DISCUSSION PAPER SERIES

DP15973

An Experimental Analysis of the Prize-Probability Tradeoff in Stopping Problems

Yair Antler and Ayala Arad

ORGANIZATIONAL ECONOMICS

CEPR

ISSN 0265-8003

An Experimental Analysis of the Prize-Probability Tradeoff in Stopping Problems

Yair Antler and Ayala Arad

Discussion Paper DP15973
Published 27 March 2021
Submitted 18 March 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

An Experimental Analysis of the Prize-Probability Tradeoff in Stopping Problems

Abstract

We experimentally examine how individuals stop risky processes such as the evolution of prices when they have commitment power. We find types who consistently choose stopping rules with large potential losses and small potential gains to induce a high winning probability (L-types), although such choices entail a considerable downside risk. A smaller proportion of types choose stopping rules with the opposite characteristics. While the latter pattern is consistent with cumulative prospect theory, the former pattern is inconsistent with prominent decision theories. We suggest that L-types solve the prize-probability tradeoff in a qualitative manner, putting more emphasis on the winning probability.

JEL Classification: N/A

Keywords: commitment, risky processes, stopping problems, type classification

Yair Antler - yair.an@gmail.com
Tel Aviv University, Coller School of Management and CEPR

Ayala Arad - aradayal@tauex.tau.ac.il
Tel Aviv University, Coller School of Management

An Experimental Analysis of the Prize-Probability Tradeoff in Stopping Problems*

Yair Antler† and Ayala Arad‡

March 2021

(First version: March 2019)

Abstract

We experimentally examine how individuals stop risky processes such as the evolution of prices when they have commitment power. We find types who consistently choose stopping rules with large potential losses and small potential gains to induce a high winning probability (L-types), although such choices entail a considerable downside risk. A smaller proportion of types choose stopping rules with the opposite characteristics. While the latter pattern is consistent with cumulative prospect theory, the former pattern is inconsistent with prominent decision theories. We suggest that L-types solve the prize-probability tradeoff in a qualitative manner, putting more emphasis on the winning probability.

Keywords: commitment, compound lotteries, downside risk, experiment, risky processes, stopping problems, type classification.

* We thank Dudi Dilenberger, Kfir Eliaz, and Ariel Rubinstein for their comments, and Tomer Baratz, Shira Bukchin, Eli Mograbi, and Hadar Weisman for research assistance. We acknowledge a research grant from The Henry Crown Institute of Business Research in Israel as well as The International Visiting Fellowship at the University of Essex.
† Coller School of Management, Tel Aviv University. yair.an@gmail.com
‡ Coller School of Management, Tel Aviv University. ayala.arad@gmail.com

1. Introduction

Stopping problems appear in numerous contexts in economics and finance, ranging from option pricing and job search to experimentation, technology adoption, and gambling. In these problems, an individual observes a sequence of realizations of a stochastic process and has to decide when to stop the process. According to several prominent theories of decision-making under risk (e.g., expected utility), optimal stopping can be described by a simple *cutoff* rule, namely, stopping the process once an individual's payoff reaches a threshold.

Remarkably, despite the apparent optimality of cutoff rules and the fact that they are quite simple and easy to describe, individuals often exhibit behavior that is inconsistent with following a cutoff rule. The dynamic nature of stopping problems triggers emotions such as regret, disappointment, and elation, which may lead to dynamically inconsistent behavior and make it difficult for individuals to implement their preferred stopping rule. For example, Strack and Viefers (2021) document history-dependent behavior that is consistent with regret aversion. Biases and departures from optimal stopping can lead to detrimental outcomes. For example, the negative feelings associated with realizing losses may lead investors to hold on to badly performing stocks (Shefrin and Statman, 1985). Moreover, these biases can make it difficult to infer individuals' ex-ante preferences from their observed behavior and hence the effect of these biases on their behavior.

A natural way to mitigate biases that arise during dynamic play is to commit to a stopping plan in advance or delegate the execution to a third party. Perhaps the most prominent manifestation of such a commitment is stop-loss and take-gain orders in stock trading. The ability to commit enables individuals to choose an optimal stopping plan without worrying about their ability to implement it. Thus, understanding how individuals choose a stopping plan when they have commitment power not only can help us understand behavior in real-world stopping problems, but it can also help us grasp the biases that arise in dynamic play, which can be a first step toward mitigating these biases.

Our main research objective is to experimentally examine how individuals make stopping plans when they have commitment power and what forces shape these plans. In particular, we shall explore whether there are circumstances in which individuals choose stopping rules with a relatively large loss and a small potential gain. This type of behavior is consistent with the disposition effect and the desire to reduce the probability of finishing the process at a loss. We also study in which circumstances individuals exhibit the opposite behavior, namely, choosing to stop the process at a relatively small loss and a large gain, which induces a small downside risk.

In order to understand our setting, consider a decision-maker (DM) who faces an infinite sequence of lotteries, where each lottery pays 1 with probability p and −1 with probability $1-p$. This binomial process can approximate the movement of real prices or the value of a stock. Under various theories of decision under risk, the DM's optimal stopping plan can be described by an upper bound $h>0$ and a lower bound $l\leq 0$ such that the DM stops the process once her payoff hits one of these bounds. The higher h is, the less likely the process is to reach h before it reaches l; the lower l is, the less likely the process is to reach l before it reaches h. Thus, when choosing these bounds, the DM trades off between the probability of winning and the size of the potential gain/loss.[1] This tradeoff is at the heart of our experimental design.

To gain intuition for the problem, consider the two cutoff rules given in Figure 1. Under both rules, the sequence stops once the DM accumulates a net loss of 20. Under a (respectively, b), the sequence stops once the DM accumulates a gain of 10 (respectively, 30). We refer to cutoff rules for which the upper bound is smaller (respectively, greater) in absolute value than the lower bound as *left-biased* (respectively, *right-biased*). The likelihood that the sequence ends at a loss is smaller under the left-biased rule a, while the potential gain is greater under the right-biased rule b. Thus, when the DM chooses between the two rules, she trades off between the potential gain and the probability of a gain. In our experiment, the participants faced these types of problems where, in each problem, the participants had to choose one rule out of five rules: two right-biased ones, two left-biased ones, and a symmetric rule.

Rule	**Lower bound**	**Upper bound**
a	-20	+10
b	-20	+30

Figure 1. Two cutoff rules with the same lower bound.

In expectation, under the left-biased rule, the DM participates in a smaller number of lotteries. Thus, when the baseline lottery is unfair (i.e., $p<0.5$), a risk-neutral/averse expected utility maximizer would choose the left-biased rule. Our participants' choices match this prediction quite well, as 63% of them consistently chose (in at least 5 out of 6 problems) left-

[1] Formally, when $p \neq 0.5$, the probability of stopping the process at a gain is $\frac{1-q^l}{1-q^{h+l}}$, where $q = \frac{1-p}{p}$.

biased rules in cases where we fixed the potential loss and varied the potential gain, whereas only 12% of them consistently chose right-biased rules. In a symmetric manner, when the baseline lottery has a positive expected value, a risk-neutral individual would choose the right-biased rule *b*. While 36% of our participants consistently chose right-biased rules in this case, remarkably, many participants (37%) consistently put a larger weight on the winning probabilities (or, according to their literal explanations, the probability of not losing) and chose left-biased rules (see Table 2).

Consider the symmetric case of an individual who has to choose between two cutoff rules that share the same upper bound as in Figure 2. As a benchmark, note that a risk-neutral expected utility maximizer will prefer the left-biased rule *d* if the baseline lottery is fair and the right-biased rule *c* if $p<0.5$. Despite the larger potential loss, roughly half of our participants consistently chose a left-biased rule in problems in which the upper cutoff was fixed: 52% when the baseline lotteries were unfair and 48% when the baseline lotteries were fair.

Rule	Lower bound	Upper bound
c	-10	+20
d	-30	+20

Figure 2. Two cutoff rules with the same upper bound.

Our main finding is a general tendency to either consistently choose left-biased rules or consistently choose right-biased ones, across qualitatively different choice problems. We find a larger proportion of participants who consistently choose left-biased rules than of participants who choose right-biased rules. The gap between these proportions is greater when the baseline lottery is unfair compared to the case of a fair baseline lottery.

Our experimental design allows us to test whether or not the above findings can be explained by various theories of decision-making under risk. We examined, for each participant, whether her behavior is consistent with prominent decision theories (expected utility theory; cumulative prospect theory, Kahneman and Tversky, 1992; disappointment aversion, Gul, 1991). Roughly speaking, we say that a participant’s behavior is consistent with a particular theory if it matches the theory’s prediction in a large share of decision problems (the precise definition appears in Section 4). We found that when facing negative value lotteries, the behavior of most of our participants cannot be explained by any of these theories.

However, when facing positive expected value lotteries, the behavior of many of our participants is consistent with cumulative prospect theory (where different degrees of loss aversion and probability distortion explain the behavior of different participants). In particular, all participants who consistently chose right-biased rules fit this theory (see Barberis, 2012; Ebert and Strack, 2015), while those who consistently chose left-biased rules cannot be explained by any of these theories.

