamble when he is winning but stops gambling once he starts accumulating losses.[17]

The naive agent and the commitment-aided sophisticate solve the same problem at time 0 because they both *think* that they will be able to maintain any plan they select at that time. The two types of agents differ, however, in what they do after they enter the casino. Since he has a commitment device at his disposal, the commitment-aided sophisticate is able to stick to his initial plan. The naive agent, on the other hand, deviates from his initial plan: after he enters the casino, he continues to gamble when is losing and stops once he accumulates a significant gain.

Now that we have identified the strategy the commitment-aided sophisticate would like to commit to, the natural question is: *how* does he commit to it? For example, in the lower part of the binomial tree, how does he manage to stop gambling when he is losing even

[17] In the same way, the sufficient condition (19) for the naive agent to be willing to enter the casino is also a sufficient condition for the commitment-aided sophisticate to be willing to enter the casino.

though he is tempted to continue? And in the upper part of the tree, how does he manage to continue gambling when he is winning even though he is tempted to stop?

In the lower part of the tree, one simple commitment strategy is for the agent to go to the casino with only a small amount of cash in his pocket and to leave his ATM card at home. If he starts losing money, he is sorely tempted to continue gambling, but, since he has run out of cash, he has no option but to go home. It is a prediction of our model that some casino gamblers will use a strategy of this kind. Anecdotally, at least, this *is* a common gambling strategy, which suggests that at least some of those who go to casinos fit the mold of our commitment-aided sophisticate.

In the upper part of the tree, it is less easy to think of a common strategy that gamblers use to solve the commitment problem, in other words, to keep gambling when they are winning even though they are tempted to go home. In a way, this is not surprising. One thing our model predicts – something which, as we will see in Section 4.5, is especially true for higher values of T – is that the time inconsistency is much more severe in the *lower* part of the tree than in the upper part. By comparing the two panels in Figure 7, we see that in the lower part of the tree, the time inconsistency, and hence the commitment problem, is severe: the agent wants to gamble at *every* node in the region of losses even though his initial plan was to gamble at none of them. In the upper part of the tree, however, the time inconsistency, and hence the commitment problem, is less acute: the agent's initial plan conflicts with his subsequent actions at only a few nodes. It therefore makes sense that the commitment strategies gamblers use in practice seem to be aimed primarily at the time inconsistency in the lower part of the tree.

Although it is hard to think of ways in which gamblers themselves commit to their initial plan in the upper part of the tree, note that here, *casinos* have an incentive to help. In general, casinos offer bets with negative expected values; it is therefore in their interest that gamblers stay on site as long as possible. From the casinos' perspective, it is alarming that gamblers are tempted to leave earlier than they originally planned when they are winning. This may explain the common practice among casinos of offering vouchers for free food and lodging to people who are winning. In our framework, casinos do this in order to encourage gamblers who are thinking of leaving with their gains, to stay longer.

In this section, we have identified some important and arguably unique predictions of our framework. For example, our model predicts the common gambling strategy of bringing only a fixed amount of money to the casino; and it predicts the common casino tactic of giving free vouchers to people who are winning. These features of gambling have not been easy to understand in earlier models but emerge naturally from the one we present here. In particular, they are a direct consequence of the time inconsistency at the heart of our model.

4 Discussion

In Section 3, we presented an explanation of why people go to casinos. An important feature of our approach is that we explain casino gambling using a theory of decision under risk – prospect theory – that already explains a lot of *other* evidence on risk-taking. As such, our paper suggests that casino gambling is not necessarily an isolated phenomenon requiring its own unique explanation, but rather one of a large family of facts that can be understood using a single model of risk attitudes.

Our result that prospect theory can explain casino gambling is initially surprising: casinos offer gambles that have low expected values and that are much less skewed than a lottery ticket. These are gambles that, one would think, loss-averse agents would find unappealing. The reason why prospect theory agents *are* sometimes willing to enter a casino traces back to probability weighting. In our model, both types of agents who enter the casino – both the naive agents and the commitment-aided sophisticates – do so with the *same* plan in mind, namely a plan in which they continue gambling when they are winning but stop gambling once they start accumulating losses. This gives their perceived *overall* casino experience a positively skewed distribution – a distribution which, with sufficient probability weighting, they find attractive.

While all the agents who enter the casino in our model do so with the same plan in mind, the set of casino gamblers nonetheless consists of two very different subgroups. Some agents – the commitment-aided sophisticates – are able to stick to their initial plan even though they are later tempted to deviate from it. Other agents – the naive agents – deviate from their initial plan: they leave the casino *earlier* than planned when they are winning and *later* than planned when they are losing.

In summary, under the view proposed in this paper, casinos are popular because they cater to two aspects of our psychological make-up. First, they cater to the tendency to overweight the tails of distributions, which makes even the small chance of a large win at the casino seem very alluring. And second, they cater to what we could call "naivete," namely the failure to recognize that, after entering a casino, we may deviate from our initial plan of action.

Of all casino games, the model in Section 3 corresponds most closely to blackjack. Nonetheless, it may also be able to explain why another casino game, the slot machine, is as popular as it is. In our model, the agents who enter the casino do so because they relish the positively skewed distribution they perceive it to offer. Since slot machines already offer a skewed payoff, they may make it easier for the agent to give his overall casino experience a significant amount of positive skewness. It may therefore make sense that they would outstrip blackjack in popularity.

Our model does not necessarily predict that *all* gamblers will prefer slot machines to blackjack, however. Precisely because it offers a simpler bet – one that approximates a 50:50 bet to win or lose some amount – it may be blackjack rather than a slot machine that makes it easier for the gambler to construct the specific payoff he most prefers – put differently, it may be blackjack that provides the more versatile "building block".

In the rest of this section, we discuss a number of other issues raised by the analysis in Section 3.

4.1 Average losses

The analysis in Section 3 shows that the set of casino gamblers is made up of two distinct types: naive agents and commitment-aided sophisticates. Which of these two types loses more money in the casino, on average?

In the context of the model of Section 3 – a model in which the basic bet offered by the casino is a 50:50 bet to win or lose $\$h$ – the answer is straightforward. Since the basic bet has an expected value of zero, the average winnings are zero for both naive agents and commitment-aided sophisticates.

Now suppose, however, that the basic bet has a negative expected value, as in actual casinos. For example, suppose that the basic bet is now

$$(\$h, 0.49; -\$h, 0.51). \tag{25}$$

An agent's average winnings are the (negative) expected value of the basic bet multiplied by the average number of rounds the agent gambles. To see which of naive agents and commitment-aided sophisticates has greater average losses, we therefore need to determine which of the two groups gambles for longer, on average. The group that gambles for longer will do worse.

For $T = 5$, $h = \$10$, and $(\alpha, \delta, \lambda) = (0.95, 0.5, 1.5)$, we compute the gambling behavior of the two types of agents when the basic bet has the form in (25). We find that the behavior of the naive agent is still that shown in the right panel in Figure 7 while the behavior of the commitment-aided sophisticate is still that shown in the left panel in Figure 7. This allows us to compute that the naive agent stays in the casino almost twice as long as the sophisticated agent, on average. His average losses are therefore almost twice as large. In this sense, the naivete of the naive agent – his failure to foresee his time inconsistency – is costly.[18]

[18]In Section 4.5, we will see that, for higher values of T, the average length of stay in a casino for a naive agent is even longer, relative to that for a commitment-aided sophisticate.

