Fooled By Randomness? Financial Decision-Making Under Tail Risk

Elise Payzan-LeNestour†

February 3, 2016

This study considers model uncertainty about tail risk, the possibility of learning and how this affects choice. I designed an experiment involving repeated risk taking where assets can yield steady streams of good outcomes but eventually inflict a major loss that annihilates all previous gains. Learning about those assets’ risk/reward profiles is crucial yet challenging. The main findings are: 1) When asked to perform a stylized version of the task, participants managed to learn in a Bayesian way; 2) However, many still chose to invest in these assets, apparently owing to an overwhelming desire to “pick pennies.” These findings suggest that the issue with tail risk is not the most commonly expected one, namely, that people cannot assess it, but rather that people cannot deal with it properly due to purely behavioral issues related to limited self-control. (JEL C91, D83, D87, G02, G11)

Keywords: Tail Risk, Learning, Self-Control, Negative Skewness, Neurofinance

*This study is the product of many discussions with colleagues and friends from economics, finance, and the neurosciences. Special thanks to Patrick Bolton for invaluable help in writing the paper, and to Andrew Caplin, Paul Glimcher, and the seminar participants at The Center for Neuroeconomics at New York University, for many insightful suggestions during the very early stages of the study. I am also indebted to James Andreoni, Tony Berrada, Peter Bossaerts, Cary Frydman, Rajna Gibson, Botond Koszegi, Ron Masulis, Maureen O’Hara, Antonio Rangel, Olivier Scaillet, Mike Woodford, seminar participants at Boulder University, Caltech, Geneva University, the JP Morgan Quantference, and the Melbourne Finance Down Under conference, for very helpful comments and suggestions. On the practitioner-turned-academic side, I am extremely grateful to Lionnel Pradier, former Head of Quantitative Research at JP Morgan, for helping me draw the implications of this study for finance practitioners. The analysis and conclusions set forth are those of the author. Mehdi Adibi and Le Zhang provided outstanding research assistance.

†UNSW Australia Business School, University of New South Wales, Sydney. E-mail: elise@unsw.edu.au
1 Introduction

This paper investigates the behavior of investors facing not only parameter uncertainty about variables given a particular assumed distribution over asset returns but also structural uncertainty about the distribution itself, i.e., model uncertainty. Model uncertainty is a key refinement in recent economic theories (Hansen and Sargent, 2001). But is it just a convenient abstraction or also a key feature of individual investors’ behavior? This study sets out to answer this general question by designing a simple experiment involving learning and choice under model uncertainty.

The importance of model uncertainty for financial practitioners is clear. Return distributions often switch between Normal and fat-tailed distributions.\(^1\) Overlooking this fact has multifarious costs, including significant economic losses in asset pricing (Donnelly and Embrechts, 2010) and asset allocation (Kacperczyk and Damien, 2011). Factoring model uncertainty about tail risk in one’s decisions is thus crucial.\(^2\) Yet it is no simple task, as it involves learning about both the distribution and its parameters. Here I investigate how investors learn about model uncertainty and how they act on this learning.

I find strikingly robust evidence that people are able to learn about model uncertainty according to Bayes’ rule, despite the difficulty. However, I further find that people fail to act on this sophisticated learning. They fall victim to a very specific and novel (to the best of my knowledge) behavioral bias, which consists of exposing oneself to extreme levels of tail risk, owing to an overwhelming desire to win additional pennies. I combine different experimental treatments to show that this behavioral bias is not an artifact stemming from a mischaracterization of subjects’ learning as sophisticated; in fact, the bias persists as such even when no learning is involved.

The experimental setting used in this study is a stylized task in which the agent has to distinguish profitable reward prospects from sure bets to lose money masquerading as glamorous lotteries. The task features a bowman who, at each trial, is shooting at a target. In each trial, the agent decides whether to attempt to win $2 by betting that the next shot will hit the mark or will be a near miss. A losing bet—in the case

---

1\(^{1}\)E.g., Mandelbrot (1957) and Gabaix et al. (2003); see Ang and Timmermann (2011) for a survey.
2\(^{2}\)See, e.g., Donnelly and Embrechts (2010), pp.16-18: “Both the Normal copula and the Black-Scholes-Merton model are based on the normal distribution. Both are easy to understand and result in models with fast computation times. Yet both fail to adequately model the occurrence of extreme events. […] The broader lesson to take away is that of model uncertainty.”
that the hit goes wide—results in a loss of $40. The alternative action to betting is to “skip,” which gives $0 for sure. The agent plays 15 independent sessions of 20 trials each, each with a different bowman. There are two types of bowmen, “master” and “apprentice.” The shots from a master bowman are normally distributed around the target. The shots from an apprentice bowman are Cauchy distributed, i.e., they markedly depart from normality. By design, the optimal behavior of an omniscient agent consists of betting on the sessions with a master bowman and skipping on the sessions with an apprentice.

At the onset of any session, the agent is not told whether he faces an apprentice or a master. However, the agent sees the realized shot for each of the 20 trials, which allows the Bayesian agent to gather evidence about the bowman’s type. Optimal behavior consists of betting if and only if the evidence that the bowman is a master is above a threshold (which I define in Section 3). If Bayesian learning proves too hard, we can assume the agent reverts instead to one of the boundedly rational alternatives that have received ample empirical support, such as the generic “Experienced-Weighted Attraction” (EWA) learning approach introduced by Camerer and Ho (1999). The EWA learning formulation subsumes a wide category of adaptive behaviors (which I describe in Section 3). Its core principle is the ingrained “law of effect” (Thorndike, 1898), whereby the value of an action merely reflects the outcomes—the actual outcome as well as the counterfactual—that the action yielded in the recent past. By automatically deeming an action good if it yielded good outcomes in the past, this approach conveniently obviates inferring the outcome probabilities. This formulation of learning has been popular in experimental psychology and experimental finance (Pouget, 2007), and it has solid neurobiological foundations (e.g., Schultz et al. (1997) and Lohrenz et al. (2007)).

University students playing for very significant monetary payoffs performed the task. I find that the Bayesian model fits their behavior markedly better than all the boundedly rational contenders. In follow-up investigations, direct elicitation of the subjects’ beliefs regarding the type of the bowman they face strengthens the evidence for Bayesian learning in the task: the subjects’ beliefs closely match those of the Bayesian agent. I document these findings in Section 4.

The second main result of this study cropped-up unexpectedly in the process of studying learning. I find that as much as 41% of the task participants bet after encountering a large outlier whose occurrence made it clear that the bowman was
an apprentice (so betting had a highly negative expected value). This behavioral pattern is inconsistent with Bayesian behavior, but appears to coexist with Bayesian learning. The simplest explanation for this apparently puzzling coexistence would be that I mistakenly identified Bayesian learning when in reality subjects did not learn successfully which type of bowman they were facing. This is not the case though. Notably, even when subjects are told the type of bowman they are facing before each session, they act in the same way (this was done as a follow-up experimental treatment). This behavioral bias is reminiscent of a widespread behavioral pitfall which finance practitioners refer to as “picking pennies in front of a steamroller.” The financial history is strewed with episodes in which this pitfall—henceforth, “the picking pennies bias”—prevailed. I elaborate below.

One puzzles over the behavior of those subjects who picked pennies (“the penny-pickers”), inasmuch as apart from the foregoing behavioral bias, their behavior is similar to the behavior of the other subjects. The most natural explanations of the occurrence of the picking pennies bias are all ruled out. Specifically, I find that the penny-pickers were not more risk lovers than the other subjects during the task. Their beliefs during the task were similar to those of the other subjects; the evidence suggests that most often, they knew that the bowman was an apprentice when they were picking pennies. Neither does the picking pennies bias reflect behavioral mistakes (“trembling”), “choice stickiness” (the tendency to stick to one’s ongoing course of action; see Feltovich (2000)), “Probability Matching” (the tendency to choose randomly between potential actions, where the randomization matches the probability of success of each action\(^3\)), or “the Gambler’s Fallacy” (the underestimation of the probability of sequential streaks occurring by chance\(^4\)).

However, I find that several of the penny-pickers behaved as if they perceived the loss outcome to be four times smaller than the actual loss outcome, which led them to persevere in betting after losing. This type of behavior is consistent with recent neuroscientific evidence that learning to avoid negative outcomes is hampered by overstimulation of the “reward or wanting system” in the brain [the midbrain region that signals rewards through dopamine release (Panksepp, 2004)]. See, e.g., Frank et al. (2004), Knutson et al. (2008), and Frank and Hutchison (2009). In light of these papers, it may seem natural to conjecture that the picking pennies bias emerges in

---

\(^3\)See, e.g., Vulcan (2000), Shanks et al. (2002), and Brennan and Lo (2011).

the task from the repeated exposure to good payoffs, which overstimulates the reward system, i.e., exacerbates greed—“the greed hypothesis.”

The greed hypothesis implies that those subjects who picked pennies in the task could not help betting, i.e., their behavior was at variance with deliberation. To test this implication, I ran a follow-up experimental treatment that replicated the original task except that the subjects in each trial were also given the option to either bet or skip in all of the remaining trials of the session. I find that the subjects made extensive use of the option to skip for all trials. Thus, when betting was suboptimal during the task, subjects effectively bound themselves to not betting. Such “self-binding” is reminiscent of Ulysses’ reliance on an externally imposed constraint to resist the call of the sirens (Elster, 1979). I further find that in that treatment, the subjects are overwhelmingly Bayesian, and the prevalence of the picking pennies bias is significantly reduced. This result points to greed coupled with limited self-control as the primary cause of the picking pennies bias in the task.

Finally, I ran a couple of experimental sessions that replicated the original treatment of the task except that just before each session began, the subjects were told the nature of the bowman in the session, whereby the subjects knew beyond a doubt the type of the bowman in each session. The picking pennies bias prevailed exactly like in the original experiment; it thus appears to be remarkably robust. I report those findings in Section 4.3.3 (“Robustness Checks”).

Relation to the literature and implications This study adds to the growing experimental finance literature on learning in financial markets. Broadly speaking, this literature suggests that the ability to implement Bayes’ rule varies across individuals (Kluger and Wyatt, 2004) and is situation-dependent, influenced in particular by emotions (e.g., Charness and Levin (2005), Lohrenz et al. (2007), Kuhnen and Knutson (2011)) and the valence of past outcomes (e.g., Kuhnen and Knutson (2005), Malmendier and Nagel (2011), and Kuhnen (2014)). Closest to the current study, Payzan-LeNestour et al. (2013) and Payzan-LeNestour and Bossaerts (2015) provide both neural and behavioral evidence that when given the right incentives and suffi-

5Malmendier and Nagel (2011) and Kuhnen (2014) document that after witnessing bad economic payoffs, agents form pessimistic beliefs about future stock returns, which is consistent with neuroscience evidence that the brain processes deployed when people learn from their environment differ depending on whether they are faced with positive or negative outcomes (Kuhnen and Knutson, 2005).
cient background information, people can assess unstable asset values according to Bayes’ rule, despite the difficulty. The task used there involved the combined learning of jump and outcome probabilities, whereby parameter uncertainty was maximal. However it did not involve learning about the type of return distribution, i.e., model uncertainty was absent. Here I use similar levels of incentives and increase complexity by a notch, to assess whether people are able to learn according to Bayes’ rule under model uncertainty as well. I provide evidence that they are.