In light of the negative findings discussed above, we look for an alternative explanation for the participants' behavior and especially for the tendency to choose left-biased rules. To this end, we first identified the participants who chose left-biased or right-biased rules in a large share of the decision problems and classified them as *L-types* and *R-types* (the share is identical to the one above). Second, we examined these individuals' ex-post explanations of their choices to acquire a better understanding of their underlying reasoning. Over 63% of the participants in our main treatment were classified as L- or R-types, where the vast majority of the types were L. Overall, L-types explained that they focused on reducing the probability of finishing the game with a loss, whereas R-types explained that they put a relatively large weight on the size of the potential gains and losses.[2] The focus of L-types on the stopping rule's induced probabilities, even at the expense of a greater potential loss (e.g., choosing rule *d* over rule *c*), is somewhat surprising as in one of our two treatments *the rules' induced odds* were not explicitly given to the participants (only the baseline lottery's odds p were given), which supposedly made the gains and losses more salient than their respective probabilities.

Our experimental design enables us to examine how each participant's behavior depends on the favorability of the drift. We find that when the drift becomes positive, the number of times an individual (whether classified as a type or not) chooses a left-biased rule weakly decreases. The participants' explanations suggest that when $p<0.5$ they focus on the probability of finishing the game with a loss, whereas in the case of $p>0.5$ they put a larger weight on the size of the potential gain.

The participants' behavior, together with their explanations, suggests that most of them try to solve a simple tradeoff between the likelihood of winning or losing and the size of the prizes. The particular way in which this tradeoff is solved depends on the favorability of the baseline lottery and on whether or not the potential loss is fixed. However, for most participants, the solution is virtually unaffected by the favorability or expected value of the

[2] Thus, while R types are consistent with cumulative prospect theory, their explanations suggest that a different reasoning underlies their behavior.

stopping rules' induced lotteries. This holds true when participants are not provided with the induced probabilities (in which case it is difficult to calculate the induced lotteries' odds), but also when they are provided with these odds. This finding suggests that the way in which individuals solve this tradeoff is qualitative, and that it is mainly affected by the way in which they perceive the favorability of the drift rather than from the fine details of each decision problem.

Related literature

The present paper is related to a recent strand of the literature that investigates planning in dynamic decision-making under risk. The closest papers to our paper are Heimer et al. (2020) and Dertwinkel-Kalt et al. (2020). Heimer et al. document a discrepancy between investors' initial plans and their actual behavior: most investors choose stopping rules that are right-biased ("loss exit") ex ante, but their subsequent choices follow the reverse pattern. In an online experiment, they show that when participants face fair lotteries and can commit to a stopping rule, the vast majority of participants choose a right-biased one. The aggregate behavior that they find is consistent with cumulative prospect theory. Dertwinkel-Kalt et al. (2020) conduct a lab experiment in which they test the predictions of Bordalo et al.'s (2012) salience theory in a stopping problem. They find that most participants plan to use a right-biased strategy (however, in their experiment, there is no commitment).[3]

Our experimental setting differs from the above papers in several important aspects. First, losses in our experiment were framed as losses rather than as lower gains (as framed in Dertwinkel-Kalt et al.) and were deducted from an endowment that was announced a week prior to the experiment. Second, the participants in our experiment were STEM and management students who are presumably more familiar with basic statistics and may have a better understanding of the implications of different stopping rules compared to the typical online and offline subject pool. These differences may explain some of the differences in the papers' findings. Finally, the most crucial difference between the papers is in the experimental design, which, in our paper, is tailored to examine individual rather than aggregate behavior. Our

[3] This behavior is consistent with Bordalo et al.'s theory when it is applied to the two attributes gains and losses. However, it is not necessarily consistent with their theory when the set of attributes includes the probability of a gain. In our experiment, the participants' explanations suggest that, indeed, gains and losses may be more salient than probabilities when lotteries are fair, but when the process is unfair, the probability of winning becomes more prominent than the actual gains.

participants made 36 decisions, where each was a choice between five given stopping rules, whereas in Heimer et al. and Dertwinkel-Kalt et al. the participants were free to design their own stopping rule and participated in a small number of such decision problems. Our design allows us to explore the participants' behavior at the individual level in different contexts, and match the diverse patterns in the data to various decision theories. Our results when the lotteries are fair are consistent with cumulative prospect theory as in Heimer et al. (2020). By contrast, when lotteries are unfair as in Dertwinkel-Kalt et al., our results are qualitatively different from those obtained in these papers and are inconsistent with prominent decision theories including expected utility theory and cumulative prospect theory.

Fischbacher et al. (2017) show that stop-loss and take-gain strategies mitigate the disposition effect. Key differences from our setting are (i) the lack of commitment power, (ii) the fact that each participant effectively chooses a stopping rule once, and (iii) the lack of knowledge of the baseline lottery's drift (i.e., p). Thus, while Fischbacher et al.'s (2017) elegant design allows them to investigate the effect of stop-loss and take-gain strategies on individuals' tendency to hold on to losing assets, it cannot be used to examine what type of rules people choose, what makes them choose these rules, the consistency of these choices, and how this consistency depends on the favorability of the underlying process.

Other papers study stopping decisions without planning. In Strack and Viefers (2021), the participants choose when to stop a multiplicative random walk and exhibit history-dependent behavior, which is consistent with regret aversion and inconsistent with cutoff rules. Sandri et al. (2010) examine exit decisions and find that most individuals tend to hold on to a badly performing asset longer than is consistent with real option reasoning. Aloui and Fons-Rosen (2017) find that grittier individuals have a higher tendency to over-gamble relative to their original plans. In their experiment, the lotteries are unfair and most of the individuals choose to play even though they are not obliged to do so. In none of the aforementioned papers, however, do the participants have commitment power. Moreover, the bulk of this strand of the literature focuses on dynamic play rather than planning (Aloui and Fons-Rosen, 2017, being the exception).

Stopping plans were studied indirectly in the experimental literature on dynamic inconsistency, which focuses on *deviations* from planning when individuals face a small number of lotteries. Barkan and Busmeyer (1999, 2003) and Ploner (2017) find evidence of dynamically inconsistent behavior in settings where individuals decide whether to participate in an additional lottery after experiencing one outcome. Cubitt and Sugden (2001) do not reject the dynamic consistency hypothesis when participants have to decide in how many all-or-nothing additional gambles to participate after winning in four mandatory rounds.

Finally, our work relates to the literature on skewness-seeking and prudent behavior. Skewness corresponds to our notion of left/right-biased stopping rules. The more right-biased a rule is, the more positively skewed its induced lottery. Golec and Tamarkin (1998) find evidence of skewness-seeking behavior in horse-race betting. Brunner et al. (2011), Deck and Schlesinger (2010, 2014), Ebert and Weisen (2011, 2014), Ebert (2015), Grossman and Eckel (2015), Maier and Ruger (2012), and Noussair et al. (2014) provide evidence for skewness-seeking and/or prudent behavior in lab experiments. Bleichrodt and van Bruggen (2018) find prudent behavior in the gain domain and imprudent behavior in the loss domain.

There are several differences between our setting and the typical setting in this strand of the literature. The experiments on skewness-seeking and prudent behavior examine choices between lotteries with identical means and variance. By contrast, the stopping rules in our setting induce compound lotteries with different means and variance such that prudence does not imply a tendency to choose right-biased rules (e.g., facing the two rules in Figure 1, a prudent individual may choose the left-biased rule when $p<0.5$ as it induces a greater expected value and a smaller variance than the right-biased rule). Moreover, reducing a stopping rule to its induced lottery is a daunting task as the participants know only the probability of winning a single baseline lottery. (For example, Halevy, 2007, establishes that even in simpler settings individuals often fail to reduce compound lotteries.) Thus, stopping problems may encourage reasoning in qualitative terms, which is less likely to be triggered in the typical setting of the literature on skewness-seeking and prudence.

The paper proceeds as follows. Section 2 presents our experimental design and Section 3 describes the results at both the aggregate and the individual levels. In Section 4, which is the heart of the paper, we classify the participants into theory-based types according to their choices. In Sections 5 and 6, we investigate the mechanisms that underlie our key findings, and Section 7 concludes.

2. Experimental Design

The experiment was carried out in the Interactive Decision-Making Lab at Tel Aviv University in April–May 2017. The participants were 114 Tel Aviv University undergraduate students in management and STEM, 44% of whom were women. The average age was 25. Recruitment of participants was done via ORSEE (Greiner, 2004).

Each participant received 55 NIS (roughly $15) at the beginning of the experiment. In an attempt to make the participants internalize this endowment, one week prior to the session, we notified them that they would receive this amount and could lose part of it (at most 30 NIS) or win an additional amount, depending on their choices in the experiment. A reminder of that was sent on the day before the session as well. The experiment included 57 computerized decision problems (we refer to these decision problems as Questions 1–57 or Q1–Q57), one of which was randomly selected at the end of the experiment to determine the payment for the participants. The amount won (or lost) in that game was added to (or subtracted from) the initial endowment. In practice, each participant could win at most an additional 45 NIS and could lose at most 28 NIS out of her initial endowment. All sessions were completed within an hour.