4.2 One-time gambles

The analysis in Section 3 raises the following question. Given that the agents who enter the casino in our model do so because they like the positively skewed payoff offered by an evening of gambling, why do casinos not offer them this positively skewed payoff as a single bet? For example, when $(\alpha, \delta, \lambda) = (0.95, 0.5, 1.5)$, the naive agent chooses the initial plan shown in the left panel of Figure 7 because he likes the positively skewed gamble it represents, namely

$$\widetilde{G}_s \sim (\$50, \frac{1}{32}; \$30, \frac{3}{32}; \$10, \frac{2}{32}; \$0, \frac{5}{16}; -\$10, \frac{1}{2}).$$

Why does the casino not allow him to play this gamble in one go?

At first sight, allowing gamblers to collapse their evening at the casino into a single bet seems to have two advantages. First, it saves time: the agent can complete his gambling in the space of a few minutes rather than over the course of a few hours. Second, by allowing agents to collapse their gambling into a single bet, the casino is essentially offering them an explicit way of committing to their initial plan of action. As a result, it can attract to the casino all the no-commitment sophisticates who stay away precisely because they are unable to find a way of committing to their initial plan.

On reflection, we see two reasons why casinos would *not*, in fact, want to offer one-shot bets in this way. First, such a product would be hard to administer. Depending on his values of α, δ, and λ, a given gambler will have his own preferred positively skewed payoff and there is no way for the casino to know in advance what this preferred payoff is. Instead, by offering simple bets such as the 50:50 bet available at blackjack tables, the casino leaves it to the gambler to construct the positively skewed payoff he most prefers through his own personal choice of exit strategy.

Even if a casino can overcome this first difficulty, it is still far from clear that it would want to offer one-shot bets in the way described above. At first sight, it seems that it would be profitable to do so: the casino could attract the no-commitment sophisticates who are currently staying away. However, this argument misses another consequence of offering one-shot bets. If a *naive* agent thinks that there is any chance at all that he may be time inconsistent, he will take the one-shot bet, thereby in a sense converting himself from a naive agent to a commitment-aided sophisticate. If many naive agents act in this way, the casino will lose money because, as we saw in Section 4.1, it makes more money from naive agents than from commitment-aided sophisticates. Any profits the casino makes by attracting the no-commitment sophisticates could therefore be wiped out by allowing the naive agents to commit to their initial plan.

4.3 Predictions and other evidence

Researchers have not, as yet, had much success in obtaining large-scale databases on gambling behavior. While our model matches a range of anecdotal evidence on gambling – for example, the tendency to gamble longer than planned in the region of losses, the strategy of leaving one's ATM card at home, and casinos' practice of giving free vouchers to people who are winning – there is, unfortunately, little systematic evidence by which to judge our model.

Our model does, however, make a number of novel predictions – predictions that, we hope, can eventually be tested. Perhaps the clearest prediction is that gamblers' planned behavior will differ from their actual behavior in systematic ways. If we survey people when they first enter a casino as to what they *plan* to do, we should find that they are planning to gamble for longer if they start accumulating gains than if they start accumulating losses – in our model, both naive and sophisticated agents *plan* to gamble for longer in the region of gains. If we then look at what people actually do, we should find that, on average, they exit sooner than planned in the region of gains and later than planned in the region of losses. Moreover, if gamblers who are more sophisticated in the real-world sense of the word – in terms of education or income, say – are also more sophisticated in terms of recognizing their time inconsistency, we should see a larger difference between planned and actual behavior among the less sophisticated.

Some recent *experimental* evidence gives us hope that these predictions will be confirmed in the field. Andrade and Iyer (2008) offer subjects a sequence of 50:50 bets in a laboratory setting; but before playing the gambles, subjects are asked how they *plan* to gamble in each round. Andrade and Iyer find that, consistent with our model, subjects plan to gamble more after a gain than after a loss. They also find, again consistent with our model, that subjects systematically gamble *more* than planned after an early loss. After an early gain, however, there is no statistically significant difference between planned and actual behavior.

Another prediction comes from Figure 6, which shows that people are more likely to enter a casino if they have low values of δ and λ – in other words, if they overweight the tails of distributions more and if they are less loss averse. If we estimate δ and λ for casino goers – perhaps with the help of gambles like those used by Tversky and Kahneman (1992) – we should obtain lower values than for non-casino goers.[19]

[19] A commonly heard term in the context of casino gambling is the "house money effect," the idea that people are more willing to take risk after winning some money than they were before. There is very little direct evidence of this effect from casinos, but Thaler and Johnson (1990) document it in an experimental setting. The naive agent in our model exhibits a house money effect, and he does so for the reason proposed by Thaler and Johnson (1990), namely that, after a gain, the agent moves away from the kink, the most risk averse point of the value function.

4.4 An alternative specification

In Section 3.3, we formulated the commitment-aided sophisticate's problem in the way shown in (24). This specification assumes that, while the sophisticate knows that his future "selves" will have different preferences over exit strategies, he puts no weight on those preferences when choosing a strategy to commit to at time 0. For example, at time 0, he chooses an exit strategy under which he will leave the casino if he loses in the first round even though he knows that his time 1 self would prefer to keep gambling at that point.

In our view, the specification in (24) is a reasonable way of formulating the commitment-aided sophisticate's decision problem – in other words, we think it reasonable that the sophisticate *would* ignore his future preferences even though he is aware of them. One justification for this is that the sophisticate *disapproves* of his future preferences. He knows that, if he does put weight on those future preferences, he will act more like a naive agent, which, from Section 4.1, means that he will stay in the casino longer on average and will therefore lose more money, on average. As a result, he does not like his future preferences.[20]

Nonetheless, we are also open to an alternative model, one in which the sophisticate puts at least *some* weight on his future preferences. Intuitively, this is an agent who is only mildly disapproving of his future preferences. Instead of solving (24), this agent maximizes the sum of five cumulative prospect theory terms, each one corresponding to the cumulative prospect theory value of an exit strategy as perceived at each of the five dates, $t = 0, \ldots, 4$.

Intuitively, since he puts weight on his future preferences, this sophisticate would behave more like a naive agent: he would gamble longer in the region of losses than the commitment-aided sophisticate of Section 3.3 but not as long as the naive agent of Section 3.1; and he would gamble longer in the region of gains than the naive agent of Section 3.1 but not as long as the commitment-aided sophisticate of Section 3.3. While much more complex than the framework of Section 3.3, this alternative specification has at least one advantage: it predicts that the agent enters the casino with a plan that may be closer to that of actual gamblers, namely one in which he allows himself to accumulate some losses before exiting, rather than exiting as soon as he accumulates any loss at all.