A novel contribution of this study, therefore, is to document that task participants were able to learn about the type of payoff distribution. This finding may come as no surprise in light of the “volatility skew” featured by equity options since the 1987 crash. Before the crash, agents thought the underlying stock returns followed a Normal distribution (because the empirical distribution available from past history was roughly Normal). They learned from the crash that the distribution was actually fat-tailed, and the prices of out-of-the-money options, particularly put options, increased accordingly (Bates, 2000). Farhi et al. (2014) document that after the Fall of 2008 a similar phenomenon occurred with currency options prices.

The finding that task participants could learn fairly well about tail risk is also consistent with the existing literature on the psychology of tail events. That literature indeed points to two cognitive biases concerning learning about rare events: the “availability heuristic” (Tversky and Kahneman, 1974), which causes the agent to over-estimate (resp. under-estimate) the likelihood of the rare event when the latter is salient (resp. not salient), and limited sampling of the rare event (Hertwig et al., 2004), which causes the agent to under-estimate that same likelihood. See Barberis (2013) for a survey. In the current task the loss outcome was very salient (making subjects exaggerate its probability) and sampling of the loss outcome was limited by design (making subjects underestimate its probability); thus the two antagonistic biases canceled each other out, which would explain that subjects were well calibrated on average.

Perhaps the most important contribution of the current study is to document a novel behavioral bias—the picking pennies bias. I provide evidence that task participants picked pennies despite knowing that the bowman was an apprentice, because they could not resist the temptation to do so. This study thus adds to the literature on self-control in finance. On the theoretical side, Shefrin and Statman (1984) show how self-control considerations may shape corporate decisions, and Gul et al.
(2014) conjecture that investors’ limited self-control has a significant impact on financial market prices. However, the empirical evidence has been inconclusive so far (e.g., DeJong and Ripoll (2007)). Notably, self-control is usually formalized within the “delayed gratification” framework, i.e., it concerns the temptation to liquidate one’s financial holdings to enjoy immediate consumption, whereas the present study is about the temptation to pick pennies. I show that such a temptation seems to be very real and very robust; asset pricing models could account for it to yield new predictions.

Interestingly, the prevalence of the picking pennies bias in the current experiment suggests that repeated exposure to good outcomes makes the agent like left skewness, which seems at odds with the basic prediction from prospect theory that people dislike left skewness. Together with prospect theory, the current evidence suggests that aversion for left skewness occurs with one-shot gambles and completely vanishes with sequential gambles that feature repeated exposure to good outcomes. That is, people’s attitude toward skewness would highly depend on the circumstances—as already noted in Barberis (2013).

By showing that subjects picked pennies regardless of how much they were actually aware of the risks at stake, the current findings may help flesh out the nature of the psychological forces at stake during financial crises. All historical analyses converge on the idea that “irrational exuberance” (Shiller, 2000) plays a pivotal role in financial crises. However, the root causes underlying the phenomenon of irrational exuberance are not completely understood—and Shiller (2000) warns us against “pop-psychological” theories devoid of empirical support. It has been proposed that the real reason for irrational exuberance is the public’s overoptimism (see Part 3 of Shiller (2000), Shefrin (2009), and Ubel (2009), among others). But in other accounts, it is greed that stands out as the primary force behind irrational exuberance. For instance, in the well-known Kindleberger-Alibert-Minsky paradigm, the real reason for an economic bubble is greed, fueled by aggressive bank lending (Kindleberger and Alibert, 2005). Shiller himself recognizes that the overoptimism account has some

---

6 For instance, Shiller (2000) defines irrational exuberance as a key “amplifying mechanism” which causes precipitating events (such as demographical booms and technological innovations) to have an outsized effect on the stock market.

7 For instance, in their description of the emotional frame of mind of investors during economic booms, Kindleberger and Alibert (2005) refer to “a pervasive sense that it is time to ‘get on the train before it leaves the station’ and the profit opportunities disappear” (p.12).
limits and that greed is a force to be reckoned with.\textsuperscript{8} He conjectures that when people decide to participate in the stock market at the peak of a speculative bubble, “\textit{deep down they know} that the market is highly priced, and they are uncomfortable about this fact” (Shiller (2000), p.14). The current findings provide experimental evidence for this view. I document that in the current experiment, subjects picked pennies regardless of how much they were actually aware of the risks at stake, because they could not help doing so.

Looking at the financial history over the past 40 years, there have been several episodes in which investors picked pennies like the penny-pickers in the current experiment. The high-yield bond market at the end of the 1980s is a case in point. This niche of the credit market in the 1970s had grown to a $100 billion market in the mid 1980s. At this stage, it was still a relatively new market with a short history, but following the large number of defaults in 1986, it was clear that high-yield bonds were not a prudent investment. Nevertheless, the demand for higher-yielding debt continued unabated for three years (see Altman (1987) and Altman (1992)).\textsuperscript{9} Similar behavior prevailed in 2007; some agents continued to invest in Residential Mortgage Backed Securities (RMBS) CDOs even though in the summer of 2006, distress among subprime mortgage lenders became clear, as documented in Crouchy et al. (2008).\textsuperscript{10}

The current findings may also help understand the behavior of those who invested in high-yield bonds just prior to the 2008 global financial crisis. In that period, the demand for high-yield bonds was insatiable despite record-low yields (Reilly et al., 2009)\textsuperscript{11} and despite some Wall Street analysts warning yield-starved investors that the market was ripe for a fall. Such high-yield bond mania may seem very strange at first. High-yield bonds are notorious for their high risk levels since the high-yield meltdown at the end of the 1980s, which suggests that those investors were picking

\textsuperscript{8}Specifically, Shiller documents that in surveys he ran with investors during a bubble episode, overall expectations of price increases among most investors were not overly optimistic but quite reasonable (Shiller (2000), p.54). He relates the desire to participate in the stock market during a speculative bubble to “the feeling that the stock market is the only game in town” (pp.55-56).

\textsuperscript{9}The size of the high-yield debt market grew to approximately $190 billion in 1989. The annual return spread over US Treasuries from 1978 to 1989 dropped considerably, to just under one percent per year.

\textsuperscript{10}At that time, rating agencies and banks issued warnings about the deteriorating state of the subprime market (e.g., “Early Warning Signs in Mortgage Credit,” Mortgage Strategist, UBS, August 22, 2006).

\textsuperscript{11}In 2007 high-yield bonds were yielding 2.5\textsuperscript{1}, a historical low, and would return −28\% the following year.
pennies despite knowing the risks at stake. Our findings in the uncertainty-free variant of the task (the one in which subjects were told the nature of the bowman before each session began, and yet picked pennies) point to greed coupled with limited self-control as being the root cause of their behavior.

2 Experimental Design

2.1 Betting task

The task features a bowman who, at each trial, is shooting at a target on a wall. The wall is represented by a line; the target corresponds to the zero mark on that line, and the shot realized at trial $t$, denoted by $X_t$, can be anywhere on the line. In each trial $t$, the agent must decide whether to bet that the next shot will fall up to four meters away from the target on both sides (i.e., $X_t \in [-4; 4]$). A winning bet yields $2; a losing bet ($X_t \notin [-4; 4]$) results in a loss of $40. The alternative to betting is to “skip,” which gives $0 for sure. Immediately after deciding whether to bet or skip, the agent sees the realized shot and then proceeds to the next trial. The goal of the subject is to maximize the outcomes accumulated from all trials.

2.2 Stochastic structure

The agent plays 15 independent sessions comprising 20 trials each, and each session has a different bowman. There are two types of bowmen: “master” and “apprentice.” The shots from a master bowman are normally distributed around the target (i.e., the mean is 0) with a standard deviation that is unknown to the agent. The value of the standard deviation differs across master bowmen; it is uniformly distributed between 0.1 and 2. The shots from an apprentice bowman are Cauchy distributed around the target (i.e., the center is 0), with a dispersion of 1. At the onset of each session, there is an equal chance of facing an apprentice or facing a master.

2.3 Information provided to the subjects

The text of the task instructions can be found in Appendix 7.3. The agent knows the structure of the task (15 independent sessions, each comprising 20 trials) and that in each session, there is an equal chance of facing an apprentice or facing a master.
agent further knows that the shots from a master are normally distributed with zero
mean and a standard deviation uniformly distributed between 0.1 and 2, whereas the
shots from an apprentice are Cauchy distributed with a center of 0 and a dispersion
of 1.

To ensure that the subjects grasped the stochastic structure of the task, the task
instructions did not simply tell the expected values of betting under each kind of
bowman. Rather, the instructions allowed the subjects to actually experience those
values, using Gigerenzer et al. (1988)’s method. Specifically, the subjects saw an
animation or “distribution builder” showing 300 successive sample shots from an ap-
prentice as well as three other distribution builders showing 300 successive shots from
different master bowmen—the first with the minimum level of standard deviation
(0.1), the second with the maximum level (2), and the third with an intermediate
level (1). At the end of the experimental session, subjects filled out a debriefing
questionnaire that assessed their estimate of the likelihood of losing after betting with
an apprentice and a master bowman respectively. Subjects were well calibrated over-
all, with a slight tendency to exaggerate the probability of losing with an apprentice
bowman.

In each session, the subjects were not told whether the bowman was a master, so
they faced model or distribution uncertainty (as defined in the Introduction). They
further faced parameter uncertainty inasmuch as the standard deviation of the shots
from the master bowmen was random and hidden. As such, the task is primarily a
complex learning task combining both model and parameter uncertainty.

2.4 Incentives

An ample literature suggests that one should provide participants with high monetary
incentives to allow sophisticated behavior to emerge in complex tasks. See, among
others, Wilcox (1993), Hertwig and Ortmann (2001), and Charness et al. (2010).

The method used here is grounded in the distinction between experienced vs. stated represen-
tations of randomness. Psychologically, there are important differences in the perception of randomness
that is stated in the form of summary statistics (such as reading the probability 0.25 or reading that
the distribution is normally distributed with a mean of 0 and a standard deviation of 1) and ran-
doness that is experienced (such as seeing 25 of 100 coin flips turn up heads or seeing 300 samples
drawn from the standard normal distribution); see, e.g., Gigerenzer et al. (1988) and Goldstein et al.
(2006). For example, Gigerenzer et al. (1988) allowed people to experience random sampling directly
in Kahneman and Tversky (2009)’s classic Engineer-Lawyer problem and as a result, the “base rate
neglect” bias diminished.
Given the complexity of the current task, in order to ensure that subjects were highly motivated to maximize their outcomes in the task, I used a payoff rule that neatly discriminated between high versus mediocre task performance levels. As per this rule, which I describe in Section 4.1, one-third of the subjects ended up with more than $100 from the experiment. Thus, the task participants were provided with considerable monetary incentives, higher than in similar experiments (except the one reported in Payzan-LeNestour and Bossaerts (2015), which used similar payoff scheme), and important relative to their standard of living.

2.5 Why the current experimental design

To test whether investors can learn about tail risk, other tasks could be devised; why then using this specific task? The current task has clear financial undertones. It is best understood as a stylized representation of the credit market in several periods of the financial history. Specifically, apprentice bowmen represent assets that combine high levels of tail risk and low payoffs, like many asset classes in specific periods of the financial history; examples include high-yield bonds in 1989 (see Altman (1987) and Altman (1992)) and RMBS CDOs in 2007 (see Crouchy et al. (2008)). In this analogy, master bowmen represent investment-grade bonds or government bonds.