2.1 Detailed description of the experiment

In each session, the participants were randomly assigned to two treatments, denoted by T_0 and T_p, each with four parts, which are described below. Out of the 114 participants, 67 participants were assigned to our main treatment, T_0, and 47 were assigned to T_p. The complete questionnaire can be found in the appendix. In short, Part A (respectively, Part B) examines the choice of a stopping rule when the baseline lottery has a negative (respectively, positive) expected value, and Part C explores the participants' ability to estimate the rules' induced probabilities. Part D studies the participants' behavior in a simpler setting to identify whether their choices in Parts A and B are related to a pure taste for skewed lotteries.

Part A. In this part, participants faced a sequence of computerized lotteries, each with an 18/37 probability to win 1 NIS and a 19/37 probability to lose 1 NIS. These probabilities resemble the win/loss probability in the "Red or Black" roulette game. In each decision problem, the participants were asked to choose a cutoff stopping rule. The participants faced 18 decision problems, in each of which they chose one out of five alternative cutoff stopping rules. If one of these problems was randomly selected for payment, then the stopping rule was automatically and instantaneously implemented by the computer.

The only difference between the two treatments was that in T_p the participants were informed of the probability of ending the game with a gain given each of the five stopping rules, whereas in T_0 they were not (in both treatments the participants were informed about the winning probability in the baseline lottery). This difference allows us to examine the extent to which the choice patterns observed in T_0 were affected by the participants' knowledge of the rule's induced probability of winning in T_p. While the literature on stopping problems is not insubstantial, to the best of our knowledge the difference between these two conditions is underexplored.

In order to distinguish between different theories of decision-making under risk that may explain the participants' behavior, we considered three types of decision problems, illustrated in Figure 3. In Q1–Q6 (*fixed loss*), the participants are required to choose between five stopping rules that share the size of the loss in which the process is stopped and vary in the size of the respective gain. In Q7–Q12 (*fixed gain*), the participants are required to choose between five stopping rules that share the size of the gain in which the process is stopped and vary in the size of the respective loss. The problems Q13–Q18 (*not fixed*) vary both the gain and loss. Section 2.2 describes the predictions of the theories we examine.

In each decision problem, there are two rules in which the potential loss is greater than the potential gain, two rules in which the potential gain is greater than the potential loss, and one rule in which the potential gain and the potential loss are equal. We refer to these rules as left-biased, right-biased, and symmetric rules, respectively. We shall refer to choosing the most left-biased rule (i.e., with the largest loss and the smallest gain) as *Rule 1*, the second-most left-biased rule as *Rule* 2, the symmetric rule as *Rule* 3, the most right-biased rule (i.e., with the largest gain and smallest loss) as *Rule 5*, and the second-most right-biased rule as *Rule 4*. The five stopping rules were presented to the participants either in order from the left-biased rule with the largest loss and smallest gain to the right-biased one with the largest gain and smallest loss (as in Figure 3) or in the reverse order.[4] Thus, the five stopping rules were always ordered either from the highest probability of a gain to the lowest one or the other way around.

Part B. This part consists of 18 decision problems (Q19–Q36) and is similar in structure to Part A. The main difference between the two parts is that the probabilities of gain and loss in the baseline lottery are reversed in Part B (i.e., the probability of winning in a single lottery is 19/37). In addition, we tried to diversify the problems in Parts A and B to prevent a sense of repetition. Thus, the stopping rules in Part B are similar to the ones in Part A, yet they are not identical.

[4] The randomly selected order was used consistently throughout Parts A and B. The results suggest that the order did not affect the choices in the experiment and hence we merge the data from the two variations in the analysis.

Type (i): Fixed loss

	Loss	Gain	Probability of gain
Rule 1	-21	+9	52%
Rule 2	-21	+15	35%
Rule 3	-21	+21	24%
Rule 4	-21	+27	17%
Rule 5	-21	+33	12%

Type (ii): Fixed gain

	Loss	Gain	Probability of gain
Rule 1	-20	+12	42%
Rule 2	-16	+12	39%
Rule 3	-12	+12	34%
Rule 4	-8	+12	%28
Rule 5	-4	+12	%18

Type (iii): Not fixed

	Loss	Gain	Probability of gain
Rule 1	-27	+15	38%
Rule 2	-24	+18	31%
Rule 3	-21	+21	24%
Rule 4	-18	+24	%19
Rule 5	-15	+27	%14

Figure 3. The three types of questions in Part A. The probability of a gain given each stopping rule is provided for the reader's convenience. Only participants in T_p received information on the probability of a gain and a loss given the stopping rule, which was presented in a sentence below the description of the rule's upper and lower cutoffs (see the appendix).

Part C. This part includes three problems (Q37–Q39), where each problem presents a different stopping rule. In each of the three problems, the participants were asked to consider a baseline lottery that pays 1 with probability 18/37 and -1 with probability 19/37 (as in Part A) and estimate the probability that the game will end at a gain given the stopping rule. In particular, in the first problem, they had to gauge the probability of finishing the game with a gain of 25 given that the stopping rule is (-25, +25). The second and third problems were similar except that the stopping rules were (-25, +50) and (-25, +100), respectively. The correct answers to these three questions are roughly 20.5%, 5%, and 0.3%, respectively. The payment for each of the problems in Part C (in case one of these problems was selected for payment) was 40 NIS minus the size (in absolute terms) of the error in the participant's estimation. There was no difference between the two treatments in this part.

Part D. In this part, the participants faced 18 decision problems (Q40–Q57). In each problem they chose between two binary *lotteries* with known probabilities of loss and gain (as illustrated in Figure 4). In each problem, the two lotteries were a "mirror image" of each other (i.e., $-x$ with probability p and $+y$ with probability $1-p$ vs. $-y$ with probability $1-p$ and $+x$ with probability p), and had an expected value of 0, the same variance, the same kurtosis, but different skewness. The departure from the context of a stopping problem, as well as the fact that the lotteries had the same expected value and variance, allowed us to test whether some individuals have a "pure" taste for skewness. Such preferences may be relevant for the choice of a stopping rule in Parts A and B. In each question, the order of appearance of the two available lotteries was randomly and independently determined. There was no difference between the two treatments in this part.

Part D: Game 2

Choose your preferred lottery from the following two lotteries:

a.

chance	24%	76%
amount	-25	+8

b.

chance	76%	24%
amount	-8	+25

Figure 4. An example of a decision problem in Part D.

At the end of Parts A, B, and D, the participants were asked to explain the principles that guided them in their choices. We analyzed the participants' explanations in order to obtain a better understanding of their reasoning process.

Discussion: Choosing from a fixed set of rules

In each of the decision problems in Parts A and B, the participants chose one out of five stopping rules. Alternatively, we could have allowed them to select the potential gains and losses (i.e., the stopping rule's upper and lower bounds) from an unconstrained set. We decided to constrain their choice as it enabled us to focus on the effect of the qualitative properties of the stopping rules (e.g., the effect of the fairness of the baseline lotteries, how the choices differ given fixed loss/gain) while keeping the participants' decision problems relatively simple. The main advantage of the "constrained" problems is that they differ from one another and therefore allow us to examine whether participants use rules that share similar properties in a large number of different problems. We believe that it would be significantly more difficult to make the participants perceive numerous unconstrained problems as different from one another, especially in situations where the stopping rules' induced probabilities are not specified as in our main treatment.

Our restricted set of options resemble risk questionnaires that investment banks often use to elicit investors' preferences over investment strategies. In these type of questionnaires, individual investors often have to choose pairs of cutoffs that represent the maximal loss that they are willing to bear in a given time period and the gains that they expect to obtain in that period. In practice, investors are often given a fixed set of cutoffs to choose from rather than allowed to choose the cutoffs freely. Fixing the set of cutoffs allows the bank to categorize the investors to a manageable number of categories and implement an investment strategy suitable for each category.

2.2 Theoretical predictions

First, consider the behavior of a risk-neutral expected utility maximizer. Such an individual would try to minimize the number of lotteries in which he participates in Part A where $p<0.5$ and maximize the number of lotteries in which he participates in Part B where $p>0.5$. This implies choosing Rule 1 in Q1–Q6 and Q25–Q30 in which the loss is fixed, and choosing Rule 5 in Q7–Q12 and Q19–Q24 in which the gain is fixed.

Second, consider the behavior of a risk-averse expected utility maximizer. A very small degree of risk aversion would make the individual behave as if she were risk-neutral. However, a sufficient degree of risk aversion[5] would lead the individual to minimize the expected number of lotteries in which she participates in both parts of the experiment. This implies choosing Rule 1 in Q1–Q6 and Q19–Q24, and choosing Rule 5 in Q7–Q12 and Q25–Q30.

Third, consider the behavior predicted by cumulative prospect theory (CPT). Under this theory, individuals put a higher weight on losses than on gains and distort probabilities of gains and losses. In particular, they assign a relatively high weight to low-probability events and a relatively low weight to high-probability events. As shown in Barberis (2012) and Ebert and Strack (2015), this distortion makes right-biased stopping rules more attractive for such individuals. If this effect is not too small, it leads to choosing Rule 5 in Q1–Q12 and Q19–Q30.

We also examined the predictions of Gul's (1991) disappointment aversion (DA). When the underlying utility function is linear, this theory's predictions in the above 24 problems are similar to those of expected value maximization as long as the disappointment aversion coefficient is smaller than 2.