[20]In a general discussion of non-expected utility preferences, Machina (1989) suggests another sense in which the time 0 agent may disapprove of his future preferences. When, at time $t > 0$, the agent departs from his initial plan, he is committing what Machina (1989) calls a fallacy of "consequentialism," in which he wrongly ignores branches of the binomial tree that have not been realized but whose risk has been borne in earlier stages. In this sense, the time 0 agent views his future inconsistency as an error.

4.5 Higher values of T

In our analysis so far, we have focused primarily on the case of $T = 5$. The reason is that, for higher values of T, the set of feasible plans $S_{(0,1)}$ that we search over in problems (17) and (24) becomes very large.

Can we nonetheless say more about what happens for higher values of T? It turns out that we can. We now present some results for $T = 10$. In this case, even though the set of plans $S_{(0,1)}$ is very large, we can still describe the behavior of naive agents and of no-commitment sophisticates; and with one reasonable simplification, we can also describe the behavior of commitment-aided sophisticates. In brief, we find that the results for $T = 10$ are similar to those for $T = 5$.

Throughout this section, we again set $h = \$10$ and $(\alpha, \delta, \lambda) = (0.95, 0.5, 1.5)$. Figure 11 shows the behavior of the naive agent when $T = 10$. As for $T = 5$, the naive agent is willing to enter the casino; and once he does so, he continues gambling as long as possible when he is losing but stops gambling once he accumulates significant gains.

How did we do the calculations behind Figure 11? The naive agent gambles at node (t, j) if and only if

$$\max_{s \in S_{(t,j)}} V(\widetilde{G}_s) \geq v(h(t + 2 - 2j)). \tag{26}$$

For $t \geq 5$, it is straightforward to check condition (26) because the set $S_{(t,j)}$ contains a manageable number of elements. But how can we check this condition for $t < 5$, when $S_{(t,j)}$ is vastly larger in size? We use a simple trick: so long as we can find *some* plan $s \in S_{(t,j)}$ for which $V(\widetilde{G}_s)$ exceeds $v(h(t + 2 - 2j))$, we immediately know that condition (26) holds and hence that the naive agent gambles at node (t, j). We therefore check whether $V(\widetilde{G}_s)$ exceeds $v(h(t + 2 - 2j))$ for any of a small number of plans that, intuitively, should deliver high prospect theory utility. For all nodes with $t < 5$, we were quickly able to find a plan with positive prospect theory utility. We can therefore be certain that the naive agent gambles at all of these nodes, as shown in Figure 11.

We now turn to the no-commitment sophisticate. Since this agent uses dynamic programming to determine his course of action and since the state space has a low dimension, it is straightforward to solve the problem he faces, even for large values of T. Figure 12 shows his behavior. The agent realizes that, if he does enter the casino, he will keep gambling if he is losing and will stop gambling if he accumulates some gains. This strategy generates a negatively skewed distribution of casino winnings which, under probability weighting, is unattractive. As for $T = 5$, then, the agent decides not to enter the casino.

The most difficult case is that of the commitment-aided sophisticate. This agent looks for the plan $s = s^*$ which solves

$$\max_{s \in S_{(0,1)}} V(\widetilde{G}_s) \tag{27}$$

and enters the casino if $V(\widetilde{G}_{s^*}) \geq 0$. To determine his behavior, we would need to search across all the vast number of elements in $S_{(0,1)}$. Since this is computationally infeasible, we make the following simplification. We restrict the commitment-aided sophisticate to a subset $S^{'}_{(0,1)}$ of the plans in $S_{(0,1)}$ that consists of what we call "threshold" plans. These are plans in which, either, (i) at each date $t \in [1, T-1]$, the agent leaves the casino only if his accumulated winnings fall below some level, which can depend on t; or, (ii) at each date $t \in [1, T-1]$, the agent leaves the casino only if his accumulated winnings rise above some level, which can again depend on t. For $T \leq 5$, we have always found the solution to (27) to be a threshold plan.

With this restriction, the commitment-aided sophisticate now solves

$$\max_{s \in S^{'}_{(0,1)}} V(\widetilde{G}_s). \tag{28}$$

Figure 13 shows the plan which solves this problem. As for $T = 5$, the agent plans to continue gambling as long as possible if he is winning but to stop gambling if he starts accumulating losses.[21]

5 Other Applications

The framework in this paper can be applied in contexts other than casino gambling. To illustrate this, we now briefly show how it can shed light on the trading behavior of individual investors.

To see this, we first reinterpret the binomial tree in Section 3 as capturing not the accumulated winnings in a casino, but rather the evolution of a stock price over time. Under this interpretation, each column of nodes represents the different possible stock prices on a particular date. Specifically, if the initial stock price in the left-most node is P_0 and, in each period, the stock either has a good return R_u – a step up in the binomial tree – or a poor return $R_d < R_u$ – a step down in the binomial tree – then the stock price in node (t, j) is $P_0 R_u^{t-j+1} R_d^{j-1}$. The stock return is i.i.d across time and, in each period, a good stock return and a poor stock return are equally likely.

Now suppose that, at time 0, an investor is deciding how to split his wealth between the stock we have just described and a risk-free asset that earns a net return of zero in each

[21] Figures 11 and 13 also illustrate two points we noted in earlier sections: first, that the time inconsistency is much more severe in the lower part of the tree; and second, that the naive trader stays in the casino much longer, on average, than the commitment-aided sophisticate.

period. To keep the exposition simple, we assume that if the investor takes a position in the stock at time 0, then, in future periods, he either maintains his position or else sells his entire holdings. Finally, we assume that the investor derives cumulative prospect theory utility from the gain or loss on the stock at the moment he sells his position. The question we are interested in is: if the investor buys stock at time 0, what is his *planned* selling strategy and how does this differ from his *actual* selling strategy?

The investment problem we have just described is exactly analogous to the casino problem of Section 3: just as the gambler derives prospect theory utility from his gain or loss when he exits the casino, so the investor derives prospect theory utility from his gain or loss when he exits the stock. We have analyzed the solution to this investment problem in detail. Here, we simply summarize the results without going through the analysis – this is in part for space reasons, but also because the results directly parallel those in Section 3.

Just as in Section 3, the prospect theory assumption leads to a time inconsistency. We therefore again consider three types of agents: a naive trader, a no-commitment sophisticate, and a commitment-aided sophisticate.

We find that, when the naive trader buys the stock at time 0, his initial plan parallels that of the naive gambler: he plans to keep holding the stock if its price remains above purchase price but to sell it if its price falls below purchase price. His actual selling behavior, however, follows the opposite pattern: he holds on to the stock if it falls below purchase price but sells it if it rises significantly above purchase price.

The no-commitment sophisticate recognizes that he will sell the stock if it goes up significantly and that he will hold on to it if it falls. While he dislikes the negatively skewed payoff this strategy produces, he is nonetheless willing to buy the stock at time 0 if its average return is high enough. This contrasts with the no-commitment sophisticate of Section 3 who refuses to enter the casino because, unlike stocks, the bets offered by the casino have a very low expected value.

