

1401

Discussion
Papers

Too Proud to Stop: Regret in Dynamic Decisions

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2014

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

Too Proud to Stop: Regret in Dynamic Decisions*

Paul Viefers[†] Philipp Strack[‡]

Abstract

Many economic situations involve the timing of irreversible decisions. E.g. People decide when to sell a stock or stop searching for a better price. We analyze the behavior of a decision maker who evaluates his choice relative to the ex-post optimal choice in an optimal stopping task. We derive the optimal strategy under such regret preferences, and show how it is different from that of an expected utility maximizer. We also show that if the decision maker never commits mistakes the behavior resulting from this strategy is observationally equivalent to that of an expected utility maximizer.

We then test our theoretical predictions in the laboratory. The results from a structural discrete choice model we fit to our data provide strong evidence that many people's stopping behavior is largely determined by the anticipation of and aversion to regret.

JEL-Classification: D03, C91

Keywords: Optimal stopping, Dynamic behavior, Regret.

*This version: July 17, 2014. We thank Martin Cripps, Paul Heidhues, Markus Mobius, Klaus Schmidt, Tomasz Strzalecki, and Georg Weizsäcker for helpful comments and discussions. We also like to thank seminar participants at Berlin, LMU Munich, Laval, EEA Gothenburg and Florence. The former title of this paper was "A Bird in the Hand is Worth Two in the Bush: On Choice Behavior in an Optimal Stopping Task". This paper is part of Viefers' dissertation submitted to the Humboldt University, Berlin. Part of this paper was written while Viefers was a guest at Microsoft Research and their hospitality is gratefully acknowledged. We also thank Jörn Hees for the excellent experimental software and the ERC for financial support (Starting Grant 263412).

[†]DIW Graduate Center, Mohrenstr. 58, D-10117 Berlin, Germany (email: pviefers@diw.de).

[‡]University of California Berkeley, Office 513, Evans Hall, Berkeley, 94720 California, USA.

1. Introduction

”My father invested some money in the stock market when the index was 190. A few months later the index jumped to 240, and then, after some time, fell to 200. Knowing my father I urged him to sell his position, telling him that the market was too risky at the moment. My father replied: To sell now, after I lost 40 points? No way!”

– from Gneezy (2005), p. 420

This paper is concerned with theory and behavior in stopping problems. Hence, consider the following setting: An agent observes a sequence of offers, $X = (X_1, X_2, \dots)$, which are the realizations of some stationary stochastic process. After observing the t -th offer, the agent has to decide whether to continue and thus forego the current offer or to stop and seize it. In the former case, she observes the next offer and faces the same decision again. In the latter case, the agent’s decision to stop is irreversible and she receives a net payoff $X_t - K$, where $K > 0$ is known and fixed. In considering which stopping rule is optimal, the agent has to weigh the immediate gains from stopping at X_t , against the loss of the option to stop at higher values in the future.

We address two fundamental questions in this setting. First, what do theories of dynamic behavior predict people to do in stopping problems? Toward that end, we generalize some existing results about optimal stopping under expected utility, but also provide new results under different alternative behavioral theories. Second, how well do the predicted stopping rules describe actual behavior? To answer this, we ran a laboratory experiment based on the design by Oprea et al. (2009) and tested several key properties of our predictions.

Testing stopping theory is important, because many different decision-making contexts in finance, economics and statistics are modelled as stopping problems. Most prominently in economics, optimal stopping theory has been applied to model job search frictions in macroeconomics (see *inter alia* Stigler, 1962; McCall, 1970; Rogerson et al., 2005), but also to model irreversible investment-, option pricing- and market entry decisions (McDonald and Siegel, 1986; Jacka, 1991; Dixit and Pindyck, 1994) or the optimal sample size for sequential hypothesis testing (Wald, 1945). While stopping theory is widely applied, the vast majority of the theoretical literature, let alone the experimental literature, almost exclusively considers agents with expected utility (EU) preferences, or even more narrowly: risk neutral agents.

Under EU, the optimal stopping rule turns out to be simple: stop as soon as the

payoff process hits a certain reservation level. Otherwise wait. This strategy, which we call a cut-off strategy, comprises two important properties. First, it is a reservation-level strategy. That is, the agent has a unique payoff reservation level that makes it optimal for her to seize the option which is independent of the path of X_t that lead there. Second, the agent behaves time-consistently, i.e. the process is stopped *the first time* it reaches this reservation level.

Our first contribution is to show that this prediction holds under a more general notion of EU than previously documented. The way we generalize the notion of EU relative to the literature, is by relaxing a standard monotonicity assumption to a single-crossing assumption on the Bernoulli utility function of an agent. While our single-crossing assumption holds for any strictly increasing and concave utility, the monotonicity assumption does, e.g., not necessarily hold under constant absolute risk aversion.

Second, under the single-crossing assumption we show that cut-off strategies are not only optimal for a larger class of utility functions, but also if one considers off-equilibrium optimal strategies, i.e. subgame perfect strategies. That is, cut-off strategies are not only optimal under a wider class of utility functions, but are also robust to erroneous deviations. Even if an agent reaches a point she never planned to reach ex ante, she will not reconsider her subsequent plan.

However, the gain from weakening this assumption is even more substantial. Since the single-crossing assumption does not require utility to be concave or differentiable, we are able to extend our results to value functions or gain-loss utility (Kahneman and Tversky, 1979), where utility is defined over gains and losses relative to a *fixed* reference point and agents are potentially risk-averse over gains and risk-seeking over losses.¹

When it comes to actual behavior in the lab and in the field, there is much anecdotal and empirical evidence, however, that playing cut-off strategies is not what people actually do. Instead, it is usually found that individuals' decisions are influenced by the history of events. For example, it is a well-established fact in the behavioral finance literature that individual investors have a disposition to sell winning stocks and ride losing stocks (Shefrin and Statman, 1985; Lakonishok and Smidt, 1986; Weber and Camerer, 1998; Grinblatt and Keloharju, 2000). Hence, investors' willingness to stop and sell an asset at a given price depends on prior gains and losses relative to, e.g., the purchasing price of the asset. In a similar vein, Post et al. (2008) find that subjects' choice behavior

¹Even though, the terms *value function* and *gain-loss utility* are often used interchangeably to describe an agent whose utility is defined over final wealth relative to a fixed and known reference point, we will henceforth use the term *gain-loss utility* to avoid confusion with the term *value function* found in the dynamic programming literature.

in a large-payoff TV game show is significantly less risk-averse after a series of unfavorable events, than after a series of favorable events. Gneezy and Potters (1997); Gneezy et al. (2003); Haigh and List (2005) obtain similar results in the laboratory.

Motivated by these findings, several theoretical papers have sought to rationalize the observed deviations from EU. The majority of papers argues that a version of Kahneman and Tversky's prospect theory preferences, adapted to a dynamic context, is able to better explain observed behavior in stopping problems (Barberis and Xiong, 2009; Henderson, 2012). Nonetheless, adaption of prospect theory preferences to dynamic contexts like ours remains a delicate issue and, especially in combination with probability weighting, is found to easily run into conceptual and theoretical problems (see e.g. Hens and Vlcek, 2011; Ebert and Strack, 2012).

There are several alternative forces conceivable that are likely to play in explaining the observed deviations from expected utility. For example, Shefrin and Statman discuss, via intuition and example, how a combination of prospect theory, regret aversion, mental accounting and self-control problems may rationalize the disposition effect. Our third contribution is to derive the optimal stopping rule for an agent who exhibits regret aversion. That is, we derive the optimal strategy of an agent whose utility is defined over final wealth relative to what would have been *ex post* optimal (Loomes and Sugden, 1982).² In our setting this implies a dynamic reference point at the past maximum of X_t . Even though this renders regret preferences time-inconsistent, we prove that the optimal strategy of such an agent is a cut-off strategy, i.e. time-consistent.

This does not necessarily imply that EU, gain-loss preferences and regret preferences are behaviorally equivalent. While this is true in terms of the *ex ante* or minimal optimal strategy, we show it does not hold if one considers also off-equilibrium behavior. Loosely speaking, if we admit the possibility that agents do not implement their *ex ante* strategy, their *ex post* behavior is predicted to be different. Specifically, agents with EU and gain-loss preferences will behave time-consistently and *not* reconsider their behavior *ex post*. We show that a regret agent, however, will reconsider her plan and likely raise the *ex-post* cut-off to be equal to the past maximum of the payoff process.³

²Also Harry Markowitz admitted in an often-cited quote in the January issue of the magazine *Money* in 1998 that his actual portfolio choice is largely dictated by anticipated regret: "I should have computed the historical covariance of the asset classes and drawn an efficient frontier. Instead I visualized my grief if the stock market went way up and I wasn't in it or if it went way down and I was completely in it. My intention was to minimize my future regret, so I split my [pension scheme] contributions 50/50 between bonds [risk-free assets] and equities [risky assets]."

³This provides a formal proof for the conjecture made by Gneezy (2005) about the behavior of an agent with reference point at the historical peak of the payoff process.

In our experiment, we adopt the design by Oprea et al. (2009) and give subjects the option to stop a geometric random walk. This is arguably the most standard stopping setting found in economic applications and used widely to describe, e.g., investment settings (Dixit and Pindyck, 1994). We find that subjects' stopping decisions vary substantially across 65 repetitions of the same stopping task and that the majority of decisions are not time-consistent. There seems to be a strong stochastic component in subjects' behavior. In order to account for this, we then use our closed-form expressions for the incentive to stop or continue to fit a dynamic discrete choice model to the data. Such a structural model is not only interesting because it is able to capture the stochastic component in our data, but also because it renders the regret model testable. At this point, the distinction between ex-ante and the subgame-perfect optimal strategy proves to be more than a purely academic subtlety, but in fact crucial to explain observed behavior in the lab. Given the stopping decisions of subjects, we are able to use the likelihood principle to assess whether subjects systematically base their stopping decisions on the past maximum of the process. That is, whether if subjects reach a point above the ex-ante optimal cut-off, but below the running maximum of the process, they will not want to stop, but wait for the process to return to its previous maximum. In fact, this is a choice pattern we observe to be abundant in our data and consequently a regret model performs significantly better in explaining subjects' decisions.

The paper is organized as follows. In section 2, we give an overview over the related literature. In section 3, we describe the choice setting we consider. In section 4, we present our model of expected utility and regret and derive the optimal strategy under both preferences and discuss testable predictions. In section 5, we describe how we implemented our setting in the laboratory. Section 6 presents our empirical results. Section 7 finally concludes.

2. Related Literature

The theoretical literature on optimal stopping is vast and too large to give an exhaustive overview. Peskir and Shiryaev (2006) provide an overview over the mathematical literature on optimal stopping and free-boundary problems in discrete and continuous time, whereas Dixit and Pindyck (1994) provide an overview of applications in finance and economics. Most prominently, stopping models have been applied to model job search on labor markets (see *inter alia* Stigler, 1962; McCall, 1970; Rogerson et al., 2005), but also to settings where firms face an irreversible investment or market-entry decision

(McDonald and Siegel, 1986; Dixit and Pindyck, 1994).

More recently, there has been growing interest in the implications of non-standard preferences for optimal stopping rules or more generally dynamic behavior. For example, Epstein and Schneider (2003); Nishimura and Ozaki (2007, 2004); Riedel (2009); Cheng and Riedel (2013) consider optimal stopping rules of agents who are ambiguity-averse. Hens and Vlcek (2011); Henderson (2012); Ebert and Strack (2012) consider settings where agents have gain-loss utility or prospect theory preferences à la Kahneman and Tversky (1979).

While much attention has been devoted to the impact of regret on decision making in the psychology (Gilovich et al., 1998; Zeelenberg et al., 1998, 2000; Gilbert et al., 2004) and neuroscience literature (Camille et al., 2004; Coricelli et al., 2005, 2007), the choice theoretic literature that investigates the theoretical prediction under regret – especially in a dynamic context – is still fairly infant.⁴ To the best of our knowledge the only papers that take regret theory to a dynamic setting are due to Krähmer and Stone (2012) and Hayashi (2009, 2011). While Krähmer and Stone (2012) treat the case of a finite-horizon choice problem under regret à la Loomes and Sugden, Hayashi (2009, 2011) is concerned with commitment and consistency issues in a finite-horizon stopping problem where the agent has minimax-regret preferences in the sense of Wald-Savage.

Compared to the theoretical literature, the experimental literature on optimal stopping is relatively small. The first experimental papers due to Rapoport and Tversky and others, contrast the theoretical predictions from sequential search models with individual behavior in the laboratory (see *inter alia* Rapoport and Tversky, 1966; Kahan et al., 1967; Rapoport and Tversky, 1970; Seale and Rapoport, 1997, and references therein). These experiments test theoretical predictions made by models of sequential search. In a similar vein, there is a strand of experimental literature putting a focus on testing implication for the particular class of job search models in the laboratory along various dimensions, e.g. Schotter and Braunstein (1981); Cox and Oaxaca (1989, 1992, 2000) and Brown et al. (2011).

Oprea et al. (2009) choose an experimental design, aimed at replicating the investment setting motivated in Dixit and Pindyck (1994). In their paper, subjects observe a geometric binomial random walk and have the option to earn the current value of the

⁴Nonetheless, there is a growing literature that considers the impact of emotions and counterfactual thinking on choice behavior. For example, see Rabin (2004); Battigalli and Dufwenberg (2007, 2009) or Bordalo et al. (2012). Examples for applications of minimax-regret preferences to economics are Bergemann and Schlag (2011) who consider monopoly pricing, Filiz-Ozbay and Ozbay (2007) who consider auctions and Linhart and Radner (1989) who consider bargaining under minimax-regret.

random walk less some fixed cost, or forego it in favor of future values. We replicate their setup in this paper. They find that subjects approximate the risk-neutral optimal strategy surprisingly well. While in three out of four treatments with different parameters for the evolution of the random walk, subjects stop too early, stopping decisions are nearly optimal in one treatment. Moreover, Oprea et al. show that subjects adjust their reservation levels in response to forgone earnings, i.e. regret associated with their stopping decision in the current round, leads them to reconsider their strategy in the coming rounds.