In all of the above problems (Q1–Q12, Q19–Q24, and Q25–Q30), the theories' predictions are monotone. That is, an individual whose behavior is consistent with one of these theories either prefers rule i to Rule $i+1$ for all $i \in \{1,2,3,4\}$ or prefers Rule i to Rule $i-1$ for all $i \in \{2,3,4,5\}$. Figure 5 summarizes this discussion.

Problem	**Risk neutrality**	**Risk aversion**	**CPT**	**DA**
Q1–Q6	$1 > 2 > 3 > 4 > 5$	$1 > 2 > 3 > 4 > 5$	$5 > 4 > 3 > 2 > 1$	$1 > 2 > 3 > 4 > 5$
Q7–Q12	$5 > 4 > 3 > 2 > 1$	$5 > 4 > 3 > 2 > 1$	$5 > 4 > 3 > 2 > 1$	$5 > 4 > 3 > 2 > 1$
Q19–Q24	$5 > 4 > 3 > 2 > 1$	$1 > 2 > 3 > 4 > 5$	$5 > 4 > 3 > 2 > 1$	$5 > 4 > 3 > 2 > 1$
Q25–Q30	$1 > 2 > 3 > 4 > 5$	$5 > 4 > 3 > 2 > 1$	$5 > 4 > 3 > 2 > 1$	$1 > 2 > 3 > 4 > 5$

Figure 5. The prediction of prominent decision theories under risk for each type of problem.

In the remaining 12 decision problems (Q13–Q18 and Q31–Q36, in which both the potential gains and the potential losses vary), the theories' predictions are more nuanced: the rules' ranking need not be monotone and may change across decision problems (this depends on the functional form and the parameters used). These problems allow us to distinguish between

[5] We allow for various degrees of risk aversion when we classify the participants into types.

theories that cannot be distinguished by the participants' choices in Q1–Q12 and Q19–Q30. In particular, the predictions for disappointment aversion and risk neutrality may differ from one another in these 12 problems even for degrees of disappointment aversion smaller than 2.

3. General Description of the Participants' Choices

We shall now focus on the main treatment, T_0, in which the participants were not provided with the rules' induced probabilities. In Section 5.2 we present the results obtained in T_p, in which the rules' induced probabilities were provided, and compare them to the results in T_0. We present the results in a sequential manner in order to avoid repetition and focus on the differences between the results in the two treatments.

Recall that, in each problem, the participants chose between five stopping rules: two left-biased ones, a symmetric rule, and two right-biased rules. The behavior in T_0 provides indications of two general patterns, both at the aggregate level and at the individual level. First, in Part A, where the baseline lottery is unfair, there is a strong tendency to choose Rules 1 and 2, namely, *left-biased* stopping rules. Second, in Part B, this tendency is weaker compared to Part A. Nonetheless, most of the participants make more left-biased choices than right-biased ones in Part B.

3.1 Aggregate-level analysis

At the aggregate level, in each part of the experiment there were 1,206 choices (67 x 18). We found that 66% of the choices in Part A were of left-biased rules and only 25% were of right-biased ones. In Part B, 46% of the chosen rules were left-biased whereas 35% were right-biased (see Table 1).

	Part A	**Part B**
Rule 1	31%	23%
Rule 2	35%	23%
Rule 3	9%	19%
Rule 4	10%	19%
Rule 5	15%	16%

Table 1. The proportions of choices in T_0, out of 1,206 (67 x 18) choices that were made in each part.

3.2 Individual-level analysis

Examining the participants' choices at the individual level reveals that many of them were consistent in their tendency to choose either left-biased rules or right-biased rules, but, at the same time, diversified between the extremely biased rule and the moderately biased rule. In both parts, A and B, the differences between the number of times the individuals chose Rule 1 and the number of times they chose Rule 2, as well as between the number of times they chose Rule 4 and the number of times they chose Rule 5, are small and insignificant.

Thus, we focus here on the choice categories of left-biased rules and right-biased rules and measure the tendency to consistently choose one of them. To this end, we consider the number of times each individual chose a left-biased rule, which ranges from 0 to 18 in each part of the experiment. We refer to this measure as the *number of left-biased choices*. In a similar manner, we consider the number of times each individual chose a right-biased rule and refer to this measure as the *number of right-biased choices*.

The number of left-biased choices is higher on average in Part A than it is in Part B, according to a paired sample t-test (11.9 vs. 8.3, $t(66)=4.44$, $p<0.001$). Figure 6a shows that the cumulative distribution of the number of left-biased choices per individual in Part A stochastically dominates the one in Part B. The number of right-biased choices is higher on average in Part B than it is in Part A (6.31 vs. 4.54, $t(66)=-2.57$, $p=0.012$). Figure 6b shows that the cumulative distribution of the number of right-biased choices per individual in Part B stochastically dominates the corresponding one in Part A.

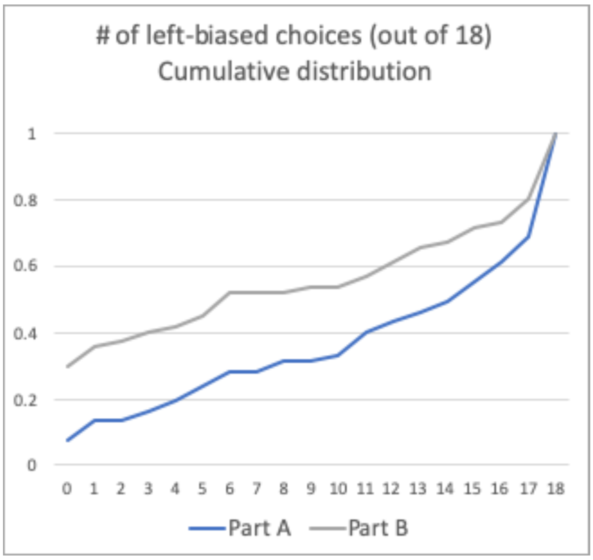

Figure 6a. Cumulative distribution of the number of left-biased choices per participant in Part A vs. Part B.

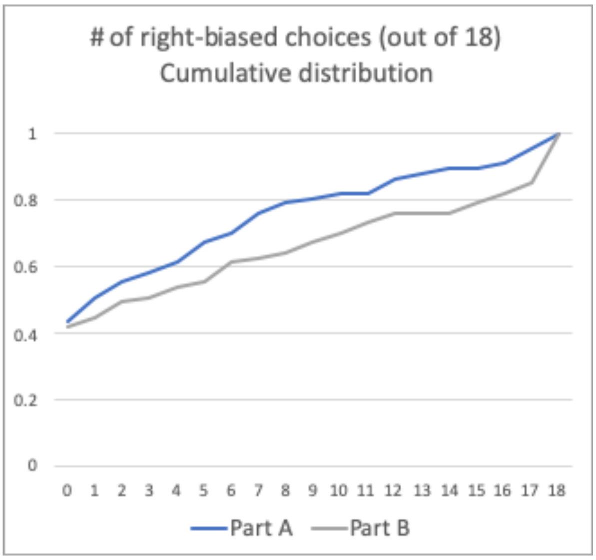

Figure 6b. Cumulative distribution of the number of right-biased choices per participant in Part A vs. Part B.

Despite the substantial differences in behavior in Parts A and B, the individuals' choices in these two parts are correlated according to the number of left-biased choices (Pearson's r=0.56, p<0.001) and according to the number of right-biased choices (Pearson's r=0.64, p<0.001). The combination of these findings suggests that there exists an individual tendency either to choose left-biased rules or to choose right-biased rules. The favorability of the baseline lottery's odds reduces the common tendency toward the left-biased rules.

3.3 Analysis of the three types of problems

Examining each of the 36 problems in Parts A and B separately suggests that left-biased choices are more prevalent than right-biased ones in all but two of them. In Part A, the median choice was 2 in all 18 of the problems and the average choice was in the range 2.06–2.87. In Part B, the median choice in most problems was 3 (and 2 in the rest) and the average choice was in the range 2.54–3.13. The range of the average choice suggests that there are some differences between the problems in the extent of choosing left-biased rules. We now examine how the type of problem (i.e., whether the loss or gain is fixed) affects the tendency to choose left-biased rules. Table 2 presents two indications of this tendency in each of the three types of problems in Parts A and B (1–6: fixed loss, 7–12: fixed gain, and 13–18: not fixed) and compares it to the tendency of choosing right-biased rules. These results establish that left-biased choices are more common in Part A than in Part B, regardless of the type of problem. Further, we find that the tendency toward left-biased choices in problems 1–6 of Part A is greater than that in problems 7–12 of Part A (the average number of left-biased choices is 4.4 vs. 3.67, $t(66)$=2.78, p=0.007), while the opposite pattern occurs in Part B (2.58 vs. 3.16, $t(66)$=-1.95, p=0.055).

Considering the aggregate data, the large proportion of choices of left-biased rules in all types of problems is inconsistent with the theories described in Section 2.2. As Table 2 suggests, and as Section 4 shows, the behavior of a large share of the participants is inconsistent with these theories. We now suggest a mechanism that may explain these results, and in Sections 4 and 5 we provide some supporting evidence for it. In a stopping problem, participants essentially choose a potential gain and a potential loss. The larger the gain is, the less likely a participant is to finish the game at a gain, and the larger the loss is, the more likely she is to finish the game at a gain. Thus, when facing a stopping problem, individuals trade off between prizes and probabilities. As this qualitative feature is intuitive and easy to grasp (as we establish in the discussion of Part C), we suggest that this tradeoff is solved in a qualitative and consistent manner, though the solution may be affected by several factors as explained below.