Despite its financial undertones, the task is devoid of any financial connotation, for two reasons. First I wanted to avoid subjecting the task participants to framing effects. Second, I wanted to limit the extent of confounding factors in the analysis. The idea here is that given the documented extent of financial illiteracy (e.g., Bernheim (1995) and Lusardi and Mitchell (2007)), any factors associated with financial illiteracy and cognitive limitations would masquerade as each other in a financial task. Here I isolate the cognitive capabilities of people to cope with tail risk, irrespective of their mastery of financial concepts.

The main motivation for using the Cauchy distribution to model fat-tailed payoffs is tractability. Using other fat-tailed distributions (the Lévy distribution in particular, see Mandelbrot (1960)) would not change the key features of the theory described in the next Section, but it would make it more complicated.

The motivation for using the current parameters (for the payoff levels, standard deviations, dispersion parameter, etc.) is that the resulting task naturally lends itself to a model comparison analysis between the Bayesian and boundedly rational
learning models, inasmuch as the behaviors of the models unambiguously differ under this specification of the task (in a sense that will become clear in the next Section). It is thus possible to neatly distinguish between Bayesian and non-Bayesian behavior in the subjects. Other parameter specifications were tested in Monte Carlo simulations. The current specification is the one that turned out to provide the best statistical power for the model comparison analysis.

3 The Behavioral Models

The optimal approach to solving the task consists of inferring the hidden state of the world in each session (whether the bowman is a master or an apprentice) by observing the realized shots. Bayesian learning offers a parsimonious, principled method to make such an inference; see Section 3.1.

Alternatively, the task participants could use two boundedly rational (adaptive) strategies. The first is the “Experience-Weighted Attraction” (EWA) approach, which evaluates each possible action based on its observed consequences (outcomes). The behavior of the EWA model is governed by the principle that actions that previously led to successful outcomes—either successfully chosen actions or unchosen actions that would have been successful—will be chosen more often in the future. I describe this approach in detail in Section 3.2. The second boundedly rational approach consists of forecasting for each trial whether the next shot will fall within the winning range \([-4; 4]\) only by observing whether this event occurred in the previous trials, using an algorithm borrowed from the Statistical Calibration literature (Foster and Vohra (1998), Carvajal (2009)). This algorithm, which I refer to as the “Adaptive Forecaster model,” is purely adaptive in that it is completely ignorant of the real probability that a shot will occur within \([-4; 4]\) or the reasons why such an event occurs or fails to occur. Moreover, the model does not learn anything about the future from the shot history; it merely considers its own past inaccuracies and tries to compensate for them.

The Adaptive Forecaster model implements a “regret-minimization” algorithm (see Foster and Vohra (1998)). As such, it has solid neurobiological foundations (it has been shown that the brain implements this kind of algorithm; e.g., Camille et al. (2004)). Yet, the Adaptive Forecaster model does not fit subject behavior as well
as EWA does in the current experiment,\textsuperscript{13} so the results obtained with the Adaptive Forecaster model are not discussed further in this paper. However, for the sake of completeness and for replication purposes, I provide all the necessary details on the Adaptive Forecaster model in Appendix 7.1.

The fundamental difference between the Bayesian and boundedly rational learning approaches is that the Bayesian agent infers the hidden model of the world in each session of the task—i.e., it is “Model-Based” (MB)—whereas the boundedly rational agent does not; it is purely adaptive or “Model-Free” (MF). There is ample neuroscience evidence that the brain contains separate, competing systems for implementing both MB and MF learning (see Doll et al. (2012) for a review).

The behaviors of the two kinds of learners (Bayesian versus boundedly rational) are markedly different in the current task. Specifically, there are three behavioral signatures of Bayesian behavior in the task. First, The Bayesian agent skips in the first trials of each session. Second, he begins betting only when/if the evidence that the bowman is a master becomes sufficiently large. Third, the Bayesian agent does not systematically skip after observing a loss; it usually perseveres and bets despite losing $40 when the shot that led to the loss falls sufficiently near the target. In contrast, the boundedly rational agent appears to be obsessed by recent past outcomes, in the sense that a stream of good outcomes mechanically leads the agent to bet, because the agent focuses exclusively on the opportunity costs entailed by skipping, instead of focusing on the evidence that the bowman is a master. For the same reason (focusing on prior outcomes rather than on the evidence that the bowman is a master), and contrary to that with the Bayesian agent, the occurrence of a loss outcome leads the adaptive agent to skip systematically, irrespective of the value of the shot that led to the loss. I formally show the behavioral differences between the Bayesian and boundedly rational models in the following.

3.1 Bayesian model

Inferring the model of the world At each trial \( t \) of a given session, the Bayesian agent compares the plausibility (likelihood) of the two possible models of the world (Normal versus Cauchy) given the data available until (included) trial \( t X_t = (X_1, X_2, \cdots, X_t) \):

\textsuperscript{13}To be more specific, in the model comparison analysis (using the methods described in Section 4.2), the Adaptive Forecaster model was unambiguously trumped by the best-fit EWA model according to a paired t-test based on the differences in the individual fits (p-value: 0).
• Under the Normal Model \((M_1)\) (i.e., if the bowman is a master), the data \(X_t\) have density \(f_1(X_t | \sigma)\):

\[
f_1(X_t | \sigma) = \prod_{k=1}^{t} \frac{1}{\sqrt{(2\pi)\sigma}} \exp\left\{ -\frac{X_k^2}{2\sigma^2} \right\},
\]

where \(\sigma\) (the unknown standard deviation) is uniformly distributed between 0.1 and 2.

• Under the Cauchy Model \((M_2)\) (i.e., if the bowman is an apprentice), the data \(X_t\) have density \(f_2(X_t)\):

\[
f_2(X_t) = \prod_{k=1}^{t} \frac{1}{\pi(X_k^2 + 1)}.
\]

At each trial, the Bayesian model assesses how likely it is that the session is Normal versus Cauchy in light of the available data \(X_t\). The metric used is the “marginal or predictive density” (Berger and Pericchi, 2001) of \(X_t\) under each model \((M_1\) and \(M_2)\):

\[
m_1(X_t) = \int_{0.1}^{2} f_1(X_t | \sigma) \times \pi(\sigma)d\sigma,
\]

\[
m_2(X_t) = f_2(X_t),
\]

where \(\pi(\sigma)\) denotes the prior used for \(\sigma\). The uniform prior \((\pi(\sigma) = \frac{1}{2-0.1})\) may appear to be a natural choice, although the “reference” (noninformative) prior \(\pi(\sigma) = \frac{1}{\sigma}\) may actually be preferred in reference to prior statistics studies that commonly advocates its use for model selection (e.g., Berger and Pericchi (1996) and Betolino and Racugno (1997)). Both priors were tested in simulations, and the behavior of the model does not appear to be sensitive to the choice of the prior.

The evidence for the Normal model \((M_1)\) against the Cauchy model \((M_2)\) is provided by the Bayes Factor of \(M_1\) to \(M_2\):

\[
B_{12}(t) = \frac{m_1(X_t)}{m_2(X_t)}.
\]
For instance, if $B_{12} = 3$, the Normal hypothesis is favored over the Cauchy one at odds of 3 to 1. To compute the expected value of each action, the Bayesian model first assesses the posterior probability of each model given the data $X_t$, which turns out to be a simple transform of $B_{12}$ under the assumption that the prior model probabilities, denoted by $P(M_1)$ and $P(M_2)$, equal $1/2$:

$$P(M_1 | X_t) = \frac{P(M_1) m_1(X_t)}{P(M_1) m_1(X_t) + P(M_2) m_2(X_t)} = \frac{1}{1 + B_{12}^{-1}},$$  \hfill (6)

$$P(M_2 | X_t) = \frac{P(M_2) m_2(X_t)}{P(M_1) m_1(X_t) + P(M_2) m_2(X_t)} = \frac{1}{B_{12} + 1}. \hfill (7)$$

**Decision** From these posterior probabilities, the Bayesian model directly derives the expected value of Action Bet at trial $t$, which is denoted by $V_t$:

$$V_t = P(M_1 | X_{t-1}) \left[ p_g \times 2 - (1 - p_g) \times 40 \right] + P(M_2 | X_{t-1}) \left[ p_c \times 2 - (1 - p_c) \times 40 \right], \hfill (8)$$

where $p_g$ (resp. $p_c$) denotes the probability of a winning bet ($X_t \in [-4, 4]$) when the session is Normal (resp. Cauchy).\(^\text{15}\) $p_g$ can be assessed in two ways, one integrative and the other dynamic. Under the integrative approach, $p_g$ reflects the likelihood of a winning bet for each possible value of the standard deviation:

$$p_g = \frac{1}{2 - 0.1} \int_{0.1}^{2} 2 \left( 1 - \Phi \left( \frac{4}{\sigma} \right) \right) d\sigma = 0.993, \hfill (9)$$

where $\Phi$ denotes the c.d.f. of the standard Normal distribution. Under the dynamic approach, $p_g$ is computed for each trial $t$ as a function of the standard deviation estimate $\tilde{\sigma}(t) = \sqrt{\frac{1}{t} \sum_{k=1}^{t} X_k^2}$. Using this estimate, the Bayesian agent assesses the likelihood of a winning bet in a Normal world to be

\(^{14}\)This assumption reflects the fact that at the onset of each session, the subjects knew that they had an equal chance of being in a session with a master (Normal) or a session with an apprentice (Cauchy).

\(^{15}\)As an alternative to Equation (8), the agent could compute $V_t$ by putting full weight on the world that she currently thinks to be more likely, rather than weighting the two worlds based on their evidence. Such “inductive learning” model does not fit subject behavior as well as the Bayesian model does. See Appendix 7.2 for details.
\[ p_g \equiv p_g(t) = 2 \left( 1 - \Phi \left( \frac{4}{\sigma(t)} \right) \right). \] (10)

The integrative and dynamic versions of the model earned similar amounts in Monte Carlo simulations of the task (results available on request). The Bayesian solution is thus robust to slightly different specifications of the model. The dynamic version fitted the subjects’ behavior marginally better.

The probability of a winning bet \((X_t \in [-4, 4])\) when the world is Cauchy is as follows (using the definition of the Cauchy density):

\[ p_c = \frac{1}{\pi} \int_{-4}^{4} \frac{1}{x^2 + 1} \, dx = \frac{1}{\pi} \left[ \tan^{-1}(4) - \tan^{-1}(-4) \right] = 0.844. \] (11)

Because Action Skip always yields 0 regardless of the state of the world, the optimal decision simply consists of betting if the metric \(V\) (the expected value of Action Bet, as defined in Equation (8)) is positive and skipping otherwise.

Preferences The foregoing calculation of \(V\) assumes risk neutrality. The behavior of the Bayesian agent may alternatively reflect loss aversion and the distortion of the probabilities, as in Prospect Theory (Kahneman and Tversky, 1979). The expected value of Action Bet for the Prospect Theory Bayesian agent is

\[
V_t = P(M_1 \mid X_{t-1}) \left[ w(p_g) \times 2^{\alpha_1} - \lambda w(1 - p_g) \times 40^{\alpha_2} \right] + P(M_2 \mid X_{t-1}) \left[ w(p_c) \times 2^{\alpha_1} - \lambda w(1 - p_c) \times 40^{\alpha_2} \right],
\] (12)  (13)

where \(\alpha_1\) and \(\alpha_2\) govern the shape of the value function, \(\lambda\) is the loss aversion parameter, and \(w(\cdot)\) is the probability weighting function introduced by Prelec (1998): \(w(p) = \exp \{ -(-\ln p)^{\alpha_3} \}\). \(\alpha_1, \alpha_2, \alpha_3\) and \(\lambda\) are free parameters in the model comparison analysis.