Finally, the commitment-aided sophisticate recognizes that he will want to deviate from his initial plan and therefore finds a way to commit to it in advance. In other words, perhaps with the help of his broker, he commits to sell the stock if it falls but to hold on to it if it rises.

Earlier papers have also analyzed the trading behavior of an investor who derives prospect theory utility from the realized gain or loss when he sells a stock (see Barberis and Xiong, 2009, and the references therein). These papers show that the investor exhibits a "disposition effect," the name given to the propensity of actual individual investors to sell stocks in their portfolios that have *risen* in value since purchase, rather than fallen in value, a phenomenon that is hard to explain in more standard models of trading behavior.

The crucial difference between these earlier papers and our discussion here is that the prior work ignores probability weighting. Our discussion in this section shows that, as soon as we take probability weighting into account, we obtain a much richer model of the disposition effect.

First, while the earlier papers which ignore probability weighting predict that *all* investors exhibit a disposition effect, the framework in this section predicts a more interesting heterogeneity: *some* traders – specifically, the naive traders and the no-commitment sophisticates – exhibit a disposition effect, while others, namely the commitment-aided sophisticates, exhibit the opposite of the disposition effect. This is significant because while, in reality, most individual investors exhibit a disposition effect, some do not. Moreover, in our framework, naive agents always exhibit the disposition effect while sophisticated agents may or may not exhibit it, depending on their ability to commit. If investors who are sophisticated in the real-world sense of the word are also sophisticated in the sense of recognizing their time inconsistency, our framework predicts that investors who are more sophisticated will exhibit the disposition effect less. This is exactly what we find in the data (Dhar and Zhu, 2006).

Second, in the earlier papers which ignore probability weighting, the disposition effect always represents time-consistent behavior. By contrast, our framework raises the interesting possibility that it can represent time-inconsistent behavior. Naive agents, for example, plan to exhibit the opposite of the disposition effect; but in their actual trading, they exhibit the disposition effect itself. Meanwhile, sophisticated agents who recognize their inconsistency will try to commit in advance to their initial plan. There is evidence that fits with this view. Many asset management firms set in place formal rules that require a position to be unwound if it loses a certain amount of value – 15% of its value, say. This is consistent with a framework in which, while, ex-ante, traders would like to sell a stock if it goes down, they nonetheless, ex-post, find themselves reluctant to execute this plan. They therefore put in place a commitment device to ensure that the sale takes place.

6 Conclusion

Casino gambling is a hugely popular activity around the world, but there are still very few models of why people go to casinos or of how they behave when they get there. In this paper, we show that prospect theory can offer a surprisingly rich theory of gambling, one that captures many features of actual gambling behavior. First, we demonstrate that, for a wide range of parameter values, a prospect theory agent would be willing to gamble in a casino, even if the casino only offers bets with zero or negative expected value. Second, we show that prospect theory predicts a plausible time inconsistency: at the moment he enters a casino, a prospect theory agent plans to follow one particular gambling strategy;

but *after* he enters, he wants to switch to a different strategy. The model therefore predicts heterogeneity in gambling behavior: how a gambler behaves depends on whether he is aware of the time-inconsistency; and, if he *is* aware of it, on whether he is able to commit, in advance, to his initial plan of action.

7 References

Andrade E. and G. Iyer (2008), "Planned and Actual Betting in Sequential Gambles," *Journal of Marketing Research*, forthcoming.

Barberis N. and M. Huang (2008), "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," *American Economic Review* 98, 2066-2100.

Barberis N. and W. Xiong (2009), "What Drives the Disposition Effect? An Analysis of a Long-standing Preference-based Explanation," *Journal of Finance*, forthcoming.

Benartzi, S. and R. Thaler (1995), "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics* 110, 75-92.

Conlisk, J. (1993), "The Utility of Gambling," *Journal of Risk and Uncertainty* 6, 255-275.

Della Vigna S. and U. Malmendier (2006), "Paying Not to Go to the Gym," *American Economic Review* 96, 694-719.

Dhar, R. and N. Zhu (2006), "Up Close and Personal: An Individual Level Analysis of the Disposition Effect," *Management Science* 52, 726-740.

Feller W. (1968), *An Introduction to Probability Theory and Its Applications*, Wiley.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47, 263-291.

Laibson, D. (1997), "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics* 62, 443-477.

O'Donoghue, E. and M. Rabin (1999), "Doing it Now or Later," *American Economic Review* 89, 103-124.

Thaler, R. and E. Johnson (1990), "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes in Risky Choice," *Management Science* 36, 643-660.

Tversky, A. and D. Kahneman (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297-323.

8 Appendix

Proof of Proposition 1: Through extensive numerical analysis, we find that when the naive agent enters the casino, he almost always chooses the following strategy or one similar to it: he exits (i) if he loses in the first round; (ii) if, after the first round, his accumulated winnings ever drop to zero; and (iii) at time T, if he has not already left by that point. Condition (19) simply checks whether the cumulative prospect theory value of this exit strategy is positive. If it is, we know that the agent enters the casino.

If the agent exits because he loses in the first round, then, since the payoff of $-\$h$ is the *only* negative payoff he can receive under the above exit strategy, its contribution to the cumulative prospect theory value of the strategy is

$$-\lambda h^{\alpha} w(\frac{1}{2}).$$

If he exits because, at some point after the first round, his accumulated winnings equal zero, this contributes nothing to the cumulative prospect theory value of the exit strategy, precisely because the payoff is zero. All that remains, then, is to compute the component of the cumulative prospect theory value of the exit strategy that stems from the agent exiting at date T.

Under the above exit strategy, there are $T - [\frac{T}{2}]$ date T nodes with positive payoffs at which the agent might exit, namely nodes (T, j), where $j = 1, \ldots, T - [\frac{T}{2}]$. The payoff in node (T, j) is $(T + 2 - 2j)h$. We need to compute the probability that the agent exits at node (T, j), in other words, the probability that he moves from the initial node $(0, 1)$ to node (T, j) without losing in the first round and without his accumulated winnings hitting zero at any point after that. With the help of the reflection principle – see Feller (1968) – we compute this probability to be

$$2^{-T} \left[\binom{T-1}{j-1} - \binom{T-1}{j-2} \right].$$

The probability weight associated with node (T, j) is therefore

$$w(2^{-T}\binom{T-1}{j-1}) - w(2^{-T}\binom{T-1}{j-2}).$$

In summary then, the exit strategy we described above has positive cumulative prospect theory value – and hence the naive agent is willing to enter the casino – if

$$\sum_{j=1}^{T-[\frac{T}{2}]} ((T+2-2j)h)^{\alpha} \left(w(2^{-T}\binom{T-1}{j-1}) - w(2^{-T}\binom{T-1}{j-2}) \right) - \lambda h^{\alpha} w(\frac{1}{2}) \geq 0.$$

This is condition (19).