	Part A			Part B		
	Fixed loss	Fixed gain	Not fixed	Fixed loss	Fixed gain	Not fixed
Range of proportion of participants choosing left-biased rules	66%–79%	55%–64%	55%–69%	34%–51%	48%–54%	42%–49%
Proportion of participants choosing at least 5 (out of 6) left-biased rules	63%	52%	60%	37%	48%	37%
Proportion of participants choosing at least 5 (out of 6) right-biased rules	12%	24%	19%	36%	28%	27%

Table 2. The tendency toward left-biased rules in the three types of questions in Parts A and B of T_0. The third row presents the proportion of participants who tended to choose right-biased rules.

In Part A, left-biased rules are common in the fixed-loss problems. Intuitively, as the participants cannot control the severity of the loss, they focus on the likelihood of the loss. In the fixed-gain problems, participants can also control the size of the potential loss and trade off the probability of winning with a smaller potential loss; thus, a smaller proportion of participants choose left-biased rules. Nonetheless, the majority of participants choose left-biased rules in all types of problems in Part A. This finding suggests that when the odds are not on their side, many participants focus on reducing the probability of the potential loss rather than reducing the size of that loss.

By contrast, in Part B, where the odds are on the participants' side, it seems that the participants' focus shifts toward the potential gain.[6] We suggest that when participants face the fixed-gain problems, they focus on increasing the probability of winning (as they cannot control the size of the gain) by choosing left-biased rules. In the fixed-loss problems, they can also control the size of the potential gain and hence opt for right-biased rules more often.

In the next section, we dig deeper into the individual-level behavior, examine each participant's consistency with prominent decision theories and with the above-suggested mechanism, and look for additional insights into their reasoning by analyzing their explanations.

[6] The participants' explanations provide some indications that, in Part B, they shift attention and focus more on the potential gains rather than on the potential losses. For example, keywords were classified into the following categories: probability of winning, probability of losing, gain size, and loss size. Accounting for the use of these categories in participants' explanations suggests that the ratio of the probability of winning vs. that of losing increases in Part B (54:29 in A and 64:14 in B) as does the size of the gain vs. that of the loss (66:65 in A and 54:39 in B).

4. Theory-Based Type Classification

In this section, we examine to what extent the participants' choices can be reconciled with the theories of decision-making under risk we considered in Section 2.2: expected utility (EU), cumulative prospect theory (CPT), and disappointment aversion (DA). For each theory and each problem, we sort the five rules by the value that they induce according to the theory and say that a participant's choice is consistent with the theory if it matches one of the top two rules (for some parameter).[7] This consistency definition reflects the idea that participants who are guided by the theory may, at times, consider additional factors that are not captured by the theory. Another way to interpret our definition is to say that we allow for some structured noise in the participants' behavior.

To account for the possibility that individuals' behavior changes systematically when the baseline lottery becomes more favorable, we examine the participants' consistency with each theory according to their choices in Part A (18 choices), their choices in Part B (18 choices), and their choices in Parts A and B together (36 choices). Overall, we present three consistency tests per theory. We classify a participant as consistent with a theory in Part A (respectively, Part B) if her choices are consistent with the theory in at least 13 out of 18 problems. When considering all 36 of the problems together, we define a participant as consistent with the theory if her choices are consistent with the theory in at least 22 cases. This definition is in the spirit of Ebert and Wisen's (2011) definition of skewness-seeking. We chose these numbers so that the probability of being defined as a type under random choice is less than 1%.

Expected utility. We allowed different agents to hold different degrees of risk aversion and classified a participant as risk-averse if there is some degree of risk aversion that matches the participants' choices. We assumed a constant relative risk aversion representation ($u(x) = \frac{x^{1-\sigma}}{1-\sigma}$ for $\sigma \neq 1$ and $u(x) = \log(x)$ for $\sigma = 1$) and classified an individual as risk-averse if her choices were consistent with the predictions under this functional form for some $\sigma \in [0,2]$. We obtained similar results for other functional forms (e.g., constant absolute risk aversion) and for other ranges of parameters. According to this criterion, we classified 15% of the 67 participants as risk-averse in Part A, 9% of the participants as risk-averse in Part B, and 13% of them as risk-averse when considering the 36 choices combined.

[7] If instead we were to define a participant's choice as consistent with the theory only if it matches the highest-valued rule, then the absolute performance of the various theories would change substantially. Nonetheless, the relative performance of these theories would be similar.

Cumulative prospect theory. To classify the participants, we used a piecewise-linear value function

$$u(x) = \begin{Bmatrix} -\lambda x & for & x < 0 \\ x & for & x \geq 0 \end{Bmatrix}$$

and, consistent with Kahneman and Tversky (1992), the following probability weighting function

$$w(p) = \frac{p^{\delta}}{(p^{\delta}+(1-p)^{\delta})^{\frac{1}{\delta}}}.$$

A participant was classified as consistent with CPT if her choices were consistent with some parameters $0.3 < \delta \leq 1$ and $1 \leq \lambda \leq 3$. This wide range of parameters includes the estimated parameters in Kahneman and Tversky (1992). Modifying this range or allowing for diminishing sensitivity to gains and losses did not change our results substantially. We were able to classify as consistent with CPT 24% of the 67 participants in Part A, 39% of the participants in Part B, and 42% when considering all 36 of the problems.

Disappointment aversion. According to this theory, an outcome creates disappointment if it is worse than the certainty equivalent of the lottery and it creates elation if it is better than the certainty equivalent. Formally, we used Gul's (1991) representation with a linear utility function. Thus, a stopping rule that induces $x_1 < 0$ with probability q_1 and $x_2 > 0$ with probability $1 - q_1$ yields the function

$$\left(1 + \frac{(1-q_1)\beta}{1+q_1\beta}\right) q_1 x_1 + \left(1 - \frac{q_1\beta}{1+q_1\beta}\right)(1-q_1)x_2.$$

To capture disappointment aversion, we set $\beta \geq 0$ (smaller values of β do not capture disappointment aversion). We classified a participant as disappointment-averse if her choices were consistent with some parameter $2 \geq \beta \geq 0$. We were able to classify 15% of the 67 participants as disappointment-averse in Part A, 13% of the participants as disappointment-averse in Part B, and 28% of them as disappointment-averse when considering all 36 problems.

Table 3 summarizes the above findings. The findings suggest that in Part A, only a small share of the participants exhibit behavior that is consistent with the prominent theories we considered, whereas in Part B, CPT is consistent with the behavior of a large share of the participants. A natural question that arises is, what can explain the behavior of the rest of the participants? In what follows, we suggest a new decision rule that may explain their behavior, analyze it, and reconcile it with the participants' explanations. We conclude by identifying, for each participant, the decision rule(s) most suitable to describe her behavior.

N=67	Part A	Part B	Overall
EU	15% (10)	9% (6)	13% (9)
CPT	24% (16)	39% (26)	42% (28)
DA	15% (10)	13% (9)	28% (19)

Table 3. The proportion and the number (in parentheses) of participants who exhibit behavior consistent with each of the three prominent decision theories. The proportion is estimated for each part of the experiment and overall.

4.1 A qualitative resolution of the prize-probability tradeoff

In Section 3, we noted an individual tendency to either consistently choose left-biased rules or consistently choose right-biased rules. Left-biased rules may reflect a qualitative resolution of the prize-probability tradeoff that emphasizes the winning probability. Right-biased rules may reflect a resolution that focuses on the potential of obtaining a high prize. We now perform a similar classification exercise to the one above to test the extent to which the mechanism we suggested can explain the data. In line with our above definition of consistency with a theory, we say that a participant is an L-type in Part A (respectively, Part B) if she chooses one of the two left-biased rules in at least 13 out of 18 problems, and say that she is an R-type if she chooses one of the two right-biased rules in at least 13 out of 18 problems. When considering all 36 problems, we say that a participant is an L-type (respectively, R-type) if she chooses left-biased (respectively, right-biased) rules in at least 22 out of 36 problems. Participants who could not be classified as L or R were classified as "other."

We found that in both parts, types are common (63%–70%). The proportion of L-types in Part A (57%) is significantly higher than the respective proportions in Part B (39%), and the proportion of R-types in Part A (13%) is significantly lower than the respective proportions in Part B (24%), according to a McNemar test (p=0.023 and p=0.039, respectively). Nonetheless, a participant's type in Part A is correlated with her type in Part B (Spearman's rho=0.5, p<0.001). In particular, an L-type in Part A is unlikely (only 13%) to be classified as an R-type in Part B and vice versa for an R-type in Part A (none of the R-types was classified as an L-type in Part B; some types switched to "other" in Part B). Considering all 36 of the choices, 48% of the participants are classified as L-types and 18% of them as R-types. Thus, according to each of the three classification exercises, about two-thirds of the participants are classified as types, where the majority of them are L-types.

How far from random choice are the patterns observed in the data? Figure 7 (respectively, Figure 8) shows the experimental distribution of the number of left-biased (respectively, right-biased) choices in Parts A and B, next to the probability of observing each number given a binomial process with 0.2 probability of choosing each of the five rules in each problem. Figure 9 displays these distributions for Parts A and B together. The graphs illustrate the polarization of choices in the experiment.