It may also be that the subjects felt disappointed after a loss and accounted for their disappointment ex ante in their valuation of Action Bet. In Disappointment Aversion, the possible outcomes of a prospect are evaluated relative to the
Expected Utility (EU) certainty equivalent (Bell (1985), Loomes and Sugden (1986), Gul (1991)). The EU certainty equivalent of Action Bet at trial \( t \) is the value \( C_t \), satisfying

\[
U(C_t) = p_t U(2) + (1 - p_t) U(-40),
\]

where \( U \) denotes the utility function, and \( p_t \) denotes the probability of a winning bet as estimated by the subject at trial \( t \):

\[
p_t = p_g P(M_1 \mid X_{t-1}) + p_c P(M_2 \mid X_{t-1}).
\]

For simplicity, risk neutrality is assumed (i.e., \( U \) is the identity function), so \( C_t = p_t \times 2 - (1 - p_t) \times 40 \). Taking \( C_t \) as the reference point, the reference-dependent value of Action Bet is

\[
V_t = p_t \left\{ 2 + \nu(2 - C_t) \right\} + (1 - p_t) \left\{ -40 + \nu(-40 - C_t) \right\},
\]

where \( \nu(.) \) represents disappointment-elation value relative to the reference \( C_t \) (Bell (1985) and Loomes and Sugden (1986)). Following Koszegi and Rabin (2006), a piecewise-linear disappointment-elation function \( \nu(.) \) is assumed:

\[
\nu(x) = \begin{cases} 
  x & \text{if } x \geq 0 \\
  \lambda_1 x & \text{if } x < 0
\end{cases}
\]

where \( \lambda_1 \) is a free parameter. \( \lambda_1 > 1 \) indicates disappointment aversion. Under this specification, the value of Action Bet can be rewritten as

\[
V_t = p_t \left\{ 2 + (2 - C_t) \right\} - (1 - p_t) \left\{ 40 + \lambda_1 (40 + C_t) \right\}.
\]

The behavior of loss-averse and disappointment-averse agents is not markedly different from the benchmark (risk neutrality). The three agents share three primary
behavioral signatures. First, all three Bayesian agents skip in the first trials of each session. Second, the agents begin betting only when/if the evidence for the Normal hypothesis \( B_{12} \) becomes sufficiently large, which happens in Normal sessions as well as in Cauchy sessions in which the true distribution is camouflaged by a series of seemingly Normal realizations. In such instances, the model is deluded into behaving as if the session is Normal (Figure 1). Third, the Bayesian does not systematically skip after observing a loss; it usually perseveres and bets despite losing $40 when the shot that led to the loss falls sufficiently near the target (see Figure 2). No boundedly rational model can simultaneously generate all three signatures of Bayesian learning in the current task for reasons that will become apparent in Section 3.2.

The only difference between the disappointment/loss-averse model and the risk-neutral counterpart is that under Prospect Theory or Disappointment Aversion, the agent is more reluctant to take risks. This reluctance is revealed by the threshold value of \( B_{12} \) at which the model decides to bet, which is higher than that in the benchmark case. For instance, Figure 3 indicates that a simulated subject who is quite disappointment averse (\( \lambda_1 \) equals 2 in that simulation) does not bet until \( B_{12} \) reaches 7.4 (i.e., until the Normal hypothesis is favored over the Cauchy one at odds of 7.4 to 1). In contrast, under risk neutrality, odds of 2.7 to 1 in favor of the Normal hypothesis constitute sufficient evidence for the agent to begin betting.

### 3.2 Boundedly rational model: EWA Model

The value or “attraction” of each action \( i \in \{B; S\} \) (where \( B \) and \( S \) stand for “Bet” and “Skip,” respectively), denoted by \( A_i \), depends on the past payoff experienced with the two actions. Specifically, let \( \pi_i(t) \) denote the payoff of Action \( i \) at trial \( t \) of a given session:

\[
\begin{align*}
\pi_B(t) &= \begin{cases} 
2 & \text{if } X_t \in [-4, 4] \\
-40 & \text{otherwise}
\end{cases}, \\
\pi_S(t) &= 0.
\end{align*}
\]

Within each session, the attractions are updated at each trial \( t \) as follows:

\[
A_i(t) = \frac{\phi N(t - 1) A_i(t - 1) + [\delta + (1 - \delta) I(i(t))] \pi_i(t)}{N(t)},
\]

18
Figure 1: **Two simulated runs of a Cauchy session with the Bayesian model.** Each figure reports a simulated session with an apprentice Bowman (“Cauchy session”). In the simulation reported in the top figure, the simulated shots till trial 16 fell all within the winning range ($X_{16} = [1.3; -0.6; 0.08; -1.7; 0.3; -0.9; 0.5; 0.6; 0.9; -0.1; -3.7; -1.1; 3.5; -1.9; 1.6; 0.1]$). At trial 17 an outlier occurred ($X_{17} = 82.3$). The three last trials were back into the winning range. In the simulation reported in the bottom figure, all the simulated shots fell up to three meters away from the target until the last trial, where an outlier occurred ($X_{20} = 19.2$). In both simulations, the risk-neutral version of the Bayesian model was used. 

**Legend:** Each figure shows $B_{12}$, the evidence for the Normal hypothesis against the Cauchy one (top) as well as corresponding choice made by the model (bottom) and ensuing accumulated outcomes (middle). On the bottom graph, both the actual and foregone outcomes are indicated. For instance, in the top figure, the model skipped at trial 9 and would have won $2 if he had bet; at trial 17, the model skipped and would have lost $40 if he had bet. On the top graph, the dotted line represents the threshold value of $B_{12}$ from which the model bets.

In the simulation reported in the top figure, the evidence that the master was an apprentice was low in the first part of the session; consequently, the agent bet at trials 10 and 11. The agent did not fully realize that the Bowman was an apprentice until the outlier occurred at trial 17, at which point the evidence that the bowman was an apprentice became overwhelming ($B_{12}$ extremely negative). In the simulation reported in the bottom figure, the data deluded the model that the bowman was a master. Consequently, the model bet from trial 8 on and lost $40 at trial 20.
Figure 2: Simulated run of a volatile Normal session with the Bayesian model. In this simulation of a session with a master bowman (“Normal session”), the standard deviation of the shots was 1.5. All of the simulated shots fell within the winning range \([-4; 4]\) except for the shot in trial 16, which fell at approximately five meters away from the target \(X_{16}\) neared 5). The risk-neutral version of the Bayesian model was used in this simulation.

Legend: The figure shows \(B_{12}\), the evidence that the bowman is a master (top), as well as the corresponding choice made by the model (bottom) and the ensuing accumulated outcomes (middle). In the bottom graph, both the actual and the foregone outcomes are indicated. For instance, at trial 5, the model skipped and would have won $2 if the agent had bet. In the top graph, the dotted line represents the threshold value of \(B_{12}\) from which the model bets.

The model learned that the session was Normal and hence consistently bet from trial 8 on. Notably, the model persevered with Action Bet despite losing at trial 16 because the data from trial 16 on continued to favor the hypothesis that the bowman was a master.
Figure 3: Simulated run of a Normal session with a risk-neutral Bayesian agent (top) and with a disappointment-averse Bayesian agent (bottom). In this simulation of a session with a master bowman (“Normal session”), the standard deviation of the shots was 1. All of the simulated shots fell well within the winning range $[-4; 4]$. For the disappointment-averse agent, $\lambda_1$ (as defined in the main text) was set to 2.

Legend: Each figure shows $B_{12}$, the evidence for the Normal hypothesis (top), as well as the corresponding choice made by the agent (bottom) and the ensuing accumulated outcomes (middle). On the bottom graph, both the actual and the foregone outcomes are indicated. For instance, at trial 1, both agents skipped and would have won $2 if they had bet. On the top graph, the dotted line represents the threshold value of $B_{12}$ from which the model bets. Note the threshold value is higher for the disappointment-averse agent than for the risk-neutral one.
where \( I(i(t)) = \begin{cases} 
1 & \text{if } i \text{ is chosen at trial } t \\
0 & \text{if } i \text{ is not chosen at trial } t 
\end{cases} \).

The parameters \( \phi \), \( \kappa \), and \( \delta \), respectively determine how quickly previous experience is discarded, the growth rate of attractions, and the relative weight given to the actual payoff compared to the foregone one (more on this below). \( N(t) \) measures the amount of experience that the subject has accumulated after trial \( t \) has taken place. \( N(t) \) is updated as follows:

\[
N(t) = (1 - \kappa) \phi N(t - 1) + 1.
\] (22)

Under this approach, the behavior consists of choosing Action Bet at trial \( t + 1 \) if and only if \( A_B(t) > A_S(t) \) and choosing Action Skip otherwise. The variables \( N(t) \) and \( A(t) = (A_B(t), A_S(t)) \) take the initial values \( N(0) \) and \( A(0) \). A plausible benchmark for the participants’ initial attractions, \( A(0) \), is the prior expected value of each action, which could be inferred from the task instructions. The prior value of Action \( S \), \( A_S(0) \), is zero because Action \( S \) yields $0 irrespective of the state of the world. Regarding \( A_B(0) \), the agent knows that there are equal chances of being in a Cauchy session versus a Normal one and that the chance of a winning bet is, on average, approximately 99% in a Normal session and 84% in a Cauchy session. The prior expected value of Action \( B \) is therefore

\[
\frac{1}{2} \left[ 0.84 \times 2 + (1 - 0.84) \times (-40) \right] + \frac{1}{2} \left[ 0.99 \times 2 + (1 - 0.99) \times (-40) \right] = -1.57.
\]

In the main part of the analysis reported in Section 4.2, \( A_B(0) \) was a free parameter that was optimized to fit the behavior of each subject individually. The parameter could take values in the range of \(-750\%\) to \(+750\%\) from the benchmark value \(-1.57\), whereby it was allowed to depart significantly from the benchmark.

In learning the attractions, \( N(0) \) measures the learning rate (the weight placed on prior beliefs compared to the payoffs). When \( N(0) \) is large, the agent sticks to his initial beliefs—a case of belief inertia or sticky priors. Cases of belief inertia have been documented in experiments where subjects believed that they had some prior experience with the task (Biyalogorsky et al., 2006). When \( N(0) \) is null, only actual experience matters; therefore, the agent responds strongly to payoffs and learns extremely quickly. \( N(0) = 1 \) means that equal weight is placed on initial attraction
and payoffs. By design, the subjects had weak priors in the current task (i.e., \( N(0) \leq 1 \)) because the task was not reminiscent of any familiar context outside of the lab\(^{16}\) and because the subjects knew that the sessions were independent, so their beliefs about the action value at the beginning of a session could not be influenced by the experience gained in the previous sessions.