3. The Setting

Time is discrete and indexed by $t \in \{0, 1, \dots\}$. The agent observes a sequence X_0, X_1, \dots of realizations of a multiplicative binomial random walk. For a given starting value $X_0 > 0$, future values of X_t are drawn according to the transition rule

$$X_{t+1} = \begin{cases} h X_t & \text{with probability } p \\ \frac{1}{h} X_t & \text{with probability } 1 - p \end{cases}.$$

We call $h > 1$ the step size and $p \in [0, 1]$ the uptick probability. We denote by $\mathcal{X} = \{h^k X_0 : k \in \mathbb{Z}\}$ the set of possible states of the process X_t . At the end of any period t there is a fixed exogenous probability $1 - \delta \in [0, 1]$ that the game ends and the agent receives a payoff of zero. We denote by $T \geq 0$ the random time the game ends. At any time $t < T$ before the game ended the agent observes the realization of the random walk X_t and decides whether to ‘continue’ or to ‘stop’.

If the agent chooses to stop in period t , she receives the value of the random walk X_t minus a constant transaction cost $K > 0$, such that her material pay-off equals $X_t - K$. Without loss of generality we assume that $K \leq X_0$.⁵ After an agent decided to stop, she continues to observe the realization of the process until the process jumps to zero in period T .

If the agent chooses to continue, the game ends with probability $1 - \delta$ and the agent gets a payoff of zero. With probability δ , the game does not end in period t , but period $t + 1$ starts and the agent observes the next realization of the random walk X_{t+1} .

⁵Otherwise it is always optimal for the agent to at least wait until he reaches K or the game ends.

The expected gain from stopping in period $t + 1$ instead of period t equals

$$\delta (\mathbb{E} [X_{t+1} | X_t = x] - K) - (X_t - K) = x (\delta (ph + (1 - p)h^{-1}) - 1) + (1 - \delta)K .$$

If $\delta(ph + (1 - p)h^{-1}) \geq 1$ this gain in expected payoff positive for all $x \in \mathcal{X}$ and an expected value maximizing agent never stops. This strategy however leads to a payoff of zero and thus no optimal strategy exists. The following assumption ensures that the expected value maximizing strategy is always well defined.

Assumption A1. *To make sure the problem is well-posed for the risk-neutral agent, we assume that $\delta(ph + (1 - p)h^{-1}) < 1$.*

4. Theories of Dynamic Behavior

In this section, we derive our theoretical predictions under different theories of dynamic behavior. The two broad classes of preferences we consider are (i) expected utility and (ii) minimax regret preferences.

For each class, we describe the underlying model assumptions and derive key properties about predicted behavior. It turns out that both theories predict that agents will use cut-off strategies. A cut-off strategy, is a strategy which satisfies the following definition

Definition 1 (Cut-off Strategy). *The cut-off strategy $\tau(b)$ prescribes that the agent stops at time t if the value of the process X_t exceeds the cut-off b and continues otherwise. That is*

$$\tau(b) = \min\{t \geq 0 : X_t \geq b\} . \tag{4.1}$$

It is worth to stress that a cut-off strategy comprises two important features. First, it is a reservation level strategy, i.e. it says that there exists a unique level at which it is optimal to stop for the agent. Second, it postulates that the process is stopped the *first* time the process reaches this level. This second property essentially is a time-consistency property. A strategy that satisfies both, the reservation-level property and the time-consistency property is what we call a cut-off strategy. A violation of the second property, i.e. stopping at a point X_t with $X_t < \max_{s \leq t} X_s$, indicates time-inconsistent behavior and cannot be rationalized by expected utility. What is surprising, however, is that it is also not rationalizable by certain classes of path-dependent preferences. We demonstrate this for the case of regret preferences below.

4.1. Expected Utility

An expected utility agent evaluates outcomes according to the strictly increasing (and not necessarily concave) utility function $u : [-K, \infty) \rightarrow \mathbb{R}$. Denote by $\mathbf{1}_A$ the indicator function that takes the value one on the event A and zero otherwise. The EU agent then chooses the stopping time τ that maximizes

$$\mathbb{E} [\mathbf{1}_{\{\tau < T\}} u(X_\tau - K) + \mathbf{1}_{\{\tau \geq T\}} u(0) \mid X_0 = x] . \quad (4.2)$$

Because preferences over stopping times are invariant under additive translations of the utility u , we can without loss of generality assume $u(0) = 0$. To shorten notation we denote conditional expectations by

$$\mathbb{E}_{t,x} [\cdot] = \mathbb{E} [\cdot \mid X_t = x, T > t]$$

$\mathbb{E}_x[\cdot] = \mathbb{E}_{0,x}[\cdot]$ and conditional probabilities by $\mathbb{P}_{t,x}[\cdot] = \mathbb{P}[\cdot \mid X_t = x, T > t]$ and $\mathbb{P}_x[\cdot] = \mathbb{P}_{0,x}[\cdot]$. Moreover, we introduce $V(\tau, x)$ as the expected utility of the agent when she uses the strategy τ and the initial value of the process is x

$$V(\tau, x) = \mathbb{E}_{t,x} [\mathbf{1}_{\{\tau < T\}} u(X_\tau - K)] .$$

Let us denote by $V^* : \mathcal{X} \rightarrow \mathbb{R}$ the value of the agent when he uses the optimal strategy

$$V^*(x) = \sup_{\tau} V(\tau, x) .$$

The following lemma proven in the Appendix establishes a probability theoretic result that will be useful to derive the optimal strategy.

Lemma 1 (Probability to Stop before the Deadline). *When using the cut-off strategy $\tau(b)$ as a continuation strategy at a given level $X_t = x$, the probability of stopping before the game ends, $\tau(b) < T$, is given by*

$$\mathbb{P}_{t,x} [\tau(b) < T] = \begin{cases} \left(\frac{x}{b}\right)^\alpha & \text{for all } b \geq x \\ 1 & \text{else} \end{cases} ,$$

where α is given by $\alpha = \frac{1}{\log(h)} \log \left(\frac{1}{2p\delta} + \sqrt{\frac{1}{4p^2\delta^2} - \frac{1-p}{p}} \right) > 1$.

As a consequence of Lemma 1 the expected utility from using the cut-off strategy $\tau(b)$

as a continuation strategy from $x \leq b$, equals

$$\begin{aligned} V(\tau(b), x) &= \mathbb{E}_{t,x} [\mathbf{1}_{\{\tau(b) < T\}} u(X_{\tau(b)} - K)] = \mathbb{P}_{t,x} [\tau(b) < T] u(b - K) \\ &= \left(\frac{x}{b}\right)^\alpha u(b - K). \end{aligned}$$

At any point $x > b$, the cut-off strategy $\tau(b)$ stops immediately and therefore

$$V(\tau(b), x) = \begin{cases} \left(\frac{x}{b}\right)^\alpha u(b - K) & \text{for } x \leq b \\ u(x - K) & \text{for } x > b \end{cases}. \quad (4.3)$$

If the agent decides to stop at a point x his payoff equals $u(x - K)$ if she decides to continue until either the process reached xh or the game ended he gets an expected payoff of

$$v(\tau(hb), x) = h^{-\alpha} u(xh - K).$$

Definition 2. We denote by $\Gamma : \mathcal{X} \rightarrow \mathbb{R}$ the expected gain from waiting until the process reached xh instead of stopping at x

$$\Gamma(x) = h^{-\alpha} u(xh - K) - u(x - K).$$

Γ describes the expected gain from waiting until the process makes one uptick. The following lemma shows that the gain from any other cut-off strategy can be expressed in terms of Γ .

Lemma 2 (Expected Payoff of a Cut-off Strategy). *The expected gain from using the cut-off strategy $\tau(xh^n)$ instead of stopping at x is given by*

$$V(\tau(xh^n), x) = u(x - K) + \sum_{j=1}^n h^{-(j-1)\alpha} \Gamma(xh^j).$$

Proof. We show the result inductively using the fact that once the agent reaches xh^{n-1} the continuation value is given by the expected value of waiting for one uptick

$$\begin{aligned} V(\tau(xh^n), x) &= \mathbb{E}_x [\mathbf{1}_{\{\tau(xh^n) < T\}} u(xh^n - K)] = \mathbb{E}_x [\mathbf{1}_{\{\tau(xh^{n-1}) < T\}} V(\tau(xh^n, xh^{n-1}))] \\ &= \mathbb{E}_x [\mathbf{1}_{\{\tau(xh^{n-1}) < T\}} (\Gamma(xh^{n-1}) + u(xh^{n-1} - K))] \\ &= V(\tau(xh^{n-1}), x) + \mathbb{P}_x [\tau(xh^{n-1}) < T] \Gamma(xh^{n-1}) \\ &= V(\tau(xh^{n-1}), x) + h^{-(n-1)\alpha} \Gamma(xh^{n-1}). \end{aligned}$$

The result follows inductively in combination with the fact that $V(\tau(x), x) = u(x - K)$. \square

Define the point $b^u \in \mathcal{X}$ as the smallest point such that it is not profitable to wait until the process reaches $b^u h$, i.e.

$$b^u = \min\{x \in \mathcal{X} : \Gamma(x) \leq 0\}.$$

By Definition of b^u it is never optimal to stop below b^u . If $\Gamma(b^u) = 0$ the agent is indifferent between stopping at b^u and waiting for one more uptick. Then, $\tau(b^u)$ can not be the unique optimal strategy.⁶

Definition 3 (Expected Change). *For every function $w : \mathcal{X} \rightarrow \mathbb{R}$, we denote by $\mathcal{L}w : \mathcal{X} \rightarrow \mathbb{R}$ the expected change in w from period t to period $t + 1$, conditional on being at x*

$$\begin{aligned} \mathcal{L}w(x) &= \mathbb{E}_{t,x} [\mathbf{1}_{\{t+1 < T\}} w(X_{t+1} - K) - w(X_t - K)] \\ &= \delta (p w(xh - K) + (1 - p) w(uh^{-1} - K)) - w(x - K). \end{aligned}$$

The following assumption ensure that the optimal strategy always stops above b^u .

Assumption A2 (Single Crossing). *The expected change in utility $\mathcal{L}u(x - K)$ is negative for all $x > b^u$.*

Assumption A2 ensures that stopping immediately is better than continuing and stopping in the next period for all $x > b^u$. The next Lemma shows Assumption A2 is a necessary condition for optimal strategies to be cut-off strategies

Lemma 3. *If Assumption A2 is violated and u is concave no optimal strategy is a cut-off strategy.*

It can be shown that if u is not concave and Assumption A2 is violated at least one optimal strategy is not a cut-off strategy. We say that an agent has constant absolute risk-aversion if $u(x) = 1 - \frac{\theta}{\exp}(-\theta)$ for some $\theta \geq 0$ and has constant relative risk-aversion if $u(x) = \frac{1}{\theta}((x + K)^\theta - K^\theta)$ for some $\theta \in (0, \alpha)$.⁷ The following Lemma is proven in the Appendix.

⁶As this case is non-generic under random small perturbations of u we assume throughout the paper that $\Gamma(b^u) \neq 0$.

⁷To ensure the utility of negative outcomes is well defined we look at constant relative risk-aversion relative to the wealth level $(x + K)$.

Lemma 4. *Assumption A2 is satisfied if u exhibits constant absolute or relative risk-aversion.*

As the next Proposition shows Assumption A2 is sufficient to ensure that stopping is better than any continuation strategy for all $x \geq b^u$.

Proposition 1 (The Optimal Strategy). *The unique subgame perfect optimal strategy continues for all values $x < b^u$ and stops for all values $x \geq b^u$.*

Proof. $\tau(b^u)$ is an optimal strategy:

In the first step we prove that stopping above b^u is optimal. To ease notation let us denote by $W : \mathcal{X} \rightarrow \mathbb{R}$ the continuation value from using the cut-off strategy $\tau(b)$ derived in (4.3)

$$W(x) = V(\tau(b^u), x) = \begin{cases} \left(\frac{x}{b^u}\right)^\alpha u(b^u - K) & \text{for } x \leq b^u \\ u(x - K) & \text{for } x > b^u \end{cases}.$$

By the dynamic programming principle (cf. Peskir and Shiryaev, 2006, Theorem 1.11), $\tau(b^u)$ is an optimal strategy if and only if the function $W(x)$ satisfies the dynamic programming equation for all $x \in \mathcal{X}$

$$\max\{\mathcal{L}W(x), u(x - K) - W(x)\} = 0. \quad (4.4)$$

We have that $W(x) = u(x - K)$ for all $x \geq b^u$. Hence, $\mathcal{L}W(x) = \mathcal{L}u(x - K) < 0$ for all $x > b^u$ and (4.4) is satisfied for all $x > b^u$. Let $n = \frac{\log(b^u/x)}{\log(h)}$. By Lemma 2 and the definition of b^u we have that for all $x < b^u$

$$u(x - K) - W(x) = - \sum_{j=1}^n h^{-(j-1)\alpha} \Gamma(xh^j) < 0.$$

For all $x < b^u$ it holds that $\mathbb{E}_{t,x}[\mathbf{1}_{\{t+1 < T\}} W(X_{t+1})] = W(x)$, thus $\mathcal{L}W(x) = 0$ for all $x < b^u$, and hence (4.4) is satisfied for all $x < b^u$. It remains to verify that (4.4) is satisfied for $x = b^u$. By definition $W(b^u) = u(b^u - K)$ and thus, it remains to prove that

$$\mathcal{L}W(b^u) \leq 0$$

$$\begin{aligned} \mathcal{L}W(b^u) &= \mathbb{E}_{t,b^u} [\mathbf{1}_{\{T>t+1\}} W(X_{t+1}) - W(X_t)] \\ &= \delta \left[p W(b^u h) + (1-p) W(b^u h^{-1}) \right] - u(b^u - K) \\ &= \delta \left[p u(b^u h - K) + (1-p) \left(\frac{b^u h^{-1}}{b^u} \right)^\alpha u(b^u - K) \right] - u(b^u - K) \\ &= u(b^u - K) \left[\delta \left(p \frac{u(b^u h - K)}{u(b^u - K)} + (1-p) h^{-\alpha} \right) - 1 \right]. \end{aligned}$$

By definition of b^u and as $\Gamma(b^u) \neq 0$ we have $u(b^u h - K)/u(b^u - K) < h^\alpha$. As h is the larger solution to the equation $\delta(p h^\alpha + (1-p) h^{-\alpha}) = 1$ the expected change of W at b^u is negative

$$\mathcal{L}W(b^u) < u(b^u - K) (\delta [p h^\alpha + (1-p) h^{-\alpha}] - 1) = 0.$$

$\tau(b^u)$ is the unique optimal strategy:

By Definition of b^u it is never optimal to stop at $x < b^u$. We have also shown that $V^* = W$ and $\mathcal{L}W(x) = \mathcal{L}u(x - K) < 0$ for all $x > b^u$ and thus it is never optimal to continue for any $x > b^u$. Finally, we also have that $\mathcal{L}W(b^u) < 0$ and hence it is not optimal to continue at b^u . \square

Note that Proposition 1 does not require the utility function u to be differentiable or concave as long as Assumption A2 is satisfied. It therefore covers cases where u has a kink at a reference point r . Where this reference point lies is immaterial to our results, as long as r is determined *a priori* and constant.