Figure 7. The percentage of participants with each number of left-biased choices in Parts A and B and the probability of observing each number (per participant) given a binomial process with 0.4 probability of a left-biased choice in each problem.

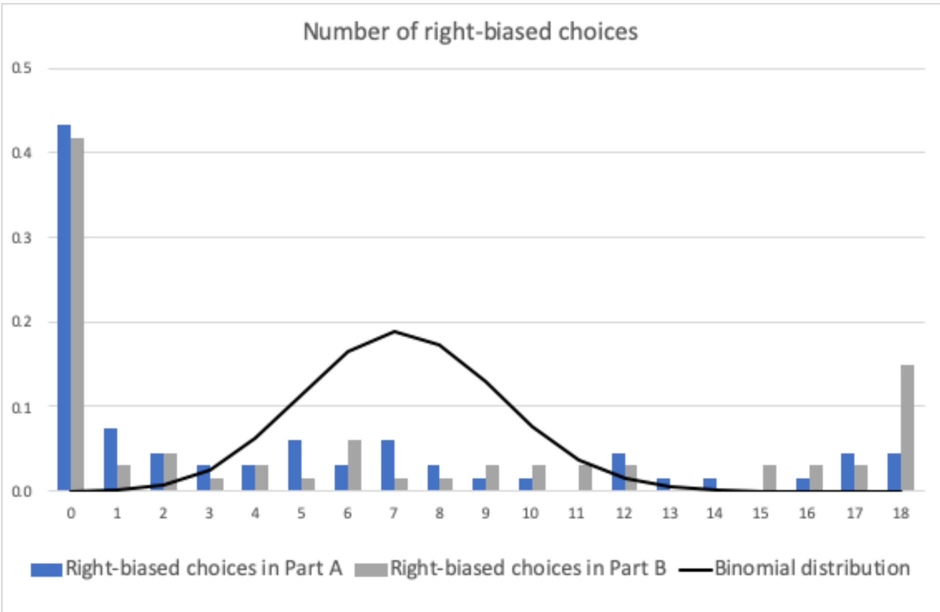

Figure 8. The percentage of participants with each number of right-biased choices in Parts A and B and the probability of observing each number (per participant) given a binomial process with 0.4 probability of a right-biased choice in each problem.

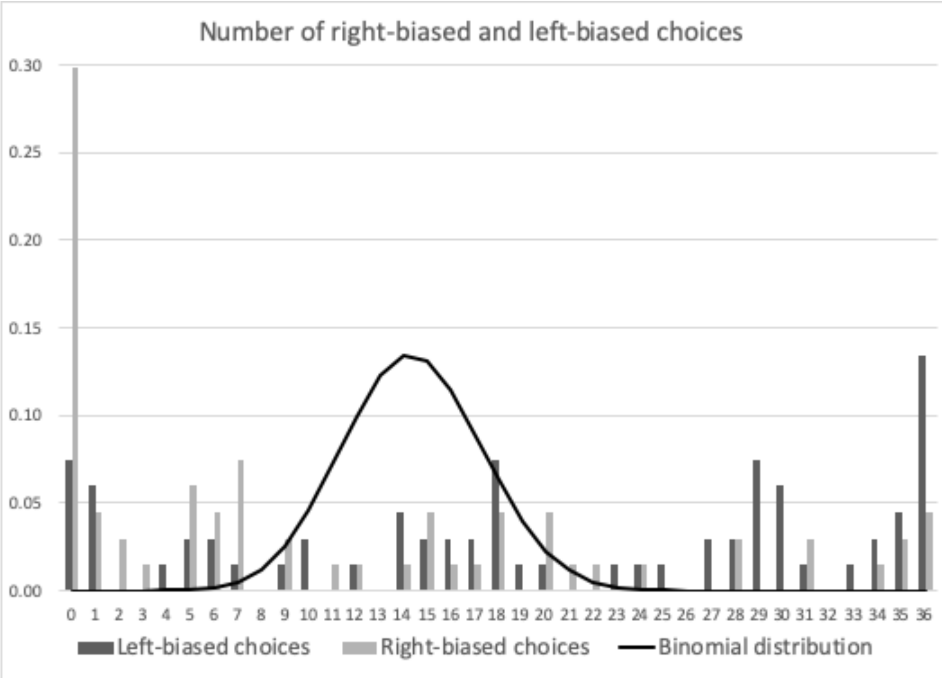

Figure 9. The percentage of participants with each number of right-biased and left-biased choices in Parts A and B together and the probability of observing each number (per participant) given a binomial process with 0.4 probability of a right- (left-)biased choice in each problem.

Consistency with the participants' explanations

At the end of each part, the participants were asked (i) how they would guide someone else to play on their behalf and (ii) what the main considerations in their choices were. To support our interpretation of the decision procedure of L- and R-types, two research assistants examined the explanations' fit with our classification of types using the following protocol. First, each assistant examined the explanations separately and classified them into four categories: (1) Literal – the explanation included a description of choosing consistently right/left-biased stopping rules, (2) Partial – the explanation included a description of a general rule of right/left-bias but mentioned exceptions in which the tradeoff is solved differently, (3) Inconsistent – the explanation contradicted the type identification, and (4) NA – the text did not provide enough information to determine whether the participant's type identification was consistent with her explanation of her reasoning. Subsequently, the assistants met to decide on cases in which their classifications disagreed (their independent classifications were identical for almost all of the participants).

The assistants' final classification suggests that only a very small number of participants whom we identified as types described a decision procedure that was inconsistent with their type classification. In fact, the vast majority included a literal description of a right or left tendency. In particular, in Part A, only **one out of the 69 L- and R-types** (overall in the two treatments) was classified as inconsistent, while six other types were classified as NA. Similarly, in Part B, **two out of 63 types** were classified as inconsistent, while eight other types were classified as NA.

The following are examples of participants' explanation in T_0 in all four categories:

Participant 57, Part A (L-type) – Literal: *I chose, each time, the option in which the loss was higher than the gain so that the chances of reaching a gain would be higher. Nevertheless, I chose the option in which the gap between the loss and the gain was relatively small. For example, a gain of 6 and a loss of 15.*

Participant 45, Part A (L-type) – Partial: *The probability that you win increases if the negative stopping rule [loss] is high and the positive stopping rule [gain] is low: you need to play safe but not too safe … because there are cases in which the conditions do not pay off.*

Participant 89, Part A (L-type) – Inconsistent: *I tried to obtain a maximal gain given that a loss was most likely, even if it meant risking in a loss.*

Participant 4, Part A (L-type) – NA: *I chose according to the probability of winning relative to the probability of losing.*

Participant 54, Part B (L-type) – Literal: *The instructions are almost identical [to Part A]. The probability of winning is a bit higher than in the previous part and hence I took a chance such that the stopping rule's gain and loss would be closer [compared to Part A] but the loss would still be further away from 0 compared to the gain in order to increase the chance of winning.*

Participant 65, Part B (R-type) – Literal: *The instructions are similar to those of Part A, just the opposite since the probability of winning is higher. Thus, I looked for the smallest bias toward the "positive [gain]."*

In conclusion, many of the participants' explanations focused on a qualitative resolution of the tradeoff between the probability of a gain/loss and the size of the potential gain/loss. Most of the participants who were classified as types (either L or R) explicitly explained their decision of consistently choosing what we refer to as left/right-biased stopping rules. They rarely mentioned calculations related to variance or expected value/utility.

4.2 Conservative classification

In the previous section, we showed that a qualitative resolution of a prize-probability tradeoff can account for the behavior of a large number of our participants. A natural question that arises is how many of the participants that we classified as R-types or L-types exhibit behavior that is also consistent with the other theories we examined. For such individuals, it is essential to understand whether, based solely on their choices, the classification as an L/R-type is more suitable than the conventional theories we examined. Similar questions arise for participants who were classified as consistent with one or more of the other theories that we examined. In this section, we answer these questions.

We now perform a more *conservative* classification into types that corresponds to each of the theories that we examined as well as to the qualitative mechanism suggested above. For each participant, we first count the number of matches to each theory and identify those that she is consistent with, as defined above. Among these theories: (1) we consider the theory with the highest number of matches, and denote that number by X, and (2) we say that the participant is a type according to any theory for which the number of matches is at least X-2. The results of this exercise are summarized in Table 4 (the number of matches to each theory per participant appears in the appendix).

Table 4 confirms that in the conservative classification, a large number of participants are classified solely as L-types in Part A. Thus, their behavior could not be explained by the other theories that we examined. In Part B, we find a smaller number of such participants. When looking at the 36 choices overall, we find again that a large share of the participants are classified as L-types and cannot be explained by any other theory.

N=67	**Part A**	**Part B**	**Overall**
CPT	1% (1)	1% (1)	15% (10)
CPT & R	10% (7)	24% (16)	12% (8)
CPT & (DA /EU)	12% (8)	12% (8)	7% (5)
L	57% (38)	36% (24)	46% (31)
DA /EU	-	1% (1)	-
CPT & L & (DA /EU)	-	1% (1)	1% (1)
No Type	19% (13)	24% (16)	18% (12)

Table 4. The proportion and the number (in parentheses) of participants in T_0 who exhibit behavior consistent with each of the decision theories and who are classified as L- or R-types, according to the *conservative* classification. The proportion is estimated for each part of the experiment and overall.