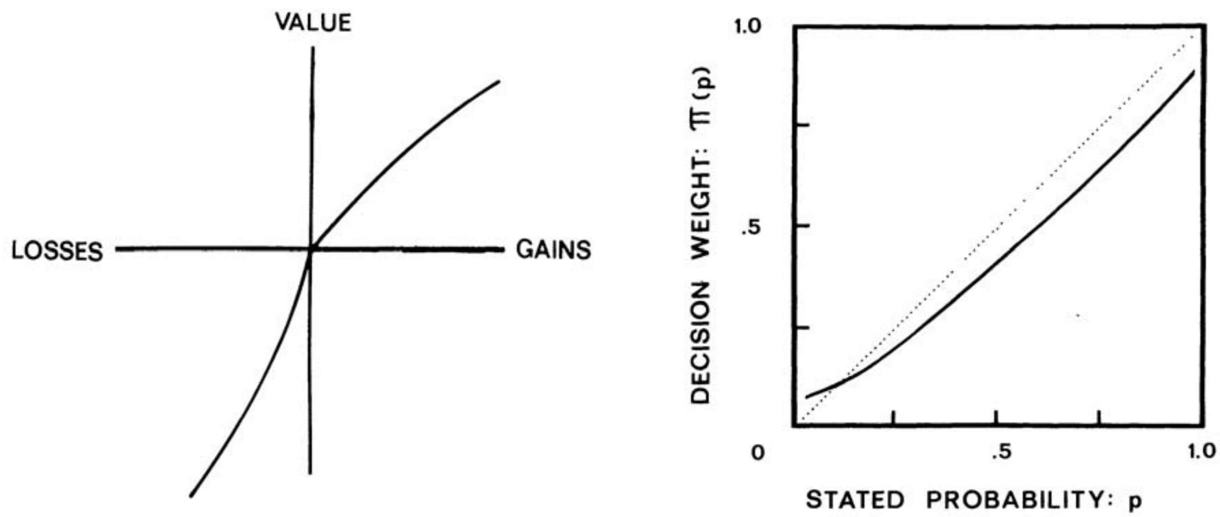

Figure 1. The left panel shows the value function proposed by Kahneman and Tversky (1979) as part of prospect theory, their model of decision-making under risk. The right panel shows the probability weighting function they propose.

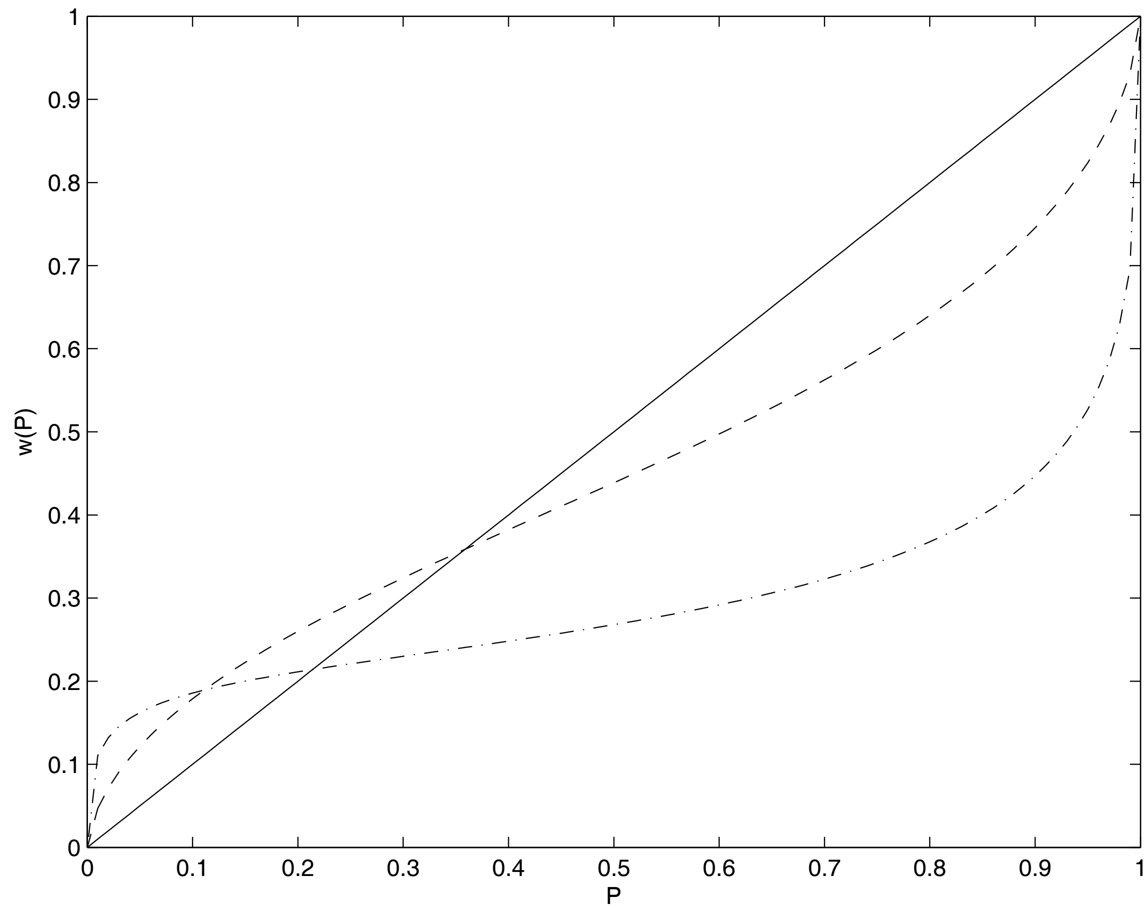

Figure 2. The figure plots the probability weighting function proposed by Tversky and Kahneman (1992), namely $w(P) = P^\delta/(P^\delta + (1-P)^\delta)^{1/\delta}$, for three different values of δ. The dashed line corresponds to $\delta = 0.65$, the dash-dot line to $\delta = 0.4$, and the solid line to $\delta = 1$.

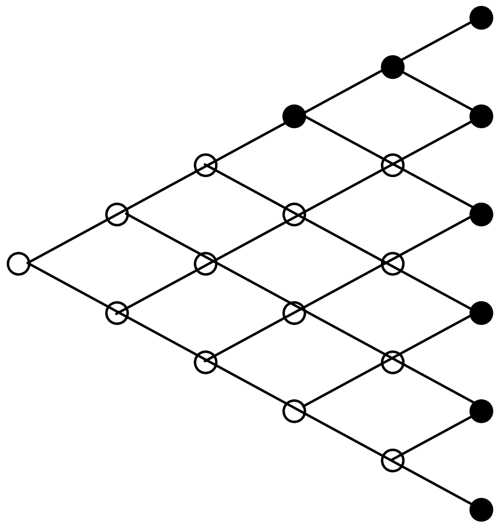

Figure 3. The figure shows how a casino can be represented as a binomial tree. Each column of nodes corresponds to a particular moment in time. Within each column, the various nodes correspond to the different possible accumulated winnings or losses at that time. A solid black node indicates that, if the agent arrives at that node, he does not gamble. At the remaining nodes, the agent does gamble.