The parameter \( \phi \) measures how quickly previous experience is discarded. In the present task, \( \phi \) is best interpreted as an index of forgetting; subjects may not recall the previous experience perfectly.\(^{17}\)

The parameter \( \kappa \) determines the growth rate of attractions, which, in turn, affects the extent to which the subject will commit to one choice over time. When \( \kappa = 1 \), attractions can grow,\(^{18}\) and the differences in action attractions can become very large. In contrast, when \( \kappa = 0 \), attractions cannot grow outside of the payoff bounds, so learning does not stop.

The parameter \( \delta \) measures the relative weight given to the foregone payoff compared to the actual one in updating attractions. \( \delta \) therefore reflects the importance of counterfactual learning in the task. Reinforcement learning entirely ignores foregone payoffs (\( \delta = 0 \)). This assumption is implausible when the forgone payoffs are known perfectly, as they are in the current task. Clairvoyant (“belief-based”) learning weights the foregone payoff as strongly as the actual payoff (\( \delta = 1 \)). \( \delta \) can be anywhere between 0 and 1, whereby EWA hybridizes the features of reinforcement and belief-based (counterfactual) types of learning.

When \( \kappa = 0 \), \( \delta = 1 \), and \( \mathbf{A}(0) \) equals the prior expected values of the two actions, the EWA attractions are the same as the expected payoffs according to “weighted fictitious play,” which, in many contexts, is indistinguishable from Bayesian learning (Camerer and Ho, 1999). Notably, this is not the case in the current task: in Monte Carlo simulations the economic performance of the weighted fictitious play agent is markedly below the one of the Bayesian (the average earnings from weighted fictitious

\(^{16}\)Not framing the task as a financial investment problem—or any other decision problem the subject might feel familiar with—ensured that was the case.

\(^{17}\)In changing environments, \( \phi \) can also be interpreted as an index for discounting the old experience related to the perceived change (Ho et al., 2007). In the current task, however, the subject knows that the data-generating process is unchanged within a session, so the discounting of past attractions should not be deliberate.

\(^{18}\)To see this, note that when \( \kappa = 1 \), then \( N(t) = 1 \), and the denominator in the attraction updating equation (Equation (21)) disappears.
play are inferior, earnings variance is higher)\textsuperscript{19}

To intuitively show the effect of each parameter and how the parameters can combine to generate the observed behavior, Figure 4 reports the behaviors of three simulated EWA models in one Normal session in which all of the shots fell within \([-4;4]\), except for the shot at trial 17. The value of the attractions as well as the corresponding choice and the ensuing outcome are indicated for each trial of the session. In the simulation reported in the top graph, the specification of the EWA model was as follows: \(\phi = \delta = 1, \kappa = 0, A_B(0) \approx -6\), and \(N(0) = 1\). In the middle graph, the specification of the EWA model was the same as in the top graph except for \(N(0)\), which was down to \(1/2\). As shown, the effect of reducing \(N(0)\) by half is to speed up learning (the curve for \(A_B\) is steeper in the middle graph than in the top one). As a result, the model began betting from the third trial, whereas it did not bet until the fifth trial when \(N(0) = 1\). In the bottom graph, the specification of EWA was as in the top graph except for \(\delta\), which was reduced by half. The effect of decreasing \(\delta\) is to reduce the extent of counterfactual learning, so it took longer for the model to build the attraction value \(A_B\) in the early stage of the session. As a result, the model did not bet until the ninth trial. A similar method used to successively vary the value of \(\kappa\), \(A_B(0)\), and \(\phi\), reveals the following:

i) EWA behavior does not appear to be sensitive to the value of \(\kappa\) in the present task, so long as \(\phi\) and \(A_B(0)\) are not too low;\textsuperscript{20}

ii) decreasing \(A_B(0)\) is effective in making EWA bet later in a given session;

iii) decreasing the value of \(\phi\) (i.e., increasing forgetting) is effective in making EWA bet after/despite experiencing a loss as long as \(A_B(0)\) is close to the aforementioned benchmark value of -1.57. However, if \(A_B(0)\) markedly departs from the benchmark to take on a very negative value, there is no combination of the parameters that can make the model bet after experiencing a loss.\textsuperscript{21}

The foregoing thus shows that there is no combination of the parameters that can make the EWA agent behave like the Bayesian model. In particular, in contrast to the EWA agent, the Bayesian model both skips early in each session and perseveres in betting after experiencing a loss if the shot that led to that loss was a near miss, as described in Section 3.1.

\textsuperscript{19}The results obtained in those simulations are available on request.

\textsuperscript{20}When \(\kappa = 0\), and \(\phi\) and \(A_B(0)\) are very small, the EWA agent commits to skipping in all sessions.

\textsuperscript{21}The graph for each case (i-iii) is not shown here for space reasons but is available on request.
Figure 4: **Simulated run of a Normal session with three types of EWA agents.** In this simulation, the standard deviation of the simulated shots was 1.5. All of the simulated shots fell within $[-4; 4]$, except for the shot at trial 17. In the simulation reported in the top figure, the specification of the EWA model was as follows: $\phi = \delta = 1$, $\kappa = 0$, $A_B(0) \approx -6$, and $N(0) = 1$. In the middle figure, the specification of the EWA was the same except that $N(0) = 0.5$ (i.e., the agent learned quickly). In the bottom figure, the specification of the EWA was as in the top graph except that $\delta = 0.5$ (i.e., the agent hybridized features of counterfactual and reinforcement learning).

**Legend:** Each figure shows the attraction value of Action Bet ($A_B$) and Skip ($A_S$) (top) as well as the corresponding choice made by the agent (bottom) and the ensuing accumulated outcomes (middle).
4 Results

4.1 Experimental procedure

The participants were 184 undergraduate students (64% males) from the University of New South Wales. The participants were directed to watch the set of online instructions for the task before the experimental session. These instructions informed the participants that they would be performing a demanding decision task, described the task, and emphasized that all of the accumulated outcomes from the task would be added to a starting account balance to determine each participant’s payoff up to an amount of $110. The participants were also told that the account balance could be anywhere between $-500 and $500 and that it would be revealed to the participants after the experimental session. So by design, on each trial the subjects did not know the current value of their wealth (their net accumulated outcomes + the amount of the account balance).

I set the amount of the account balance before the subjects started to perform the task. I wrote the amount on a sheet of paper and put the sheet in an enveloppe in the middle of the lab room, all this in front of the subjects.\footnote{This procedure was important to ensure that the experimenter was not suspected of “cheating” (changing the amount after seeing subjects’ performance).} The amount was set to ensure that the average payoff in the session would not exceed 70$, the maximum cap achievable given budgetary constraints. To forecast payoffs I used the payoffs from the previous sessions or, for the very first sessions, simulated earnings (from simulating the Bayesian model in Monte Carlo simulations of the task).

Payoffs ranged from $5 (the “show-up reward” given independently of subject performance as per the lab protocol) to $110 (the cap set by the payoff rule), with a mean of $63 (standard deviation was 42). Fifty-nine subjects got more than $100, and forty subjects ended up with the minimum payoff (i.e., $5).

The motivation for using this particular payoff rule was twofold. First, as explained earlier, given the high complexity of the task it was key to provide task participants with significant monetary incentives (see Section 2, “Incentives”). Second, by making their current wealth level unknown to the subjects during the task, the account balance feature was meant to prevent potential wealth effects from occurring during the experiment. One worry in design was indeed that subject decisions throughout the task may reflect wealth effects. For instance, subjects may be suscep-
tible to increased risk taking after prior good outcomes (Thaler and Johnson, 1990). Another concern was the possibility that after prior losses, when subjects would find themselves “in the red,” they might decide to bet unconditionally in the hope of recouping prior losses.\footnote{The latter tendency has been observed in pathological gamblers (e.g., \url{http://www.wrc.noaa.gov/wrso/security_guide/gamble.htm}). Its prevalence in the normal population has yet to be determined.} Making their current wealth level unknown to the subjects till after the task was completed should limit the extent of such wealth effects. The results reported in Section 4.3.3 validate this design choice (subject behavior does not appear to be driven by wealth effects).

Upon arrival in the lab, the participants again watched the online instructions for 30 minutes, after which they completed a multiple-choice questionnaire that confirmed their understanding of the task. After 15 minutes, the participants reviewed the answers with the experimenter. The participants were also briefed again on the stochastic structure of the task and the payment procedure. Subsequently, the participants completed a run of the task comprising 15 sessions of 20 trials each and lasting about 35 minutes.

Three experimental treatments were conducted. In the first treatment (N=70), in each trial, the subjects made a decision to either bet or skip, observed the shot realized by the bowman at the trial, and then were told their payoff. Figure 14 in Appendix 7.4 shows the timeline of a trial.

The second treatment (N=54) was similar to the first except that the participants were not told their payoff at the end of each trial. That is, whereas in the first treatment, the subjects were explicitly told the outcome of their decision on each trial (e.g., “you have earned $2”), in the second they were not; the outcome earned on each trial had to be inferred from the action taken and the realized shot (e.g., “I bet, and the shot is within the winning range, so I have earned $2”). Thus, the outcome dimension was more salient in the first treatment than in the second.

Note that the two treatments are strictly equivalent for the rational agent inasmuch as one’s payoff is directly inferred by remembering one’s action and observing the bowman’s shot. However, I conjectured that they may not be equivalent for human subjects for two reasons. The first is that the outcome earned on each trial is much more salient in the first treatment than in the second one, and the evidence suggests that rational learning is hampered by emotional reactions to outcomes (e.g.,
Thus, rational (Bayesian) learning may prevail to a greater extent in the second treatment than in the first. The second reason relates to the “compatibility effect” originally reported in Lichtenstein and Slovic (1971). According to the compatibility effect, the information format—specifically, whether the emphasis is primarily on the outcome dimension or the probability dimension of the decision problem—greatly influences subjects’ focus and, ultimately, their decision-making. This theory predicts that task participants are more likely to implement EWA, which learns about outcomes—as opposed to learning about outcome probabilities, such as in Bayesian learning—when they are induced to focus on the outcome dimension, i.e., in the first treatment.

Finally, I ran a third treatment (N=60) that was similar to the first except that at each trial in each session, the participants were given the option to skip in all of the remaining trials or bet in all of the remaining trials of the session. This decision was irreversible. The instructions emphasized to the participants that exercising the option would not accelerate the pace of the task in any way because the sequence of events displayed on the screen would be exactly the same regardless of whether the option was exercised. The experimenter further informed the participants that the decision to bet or skip in all remaining trials was merely an option (i.e., if the participant considered it useless, he or she should disregard it altogether).

Note that the rational agent would never use the option because it is strictly dominated by the default course of action, which allows the agent to change his mind from one trial to another. However, human subjects may use the option for two different reasons. The first is laziness; choosing once for the rest of the session is easier than having to deliberate for each trial. The second reason relates to self-control and the desire to use commitment devices; when subjects feel tempted to bet even though betting exposes them to mindless risks, they may want to use the option to skip for the rest of the session to ensure that they do not succumb to temptation and to implement the right decision. Fortunately, the current design allowed me to disentangle the two explanations, as will become apparent in Section 4.3.2.