As no concavity of u is required Proposition 1 furthermore covers cases of S-shaped utility as in Kahneman and Tversky (1979), i.e. risk-seeking behaviour below and risk-averse behaviour above the reference point.⁸

When the reference point is the *a priori* expected utility of the agent from stopping the process, then this may be viewed as a model of disappointment à la Loomes and Sugden (1986). It then follows from Proposition 1 that a model of elation or disappointment does not predict path-dependent behavior in our setting. Interestingly, experimental evidence seems to support this prediction. For example, Summers and Duxbury (2007) find that in an experiment where subjects do not actively trade fictitious assets, the disposition effect does not appear, while it does so when subjects had to actively choose their portfolio.

⁸For a detailed discussion of the stopping behaviour of prospect theory agent with probability distortion and naive (Ebert and Strack, 2012), with probability distortion and commitment (Xu and Zhou, 2013) or without probability distortion see Henderson (2012).

They conclude that regret and self-blame as opposed to disappointment, which lacks the self-blame component, is a key building block in explaining the disposition effect. Our theoretical model provides a rigorous argument for this finding.

4.2. Regret Preferences

In this section we examine the model predictions if the agent experiences regret. For a regret agent, the utility associated with the consequences of her action, is not solely a function of final wealth, but the difference between final wealth and the *ex post* optimal outcome. If the action chosen by the agent is *ex post* suboptimal, the agent feels regret. This makes choice context dependent, because the expected utility associated with a given act, depends on what are the counterfactual outcomes of the remaining acts available to the agent. In our setting, it is always *ex-post* optimal for the agent to stop when the process was at its maximum. Therefore regret preferences are preferences where the historical peak of the process is the reference point. The objective of the agent is then to minimize the expected regret.

To model regret aversion in our setting, we assume that the intensity with which the agent feels regret is linear in the utility difference between his strategy and the strategy that turns out to be *ex-post* optimal. In our setting there are generally two possibilities: regret relative to past and future decisions or regret relative only to past decisions.

4.2.1. Regret over Past and Future Decisions

Several authors argue that the anticipation of future regret affects choices in the present (see i.e. Loomes and Sugden, 1982). More precisely, because the agent does observe the process even after she stopped, she might also anticipate to feel regret relative to the maximum attained not only prior to stopping, but over the whole time horizon until T .

Let us denote by $S_t = \max_{r \leq t} X_r$ the maximal value of the process prior to time t . Clearly, the ex-post optimal decision for the agent is to stop when the process reaches its maximal value S_T , which would have given her a utility of $\max_t u(X_t - K) = u(S_T - K)$.⁹ The regret experienced by the agent is linear in her loss of utility due to taking a suboptimal decision

$$u(S_\tau - K) - \mathbf{1}_{\{\tau < T\}} u(X_\tau - K).$$

Note, that the agent enjoys the benefits of stopping $u(X_\tau - K)$ only if she stops before

⁹Note, that we assumed that the process starts above K and thus stopping at the maximal value $S_T \geq K$ is always better than not stopping at all.

the deadline $\tau < T$, while she suffers the regret also if she does not stop before T . Thus, when being at a point (x, s) in period t , the regret value associated with the stopping strategy $\tau \geq t$ equals

$$\begin{aligned} V(\tau, x, s) &= (1 - \kappa) \mathbb{E}_{t,x,s} [\mathbf{1}_{\{\tau < T\}} u(X_\tau - K)] - \kappa \mathbb{E}_{t,x,s} [u(S_T - K) - \mathbf{1}_{\{\tau < T\}} u(X_\tau - K)] \\ &= \mathbb{E}_{t,x,s} [\mathbf{1}_{\{\tau < T\}} u(X_\tau - K)] - \kappa \mathbb{E}_{t,x,s} [u(S_T - K)] , \end{aligned} \quad (4.5)$$

where $\kappa \in [0, 1)$ denotes the intensity of regret. Regret preferences contain expected utility preferences for $\kappa = 0$.

However, the case where the agent feels regret relative to past and future decisions does not yield behaviorally different predictions from expected utility. To see this, note that the expected regret $\mathbb{E} [u(S_T - K)]$ is independent from the agent's stopping strategy τ and thus an agent who exhibits regret over past and future decisions behaves exactly as the corresponding EU agent.

Proposition 2. *The cut-off strategy $\tau(b^u)$ is the unique optimal continuation strategy for the agent experiencing regret over past and future decisions. That is, $\tau(b^u)$ maximizes the regret functional defined in Equation (4.5) after every history.*

Thus, there is no way to distinguish regret over past and future decisions from EU. We therefore consider only regret over past decisions below.

4.2.2. Regret Only over Past Decisions

If the agent feels regret only relative to past decisions, the ex-post optimal decision for the agent is to stop at the time t before τ when the process reached its maximal value. This strategy would have given her a utility of $\max_{t \leq \tau} u(X_t - K) = u(S_\tau - K)$. Thus, the regret functional equals

$$V^r(\tau) = \mathbb{E} [\mathbf{1}_{\{\tau < T\}} u(X_\tau - K)] - \kappa \mathbb{E} [u(S_\tau - K)] . \quad (4.6)$$

As in the expected utility case we denote the value from using the continuation strategy τ at the point $X_t = x$ and the past maximum equals $S_t = s$ by $V^r(\tau, x, s)$.

An important feature of regret preferences over past decisions is their history-dependence. At first glance such preferences seem to have the potential to rationalize behavior that under EU would have been classified as time-inconsistent. In an asset-selling setting the incentive to sell the asset, is higher the lower the historical peak relative to the current price. Given that the reference or aspiration level of an agent changes with the history of

the process, one might believe that, because the agent is reluctant to realize a loss, she adopts a path-dependent strategy. The results below, however, show that if the agent always behaves optimally this is *not* the case.

We begin by establishing that the regret agent never stops after the agent maximizing expected utility.

Lemma 5. *It is optimal for the regret agent to stop at all points $x \geq b^u$.*

Proof. We first show that the regret agent always stops above the expected utility cut-off b^u . Note, that by definition of b^u any strategy that continues at a point $x \geq b^u$ yields a change in utility which is negative in expectation

$$\sup_{\tau > t} \mathbb{E}_{t,x} [\mathbf{1}_{\{\tau < T\}} u(X_\tau - K)] < u(x - K).$$

Furthermore, the maximum S_t can only increase over time and because u is increasing, regret can only increase, i.e. $\mathbb{E}_{t,x,s} [u(S_\tau - K)] > u(s - K)$ for all stopping times $\tau > t$. Hence, it is optimal to stop for all values $x \geq b^u$

$$\sup_{\tau > t} \mathbb{E}_{t,x,s} [\mathbf{1}_{\{\tau < T\}} u(X_\tau - K) - \kappa u(S_\tau - K)] < u(x - K) - \kappa u(s - K). \quad \square$$

Intuitively, regret can only make it less attractive to continue, as continuing always entails the risk of increased regret. As continuing above b^u is not optimal without regret it can never be optimal with regret. In the next step we show that below the expected utility cut-off is never optimal for the regret agent to stop below the past maximum.

Lemma 6. *It is never optimal for the regret agent to stop when $X_t < S_t \leq b^u$.*

Proof. Let $X_t = x < s = S_t \leq b^u$. By Lemma 2 the expected change in utility from the cut-off strategy $\tau(s)$ that only stops once the past maximum is reached is positive

$$\mathbb{E}_{t,x} [\mathbf{1}_{\{\tau(s) < T\}} u(X_{\tau(s)} - K)] > u(x - K).$$

As the strategy $\tau(s)$ never stops above the past maximum s it follows that the regret never increases $u(S_{\tau(s)} - K) = u(s - K)$. Thus, it is always better to wait until the process is back at its past maximum s than stopping at a value $x < s$

$$\mathbb{E}_{t,x,s} [\mathbf{1}_{\{\tau(s) < T\}} u(X_{\tau(s)} - K) - \kappa u(S_{\tau(s)} - K)] > u(x - K) - \kappa u(s - K). \quad \square$$

Intuitively as regret is sunk and does not change until the process reaches its past maximum again it can never be optimal to stop the process below its past maximum. Note, that as a consequence of Lemma 6 the agent effectively never experiences regret when stopping below b^u .

We define $\Gamma^r : \mathcal{X} \rightarrow \mathbb{R}$ as the expected change in value from waiting until the process reaches xh instead of stopping at $x = X_t = S_t$:

$$\begin{aligned}\Gamma^r(x) &= (\mathbb{P}_x[\tau(xh) < T] (1 - \kappa)u(xh - K) - \mathbb{P}_x[\tau(xh) \geq T] \kappa u(x - K)) \\ &\quad - (1 - \kappa)u(x - K) \\ &= h^{-\alpha}(1 - \kappa)u(xh - K) - \kappa(1 - h^{-\alpha})u(x - K) - (1 - \kappa)u(x - K) \\ &= h^{-\alpha}(1 - \kappa)u(xh - K) - (1 - \kappa h^{-\alpha})u(x - K).\end{aligned}$$

Define the cut-off b^r by

$$b^r = \min\{x \in \mathcal{X} : \Gamma^r(x) \leq 0\}. \quad (4.7)$$

As in the expected utility case we assume that $\Gamma^r(b^r) \neq 0$. Furthermore, we assume that the expected gain in payoff from waiting for an up-tick Γ^r changes its sign once.

Assumption A3. *We assume that $\Gamma^r(x) < 0$ for all $x \geq b^r$.*

This is a weak assumption and the next Lemma establishes that it is satisfied for all risk-averse agents.

Lemma 7. *If u is concave than Assumption A3 is satisfied.*

Proof. First, we establish existence of b^r . For that we need to show that Γ^r crosses zero at least once. The change of Γ^r in x equals

$$\begin{aligned}\Gamma^r(xh) - \Gamma^r(x) &= (h^{-\alpha}(1 - \kappa)u(xh^2 - K) - (1 - \kappa h^{-\alpha})u(xh - K)) \\ &\quad - (h^{-\alpha}(1 - \kappa)u(xh - K) - (1 - \kappa h^{-\alpha})u(x - K)) \\ &= (1 - \kappa) h^{-\alpha} (u(xh^2 - K) - u(xh - K)) \\ &\quad - (1 - \kappa h^{-\alpha}) (u(xh - K) - u(x - K))\end{aligned} \quad (4.8)$$

As u is concave, we know that

$$\frac{u(xh - K) - u(x - K)}{x(h - 1)} \geq \frac{u(xh^2 - K) - u(xh - K)}{xh(h - 1)}$$

Multiplying by both sides $x(h - 1)$ and as $\alpha > 1$ it follows

$$\begin{aligned} u(xh - K) - u(x - K) &\geq h^{-1} (u(xh^2 - K) - u(xh - K)) \\ &> h^{-\alpha} (u(xh^2 - K) - u(xh - K)) , \end{aligned}$$

As $1 - \kappa < 1 - \kappa h^{-\alpha}$ Equation (4.8) is negative. Consequently, Equation Γ is strictly decreasing and changes sign at most once. \square

The next proposition shows that it is optimal to stop if and only if the past maximum of the process S_t is at least b^r and the process is at its maximal value $X_t = S_t$ or the process x_t is above the expected utility cut-off b^u .

Proposition 3. *The optimal strategy stops if and only if $X_t \geq b^r$ and $X_t \geq \min\{b^u, S_t\}$.*

Proof. We already established that it is optimal to stop $X_t \geq b^u$ and optimal to continue for $X_t < S_t \leq b^u$. By definition of b^r it is optimal to wait for an up-tick at $X_t = S_t < b^r$. It remains to show that it is optimal to stop for $X_t = S_t \geq b^r$.

We prove this result by induction. As show it is strictly optimal to stop at $X_t = S_t = b^u$. Let $X_t \geq b^r$ and suppose it is strictly optimal to stop once the process reaches $X_t h$. Then, the change in payoff from waiting for this uptick is given by

$$V^r(x) - (1 - \kappa)u(x - K) = \Gamma^r(x) ,$$

which is negative by definition of b^r . Hence, it is strictly optimal to stop at X_t . \square

We have plotted the subgame perfect optimal stopping strategy of the regret agent in Figure 1. For $s < b^r$, the agent's optimal continuation strategy is the cut-off strategy $\tau(b^r)$, especially it is independent of s . The optimal continuation strategy changes if the agent misses to fulfill her initial plan, i.e. to stop at b^r finds herself in a history where s exceeds b^r . The optimal continuation strategy now prescribes the agent to wait for the process to return to s .