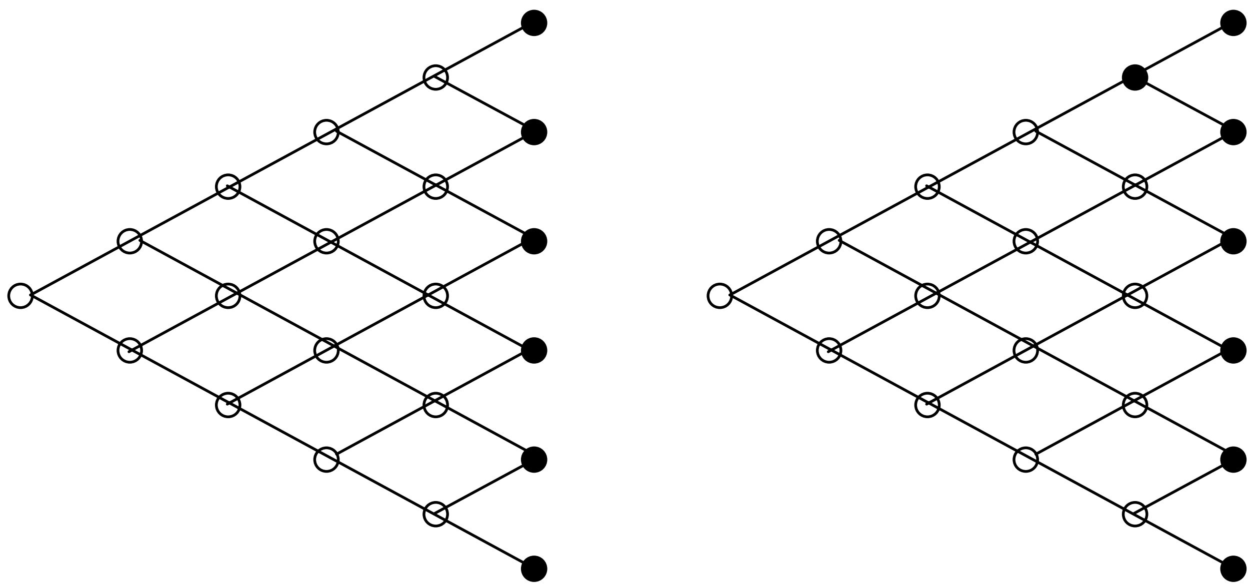

Figure 4. The figure shows two possible gambling strategies a prospect theory agent could follow. If a node has a black color, the agent plans not to gamble at that node. At the remaining nodes, he plans to gamble.

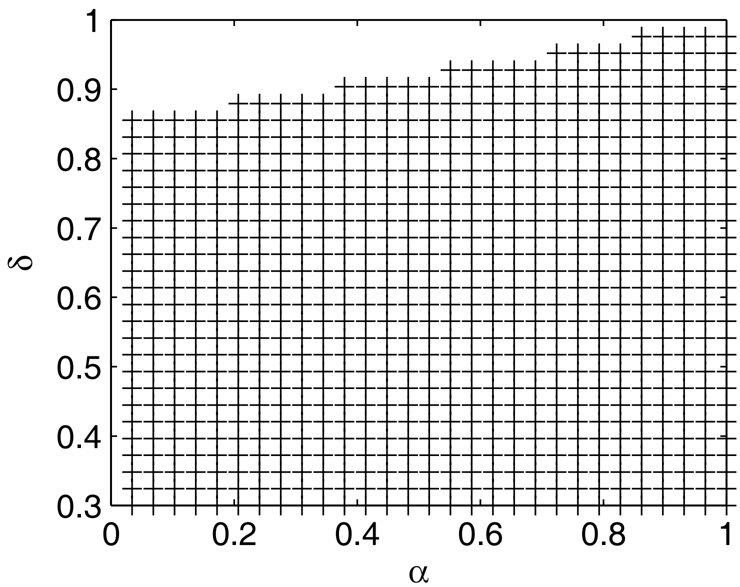

Figure 5. The figure shows the range of values of the preference parameters α and δ for which a prospect theory agent who, at time 0, is thinking about his optimal strategy at a casino, would like to keep gambling later on once he has accumulated a substantial gain. The lower the value of α, the lower the marginal utility of additional gains. The lower the value of δ, the more the agent overweights the tails of distributions.

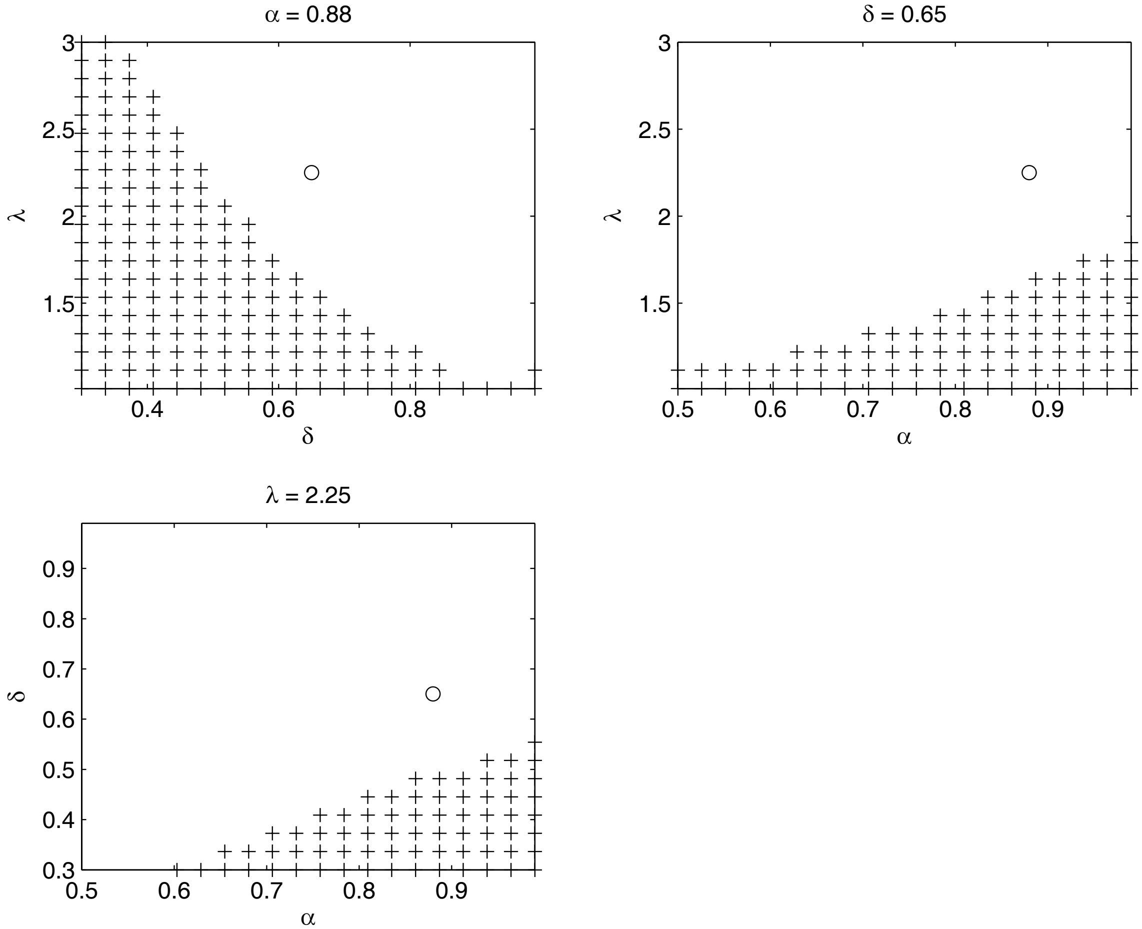

Figure 6. The "+" signs in the graphs show the range of values of the preference parameters α, δ, and λ for which an agent with prospect theory preferences would be willing to enter a casino offering 50:50 bets to win or lose a fixed amount $\$h$. The agent is naive: he does not realize that he will behave in a time-inconsistent way. Each of the three panels sets one of the three preference parameters to Tversky and Kahneman's (1992) median estimate of its value and shows the range of the other two parameters for which the agent enters the casino. The circles mark the median parameter estimates, namely $(\alpha, \delta, \lambda) = (0.88, 0.65, 2.25)$.