The other prominent explanation of the participants' behavior is CPT, which accounts for the behavior of 23%–38% of the participants. Importantly, these are participants who are not classified as L (more than 1/3 of them are classified as conservative R-types). Thus, CPT and the qualitative resolution of the prize-probability tradeoff in favor of probabilities together can explain the behavior of 74%–82% of the participants.

5. The Absence of the Rules' Probabilities and Its Implications

In this section, we examine to what extent the choices of the participants in the main treatment were affected by not knowing the rules' induced winning probabilities. Not knowing the induced probabilities should have no effect if the participants can infer these probabilities from the likelihood of winning a single baseline lottery. Thus, the first step of the analysis must examine the participants' ability to make such an inference. Part C of the experiment explores this question and shows that the participants' inferences are very far from the true winning probabilities (consistent with Gneezy, 1996). In the second part of this section, we present the results of our second treatment, in which the induced probabilities were explicitly given to the participants. A comparison of the two treatments sheds light on the effects of the unknown probabilities on the participants' behavior.

5.1 Can the participants infer the rules' induced probabilities? (Part C)

In each of the three problems in Part C, we presented the participants with a stopping rule. The rules were (-25,+25), (-25,+50), and (-25,+100) in the first, second, and third problems, respectively. The participants were asked to assess the rules' induced winning probabilities given that the probability of winning a single baseline lottery is 18/37, as in Part A. The correct induced winning probabilities were 20.5%, 5%, and 0.3%, respectively.

The participants' average estimates in T_0 were 39.6%, 24.3%, and 17.4%. The expected errors in absolute terms were 23.2%, 20.6%, and 17.4%. Moreover, only 26.8% of the answers were within a range of 5% from the correct answer (e.g., an estimate of 15.6%–25.6% in the first problem in Part C).[8] While most of the participants failed to estimate the winning probabilities correctly, they did exhibit a qualitative understanding of the prize-probability tradeoff, where 86.8% of them provided monotone estimates (an estimate is monotone if the estimate for (-25,+25) is weakly greater than the estimate for (-25,+50) and the latter is weakly greater than the estimate for (-25,+100)). The fact that the vast majority of the participants failed to estimate the induced winning probabilities provides additional motivation for our investigation of a treatment in which the participants were provided with the rules' induced probabilities.

---

[8] In T_p, where the participants observed the probabilities of the stopping rules in Parts A and B, the average estimates in Part C were 37 %, 22.7%, and 12.3%, and the average error size slightly decreased in all three questions, though the reduction in error size was marginally significant only in the first problem (error difference=3.83, $t(112)=1.89$, $p=0.062$). Only 29.8% of the answers in T_p were within a range of 5% from the correct answer.

Our findings in Part C complement Gneezy's (1996) findings, which relate to positive expected value lotteries. He finds that individuals use the stage-by-stage probability as an anchor and adjust insufficiently: estimations are biased toward the direction of the single lottery probability, resulting in an underestimation of the overall probability of winning. The combination of these findings and our results can have significant implications for situations in which processes are perceived to be "almost fair." It could lead to over-optimism and over-participation in situations where the baseline drift is slightly negative (e.g., casino gambling) and over-pessimism and under-participation in situations where the baseline drift is slightly positive (e.g., stock market trading).

5.2 Known vs. missing induced probabilities (T_p vs. T_0)

We shall start with a brief description of the behavior in T_p, in which the participants were provided with the stopping rules' induced probabilities of winning and losing. Subsequently, we shall compare between the two treatments.

At the aggregate level, the behavior patterns that are exhibited by the T_p participants are similar to the ones observed in T_0. First, when the baseline lottery is unfair, there is a tendency to prefer left-biased stopping rules to right-biased ones. Second, this tendency is weaker when p>0.5. We found that in Part A, 62% of the 846 choices (47x18) are of left-biased rules and 28% are of right-biased ones, whereas in Part B, 49% of the choices are of left-biased rules and 37% are of right-biased ones. It should be noted that the number of left-biased choices in Part A is greater than the corresponding number in Part B and the number of right-biased choices in Part A is less than the corresponding number in Part B (see Table 5).

	Part A	**Part B**
Rule 1	44%	32%
Rule 2	18%	17%
Rule 3	11%	15%
Rule 4	10%	13%
Rule 5	18%	24%

Table 5. The proportions of choices in T_p out of the 846 choices (47*18) that were made in each part.

At the individual level, the mean number of choices of left-biased rules in Part A is higher than that in Part B, according to a paired sample t-test (11.09 vs. 8.74, t(46)=2.96, p=0.005).

Nonetheless, the participants' choices in Part A and Part B are correlated according to the "number of left-biased choices" measure (Pearson's r=0.62, p<0.001).

Extreme vs. moderate choices. In contrast to T_0, the differences between the number of choices in a particular direction (i.e., choices of Rule 1 vs. Rule 2 and choices of Rule 5 vs. Rule 4) are significant. In particular, participants tended to choose the extreme stopping rules more often. Rule 1 was chosen more than Rule 2: in Part A the difference is 4.7, $t(46)$=4.14, p<0.001, whereas in Part B the difference is 2.7, $t(46)$=2.97, p=0.005. Rule 5 was chosen more than Rule 4: in Part A the difference is 1.49, $t(46)$=-2.13, p=0.039, whereas in Part B the difference is 1.96, $t(46)$=-2.99, p=0.005. Thus, the uncertainty over the induced lotteries in T_0 mitigated the individuals' extreme choices.

Table 6 presents two measures of the participants' tendency to choose left-biased rules in each of the three types of problems and compares it to the tendency to choose right-biased rules. The results suggest that left-biased choices are common in T_p as well. Furthermore, we observed a "mirror" pattern similar to that observed in T_0: in Part A the tendency toward left-biased choices in fixed-loss problems is greater than this tendency in fixed-gain problems (the average number of left-biased choices is 4.87 vs. 3.02, $t(46)$=4.8, p<0.001), whereas in Part B the reverse tendency occurs (1.96 vs. 3.94, $t(46)$=-5.09, p<0.001).

	Part A			**Part B**		
	Fixed loss	Fixed gain	Not fixed	Fixed loss	Fixed gain	Not fixed
Range of proportion of participants choosing left-biased rules	68%–87%	43%–57%	47%–55%	23%–45%	57%–75%	40%–57%
Proportion of participants choosing at least 5 (out of 6) left-biased rules	77%	40%	45%	23%	64%	38%
Proportion of participants choosing at least 5 (out of 6) right-biased rules	2%	26%	23%	47%	15%	15%

Table 6: Two measures of a tendency toward left-biased rules in the three types of questions in Parts A and B of T_p. The third row presents the proportion of participants who tended to choose right-biased rules.

The above observations suggest that the patterns of behavior in T_p are similar to those found in T_0. A comparison of the two treatments indicates that in both parts, our measures of the number of the individual's left-biased or right-biased choices are not significantly different between the treatments. Furthermore, there are no significant differences between the treatments in the number of left-biased or right-biased choices in any of the six types of questions.[9]

Table 7 presents the classification into conservative types in T_p. It appears that the main difference between the two treatments is that in T_p a somewhat larger share of the participants can be classified according to CPT, while a smaller share of the participants can be classified as R or L (though the differences between the treatments are not significant). A possible interpretation is that when the participants' decision problems are more "standard" in the sense that they know the probabilities of gains and losses, they are better able to recognize situations in which the tradeoff between prizes and probabilities is relatively small and adjust their choices accordingly. For example, in situations where a minor deduction of a winning probability leads to a major increase in prizes, they tend more to opt for right-biased rules. Thus, a larger share of them behave in a manner that is consistent with CPT.

N=47	Part A	Part B	Overall
CPT	11% (5)	2% (1)	9% (4)
CPT & R	0% (0)	15% (7)	11% (5)
CPT & (DA /EU)	19% (9)	34% (16)	26% (12)
L	45% (21)	23% (11)	36% (17)
DA /EU	2% (1)	2% (1)	0% (0)
CPT & L & (DA /EU)	0% (0)	4% (2)	6% (3)
No Type	23% (11)	19% (9)	13% (6)

Table 7. The proportion and the number (in parentheses) of participants in T_p who exhibit behavior consistent with each of the decision theories and who are classified as L- or R-types, according to the conservative classification. The proportion is estimated for each part of the experiment and overall.

In conclusion, although the behavior in T_p is not identical to that in T_0, the observed patterns are quite similar. Thus, knowing the stopping rules' induced probabilities had a minor effect on the participants' behavior. This finding suggests that, in this context, individuals use a

[9] The only significant difference in behavior between the treatments is the tendency mentioned above of choosing more extreme stopping rules (i.e., rules 1 and 5 are more common than 2 and 4).

qualitative decision procedure even when the probabilities are known: they consistently focus either on a relatively high winning probability or on the potential of obtaining a relatively high prize.