Consider a path where the process reaches b^r a second time, i.e. $X_t = b^r$ and $s > b^r$. Under EU, the situation has not changed relative to the first visit, because the running maximum s is immaterial to the EU agent. For the regret agent, this is different: Compared to the situation where $s = x = b^r$, the regret associated increased. The higher regret also enters the continuation value, but to lesser extent, because continuation includes the prospect of making up for the current regret. The prospect of making up

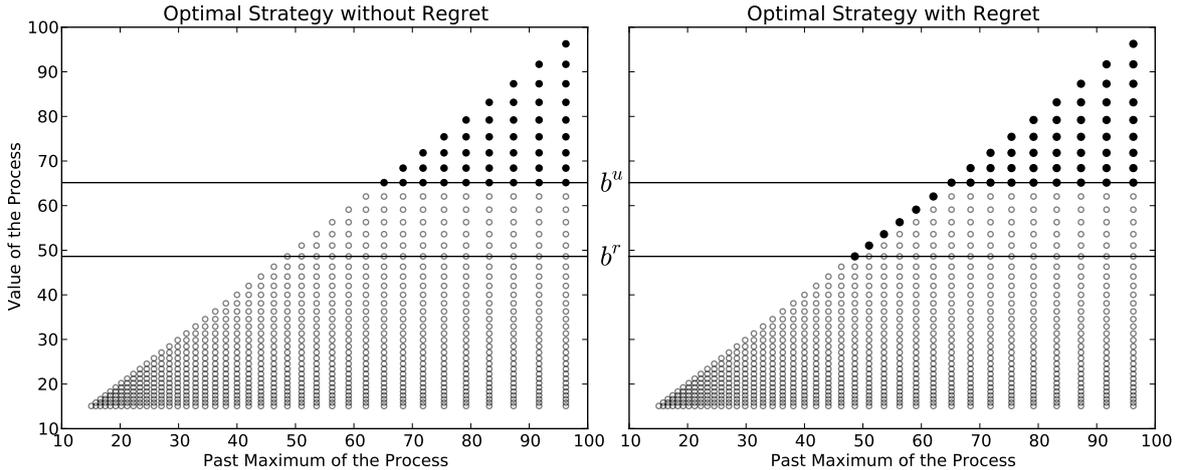


Figure 1: This shows the subgame perfect cut-off of an agent without regret on the left and with regret on the right. Empty dots mark the points where the agent continues and filled large dots show where it is optimal for the agent to stop.

for the current regret motivates the agent to continue until the process reach s again.¹⁰

Note, that while the optimal strategy for the EU and the regret agent both lead to the same observed behavior. As the regret agent always stops at b^r the realized (observed) stopping time will be equal to the realized stopping time resulting from the cut-off strategy $\tau(b^r)$.

Proposition 4. *Let τ^r be the strategy which is optimal for the regret agent. We have that $\tau^r = \tau(b^r)$ almost surely.*

Hence, regret is observationally equivalent to expected utility. Note, that this equivalence relies on the premise that the agent makes no mistakes. If the agent makes mistakes, behavioral differences arise and regret and expected utility can be told apart from observed decisions. We will use laboratory data in combination with this identification strategy in the following section to identify regret preferences.

5. A laboratory experiment

In order to investigate individuals' behavior and to test our predictions about behavior, we implemented exactly the theoretical setting described in the previous section in the

¹⁰This behavior resembles the gambling for resurrection described in the finance literature. While gambling for resurrection in this literature occurs, because the agent is insured against losses, in the regret model the agent continues because further losses do not translate into higher regret.

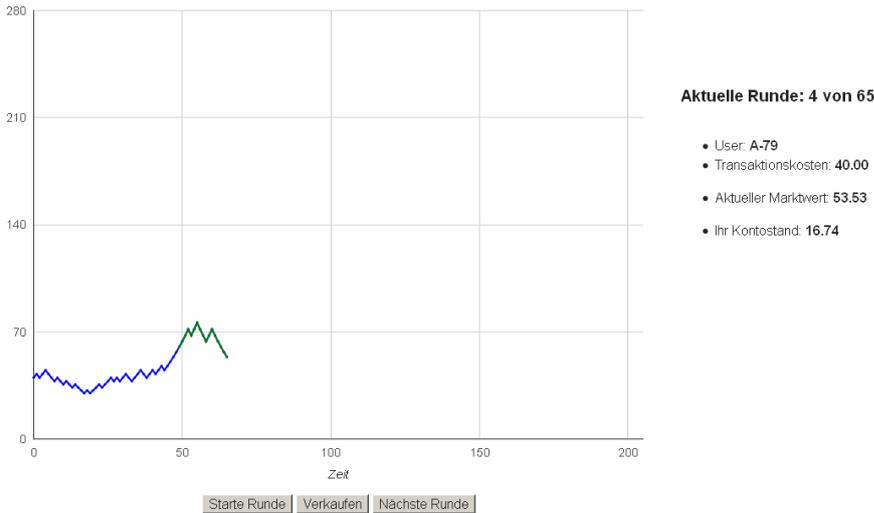


Figure 2: The main experimental screen (in German).

Table 1: Parameters for the binomial random walk in the experiment.

Cost K	Stepsize h	Uptick prob. p	Exp. prob. $1 - \delta$
40	1.06	52 %	0.7 %

laboratory. Testing individual behavior in a controlled laboratory setting has several advantages over using field data. First, we can ensure that the environment in which subjects make their decisions is truly stationary, i.e. the probability law driving the process is known and time-homogeneous. Second, we also have full discretion over the payoff-relevant state variables of a subject. With field data, we can never be entirely sure to observe all relevant state variables an individual integrates into her decision-making process. In the laboratory, we have full discretion over all payoff-relevant state variables. Especially for testing the time-consistency property, both points are crucial.

The experiment was conducted as a computer-based experiment at the laboratory of the Technical University Berlin (TU) and the WZB Berlin. The experimental software was programmed using Java and Python and ran in a browser. We ran two sessions, each session with 22 students that were randomly recruited from the ORSEE pool of the TU and WZB.¹¹ Each of the 44 subjects in our sample, played 65 rounds for pay. For each round played by subjects, we either observe a stopping decision or that the process jumped to zero before a subjects decided to stop. The latter is true for about 56%, i.e. 1581 out of the total 2860, in our sample.

¹¹See: <https://experimente.wzb.eu/>

Before the actual experiment, subjects received a four-page instruction explaining the upcoming experiment. In the instructions, we framed the optimal stopping task as an asset-selling problem. Subjects were explained that they own a fictitious stock and that they have the opportunity, but not the obligation, to sell it. The instructions then explained in detail the setting discussed in section 3 and the meaning of the relevant parameters, e.g. the uptick probability p and the step size h . The actual values for the parameters were given to subjects and are listed in table 1. Subjects were hence fully informed about how processes were generated.¹²

After subjects finished reading the instructions, they were prompted to login and begin the actual experiment. The experiment consisted of 65 rounds in which subjects had the option to sell their stock. In each round, subjects observed the path of the market price in a diagram (see Figure 2). At the beginning of each round, the computer loaded the screen with an empty diagram. At the bottom of the screen there were three buttons available. Upon pressing the left-hand button, labeled 'Start round', subjects started a given round.¹³ That triggered the market price to be displayed as a jagged blue line until the jump to zero in period T . Each second there were two ticks of the price process. Additionally, subjects were displayed several other details about a round in a panel to the right of the diagram. As soon as a round was started, the middle button, labeled 'Sell', became active. Pressing this button, a subject sold the stock at the current value of the price process. Future values of the process were then displayed in green as to visualize that selling had occurred. The right-hand button, which was inactive until the jump to zero, gave subjects the opportunity to move to the next round.¹⁴ The paths of the 65 random walks were generated prior to the experiment and were the same for all subjects. However, based on an individual login printed on the instructions, the order in which the set of 65 paths was shown to subjects was shuffled randomly.

At the end of the experiment one round was randomly selected with equal probability to determine a subject's payoff. Subjects were informed about which round was drawn on a final screen that listed their performance in each round played together with their final payoff. Subjects earned 0.15 times the number of points they had obtained in

¹²We convinced ourselves that subjects had indeed understood (i) how payoffs are computed, (ii) that the increments of the process are iid and (iii) what is the risk that a round ends before the next period, through a questionnaire with control questions that subjects had to answer prior to the experiment (see appendix). 95% of the time subjects answered our control questions correctly.

¹³The remaining two buttons were disabled before the round was started.

¹⁴We could have given subjects the option to skip to the next round immediately after the stopping decision. This, however, may provide incentives to impatient subjects to stop early and reduce lab time.

the round that was drawn by the computer plus 10 Euros show-up fee.¹⁵ The average duration of the overall experiment was 80 minutes, and the mean earnings for subjects was 12.30 Euros (median=10 Euros), where the minimum and the maximum payment were 10.00 Euros and 19.00 Euros respectively.

6. Experimental results

Following the literature, we call the value X_τ at which the agent decided to stop her reservation level. As shown in Proposition 1 and 3 an agent maximizing expected utility or minimizing regret should not vary her reservation level between rounds. We thus first inspect to which extent subjects do have a constant reservation level across different rounds. We then inspect in how far their reservation level is constant or time-consistent within rounds.

6.1. Do People have a constant reservation level?

The histogram of reservation levels in the top chart in Figure 3 shows that there is a large variation in reservation levels across rounds, ranging from slightly above the starting value of $X_0 = 40$ to the optimal risk-neutral level 128.3. The boxplots in the bottom chart of Figure 3 decompose this further into between- and within-subject variation of reservation levels. The within-subject variation accounts for 33.2% of the total variation and there is no subject for which the distance between the minimal and maximal reservation level is less than 50.41. Thus, dispersion of reservation levels is not solely due to individual differences between subjects and quite substantial in absolute terms. This shows that subjects do not play unique reservation level strategies.

It is still possible, however, that within-subject variation is due to the gradual convergence of reservation levels: subjects could successively adapt their reservation level towards a unique level after several rounds of experimentation. If that were the case, we would expect the variation in reservation levels to decrease in the number of rounds played and to settle at some constant level. Figure 8 in the appendix shows the observed reservation levels for each subjects over all rounds played in our experiment and from Oprea et al. (2009). In both experiments, the observed variation does not decrease in the number of rounds for the vast majority of subjects. In Table 2 we report the average variance of reservation levels across subjects for different blocks or rounds.

¹⁵We rounded to the nearest Euro.

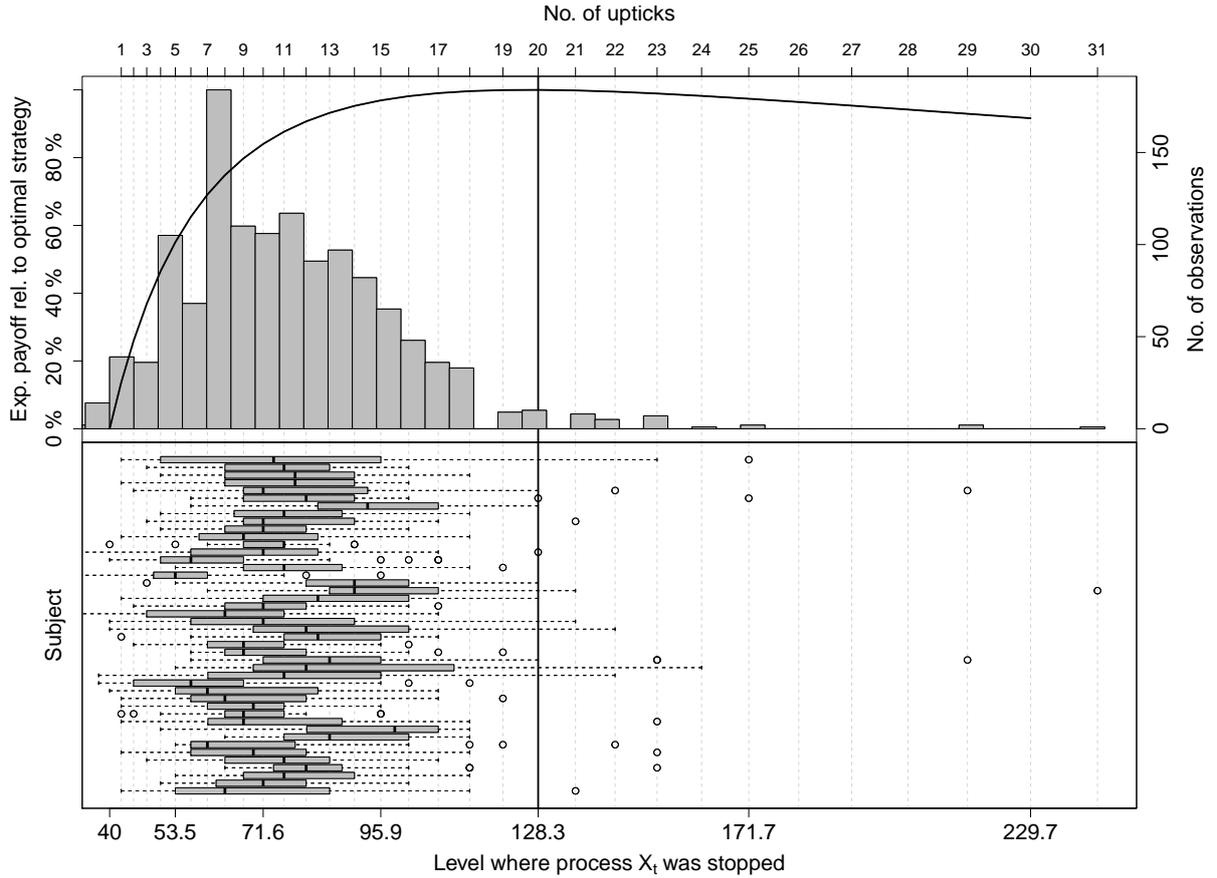


Figure 3: Line in the upper chart shows how the expected payoff from using different cut-offs changes as a percentage of the payoff under the optimal strategy (left scale), grey bars show a histogram of actual reservation levels (right scale). Boxplots in bottom chart illustrate the variation in reservation levels at the subject level. The vertical line indicates the optimal cut-off for a risk neutral agent.

Variances do not decrease on average over the course of the experiment, but fluctuate unsystematically.

Finding 1. *Subjects vary their reservation levels substantially over different rounds of the same stopping task and do not appear to converge to a unique level.*

Finding that subjects' reservation levels do not settle to a constant level does not necessarily mean that their variation is entirely unsystematic. Following Oprea et al. (2009), we therefore estimated a model on the pooled data, where subjects use a cut-off strategy $\tau(b^j)$ in every round j and adapt their reservation level b^j in response to forgone earnings in the previous round. More specifically, Oprea et al. assume that the reservation level b^j in round j follows a simple linear model, which makes the difference

Table 2: Average variance of reservation levels across subjects over blocks of 10 rounds.

Rounds	1-10	11-20	21-30	31-40	41-50	51-60	60-65
Variance	366.94	503.81	232.03	311.84	481.02	335.05	123.78

Table 3: Estimated effects of losses on subsequent stopping choices.