All the material used during the sessions (task instructions and questionnaire) as well as short movies showing the different experimental treatments, are available at

---

24 This comment was to avoid “experimental demand” effects whereby the subjects would artificially use the option only because it was proposed as a choice.
4.2 Structural estimation

The estimation procedure is described in detail in Appendix 7.5. The procedure proceeded in two steps, a first “in-sample” calibration step followed by an “out-of-sample” validation test. In the first step, the free parameters of each model were chosen separately for each subject to maximize the proximity of the model predictions to the actual choices. Specifically, the parameters were optimized by minimizing the total squared prediction error compounded over the set of trials for the first eight sessions. The ensuing parameter estimates were then used in a second stage of the analysis to predict the choices in the last seven sessions (out-of-sample validation). The predictive accuracy of each model was measured by the percentage of trials in which the model successfully predicted the subject’s choice in the last seven sessions (henceforth, the “success rate”). For each subject, the out-of-sample success rate of the best-fit boundedly rational model was compared to the one of the best-fit Bayesian model—loss-averse or disappointment averse—and (for reference) to the one of the risk-neutral version of the Bayesian model (for which there was no free parameter, i.e., no in-sample calibration was needed).

This type of two-step procedure is commonly used in model comparison analyses involving very general models such as the EWA model, to guard against overfitting. See, e.g., Camerer and Ho (1999), Ho et al. (2007), and Ho et al. (2008). The basic idea is that in comparing the fits of the EWA model versus competitors, modifying the standard goodness-of-fit criteria to account for differing degrees of freedom across models (cf. BIC, AIC, etc.) is not enough. Out-of-sample validation is absolutely required.

4.3 Empirical results

4.3.1 Main findings

Prevalence of Bayesian learning The majority of the subjects acted as if they were Bayesians. This finding was immediately apparent both from i) the way that the subjects systematically skipped in the early stage of a session, such as the Bayesian agent (see, e.g., Figure 1, p.19), and ii) the way that they consistently bet when there
was sufficient evidence that a session was Normal, even if some realizations fell outside of the winning range (this behavioral pattern is a signature of Bayesianism in the task too; see Figure 2, p.20). Specifically, 93% of the subjects skipped, at minimum, the three first trials in every session of the task. The mode of the distribution of the first trial when the subjects bet (in those sessions in which they did eventually bet) is 6; see Figure 5. At the same time, 76% of the subjects bet after encountering an ambiguous realization (defined as a hit falling between four and seven meters away from the target) in the Normal sessions. As explained in Section 3, skipping systematically early in each session combined with betting despite losing in the aftermath of an ambiguous realization is a tag of Bayesian behavior (it is incompatible with all kinds of adaptive behaviors).

![Figure 5: Distribution of the first time the subjects bet in the task. The distribution was derived across all subjects and sessions. N: cases in which the subjects skipped in all trials of a session.](image)

The conclusions drawn from the formal model comparison fully support the idea that subjects acted like Bayesians in the task. The best-fit Bayesian model fits better than the best-fit EWA model for 73% of the subjects in the first treatment and for 87% of the subjects in the second treatment.

The fraction of Bayesian subjects is significantly higher in the second treatment than in the first one according to a two-sample Z test (p value: 0.002). As explained earlier, this finding is consistent both with the notion that people are subject to the “compatibility effect” and with recent evidence that affective reactions to payoffs trump rational reasoning (e.g., Kuhnen and Knutson (2011)). As such, it raises the
intriguing possibility that the misapprehension of tail risk in some people might at least partly be attributed to the fact that focusing on the realized payoff diverts the agent from learning about the hidden probability of this payoff.

In the following analysis, the data from the two treatments are pooled together, but all results hold in each subsample. Over all of the subjects and trials, the average success rate is approximately 89% for the Bayesian model and slightly below 75% for the EWA model. A paired t-test leads to the rejection of the null hypothesis that the fits are equal with a p-value lower than $10^{-5}$. Figure 6 indicates good fits for the Bayesian model, with prediction success in the 85% \( \sim \) 90% range for the large majority of the subjects (the median is 90%). The fits of the best-fit EWA model are not quite as good.

![Figure 6](image-url)

**Figure 6: Comparative goodness-of-fit of the Bayesian and EWA models.** X-axis: The success rate (fraction of trials in which the model predicted the actual choice) of the best-fit Bayesian model; Y-axis: The success rate of the best-fit EWA model. Each data point corresponds to one subject (N=124). The Bayesian model fits better when the data point is below the 45 degree line. The points in the bottom right quadrant correspond to subjects whose behavior closely approximated the Bayesian model and clearly departed from EWA.

The best-fit EWA model turns out to be the specification used in Pouget (2007) ($A_B(0)$ equals the expected value of Action Bet, $-1.57$, $\kappa=1$, $N(0) = 1$, $\phi=1$, $\delta$ free parameter), which marginally outperforms the version of EWA in which all the
parameters are free. For the Bayesian model, the best fit is achieved with the Prospect Theory model, which is marginally better than the disappointment-averse version. As in Trommershauser et al. (2008), there is no evidence of probability distortion in the data. The average fitted values of $\alpha_1$, $\alpha_2$, and $\lambda$ are 0.4, 0.7, and 2, respectively, which is broadly consistent with the values derived in prior empirical studies (Neilson and Stowe, 2002). For the disappointment-averse model, $\lambda_1$ is above 1 (near 1.5), in accordance with the notion that people are disappointment averse.

Notably, despite the fact that assuming risk neutrality is an oversimplification of actual behavior, the risk-neutral Bayesian model, which has no free parameters, also fits the behavior significantly better than the best-fit EWA model for the majority of the subjects (p value: 0.007).

Therefore, subject behavior reflects Bayesian learning. Figure 7 shows the choice made by one of the subjects at each trial in each session of the task, along with the choice predicted by the best-fit Bayesian model. Figure 8 shows the Action Bet values that are estimated by the Bayesian model at each trial, along with the value of the shot observed at the trial. In most trials, the subject chose to skip unless the evidence that the Bowman was a master was sufficiently large that Action Bet value was positive. It appears that the Bayesian model fits the subject’s choices well, albeit not perfectly: the model turns out to be more loss averse than the subject in some trials (see, e.g., trials 6–9 in session 3). The reverse is true for other subjects who appeared to be more loss averse than the model; see Appendix 7.6 for an example.

The evidence that subject behavior is well described by the Bayesian model can be further strengthened by focusing on the subjects’ net accumulated outcomes at the end of the task (henceforth, “subject earnings”). Subject earnings averaged 81$ (standard deviation: 94) which is below the 112$ that the best-fit Bayesian model earned on average when simulated in the exact same instances of the task as the ones the subjects faced, and above the 51$ that the best-fit EWA model earned on average. However, when excluding from the analysis the group of irrational subjects who “picked pennies” as defined in the next paragraph, we find that subject earnings closely resemble those of the Bayesian model: they average 120$ (standard deviation: 58) versus 135$ for the Bayesian model (standard deviation: 68). One cannot reject the null hypothesis that the two means are equal (p value of t test: 0.08). Furthermore,

---

25That is, the Prospect Theory Bayesian model with $\alpha_3$ fixed to 1 fits the data as well as the Prospect Theory Bayesian model in which $\alpha_3$ is free (see Section 3.1 for the definition of $\alpha_3$).
Figure 7: **Subject behavior versus Bayesian behavior in one instance of the task.** Each graph shows the choice made by one of the subjects at each trial in one of the sessions of the task (sessions 1–15, from the top and the left). Legend for each graph: X axis: trial number. Y axis: ‘0’ Action Skip was chosen; ‘1’ Action Bet was chosen. Both the choice made by the subject (cross) and the choice made by the fitted Bayesian model (circle) are displayed.
Figure 8: Value of Action Bet according to the Bayesian model in the same instance of the task as the one shown in Figure 7. Each graph shows the estimated value of Action Bet according to the fitted Bayesian model at each trial in one of the sessions of the task (sessions 1–15, from the top and the left). Both the estimated value of Action Bet (blue line) and the value of the shot realized at the trial (blue dot) are displayed. Legend for each graph: X axis: trial number. Y axis, left: shot value. Y axis, right: Action Bet value.
the distribution of the earnings across subjects and the corresponding distribution of the simulated Bayesian earnings look alike (see Figure 9). Formally, one cannot reject the hypothesis that the two distributions are alike in a Kolmogorov-Smirnov test at a threshold of 5%. In contrast, similar Kolmogorov-Smirnov test unambiguously leads to reject the hypothesis that the earnings of the EWA model are distributed like subject earnings (p value is close to 0).

Figure 9: **Subject earnings versus simulated Bayesian earnings.** The figure reports the distribution of subject earnings versus the distribution of simulated earnings of the best-fit Bayesian model. Earnings are defined as the net accumulated outcomes at the end of the task. For each subject, the Bayesian model was simulated in exactly the same instance of the task as the one the subject faced, and the earnings were recorded. X-axis: earnings ranges. Y-axis: for each range, the left (resp. right) bar shows the number of observations in the range for the Bayesian model (resp. the subjects).

**Picking Pennies bias** Fifty-one subjects (i.e., 41% of the pool) picked pennies in the sense that they bet after encountering an outlier. Here, an outlier is defined as a realized shot located more than seven meters away from the target ($|X| > 7$).
The threshold value of 7 was chosen so that the probability that a shot falls beyond the threshold is zero with any kind of master bowman, even one whose shots have a standard deviation as high as 2 (the maximum standard deviation level that could be encountered in the task). It is important to note that the results still held when I increased the value of the threshold to allow for the possibility that the subjects misunderstood the instructions and took the maximum standard deviation level to be above 2. See “Robustness Checks”, p.44.

The deviations from Bayesian learning that I identified in the group of “penny-pickers” are restricted to the behavior following the loss outcome in the Cauchy sessions. Specifically, the success rate of the Bayesian model during Cauchy sessions is inferior to 60% for eight penny-pickers and inferior to 65% for 13 penny-pickers. In contrast, in the Normal sessions, the success rate is never below 60%, and it is below 65% for five penny-pickers only. (For reference, in the normal subjects, the success rate is never below 60% regardless of the session type, and it is below 65% for two subjects only.)

The behavior of this group can be described by a typology with three types. Figure 10 pictures the three types in one exemplar session of the task featuring an outlier at the 14th trial of the session. The first type of penny-picker bets immediately following the occurrence of the outlier and skips after a few trials. The second type skips following the occurrence of the outlier and returns to betting after a few trials. The third type bets on all or almost all trials following the occurrence of the outlier. A total of 22 penny-pickers are of the first type, 10 are of the second, and 19 are of the third.

Not surprisingly, the economic performance of this group is extremely poor by a mean-variance standard: their net accumulated outcomes at the end of the task averaged 29, with a standard deviation of 108, which indicates that the behavior of this group was not only detrimental to the average performance but also highly risky. For reference, the Bayesian model simulated in the same conditions (i.e., for each subject, the Bayesian model was simulated in a run of the task that exactly replicated the shot realizations seen by the subject) earned, on average, 143 (standard deviation: 47). Even the EWA model that performed worst (according to my simulations) performed

\[ \Phi_{\sigma=2}(X) = 1 \iff X > 6.72 \]

One can check that, where \( \Phi_{\sigma=2} \) denotes the c.d.f. of a normal variable with mean 0 and standard deviation 2. In plain English, for a master bowman with a standard deviation of 2, 6.72 is the minimum value of X for which the probability of seeing a realization above that value is zero; in the analysis I have rounded this value up to 7.
Figure 10: **Three types of penny-pickers in the task.** The figure pictures the three types of penny-pickers encountered in the experiment in an exemplar session. An outlier occurred in trial 14 of the session, and the behavior is described from that trial on. The first type bets immediately after seeing the outlier and skips after a few trials. The second type skips immediately after seeing the outlier and returns to betting after a few trials. The behavior of the third type combines the two first types: the agent bets on all or almost all of the trials following the occurrence of the outlier.
significantly better on both the mean and variance dimensions (average earnings: 78; standard deviation: 77).