6. Directional Bias in a Simple Context (Part D)

In Part D, the participants faced 18 problems, in each of which they had to choose between a pair of binary lotteries. In each pair, the two lotteries had an expected value of zero, the same variance, and the same kurtosis. The key difference between each pair of lotteries was that one lottery was negatively skewed (left-biased) and the other was positively skewed (right-biased).[10] We chose the prizes in the lotteries to reflect two stopping rules with a baseline lottery's winning probability of 0.5. The participants' decisions in this part of the experiment are simpler than those in Parts A and B in two main dimensions: the winning probabilities are given and the lotteries are not presented as stopping rules. Equating the lotteries' features (except for their skewness) and simplifying the problem enables us to better understand the participants' preference for skewed prospects and connect it to their choices between stopping rules in T_0.

At the aggregate level, 49% of the choices in Part D are of negatively skewed lotteries. In almost all 18 of the problems, the distribution of choices is quite balanced: between 40% and 60% of the choices are of negatively skewed lotteries, where the most extreme frequency of choices of a negatively skewed lottery is 71% (in the first problem in Part D). At the individual level, the number of choices of negatively skewed lotteries (which range from 0 to 18) is, on average, 8.75, and its median is 8. Among the 114 participants in the two treatments, we can classify 31% of the participants as L-types and 38% as R-types by the definitions that we used in Parts A and B. Despite the slightly different choice pattern, in T_0 the number of choices of negatively skewed lotteries in Part D correlates with the number of left-biased choices in Parts A and B (Pearson's r=0.23, p=0.06 and Pearson's r=0.32, p=0.009, respectively). Similarly, the number of choices of positively skewed lotteries in Part D correlates with the number of right-biased choices in Parts A and B (Pearson's r=0.23, p=0.06 and Pearson's r=0.33, p=0.007, respectively). In T_p, there is a higher correlation between the behavior in Part D and that in Parts A and B.

[10] Our notion of biasedness corresponds to the skewness of the stopping rules' induced lotteries: the more right-biased a rule is, the greater the skewness of its induced lottery is. It should be noted that a left-biased rule can induce a positively skewed lottery when p<0.5 and a right-biased rule can induce a negatively skewed lottery when p>0.5.

In conclusion, the participants did not exhibit a taste for negatively skewed lotteries to the extent that could explain the strong general tendency to choose left-biased rules in Parts A and B. Nonetheless, the significant correlation suggests that a pure preference for negative or positive skewness is related to the tendency to choose left- or right-biased stopping rules. In general, the participants' choices become more balanced when we depart from the context of stopping problems and ambiguous winning probabilities. One possible interpretation of this finding is that the context of a stopping rule and the unknown probabilities encourage the participants to overweight the winning probabilities in instances where the true differences in probabilities (between the five stopping rules) are rather small. This interpretation is also consistent with the results of Part C, namely, with a tendency to estimate the rules' induced lotteries as if the baseline lottery's winning probability were closer to 0.5 than it really is.

Comment: Relation to the literature on skewness-seeking and prudence

The literature on skewness-seeking and prudence documents a taste for positively skewed lotteries. Typically, the proportion of positively skewed choices ranges between 60%–80%. The results in Part D are closer to the results in that strand of the literature than the results in Parts A and B. Nonetheless, the proportion of positively skewed choices is still lower than it is in the literature. The lower proportion of positively skewed choices may result from the different lotteries that we used (i.e., the two lotteries were mirror images of one another, which possibly emphasized the direction bias) and from order effects (Part D was played after Parts A and B).

7. Conclusion

We examined individuals' preferences over stopping rules when they have commitment power. Our main finding is that many individuals tend to trade off between the size of the prize and the probability of winning in a consistent manner, either in favor of right-biased stopping rules or in favor of left-biased stopping rules. The participants' choice patterns depend on the favorability of the baseline lottery and the tendency to choose left-biased rules cannot be explained by prominent theories of decision under risk. The tendency to choose stopping rules that induce a relatively large winning probability at the cost of taking a large downside risk is somewhat counterintuitive as the winning probabilities are relatively difficult to calculate, which could have made the prizes more salient than the probabilities.

Our analysis suggests that many individuals use qualitative decision procedures even when the stopping rules' induced probabilities are known. These individuals consistently focus either on the winning probability or on the size of the potential gains and losses. More generally, our results provide indications of qualitative reasoning: individuals think in relative terms and are not responsive to a decision problem's fine numerical details. An interesting direction for future research would be to examine whether this type of reasoning arises in stopping problems in other contexts, such as job search and experimentation in R&D.

References

1. Alaoui, L. and Fons-Rosen, C. (2017): "Know When to Fold 'em: The Grit Factor," *Barcelona GSE Working Paper*.
2. Barberis, N. (2012): "A Model of Casino Gambling," *Management Science*, 58, 35–51.
3. Barkan, R. and Busemeyer, J. (1999): "Changing Plans: Dynamic Inconsistency and the Effect of Experience on the Reference Point," *Psychonomic Bulletin & Review*, 6, 547–554.
4. Barkan, R. and Busemeyer, J. (2003): "Modeling Dynamic Inconsistency with a Changing Reference Point," *Journal of Behavioral Decision Making*, 16, 235–255.
5. Bordalo, P., Gennaioli, N., and Shleifer, A. (2012): "Salience Theory of Choice under Risk," *Quarterly Journal of Economics*, 127(3), 1243–1285.
6. Brunner, T., Levínský, R., and Qiu, J. (2011): "Preferences for Skewness: Evidence from a Binary Choice Experiment," *The European Journal of Finance*, 17, 525–538.
7. Cubitt, R. and Sugden, R. (2001): "Dynamic Decision-Making under Uncertainty: An Experimental Investigation of Choices between Accumulator Gambles," *Journal of Risk and Uncertainty*, 22, 103–128.
8. Dertwinkel-Kalt, M., Frey, J., and Koster, M. (2020): "Optimal Stopping in a Dynamic Salience Model," Mimeo.
9. Deck, C. and Schlesinger, H. (2010): "Exploring Higher-Order Risk Preferences," *Review of Economic Studies*, 77, 1403–1420.
10. Deck, C. and Schlesinger, H. (2014): "Consistency of Higher-Order Risk Preferences," *Econometrica*, 82, 1913–1943.
11. Ebert, S. (2015): "On Skewed Risks in Economic Models and Experiments," *Journal of Economic Behavior and Organization*, 112, 85–97.
12. Ebert, S. and Strack, P. (2015): "Until the Bitter End: On Prospect Theory in a Dynamic Context," *American Economic Review*, 105, 1618–1633.
13. Ebert, S. and Wiesen, D. (2011): "Testing for Prudence and Skewness Seeking," *Management Science*, 57, 1334–1349.

14. Ebert, S. and Wiesen, D. (2014): "Joint Measurement of Risk Aversion, Prudence, and Temperance," *Journal of Risk and Uncertainty*, 48, 231–252.
15. Fischbacher, U., Hoffmann, G., and Schudy, S. (2017): "The Causal Effect of Stop-Loss and Take-Gain Orders on the Disposition Effect," *Review of Financial Studies*, 30, 2110–2129.
16. Gneezy, U. (1996): "Probability Judgments in Multi-stage Problems: Experimental Evidence of Systematic Biases," *Acta Psychologica*, 93, 59–68.
17. Golec, J. and Tamarkin, M. (1998): "Bettors Love Skewness, Not Risk, at the Horse Track," *Journal of Political Economy*, 106, 205–225.
18. Greiner, B. (2004): "An Online Recruitment System for Economic Experiments," MPRA Paper, University Library of Munich, Germany.
19. Grossman, P. and Eckel, C. (2015): "Loving the Long Shot: Risk-taking with Skewed Lotteries," *Journal of Risk and Uncertainty*, 51, 195–217.
20. Gul, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59, 667-686.
21. Halevy, Y. (2007): "Ellsberg Revisited: An Experimental Study," *Econometrica*, 75, 503–536.
22. Heimer, R., Iliewa, Z., Imas, A., and Weber, M. (2020): "Dynamic Inconsistency in Risky Choice: Evidence from the Lab and Field," Mimeo.
23. Maier, J. and Rüger, M. (2012): "Experimental Evidence on Higher-Order Risk Preferences with Real Monetary Losses," Mimeo.
24. Menezes, C., Geiss, C., and Tressler, J. (1980): "Increasing Downside Risk," *American Economic Review*, 70, 921–932.
25. Noussair, C. N., Trautmann, S. T., van de Kuilen, G. (2014): "Higher Order Risk Attitudes, Demographics, and Financial Decisions," *Review of Economic Studies*, 81(1), 325–355.
26. Plonner, M. (2017): "Hold on to It? An Experimental Analysis of the Disposition Effect," *Judgment and Decision Making*, 12, 118–127.
27. Sandri, S., Schade, C., Muβhoff, O., and Odening, M. (2010): "Holding on for Too Long? An Experimental Study on Inertia in Entrepreneurs' and Non-entrepreneurs' Disinvestment Choices," *Journal of Economic Behavior & Organization*, 76, 30–44.
28. Shefrin, S. and Statman, M. (1985): "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence," *Journal of Finance*, 40, 777–790.
29. Strack, P. and Viefers, P. (2021): "Too Proud to Stop: Regret in Dynamic Decisions," *Journal of the European Economic Association,* 19, 165–199.
30. Tversky, A. and Kahneman, D. (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.