Parameter	Oprea et al.	This study
$\delta_E \times 1,000$	0.5486***	1.3873**
$\delta_L \times 1,000$	-0.9185***	-1.1227

Notes: Median estimates and p -value for the Wilcoxon signed-rank test that the distribution is centered around zero: *** $p \leq 0.01$, ** $p \leq 0.05$ and * $p \leq 0.1$.

in reservation levels between round j and $j - 1$ a linear function of previous losses

$$b^j = b^{j-1} + K [\delta_E \mathbf{1}_{\{\tau^{j-1} < T\}} + \delta_L \mathbf{1}_{\{\tau^{j-1} \geq T\}}] (S_{\tau^{j-1}}^{j-1} - b^{j-1}) . \quad (6.1)$$

The parameters δ_E and δ_L measure an individual’s sensitivity to a loss that stems from stopping below S_j and from not having stopped before the deadline, respectively.¹⁶ In Table 3 we follow Oprea et al. and report the median of the by-subject median of δ_E and δ_L . Our estimates are qualitatively similar to that of Oprea et al.. The estimate for δ_E implies that subjects increase their reservation level, if in the previous round they observed that after stopping they could have stopped at higher values. The estimate for δ_L implies that subjects reduce their reservation level, if in the previous round the process jumped to zero before they stopped and they missed the opportunity to get a positive payoff. In our sample the adjustment to the latter is not statistically significant.¹⁷

To inspect how much variation between rounds can be explained by this model, we took the first observed reservation level for each subject and forecasted their reservation levels for all remaining rounds. We have plotted the results in Figure 9 in the appendix.

¹⁶We use the same estimation method as Oprea et al. (2009). That is, we obtain an estimate of δ_E for each two consecutive rounds a subject stopped by setting $b^j = X_{\tau^j}^j$ and solving for δ_E in (6.1). For each block of consecutive rounds without observing a subjects stopping decision, we may use the two adjacent reservation levels to estimate δ_L from the losses suffered due to not stopping.

¹⁷This is actually a common finding of most of the literature on regret and counterfactual thinking, i.e. that people experience more regret over outcomes that stem from action than from equally miserable outcomes that stem from inaction (see e.g. Kahneman and Tversky, 1982; Gilovich and Medvec, 1994; Gilovich et al., 1998; Coricelli et al., 2005; Summers and Duxbury, 2007).

Table 4: Contingency table for observed decisions.

	$X_\tau = S_\tau$	$X_\tau < S_\tau$	No. of obs.
stopped first time	326 (25%)	0 (0%)	326 (25%)
not stopped first time	205 (16%)	748 (58%)	953 (75%)
No. of obs.	531 (42%)	748 (58%)	1279 (100%)

Notes: Decisions in the upper left cell are time-consistent. Other cells represent number of deviations from the time-consistency property in terms of the maximum of the process and the multiplicity of the stopping value in the history of the process. X_t denotes the value of the process and $S_t = \max_{s \leq t} X_s$. Percentage of total observations in parantheses.

The plot shows that the model has limited explanatory power in our sample. In the first 20 rounds, the adaption model tracks the development of reservation levels reasonably well, but it clearly overshoots thereafter. This shows that the entire variation captured in the estimated adjustment coefficients stems from behavior in the first 20 rounds.

Finding 2. *Subjects calibrate their reservation levels during the first 20 rounds in response to forgone earnings and then stop to systematically adjust and stick with this level for all later rounds.*

6.2. Do People use Time-Consistent Strategies?

In the previous section we found that subjects' behavior is not constant across different rounds. But do subjects play constant reservation levels within rounds? We first measure deviations from this time-consistency property along two dimensions: (i) we count the number of ticks the stopping value is below the current running maximum and (ii) we count the deviation in terms of *multiplicity*, i.e. we count the number of times a subject had seen her stopping value before stopping eventually.

Table 4 shows the results in a simple contingency table. The columns of table 4 contain observations that correspond to stopping *at* the running maximum (left column) and below the running maximum (right column). Hence, 42% of observations stop at the maximum of the process. The remaining 58% do not. The rows of tab. 4 contain observations that correspond to stopping decisions that stopped *the first time* the process reached this value (top row) and decisions that stopped afterwards (bottom row). We observe only 326 out of 1279, i.e. roughly 25%, decisions that are perfect cut-off decisions. The remaining 75% are not.

Finding 3. *Subjects stopping decisions are not time-consistent 75% of the time.*

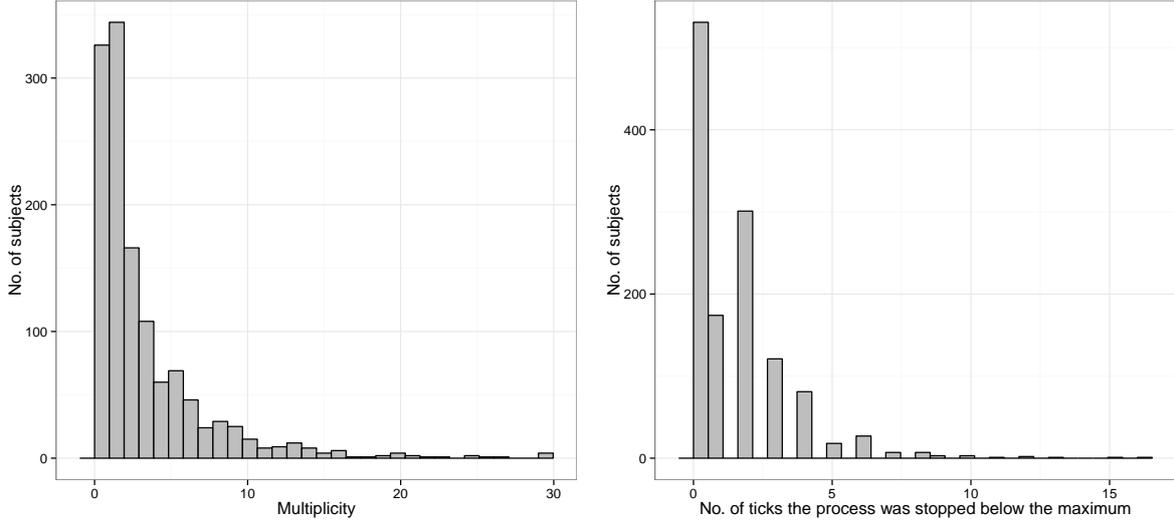


Figure 4: The empirical distribution of the multiplicity of stopped values for all subjects (left) and the number of ticks subjects stopped below the previous maximum (right).

It is worth to stress that the above finding is a strong result in our setting. In fact, this renders 75% of all observed stopping decisions are irreconcilable with any of the choice models we presented earlier.

To shed more light on the magnitude of time-inconsistency, we have plotted a histogram of the multiplicity of subjects' stopping decisions in Figure 4. The average multiplicity of subjects' stopping decisions is roughly 3, i.e. subjects on average stop at a point they have seen three times before. The maximum multiplicity is equal to 29. About 35% of the observed stopping decisions have a multiplicity greater two.

In light of these findings, subjects clearly do not play cut-off strategies. Instead they vary their reservation levels over different rounds of the same stopping task, do not behave time-consistently 75% of the time and visit the same level of the process on average three times before they eventually stop at it.

An interesting feature of the observed behavior is that stopping seldom occurs appreciably below the past maximum. That is, while 59% of the observed stopping decisions are such that stopping occurs below the past maximum. The histogram on the right of figure 4 shows that the difference between the stopped value and the past maximum is seldom more than two or three ticks.

Finding 4. *Subjects show a disposition to stop near the past maximum of the process and not to stop appreciably below.*

Given that in the experiment there are two ticks every second, it seems possible on the

one hand that subjects attempt to stop at the maximum the first time, but their limited reaction time often leads to miss the exact maximum and fall slightly below. On the other hand, more systematic forces such as an aversion to realize a loss relative to the past maximum could lead subjects to deliberately waive any smaller values. In the next section, we put forward a structural econometric model that is capable of disentangling the two possible explanations in a likelihood sense.

6.3. Dynamic stochastic choice

The observed deviations from cut-off behavior can be explained by introducing a stochastic component alongside a structural or deterministic models of choice. Not only does such a model open up a way to accommodate for randomness in our data, it also provides a way to test an expected utility model against a model with regret.

In our dynamic discrete choice (DDC) model the process that leads the agent to stop or continue in period t is influenced by the level of the process X_t and – in the case of regret – by the running maximum S_t . As in the deterministic case, the choice between stopping and continuation at a point (x, s) is determined by the difference between the stopping and continuation value. The stopping value $sv(x, s)$ is the utility from stopping at a given point (x, s) , i.e.

$$sv(x, s) = u(x - K) - \kappa u(s - K) , \quad (6.2)$$

and $cv(x, s)$ is the expected utility from rejecting the current value in favor of future values. With probability $1 - \delta$ the process jumps to zero in the coming period and continuation yields a payoff of $u(0) - \kappa u(s - K)$. With probability δ the process does not jump to zero, but increases or decreases by one tick. Hence, the agent will continue from hx with probability p , and from $h^{-1}x$ with probability $(1 - p)$. Thus, using Equation (4.3)

$$cv(x, s) = \begin{cases} \delta [pV^*(xh, s) + (1 - p)V^*(xh^{-1}, s)] - (1 - \delta)\kappa u(s - K) & \text{if } x < s , \\ \delta [pV^*(xh, sh) + (1 - p)V^*(xh^{-1}, s)] - (1 - \delta)\kappa u(s - K) & \text{if } x = s . \end{cases}$$

In contrast to the deterministic case, there are now two additional factors that affect choice. First, there is an unobserved factor ϵ_t that is an additive random shock to the current utility difference between stopping and continuation. We assume that ϵ_t is i.i.d. $\mathcal{N}(0, \sigma^2)$. An agent's choice in period t is then viewed conditional on the realization of

ϵ_t and is given by the choice function

$$\psi(x, s, \epsilon) = \begin{cases} \text{stop} & \text{if } sv(x, s) - cv(x, s) + \epsilon \geq 0 \\ \text{continue} & \text{if } sv(x, s) - cv(x, s) + \epsilon < 0, \end{cases}$$

Note that in order to calculate the expected utility from continuation with the optimal strategy V^* , we need to evaluate the expected utility from continuation for the regret agent $\mathbb{E}_{t,x,s} [\mathbf{1}_{\{\tau < T\}} u(X_\tau - K)] - \kappa \mathbb{E}_{t,x,s} [u(S_{\min\{\tau, T\}} - K)]$. In particular, this requires an expression for the anticipated regret, which is given in the following lemma proven in the appendix

Lemma 8 (Anticipated Regret). *The anticipated regret $\rho(x, s, b)$ associated with the cut-off strategy $\tau(b)$ when being at point (x, s) equals*

$$\begin{aligned} \rho(x, s, b) &= \mathbb{E}_{t,x,s} [u(S_{\min\{\tau(b), T\}} - K)] \\ &= \kappa \sum_{i=0}^{m-1} \left(\frac{x}{xh^i}\right)^\alpha (1 - h^{-\alpha}) \max\{u(s - K), u(xh^i - K)\} \\ &\quad + \kappa \left(\frac{x}{b}\right)^\alpha \max\{u(s - K), u(b - K)\}. \end{aligned}$$

where $m = \frac{\log(b/x)}{\log(h)}$.

However, simultaneous estimation of the error variance σ^2 and the curvature of the utility θ generally is a delicate issue in models of stochastic discrete choice (see e.g. Wilcox, 2011, for a discussion of such issues in static models). A very much related issue in our case is that a large variation in reservation levels across rounds leads to a relatively large error variance. A high error variance, however, translates into a low probability to reach higher levels of X_t a posteriori. In order to reconcile that we do observe subjects stopping at high and low values of X_t , the model needs to make subjects less risk averse because that increases their likelihood to reach higher values of X_t . Hence, there is a purely mechanical relationship between σ and θ which somewhat blurs the interpretation of θ as a measure of risk aversion in the sense of Pratt (1964).

We therefore incorporate a second terms that affects choice: effort, capacity or attention (Kahneman, 1973). Each tick, when viewed as a separate stimulus, first has to pass an 'attention filter' before a subject acts based on it and makes a (potentially noisy) decision. We see at least two reasons why such a pre-filter stage should be present: attentional overload and meaningfulness of each individual tick. For example, both the

bottleneck and the capacity theory of attention in psychology suggest that for stimuli that arrive in rapid succession, humans easily face an attentional overload. For example, it has been documented in the psychology literature that response times increase in the latency of the stimulus (Craik, 1947, 1948; Smith, 1969). That is, for signals that arrive in rapid succession, the human brain demands a refractory period that delays response to stimuli due to the mental capacity that is occupied by processing and responding to past stimuli. Since processes in our experiment evolved at a high rate (2 ticks per second), it seems plausible that such a refractory period lead subjects to respond at a lower rate. The second reason relates to the fact that each individual tick did not change material payoffs dramatically. For that reason subjects could have diverted attention from the task, because the individual tick was not meaningful enough to warrant perfect attention in the eyes of the subject.

We thus assume that in every period t there is a constant probability w that the agent does not pay attention and continues with probability one. We further assume that w is constant across all rounds. Since we do not observe w , we estimate it from the data. Note that under this assumption w has a more intuitive interpretation in terms of the average delay between two decisions. Because with constant w the time between two periods the agent actively decides is exponentially distributed, we may interpret $0.5(1 - w)^{-1}$ as the average number of seconds between two decisions. From a purely mechanical point of view, w provides a way to explain the variation in reservation levels without an unduly upsurge in σ and θ .¹⁸

For estimation, denote by $\hat{\psi}_t = 1$ and $\hat{\psi}_t = 0$ the observation that the agent decided to stop or continue in period t respectively and let $\gamma = [\theta, \kappa, \sigma, w]$ be the vector of parameters to estimate. The probability that an agent chooses to continue at a given point (x, s) then equals

$$\mathbb{P}[\text{continue at } (x, s) \mid \gamma] = (1 - w)\Phi_\sigma(\text{cv}(x, s \mid \theta, \kappa) - \text{sv}(x, s \mid \theta, \kappa)) + w$$

where Φ_σ denotes the normal cdf with mean zero and variance σ^2 . To shorten notation we will henceforth denote the conditioning on the parameters by \mathbb{P}_γ . Note that our model entails that mistakes are more likely to occur, if it is less costly for the agent, i.e. when the difference between stopping and continuation value is small. For estimation,

¹⁸We have fitted a model without an inattention parameter w and our results confirm that θ and σ both inflate drastically. The implied reservation levels lie above 99% of the observed decisions and thus the model has to attribute almost all observed stopping decisions to the error term.