4.3.2 Follow-up analysis

Modeling of the Picking Pennies behavior  An immediate and natural explanation of the penny-pickers' behavior would be that they were risk lovers in the experiment. But the analysis reveals that the risk appetite of the penny-pickers was not different from the one of the other subjects. In particular, like the other subjects, the penny-pickers systematically skipped in the first trials of each session (see Figure 11, top graph), which is contrary to what risk-loving agents would do.

Can the behavior of the penny-pickers be due simply to their having incorrect beliefs about the stochastic structure of the task and their having poor memory? The analysis suggests that this is not the case. I find that the behavior of the penny-pickers in the Cauchy sessions cannot be accounted for by any Bayesian model, even one with incorrect priors. And it cannot be accounted for by any EWA model, even one with an extreme degree of forgetting.

I also conjectured that the penny picking bias could emerge as a consequence of subject fatigue towards the end of the task. But it appears that the probability of occurrence of the bias is not higher in the second half of the task compared to the first half.

A number of more sophisticated explanations are ruled out as well. First, the penny-pickers' behavior cannot be explained by the “local thinking” model proposed by Gennaioli and Shleifer (2010). Local thinkers initially neglect the probability of the loss outcome and overreact after seeing an outlier. Hence they bet (resp. skip) to a greater extent than the rational agent does before (resp. after) encountering an outlier, which is contrary to what the penny-pickers did.

Second, I conjectured that “choice stickiness” (the tendency to persevere in a given course of action irrespective of the action payoff) could explain that penny-pickers bet after seeing an outlier during the Cauchy sessions. To test this conjecture I fitted to the penny-pickers' data the “Inertial” model proposed by Feltovich (2000). The Inertial model predicts that the subject will behave in the current trial exactly as in the previous trial with probability $p_{\text{same}}$ irrespective of the payoff in either trial. This class of model is denoted by $IN(p_{\text{same}})$. The members I examined are $IN(0.8)$, $IN(0.7)$, and $IN(0.6)$. I also tested $IN(0.5)$, which corresponds to completely random
play (more on this below). All these models fit the data poorly.

Furthermore, the picking pennies behavior does not merely reflect behavioral mistakes caused by subject “trembling”. The behavior of the penny pickers appears to be non-stochastic. This is already apparent from simple descriptive statistics: as noted above the penny pickers systematically skipped early in each session and bet consistently later in the Gaussian sessions. Also, the picking pennies anomaly regularly occurred during the task. Specifically, 75% of the penny pickers picked pennies in at least two sessions during the task, 41% in three sessions or more, and 30% in four sessions or more (average: 2.6 sessions; mode: 2; maximum: 6). So the picking pennies behavior was not an isolated event. All this suggests that it is not trembling that caused the picking pennies bias, which is confirmed by more formal analysis. The random play model (IN(0.5) above—the model in which probability to skip and probability to bet equal 1/2 on each trial) and a stochastic Bayesian model that uses a “softmax” (logit) decision rule with very high “inverse temperature” (the coefficient that governs the degree of choice randomness) both fail to fit the penny pickers’ behavior. That is, the benchmark Bayesian model unambiguously provides a better fit than either model in the model comparison analysis.

So randomizing behavior in the form of trembling can be ruled out as an explanation of the picking pennies bias. But other forms of randomizing behavior could underlie it. The most plausible one is Probability Matching (e.g., Vulcan (2000) and Shanks et al. (2002)), a widespread tendency across species that has been proposed to be a product of evolution (Brennan and Lo, 2011). According to the Probability Matching model, the penny-pickers bet 80% of the time when their estimated probability to win by betting is 0.8. Although that model rightly predicts that the agent sometimes bets following the occurrence of an outlier, it fits the data poorly overall. This is because a second central prediction of the model is that the agent tends to bet early in each session, which is contrary to what the actual penny-pickers did. The model also predicts a global betting frequency that is much higher than actual. In particular, under Probability Matching, the agent never skips in all trials of a session. This result is in sharp contrast with actual behavior; there were sessions in which the penny-pickers did skip in all trials. See Figure 11, which compares the behavior of the actual penny-pickers to the behavior of a simulated Probability Matching agent that bets 80% of the time in the task.\(^{27}\)

\(^{27}\)Our choice of a 80% betting frequency for the simulated Probability Matching agent reflects the
Figure 11: Distribution of the first time the subjects who picked pennies (“the penny-pickers”) bet in the actual task (top graph) versus in a simulated task in which choice followed Probability Matching (bottom graph). The distributions were derived across all penny-pickers and sessions. In this simulation, the Probability Matching agent bet with probability 0.8 on each trial. N: cases in which the subjects skipped in all trials of a session. It appears that in approximately 15% of the sessions of the actual experiment, the subjects skipped in all trials; this never occurred in the simulation with the Probability Matching agent.
I next considered an altered version of the Bayesian model in which the agent believes in the Law of Small Numbers and hence falls victim to the Gambler’s Fallacy (Tversky and Kahneman (1971), Tversky and Kahneman (1974), Gilovich et al. (1985), Chen et al. (2015)). According to the Gambler’s Fallacy hypothesis, immediately after seeing a shot go wide, the agent mistakenly reasons that “the disaster event has happened already, so now a near miss is due and hence an outlier will not happen again in the short term.” This model thus predicts that the agent bets immediately after seeing an outlier, which is a distinctive behavioral trait of the first and third types of penny-pickers in the foregoing typology.

In the current task, the Gambler’s Fallacy phenomenon can be formulated using a model that mistakenly views the stochastic structure of the task as an urn with a majority of white balls and a few black balls. In each trial, a ball is drawn without replacement. Drawing a white (resp. black) ball corresponds to the occurrence of a shot within (resp. outside) the winning range. \( p_c \) (the chance that a shot falls within \([-4;4]\) in a Cauchy session), instead of being fixed to its true value (0.84—see p.16), decreases (resp. increases) after a white (resp. black) ball is drawn. The computational details of this model, which I refer to as “the faulty Bayesian model,” are available in Appendix 7.7.

The faulty Bayesian model fits the penny-pickers’ behavior poorly. The reason for this failure is that the model predicts not only that the penny-pickers bet immediately after seeing an outlier but also that they skipped after a prolonged series of shots within the winning range. That is, the behavior of the faulty Bayesian model is fully contrarian. Actual behavior was not. As emphasized earlier, with the penny-pickers, the departures from Bayesian learning chiefly occurred in the sessions with an outlier; the behavior was normal otherwise.

I also tested an altered version of the Bayesian model in which the agent falls victim to perceptual distortions, which leads him to neglect outliers. Specifically, the agent perceives a given outlier to be within two standard deviations of the empirical mean (which is derived from the previous observations in the session). Such perceptual distortion is a direct consequence of the way the agent’s brain adapts to the mean level of the stimulus the agent is watching. See Woodford (2012a) for a detailed fact that in the actual experiment, the estimated probability to win by betting was probably around 0.8 in the Cauchy sessions if the subjects were well calibrated (the objective probability is 0.84). In the Normal sessions, this probability should be even higher (the objective probability to win by betting is near 1). As noted before subjects appeared to be well calibrated on average.
description of this effect. To formalize it, I simply reset the outlier data observed by
the agent to a fixed threshold value not too far away from the empirical mean (I tried
different values for the threshold: 1, 2, 3, or 4 standard deviations from the mean).
That model did not fit the data better than the original Bayesian model did though,
which suggests that the pecking pennies bias was probably not caused by perceptual
distortions.

Finally, I tested an altered version of the EWA model in which the loss outcome
perceived by the model is four times lower than actual ($10$ rather than $40$). One
appeal of this model is that it may be able to predict the behaviors of all three types of
penny-pickers in the typology described in Figure 10 (p.37). That model fits subject
behavior better than the original model for 27 penny-pickers. The improvement in
the fits (above 5% on average) is significant.

The altered EWA model formalizes the idea that the penny-pickers behaved as
if they were blind or insensitive to losses. Based on recent neuroscience studies,
there may be two mechanisms underlying this behavior. Possibly, the penny-pickers
did not feel “alert” or “fear” signals upon experiencing a loss—the “somatic marker
hypothesis” (Bechara et al., 1997). A second possibility is that the penny-pickers were
unable to adjust their action plans upon receiving negative feedback, due to increased
dopamine levels caused by a series of positive outcomes—the aforementioned greed
hypothesis.

The currently available neuroscientific evidence strongly supports the greed hy-
pothesis (see, e.g., Frank and Hutchison (2009)). One litmus test for this hypothesis
consists in manipulating the dopamine levels of task participants during the task
[neuropharmacological study in preparation]. The third experimental treatment of
the current study constitutes another route to test the somatic marker and greed
hypotheses against each other. The results obtained in that treatment support the
greed hypothesis, as I report next.

Preference for “self-binding” in the third experimental treatment The
greed hypothesis implicitly assumes that those subjects who picked pennies in the
task could not help betting owing to elevated levels of greed. It thus implies that the
penny-pickers had self-control issues, i.e., that their behavior was at variance with
deliberation. The third experimental treatment directly tests this implication.

In the third treatment, in each trial of each session, the subjects were given the
opportunity to bet in all remaining trials of the session (henceforth, the “bet in all” option) as well as to skip in all remaining trials of the session (henceforth, the “skip in all” option). The majority of the subjects (48 subjects out of 60) used the option to bind their decisions, which is at odds with the behavior of a rational agent, as emphasized earlier.

One explanation for this behavior is that the subjects simply found it easier to commit to a given course of action rather than having to deliberate on each trial. However, if this explanation were correct, the frequency of choice for the “bet in all” option should not differ markedly from the frequency of choice for the “skip in all” option inasmuch as the base rate frequencies of the two actions were not markedly different in the experiment. This explanation is in sharp contrast with the data, which indicate a marked asymmetry in the use of the options. Specifically, among the subjects who used the options, 26 subjects exclusively used the option to “skip in all” versus only 2 subjects exclusively used “bet in all.” Therefore, the “skip in all” option was predominantly used. Furthermore, among the 20 subjects who used both options, the “skip in all” option was used 4.85 times on average (median, 5 times; mode, 3 times), whereas “bet in all” was used only 2.55 times on average (median, 2; mode, 1). Only two subjects used “bet in all” more often than “skip in all.”

In light of these results, the account that laziness (the desire to spare one’s cognitive resources) may underlie the use of the options in the third treatment is not tenable. However the observed behavior is fully consistent with the hypothesis that the subjects lacked self-control and were tempted to bet even when they knew (if only intuitively) that they should not. Therefore, they found the opportunity to bind their choices useful to ensure that they would skip when (they believed) betting was detrimental to their performance during the task.