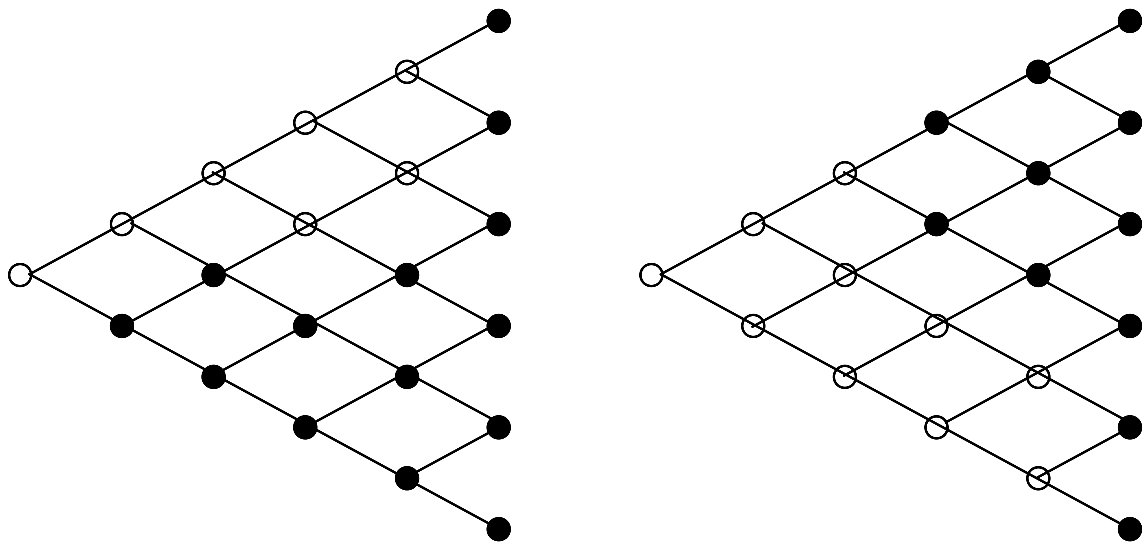

Figure 7. The left panel shows the strategy that a prospect theory agent *plans* to use when he enters a casino. The agent is naive: he does not realize that he will behave in a time-inconsistent way. If the agent arrives at a solid black node, he plans not to gamble at that node. At the remaining nodes, he plans to gamble. The right panel shows the *actual* strategy that the agent uses. If the agent arrives at a solid black node, he does not gamble. At the remaining nodes, he does gamble.

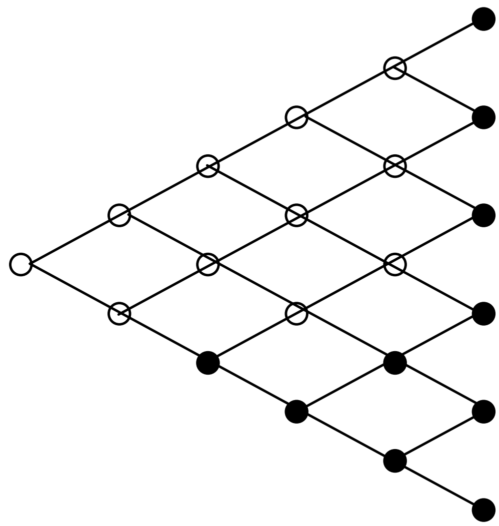

Figure 8. The figure shows a strategy under which a prospect theory agent would be willing to enter a casino and which entails a low probability of exiting in the very early rounds. The agent is naive: he does not realize that he will behave in a time-inconsistent way. If the agent arrives at a solid black node, he does not gamble. At the remaining nodes, he does gamble.