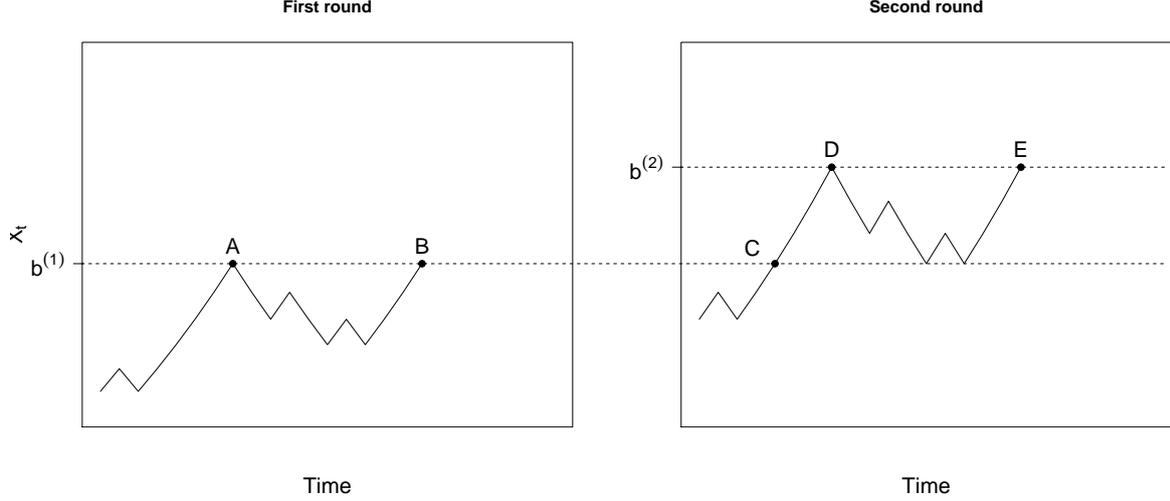


Figure 5: Illustrates how a regret model can be distinguished from an EU model over rounds of the experiment.

we assume that agents' utility is of the power-utility form

$$u(x - K) = \begin{cases} \frac{K}{\theta} \left[\left(\frac{x}{K} \right)^\theta - 1 \right] & \text{for } \theta \neq 0, \\ K \ln \left(\frac{x}{K} \right) & \text{for } \theta = 0. \end{cases} \quad (6.3)$$

If we let $Y^{(i,j)} = \{X_t^{(i,j)} = x, S_t^{(i,j)} = s, \hat{\psi}_t^{(i,j)}\}_{t=1}^T$ denote the available data for agent i in round j , the likelihood to observe that the agent stops in period τ equals¹⁹

$$p(Y^{(i,j)} | \gamma) = \mathbf{1}_{\{\tau < T\}} (1 - \mathbb{P}_\gamma [\text{continue at } (X_\tau^{(i,j)}, S_\tau^{(i,j)})]) \times \prod_{t=1}^{\min\{\tau-1, T\}} \mathbb{P}_\gamma [\text{continue at } (X_t^{(i,j)}, S_t^{(i,j)})] \quad (6.4)$$

As we mentioned before, the DDC model provides us a way to judge whether the fact that subjects tend to stop near the past maximum of the process is significantly related to regret or not. To see how this works, consider behavior over two consecutive rounds. Suppose the agent observes the two processes depicted in Figure 5 and stops at B in the first round. The agent thus missed to stop at point A, i.e. the first time the level $b^{(1)}$ was reached. Under EU, the most likely explanation for this decision in isolation is that the

¹⁹We are free to assume any value for $\hat{\psi}_t$ after stopping occurred, because these periods do not enter the likelihood function and are payoff-irrelevant. Note that the way we specify the likelihood already takes into account any effects of censoring and thus already corrects for any bias that stems from rounds that subjects did not stop before T .

agent committed just a single error at point A. Hence, fix θ such that $b^{(1)}$ is the ex ante optimal cut-off $b^u = b^{(1)}$. Similarly, fix the parameters (θ, κ) for a corresponding regret agent such that $b^r = b^{(1)}$.²⁰ In the next round, suppose that the agent stopped at point E, i.e. at a higher value than in the previous round $b^{(2)} > b^{(1)}$. For given parameters, the EU model can only reconcile this with the decision in the previous round by assuming that for all points on the segment between C and E the subject erroneously continued to play. As long as $b^{(2)} \leq b^r$, the regret agent only errs at the points on the segment from C to D. For all points on the line segment from D to E, the agent does not want to stop by Lemma 6. The regret agent's continuation value is thus larger than the stopping value. In terms of the likelihood given in equation (6.4), the latter fact implies a higher second-round likelihood for the regret model. Intuitively, the regret model is able to attribute more variation in behavior to the structural part of the model and less to noise. Thus, if the described pattern is abundant in the data, the regret model finds more support in terms of any likelihood criterion. If the fact that subjects stop close to the past peak is merely because they missed to stop at exactly the maximum due to some attentional lapse, the likelihood of the regret model will only be marginally higher. Then any criterion that sufficiently penalizes the regret model for its additional degree of freedom would prefer the more parsimonious EU model. Additionally, the parameter w provides a way for the model to attribute any such unsystematic variation to inattention. Hence, any regret aversion that we find in our model, may be regarded as being net of inattention or reaction time.

We use Bayesian inference to estimate the model. Given a prior distribution $p(\gamma, \sigma)$ and the joint posterior density is proportional to prior times likelihood

$$p(\gamma | Y) \propto p(Y | \gamma)p(\gamma) . \quad (6.5)$$

In the prior distribution we assume that θ , κ , w and σ are *a priori* independent

$$p(\gamma) = p(\theta)p(\kappa)p(w)p(\sigma) . \quad (6.6)$$

For κ and w we set the prior equal to the uniform distribution over the interval $[0, 1]$, for θ we set a normal prior that was calibrated to match the empirical findings of Holt

²⁰According to Proposition 4, there always exists such a tuple.

Table 5: Summary statistics for posterior modes across subjects.

	Expected utility ($\kappa = 0$)			Regret			
	Rel. Risk Aversion	Noise Level	Inattention Period	Rel. Risk Aversion	Regret Aversion	Noise Level	Inattention Period
	$1 - \theta$	σ	$\frac{1}{2(1-w)} - \frac{1}{2}$	$1 - \theta$	κ	σ	$\frac{1}{2(1-w)} - \frac{1}{2}$
Min.	0.82	1.96	0.00 s.	0.63	0.00	1.16	0.00 s.
1st Qu.	0.30	3.90	1.50 s.	0.22	0.00	3.33	1.58 s.
Mean	0.15	4.91	1.42 s.	0.06	0.10	4.32	1.50 s.
3rd Qu.	-0.03	5.70	5.75 s.	-0.11	0.18	5.35	7.83 s.
Max.	-0.29	8.63	16.17 s.	-0.46	0.36	7.40	49.50 s.

and Laury (2002) and for σ we set the prior equal to $1/\sigma$

$$p(\gamma) = \mathcal{N}(0.7, 0.3) \times \text{Uniform}(0, 1) \times \text{Uniform}(0, 1) \times \frac{1}{\sigma}. \quad (6.7)$$

The joint posterior density is not of any known form and thus there is no direct way to sample from it. We therefore first find the posterior modes of (6.5) with a standard hill-climbing algorithm (BFGS) and take these as starting values for a Markov Chain Monte Carlo algorithm (Metropolis-Hastings) that simulates the joint posterior of all unknowns (Metropolis et al., 1953). We have deferred further details on the model and estimation to a web appendix.

Table 5 and Figure 6 show summary statistics and histograms of the posterior modes. We find that the average subject is mildly risk-averse under the EU model and virtually risk-neutral under the regret model. This implies that with regret the observed behavior is characterized by regret aversion, whereas risk-aversion seems to play a negligible role. The probability w is found to be relatively high in our sample. In terms of attentional lapse, the numbers imply that the average subject in our sample makes decisions every 1.42 seconds in the EU model and every 1.5 seconds in the regret model. This seems to be a plausible amount of inattention in our experiment, where the process evolves relatively fast. We have also plotted the posterior means and credible intervals for each subjects separately in Figure 6. Again, posterior means for θ fluctuate around the risk-neutral level indicated by the horizontal line. The inattention intervals and their posterior intervals around the posterior mean are found to be somewhat extreme for several subjects in our sample. For the large majority, however, the posterior means and credible intervals of the inattention periods are well below ten seconds.

We also find that the bulk of posterior mass for κ is tilted away from zero in most

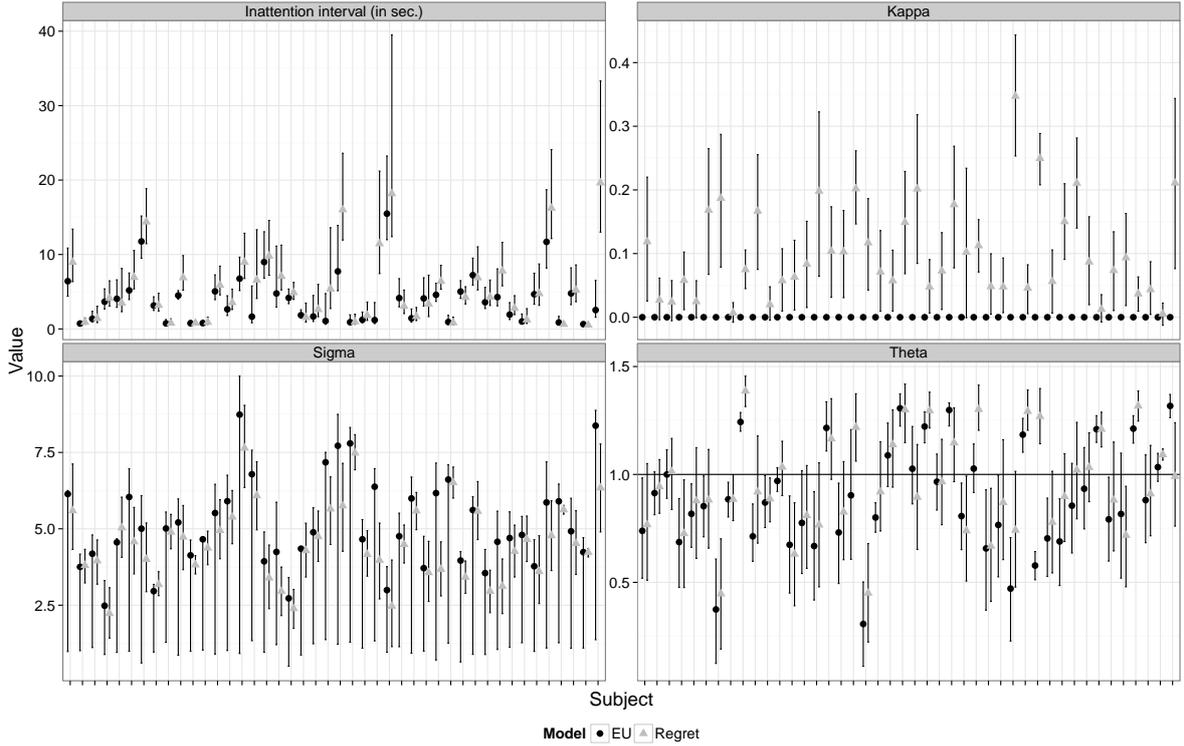


Figure 6: Posterior means of the parameters and 68% credible intervals.

cases and that the respective 68% credible intervals do not encompass zero. Does the regret model therefore fit the data better than the EU model? To have a fair comparison between the two models, we have to take into account that the regret model has one additional degree of freedom. We have therefore computed the Deviance information criterion from the Metropolis output (Gelman et al., 2004). The smaller the DIC, the better the fit of a model. Because the DIC rises in the complexity of the model, i.e. in the number of parameters, more parsimonious models are preferred. The points in Figure 7 depict the differences between DIC of the EU and the regret models for each subjects. To illustrate how the model comparison is related to the intensity of regret, the size of the points increases with the value of the posterior mean of κ . According to the DIC, the regret model is preferred for 22 of our subjects, i.e. 50%. As expected, the regret model tends to be preferred the higher the posterior mean of κ .

Finding 5. *There is strong empirical evidence that subjects vary their reservation levels in relation to the past maximum of the payoff process X_t . Subjects are reluctant to realize losses relative to the past peak and wait for the payoff process to return to it before stopping.*

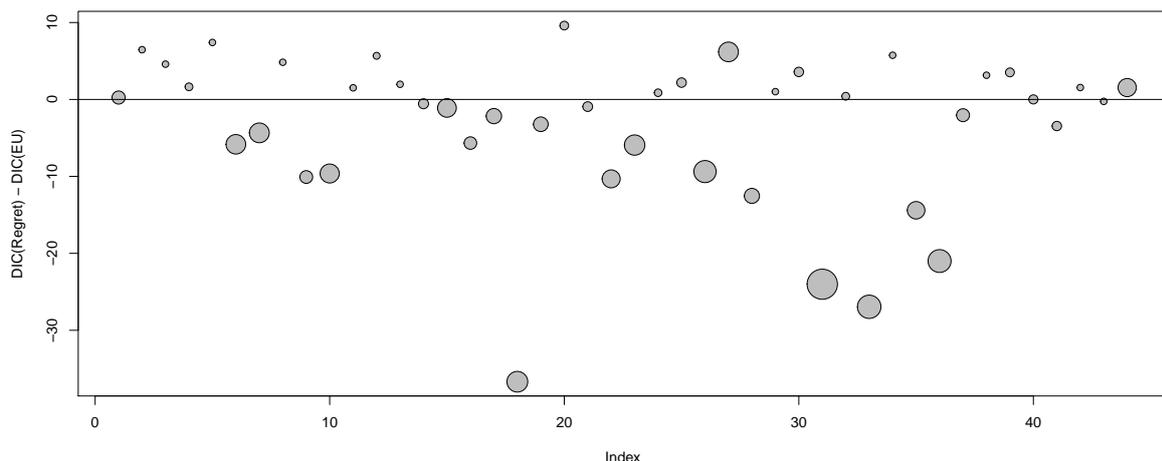


Figure 7: Shows the difference between the DIC of the EU and the regret model.

7. Conclusion

In this paper we sought to answer two questions: What do theories of dynamic behavior predict people to do in stopping problems and how well do these predictions fare when it comes to describing actual behavior? A classical theoretical from the literature is that the optimal strategy of an agent maximizing expected utility is to wait for the payoff process to reach a given reservation level and then stop immediately. That is, (i) agents are predicted to have a unique reservation level and (ii) they should never stop at a point where they chose to continue before, i.e. behave time-consistently.

We have shown that this holds more widely than previously known and thus forms a very robust prediction about behavior. It is not restricted to preferences of the EU type, but also extends to gain-loss preferences and regret preferences. If and to which extent this prediction is satisfied by subjects in the laboratory has been largely unexplored and not recorded in the data in the extant literature. And it seems that a substantial deviation from what theory predicts has therefore gone undetected. Our data confirms that subjects violate the time-consistency property in roughly 75% of all cases and thus the majority of the decisions is irreconcilable with any of the considered theories.

We have also demonstrated that only if one admits unanticipated deviations from the optimal stopping rule, regret preferences become behaviorally different to EU or gain-loss utility. For a regret agent the subgame perfect optimal strategy is as follows: Stop the process if it hits the ex ante reservation level. If you miss to stop at your ex ante level, stop the process if it is at least as large as the past maximum or the ex ante cut-off

of the EU agent. A regret agent will thus have a disposition not to exercise her wait option below the historical maximum of the process. We find this to be a prevalent pattern in our data. Subjects in our experiment seldom stopped at a value the first time they reached it, but also almost never stopped appreciably below the running maximum of the process. Estimating a structural econometric model confirms that a regret model explains the data better, even after penalizing for its additional degrees of freedom.

There are several possible avenues for future research on this topic. We plan for example to investigate whether we can find more direct evidence for the impact of regret by manipulating the channel through which it is supposed to be conveyed. For example, not showing subjects the past of the payoff process, but only its current value makes it less salient to compare the current value with the past maximum. We would then expect that subjects tend to stop at higher values than with the full feedback, since regret is a substitute for risk aversion. In a similar vein, one could redesign the experiment to elicit the willingness to pay for observing the past maximum of the process.

References

- Barberis, N. and Xiong, W. (2009). What drives the disposition effect? an analysis of a long-standing preference-based explanation. *the Journal of Finance*, 64(2):751–784.
- Battigalli, P. and Dufwenberg, M. (2007). Guilt in games. *The American Economic Review*, 97(2):170–176.
- Battigalli, P. and Dufwenberg, M. (2009). Dynamic psychological games. *Journal of Economic Theory*, 144(1):1–35.
- Bergemann, D. and Schlag, K. (2011). Robust monopoly pricing. *Journal of Economic Theory*, 146(6):2527–2543.
- Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Saliency theory of choice under risk. *The Quarterly Journal of Economics*, 127(3):1243–1285.
- Brown, M., Flinn, C., and Schotter, A. (2011). Real-time search in the laboratory and the market. *The American Economic Review*, 101(2):948–974.
- Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J.-R., and Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science*, 304(5674):1167–1170.
- Cheng, X. and Riedel, F. (2013). Optimal stopping under ambiguity in continuous time. *Mathematics and Financial Economics*, 7(1):29–68.
- Coricelli, G., Critchley, H. D., Joffily, M., O’Doherty, J. P., Sirigu, A., and Dolan, R. J. (2005). Regret and its avoidance: a neuroimaging study of choice behavior. *Nature neuroscience*, 8(9):1255–1262.
- Coricelli, G., Dolan, R. J., and Sirigu, A. (2007). Brain, emotion and decision making: the paradigmatic example of regret. *Trends in cognitive sciences*, 11(6):258–265.
- Cox, J. and Oaxaca, R. (1989). Laboratory experiments with a finite-horizon job-search model. *Journal of Risk and Uncertainty*, 2(3):301–329.
- Cox, J. and Oaxaca, R. (1992). Direct tests of the reservation wage property. *Economic Journal*, 102(415):1423–32.
- Cox, J. and Oaxaca, R. (2000). Good news and bad news: Search from unknown wage offer distributions. *Experimental Economics*, 2(3):197–225.

- Craik, K. J. (1947). Theory of the human operator in control systems i. the operator as an engineering system. *British Journal of Psychology*, 38(2):56–61.
- Craik, K. J. (1948). Theory of the human operator in control systems ii. man as an element in the control system. *British Journal of Psychology*, 38(3):142–148.
- Dixit, A. and Pindyck, R. (1994). *Investment under Uncertainty*. Princeton University Press.
- Ebert, S. and Strack, P. (2012). Until the bitter end: On prospect theory in the dynamic context. *Available at SSRN 2005806*.
- Epstein, L. and Schneider, M. (2003). Recursive multiple-priors. *Journal of Economic Theory*, 113(1):1–31.
- Filiz-Ozbay, E. and Ozbay, E. Y. (2007). Auctions with anticipated regret: Theory and experiment. *American Economic Review*, 97(4):1407–1418.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Chapman & Hall/CRC.
- Gilbert, D. T., Morewedge, C. K., Risen, J. L., and Wilson, T. D. (2004). Looking Forward to Looking Backward: The Misprediction of Regret. *Psychological Science*, 15(5):346–350.
- Gilovich, T. and Medvec, V. H. (1994). The temporal pattern to the experience of regret. *Journal of Personality and Social Psychology*, 67(3):357.
- Gilovich, T., Medvec, V. H., and Kahneman, D. (1998). Varieties of regret: A debate and partial resolution. *Psychological Review*, 105(3):602.
- Gneezy, U. (2005). Updating the reference level: Experimental evidence. *Experimental Business Research: Volume III: Marketing, Accounting and Cognitive Perspectives*, 3:263–284.
- Gneezy, U., Kapteyn, A., and Potters, J. (2003). Evaluation periods and asset prices in a market experiment. *The Journal of Finance*, 58(2):821–838.
- Gneezy, U. and Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, 112(2):631–645.
- Grinblatt, M. and Keloharju, M. (2000). The investment behavior and performance of various investor types: a study of finland’s unique data set. *Journal of Financial Economics*, 55(1):43–67.
- Haigh, M. S. and List, J. A. (2005). Do professional traders exhibit myopic loss aversion? an experimental analysis. *The Journal of Finance*, 60(1):523–534.

- Hayashi, T. (2009). Stopping with anticipated regret. *Journal of Mathematical Economics*, 45(7):479–490.
- Hayashi, T. (2011). Context dependence and consistency in dynamic choice under uncertainty: the case of anticipated regret. *Theory and decision*, 70(4):399–430.
- Henderson, V. (2012). Prospect theory, liquidation, and the disposition effect. *Management Science*, 58(2):445–460.
- Hens, T. and Vlcek, M. (2011). Does prospect theory explain the disposition effect? *Journal of Behavioral Finance*, 12(3):141–157.
- Holt, C. and Laury, S. (2002). Risk aversion and incentive effects. *The American Economic Review*, 92(5):1644–1655.
- Jacka, S. (1991). Optimal stopping and the American put. *Mathematical Finance*, 1(2):1–14.
- Kahan, J., Rapoport, A., and Jones, L. (1967). Decision making in a sequential search task. *Attention, Perception & Psychophysics*, 2(8):374–376.
- Kahneman, D. (1973). *Attention and Effort*. Prentice-Hall.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, pages 263–291.
- Kahneman, D. and Tversky, A. (1982). The psychology of preferences. *Scientific American*.
- Krähmer, D. and Stone, R. (2012). Regret in dynamic decision problems. *University of Bonn Working Paper*.
- Lakonishok, J. and Smidt, S. (1986). Volume for winners and losers: Taxation and other motives for stock trading. *The Journal of Finance*, 41(4):951–974.
- Linhart, P. B. and Radner, R. (1989). Minimax-regret strategies for bargaining over several variables. *Journal of Economic Theory*, 48(1):152–178.
- Loomes, G. and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92(368):805–824.
- Loomes, G. and Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, 53(2):271–282.
- McCall, J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, pages 113–126.

- McDonald, R. and Siegel, D. (1986). The value of waiting to invest. *The Quarterly Journal of Economics*, 101(4):707–727.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092.
- Nishimura, G. and Ozaki, H. (2007). Irreversible investment and Knightian uncertainty. *Journal of Economic Theory*, 136(1):668–694.
- Nishimura, K. and Ozaki, H. (2004). Search and knightian uncertainty. *Journal of Economic Theory*, 119(2):299–333.
- Oprea, R., Friedman, D., and Anderson, S. (2009). Learning to wait: A laboratory investigation. *Review of Economic Studies*, 76(3):1103–1124.
- Peskir, G. and Shiryaev, A. (2006). *Optimal stopping and free-boundary problems*. Birkhäuser Basel.
- Post, T., Van den Assem, M., Baltussen, G., and Thaler, R. (2008). Deal or no Deal? Decision making under risk in a large-payoff game show. *The American Economic Review*, 98(1):38–71.
- Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society*, pages 122–136.
- Rabin, M. (2004). Incorporating Fairness into Game Theory and Economics. In Camerer, C., Loewenstein, G., and Rabin, M., editors, *Advances in Behavioral Economics*, chapter 10, pages 297–325. Princeton University Press.
- Rapoport, A. and Tversky, A. (1966). Cost and accessibility of offers as determinants of optional stopping. *Psychonomic Science*, 4:145.
- Rapoport, A. and Tversky, A. (1970). Choice behavior in an optional stopping task. *Organizational Behavior and Human Performance*, 5(2):105–120.
- Riedel, F. (2009). Optimal stopping with multiple priors. *Econometrica*, 77(3):857–908.
- Rogerson, R., Shimer, R., and Wright, R. (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4):959–988.
- Schotter, A. and Braunstein, Y. (1981). Economic search: An experimental study. *Economic Inquiry*, 19(1):1–25.

- Seale, D. and Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the "secretary problem". *Organizational Behavior and Human Decision Processes*, 69(3):221–236.
- Shefrin, H. and Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance*, 40(3):777–790.
- Smith, M. C. (1969). The effect of varying information on the psychological refractory period. *Acta Psychologica*, 30:220–231.
- Stigler, G. (1962). Information in the labor market. *The Journal of Political Economy*, pages 94–105.
- Summers, B. and Duxbury, D. (2007). Unraveling the disposition effect: The role of prospect theory and emotions. *Available at SSRN 1026915*.
- Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186.
- Weber, M. and Camerer, C. F. (1998). The disposition effect in securities trading: An experimental analysis. *Journal of Economic Behavior & Organization*, 33(2):167–184.
- Wilcox, N. T. (2011). 'Stochastically more risk averse': A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics*, 162(1):89–104.
- Xu, Z. Q. and Zhou, X. Y. (2013). Optimal stopping under probability distortion. *The Annals of Applied Probability*, 23(1):251–282.
- Zeelenberg, M., Van Dijk, W. W., Manstead, A. S., and vanr de Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition & Emotion*, 14(4):521–541.
- Zeelenberg, M., van Dijk, W. W., Van der Pligt, J., Manstead, A. S., Van Empelen, P., and Reinderman, D. (1998). Emotional reactions to the outcomes of decisions: The role of counterfactual thought in the experience of regret and disappointment. *Organizational behavior and human decision processes*, 75(2):117–141.

A. Mathematical Appendix

Proof of Lemma 1. Define the running maximum $S_t = \max_{s \leq t} X_s$. Note that

$$\mathbb{P}[\tau(b) < T \mid X_0 = x] = \mathbb{P}[S_T \geq b \mid X_0 = x] .$$

For all $x < b$ the probability of reaching the level b is equal to the probability of the deadline not arriving δ times the expected probability of reaching b from period 1 perspective

$$\begin{aligned} \mathbb{P}[S_T \geq b \mid X_0 = x] &= \mathbb{P}[\tau > 1] \mathbb{E}[\mathbb{P}[S_T \geq b \mid X_1, \tau > 1] \mid X_0 = x] \\ &= \delta (p \mathbb{P}[S_T \geq b \mid X_1 = xh, \tau > 1] \\ &\quad + (1 - p) \mathbb{P}[S_T \geq b \mid X_1 = xh^{-1}, \tau > 1]) \\ &= \delta (p \mathbb{P}[S_T \geq b \mid X_0 = xh] + (1 - p) \mathbb{P}[S_T \geq b \mid X_0 = xh^{-1}]) . \end{aligned}$$

To simplify notation define $\psi_b(x) = \mathbb{P}[S_T \geq b \mid X_0 = x] = \psi_b(x)$. By definition ψ_b is a solution to the difference equation

$$\psi_b(x) = \begin{cases} 1 & \text{for all } x \geq b \\ \delta (p\psi_b(xh) + (1 - p)\psi_b(xh^{-1})) & \text{for all } x < b \end{cases} , \quad (\text{A.1})$$

taking values in $[0, 1]$. If we have two solutions $\psi_b, \hat{\psi}_b$ of Equation A.1 it holds that

$$\begin{aligned} |\psi_b(x) - \hat{\psi}_b(x)| &= \mathbf{1}_{\{x < b\}} \delta \left| p(\psi_b(xh) - \hat{\psi}_b(xh)) + (1 - p)(\psi_b(xh^{-1}) - \hat{\psi}_b(xh^{-1})) \right| \\ &\leq \delta \sup_{z < b} |\psi_b(z) - \hat{\psi}_b(z)| . \end{aligned}$$

As $\psi_b(x), \hat{\psi}_b(x)$ lie between zero and one, the supremum of the differences $\sup_z |\psi_b(z) - \hat{\psi}_b(z)|$ exists and is bounded by one. As $\delta < 1$ it follows that

$$\sup_z |\psi_b(z) - \hat{\psi}_b(z)| \leq \delta \sup_z |\psi_b(z) - \hat{\psi}_b(z)| = 0$$

and thus Equation A.1 can have at most one solution taking values in $[0, 1]$. Guessing the solution of (A.1) to be of the form $\psi_b(x) = \mathbf{1}_{\{x < b\}} (\frac{x}{b})^\alpha + \mathbf{1}_{\{x \geq b\}}$ gives $1 = \delta (ph^\alpha + (1 - p)h^{-\alpha})$. Substituting $z = h^\alpha$ yields the quadratic equation

$$0 = \delta (pz + (1 - p)z^{-1}) - 1 = z^2 - \frac{z}{\delta p} + \frac{1 - p}{p} \Rightarrow z = \frac{1}{2\delta p} \pm \sqrt{\frac{1}{4\delta^2 p^2} - \frac{1 - p}{p}} .$$