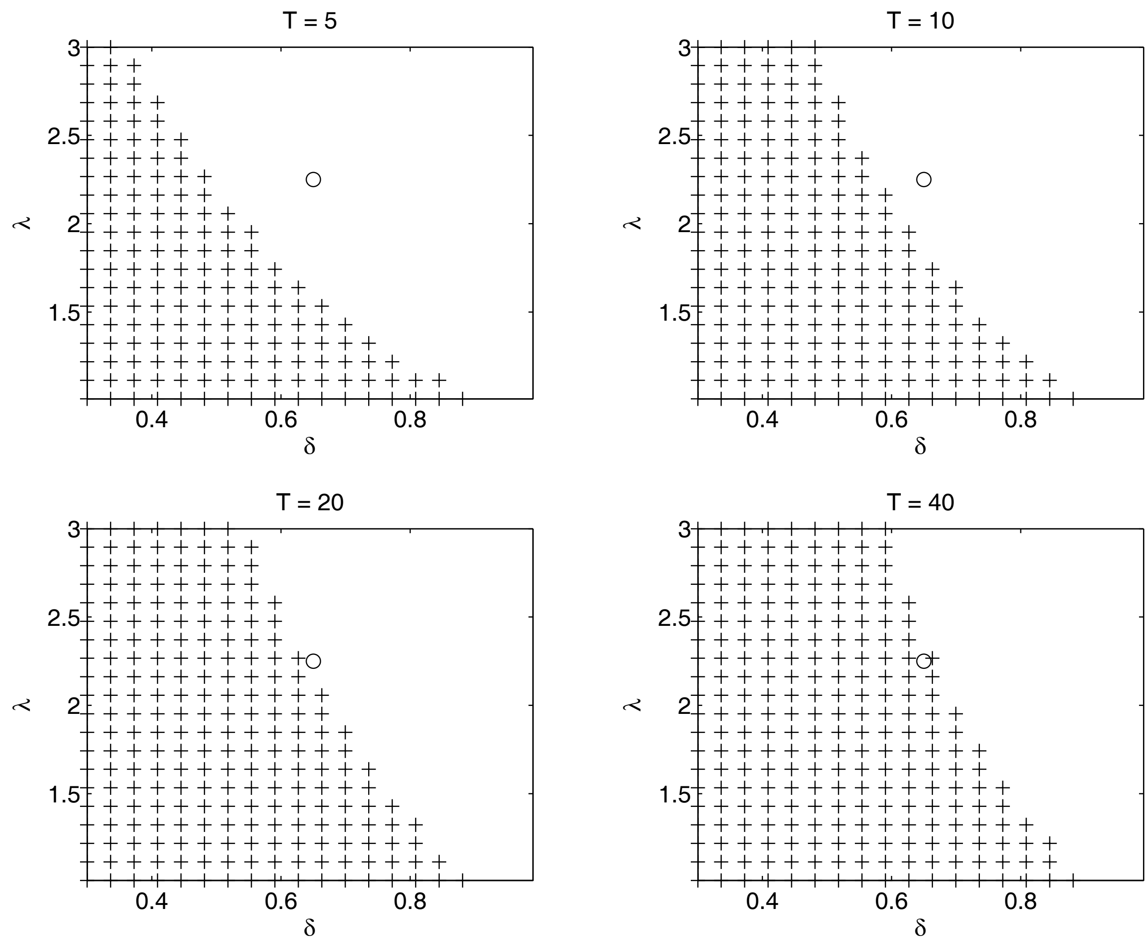

Figure 9. The “+” signs in the graphs show the range of values of the preference parameters δ and λ that satisfy a sufficient condition for an agent with prospect theory preferences to be willing to enter a casino offering 50:50 bets to win or lose a fixed amount $\$h$. The agent is naive: he does not realize that he will behave in a time-inconsistent way. The four panels correspond to four different values of T, the maximum number of rounds of gambling. In all four panels, we set the preference parameter α to 0.88. The circles mark the median parameter estimates computed by Tversky and Kahneman (1992) from experimental evidence, namely $(\alpha, \delta, \lambda) = (0.88, 0.65, 2.25)$.

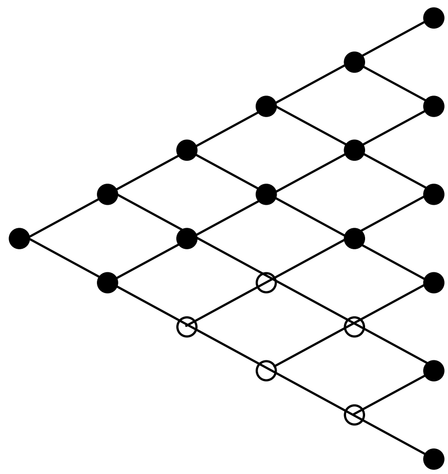

Figure 10. The figure shows the outcome of the dynamic programming procedure that a prospect theory agent uses to decide whether or not to enter a casino. The agent is sophisticated: he realizes that he will behave in a time-inconsistent way. A solid black node indicates that, if the agent were to arrive at that node, he would not gamble. If the agent were to arrive at any other node, he would gamble. The fact that the left-most node is black indicates that, in this case, the agent does not enter the casino in the first place.

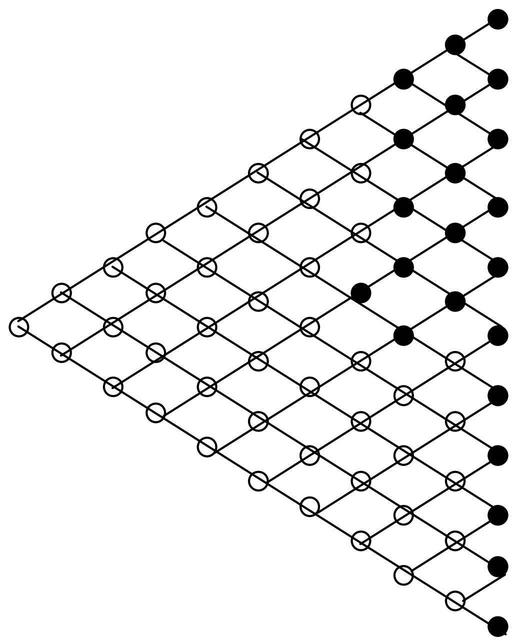

Figure 11. The figure shows how a prospect theory agent behaves after he enters a casino that offers at most ten rounds of gambling. The agent is naive: he does not realize that he will behave in a time-inconsistent way. If the agent arrives at a solid black node, he does not gamble at that node. At the remaining nodes, he gambles.

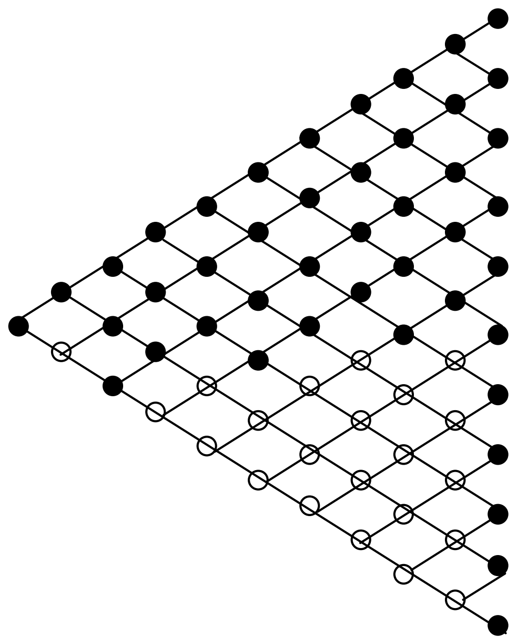

Figure 12. The figure shows the outcome of the dynamic programming procedure that a prospect theory agent uses to decide whether or not to enter a casino that offers at most ten rounds of gambling. The agent is sophisticated – he realizes that he will behave in a time-inconsistent way – but is unable to commit in advance to any particular plan of action. A solid black node indicates that, if the agent were to arrive at that node, he would not gamble. If the agent were to arrive at any other node, he would gamble. The fact that the left-most node is black indicates that, in this case, the agent does not enter the casino in the first place.

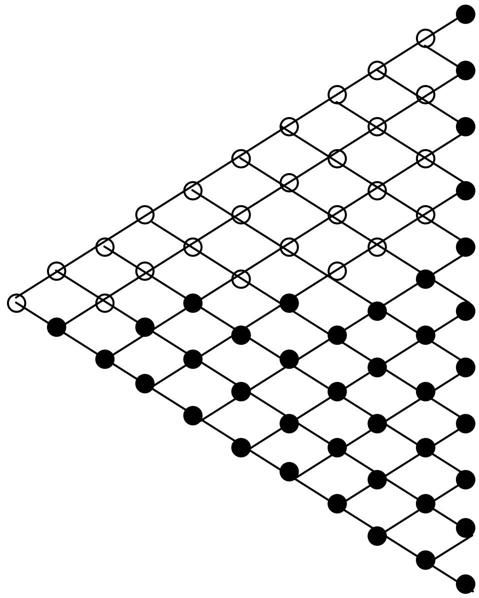

Figure 13. The figure shows how a prospect theory agent behaves after he enters a casino that offers at most ten rounds of gambling. The agent is sophisticated – he realizes that he will behave in a time-inconsistent way – but he is able to commit in advance to any plan of action he chooses. If the agent arrives at a solid black node, he does not gamble at that node. At the remaining nodes, he gambles.