By symmetry this equation has two solution of the form z, z^{-1} . Lets denote $z > 1$ the solution larger than one. For the smaller solution $\frac{1}{z}$ one it follows that $\alpha = \frac{\log(\frac{1}{z})}{\log(h)} < 0$. Hence the

resulting function $\psi(x) = \mathbf{1}_{\{x < b\}}(\frac{x}{b})^\alpha + \mathbf{1}_{\{x \geq b\}}$ is decreasing and thus takes values outside $[0, 1]$. Note that the function $z \mapsto \delta(pz + (1-p)z^{-1})$ is increasing for all $z \geq 1$. By Assumption A1, $\delta(ph + (1-p)h^{-1}) < \delta(pz + (1-p)z^{-1}) = 1$ and thus $1 \leq h < z = h^\alpha$ and $\alpha > 1$. \square

Proof of Lemma 3. By Lemma 2 the expected value of any cut-off strategy with cut-off $b = x_0 h^n$ is given by

$$V(\tau(x_0 h^n), x_0) = u(x_0 - K) + \sum_{j=1}^n h^{-(j-1)\alpha} \Gamma(x_0 h^j).$$

If u is concave Γ is monotone decreasing (as it follows by setting $\kappa = 0$ in the proof of Lemma 7). As $\Gamma(b^u) < 0$ it follows from the monotonicity of Γ that $\tau(b^u)$ is the optimal cut-off strategy. Denote by $x > b^u$ the point where $\mathcal{L}u(x - K) > 0$. Clearly, as it not optimal to stop at x if the optimal strategy is a cut-off strategy the optimal cut-off b must be greater x . As $x > b^u$ this is a contradiction. \square

Proof of Lemma 4. As shown in the proof of Proposition 1 $\mathcal{L}u(b^u - K) < 0$. Thus, if $\mathcal{L}u$ changes its sign at most once this implies Assumption A2. In the final step we show for constant absolute or relative risk-aversion $\mathcal{L}u$ changes its sign at most once.

Constant Absolute Risk Aversion: Suppose the agent accepts such a gamble at the wealth level x . Let $u(x) = -\frac{1}{\theta} \exp(-\theta x)$, i.e. assume the agent has constant absolute risk-aversion of θ . The expected change in utility from waiting one more round at x equals

$$\begin{aligned} \mathcal{L}u(x - K) &= \delta(pu(xh - K) + (1-p)u(xh^{-1} - K)) + (1 - \delta)u(0) - u(x - K) \\ &= \delta u(x - K) \left[\left(p \frac{u(xh - K)}{u(x - K)} + (1-p) \frac{u(xh^{-1} - K)}{u(x - K)} \right) + (1 - \delta) \frac{u(0)}{u(x - K)} - 1 \right] \\ &= -\frac{e^{-\theta(x-K)}}{\theta} \left[\delta(p e^{-\theta x(h-1)} + (1-p)e^{-\theta x(h^{-1}-1)}) + (1 - \delta)e^{\theta(x-K)} - 1 \right] \end{aligned}$$

We will show that the second part is monotone increasing in x . Taking derivatives of the term in square brackets gives

$$\delta(-\theta(h-1)p e^{-\theta x(h-1)} + \theta(1-h^{-1})(1-p)e^{\theta x(1-h^{-1})}) + (1-\delta)\theta e^{\theta(x-K)}$$

As $e^{-\theta x(h-1)} < 1$ and $e^{\theta x(1-h^{-1})}, e^{\theta(x-K)} > 1$ a lower is given by

$$\begin{aligned} &\geq \theta [\delta(-(h-1)p + (1-h^{-1})(1-p)) + (1-\delta)] \\ &= \theta [-\delta(hp + h^{-1}(1-p)) + 1] > 0, \end{aligned}$$

Where the last step follows as $hp + h^{-1}(1-p) < 1$ by Assumption A1. Consequently $\mathcal{L}u$ changes its sign at most once.

Constant Relative Risk Aversion: Let $u(x) = \frac{(x+K)^\theta - K^\theta}{\theta}$. The expected change in utility from waiting one more round at x equals

$$\begin{aligned}\mathcal{L}u(x - K) &= \frac{\delta}{\theta}(p(xh)^\theta - K^\theta + (1-p)(xh^{-1})^\theta - K^\theta) - \frac{1}{\theta}(x^\theta - K^\theta) \\ &= \frac{1}{\theta} \left\{ \delta \left[p(xh)^\theta + (1-p)(xh^{-1})^\theta \right] - x^\theta + (1-\delta)K^\theta \right\} \\ &= \frac{1}{\theta} x^\theta \left(\delta \left[ph^\theta + (1-p)h^{-\theta} \right] - 1 \right) + \frac{1}{\theta} (1-\delta)K^\theta.\end{aligned}$$

As $p > 1/2$ for all $\theta \geq 0$

$$\frac{\partial}{\partial \theta} \left(ph^\theta + (1-p)h^{-\theta} \right) = p \log(h)h^\theta - (1-p) \log(h)h^{-\theta} \geq p \log(h)(h^\theta - h^{-\theta}) \geq 0.$$

Thus, $ph^\theta + (1-p)h^{-\theta} < ph^\alpha + (1-p)h^{-\alpha} = \frac{1}{\delta}$ for all $\theta < \alpha$, by definition of α . As $\frac{1}{\delta}x^\theta$ is increasing in x this completes the proof. \square

Proof of Lemma 8. First, we derive the probability that the maximum of the process is at least $y \in \mathcal{X}$

$$\mathbb{P}[S_T \geq y \mid X_t = x, S_t = s] = \begin{cases} 1 & \text{if } s \geq y \\ \mathbb{P}[\tau(y) < T \mid X_t = x] & \text{if } s < y \end{cases}$$

Hence, we have that the probability that the maximum of the process is exactly $y \in \mathcal{X}$ for all $s < y$ equals

$$\begin{aligned}\mathbb{P}[S_T = y \mid X_t = x, S_t = s] &= \mathbb{P}[S_T \geq y \mid X_t = x, S_t = s] - \mathbb{P}[S_T \geq yh \mid X_t = x, S_t = s] \\ &= \left(\frac{x}{y}\right)^\alpha - \left(\frac{x}{yh}\right)^\alpha = \left(\frac{x}{y}\right)^\alpha (1 - h^{-\alpha})\end{aligned}$$

Let $b = xh^m$. Given the regret functional derived in equation 4.6 the expected value of using the cut-off strategy $\tau(b)$ equals

$$\begin{aligned}V(\tau(b), x, s) &= \mathbb{E} \left[\mathbf{1}_{\{\tau(b) < T\}} u(X_{\tau(b)} - K) - \kappa u(S_{\tau(b)} - K) \mid X_t = x, S_t = s \right] \\ &= \mathbb{P}[\tau(b) < T \mid X_t = x] u(b - K) \\ &\quad - \kappa \sum_{i=0}^m \mathbb{P}[S_T = xh^i \mid X_t = x, S_t = s] u(xh^i - K) \\ &= \left(\frac{x}{b}\right)^\alpha u(b - K) - \kappa \sum_{i=0}^{m-1} \left(\frac{x}{xh^i}\right)^\alpha (1 - h^{-\alpha}) \max\{u(s - K), u(xh^i - K)\} \\ &\quad - \kappa \left(\frac{x}{b}\right)^\alpha \max\{u(s - K), u(b - K)\}.\end{aligned}$$



B. Figures

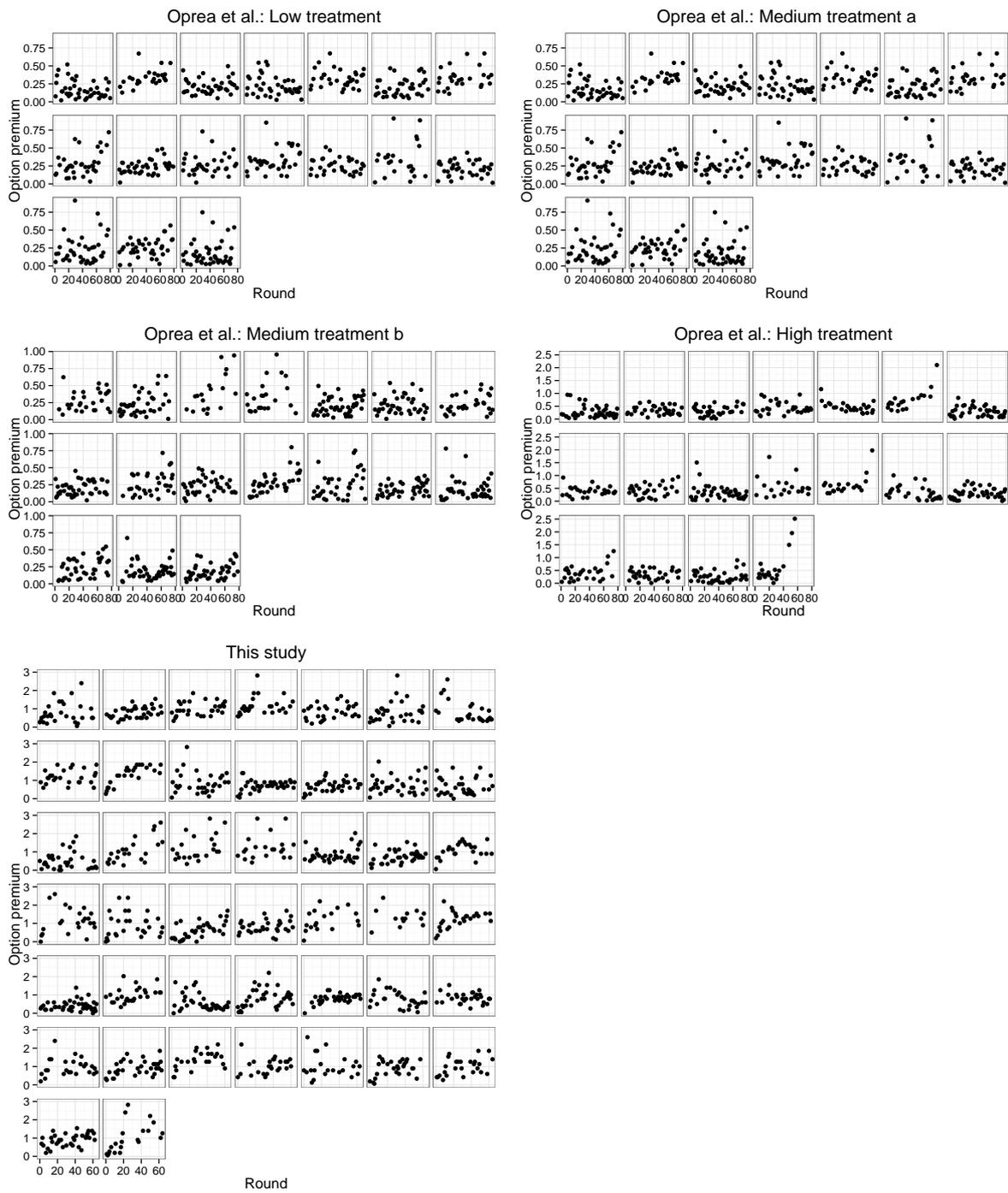


Figure 8: Shows the observed reservation levels for over rounds as played by subjects. Each panel in a given block of panels is the sequence of reservation levels for one subject.

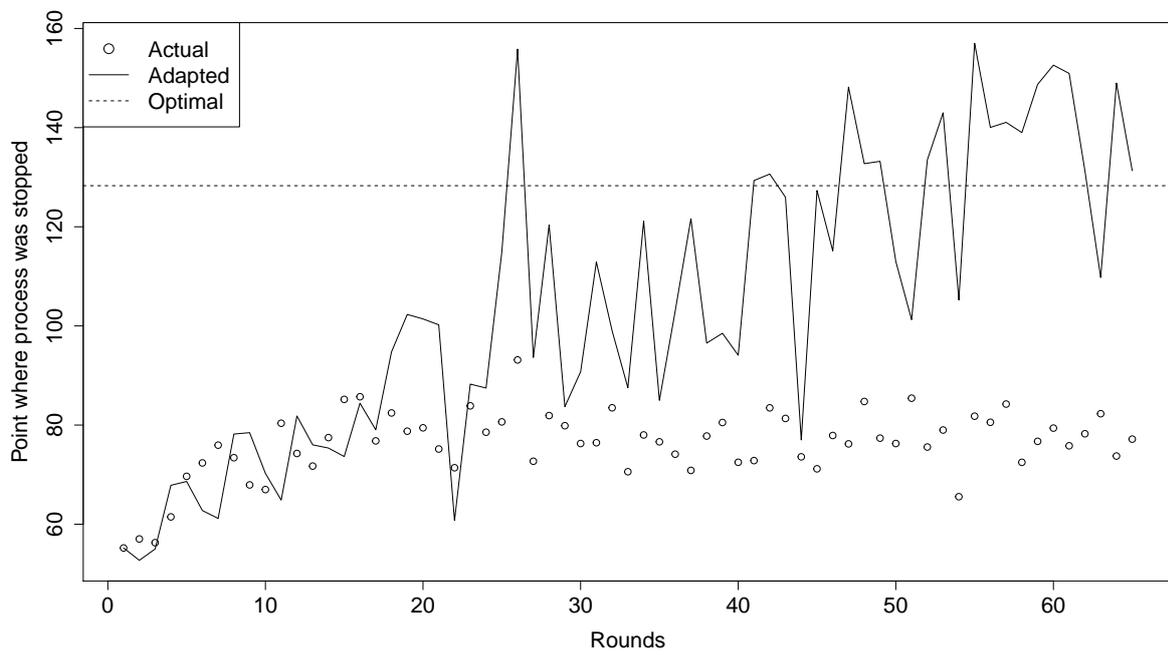


Figure 9: Shows simulation results from adaptive learning model versus actual choices.