I find that only 16 subjects picked pennies in the third experimental treatment. The proportion of penny-pickers is significantly smaller than in the previous experimental treatments according to a two-sample Z test (p value: 0.017). Furthermore, the model comparison analysis reveals that in that treatment, the Bayesian model describes subject behavior better than the best-fit EWA model for at least 90% of the subjects (see Figure 19 in Appendix 7.8, p.82), a proportion that is significantly higher than in the second treatment (p value: 0.03). These findings strengthen the evidence for the self-control hypothesis, inasmuch as they suggest that the introduction of the option to bind their decisions helped to reduce the extent of irrationality.
in the subjects’ behavior.

4.3.3 Robustness checks

Robustness of the estimation results I find that the evidence for Bayesian learning documented above also prevails in a model comparison analysis in which the EWA model is allowed to have absurd priors (unbounded initial attractions); see Appendix 7.8 for details.

Also, at one stage of the analysis I examined model fits using maximum likelihood (ML) rather than least squares techniques. It is important to note that the ML approach assumes that the subjects used a stochastic choice rule (logit, probit, or the like). This type of modeling choice can be supported when the subjects are expected to “explore” sub-optimal options by design (e.g., Manso (2011), Payzan-LeNestour and Bossaerts (2015)) or when the choice set is sufficiently large that the subjects likely “tremble” (e.g., McKelvey and Palfrey (1995), Kang et al. (2010)). However in the present task, the choice set comprised only two actions. In this context, modelling choice as stochastic is arguably implausible especially given the high monetary incentives that were provided to the subjects (see Woodford (2012b), p.46). Nevertheless, I ran the ML analysis for robustness purposes and found no notable differences in the results except that some fudge factor effect somewhat plagued the comparison of the out-of-sample log-likelihoods. The results with the ML approach are not discussed further, but they are available on request.

I also checked that the goodness-of-fit results were robust across different methodologies. I varied the length of the out-of-sample validation test (the last five sessions were used instead of the last seven sessions) and implemented alternative optimization methods to discover the optimum. See Appendix 7.5 for details.

External validity The findings reported in this paper provide evidence that subjects were lured into taking mindless risks in the task because they were greedy and had limited self-control. The task participants were undergraduates, and an important question concerns the extent to which these findings extend to populations of middle-aged investors. Inasmuch as the prefrontal cortex, which acts as “the CEO of the brain” (controlling impulses, modulating intense emotions, delaying gratification, and inhibiting inappropriate behavior), does not fully mature until near the age of 25 (Walsh, 2005), middle-aged investors might be relatively more immune to being
lured into picking pennies compared to young adults. To investigate this possibility, I performed two analyses. First, I studied gender differences in the experiment. Given that the prefrontal cortex matures markedly earlier in females than in males (Giedd et al., 1999), the propensity to pick pennies would be lower in the female task participants than in their male counterparts if the picking pennies behavior of the male participants were chiefly due to their having an immature brain. Notably, I did not find any gender difference in the propensity to pick pennies in the current experiment. Second, I replicated the second treatment of the current experiment with 42 MBA students from the Australian Graduate School of Management whose mean age was 33.3 years old. The picking pennies bias identified in the original experiment prevailed to an even greater extent in the experimental sessions with the MBA students, which strongly suggests that the current findings can be extended to populations of middle-aged agents.

The replication of the original finding in sessions with MBA students shows that it is not limited to the original sample. This replication also provides reassurance that this finding is a robust phenomenon across samples in different cultural backgrounds because only 33% of the students in the MBA sample were Australian, with 17 nationalities represented in total.

**Direct elicitation of the subjects’ beliefs** The foregoing findings suggest that when the penny-pickers were picking pennies in a given session, they knew—or at least strongly suspected—that the bowman in the session was an apprentice. This conclusion hinges on the outcome of the model comparison analysis, however (specifically, the finding that the penny-pickers acted like the Bayesian model except in those trials when they picked pennies). To double-check that this conclusion is valid, I ran additional experimental sessions that replicated the conditions of the original treatment except that at the end of each session, the subjects (N=52; same cohort as in the original experimental sessions: undergraduates at the University of New South Wales) were also asked whether, according to them, the bowman in the session was a master or an apprentice. An important rule in that treatment is that a correct answer yields $10 and an incorrect one results in a loss of $10. The absence of a reply leads to a loss of $12. This rule was well emphasized in the task instructions to ensure that the subjects were provided with the right incentives.

The subjects correctly guessed the nature of the bowman in 84% of the sessions.
Thirty-one subjects correctly assessed the type of the bowman in all but one or two sessions at most. Eight subjects managed to assess the bowman’s type in every session. Strikingly, 42 subjects answered exactly like the Bayesian agent in all but one or two sessions at most, and 23 subjects replied like the Bayesian agent in all sessions. Like the Bayesian agent, these subjects correctly identified the bowman type in all the sessions except those Cauchy sessions in which the Cauchy distribution was camouflaged by a series of seemingly Gaussian realizations (see Figure 1, p.19). Taken as a whole, the subjects answered like the Bayesian agent in the large majority (92%) of the sessions. This result both replicates and strengthens the evidence for Bayesian learning in the original treatment of the task, inasmuch as it does not rely on the structural model comparison but on direct elicitation of the subjects’ beliefs.

Notably, those subjects who picked pennies—42% of the pool, as in the original treatment—were as accurate as the others, with the exception of two subjects who ignored the nature of the bowman in most of the sessions. Specifically, the penny-pickers correctly guessed the nature of the bowman in 81% of the sessions. Among the normal subjects, the frequency of correct answers was similar. Strikingly, all of the penny-pickers knew that the bowman was an apprentice in at least one of the sessions when they were picking pennies. Most often, they knew the bowman was an apprentice each time they were picking pennies.28

This result strengthens the evidence that the penny-pickers in the task picked pennies despite knowing the true risks involved. The evidence was further strengthened in a final test run with an uncertainty-free variant of the current task.

**Uncertainty-free variant of the experiment** In the aforementioned follow-up sessions in which the subjects were asked, at the end of each session, their guess about the type of the bowman they had just faced in the session, the subjects were further asked how sure they were of their answer on a scale 1–3 (1: not quite sure; 3: certain). The subjects reported they were quite sure of their answer about the bowman type. The confidence ratings of the penny-pickers did not differ from those of the normal subjects. For both groups, the average rating was 2.4 and the mode was 3.

28The criterion that I used to infer that a given subject knew that the bowman was an apprentice when she was betting was that the subject correctly identified the bowman as an apprentice when asked at the end of the session AND encountered at least one outlier (realisation beyond seven meters away from the target) before she started to bet. The second condition is important because in sessions in which large realisations occur only at the end of the session, the subject may not yet know that the bowman is an apprentice when she is betting during the session.
Of note, the final earnings of the subjects were unaffected by the truthfulness of their confidence ratings, meaning the subjects had no incentives to be truthful. Maybe subjects were in reality uncertain regarding the bowman type. An empirical question then is to what extent such uncertainty exacerbates the propensity to pick pennies. In other words, would knowing beyond a doubt that the bowman in a given session is an apprentice effectively protect subjects against picking pennies in the session? The idea here is that there may be a psychological gap between knowing for sure and being quite convinced—a particular instantiation of the “certainty effect” (Allais, 1953).

To investigate this question, I ran a couple of sessions that replicated the original treatment of the task except that just before each session began, the subjects (N=44; same cohort as in the original experimental sessions: undergraduates at the University of New south wales) were told the nature of the bowman in the session (e.g., “The bowman in this session is a master. The standard deviation of the hits is 0.6.”). So the subjects knew beyond a doubt the type of the bowman in each session. In the task instructions, it was well emphasised that the “no deception rule” was in effect in the experiment. This was important to ensure that the subjects would not question what would be told to them during the task.29

For a risk-neutral agent, optimal behavior consists of always betting in the sessions with a master bowman and always skipping in the sessions with an apprentice. Loss-averse agents will play similarly except that they will skip in sessions with master bowmen whose standard deviation is deemed too high. Only a couple of subjects happened to skip in Gaussian sessions in which the standard deviation neared 1; sixty-six percent of the subjects happened to skip in Gaussian sessions in which the standard deviation was above 1.30

The picking pennies bias observed in the original experimental treatment of the task prevailed in this follow-up treatment as well. Half of the subjects picked pennies in the sense that in at least one of the sessions with an apprentice, they happened to bet after seeing a series of hits within the winning range. Importantly, for those penny-pickers taken as a whole, betting never occurred before the third trial in the large majority (93%) of the sessions with an apprentice.

Thus, like the penny-pickers in the original experimental treatment, the penny-

---

29*Psychology students were excluded from the pool of potential participants because psychologists commonly use deception in their experiments, so those students might have been biased.

30*This finding is consistent with the foregoing finding from the model comparison analysis that the average subject was loss averse in the task.
pickers in this new treatment systematically skipped in the first trials of the sessions with an apprentice bowman. This finding is important as it rules out the possibility that those subjects picked pennies merely because they were risk lovers in the task. Because with risk-loving subjects, the probability to bet is constant across the session whereas with the penny-pickers, the probability to bet is time-varying—nil early in the session, high later on. In contrast, the observed behavior is fully consistent with the greed hypothesis, which predicts that after the exposure to a few good—counterfactual or actually experienced—outcomes, subjects are tempted to bet despite knowing they should not.

These findings rule out the possibility that the picking pennies bias merely be an artifact stemming from a mischaracterization of subjects’ learning as sophisticated, since the bias persists as such even when no learning is involved. Furthermore, these findings rule out the possibility that subject behavior in the original treatment reflects a “disjunction effect,” as I explain in the following.

**Controlling for the Disjunction Effect**  Above, I have argued that the current findings constitute strong evidence that the subjects who picked pennies in the task did so not because they were faulty Bayesians or because they were risk-lovers in the experiment but because they were greedy and lacked self-control. There is one potential confounding factor that needs to be accounted for, however: the “disjunction effect” (see Tversky and Shafir (1992), Shafir and Tversky (1992), and Shafir (1994)). A disjunction effect occurs when people prefer a given action $x$ over an alternative action $y$ when they know that a given event $A$ obtains, and they prefer action $x$ over action $y$ when they know that event $A$ does not obtain, but they prefer $y$ over $x$ when it is not known whether or not $A$ obtains—a patent violation of the sure-thing principle.

Applied to the current task, the disjunction effect would lead risk-loving subjects to choose to bet when they know that the bowman is an apprentice and to choose to bet when they know that the bowman is a master. However, they would choose to skip when they do not know whether the bowman is an apprentice or a master (i.e., in the early stage of each session). This could explain why subjects who picked pennies in the experiment systematically skipped in the first trials of each session and often bet thereafter.

Follow-up investigations allowed me to exclude such a disjunction effect as a po-
tential explanation of the original findings. First, if the findings were driven by the disjunction effect, subject behavior (the probability to bet, in particular) would be the same in the original and third experimental treatments. In fact, the propensity to bet is markedly reduced in the third treatment, as reported. Second, findings in the aforementioned “uncertainty-free” or “beyond-a-doubt” variant of the task further rule out the disjunction effect as an explanation of subject behavior. The foregoing “disjunction effect” hypothesis would indeed predict constant betting throughout the sessions in that variant. In contrast, half of the subjects in that treatment behaved in the Cauchy sessions exactly like the penny-pickers in the original treatment of the task: they systematically skipped in the first trials and bet later in the session.

**Controlling for potential wealth effects** As emphasized in Section 2, one central design feature of the current experiment is that in each trial the subject did not know his current wealth level—the sum of his net accumulated outcomes and the starting account balance.\(^{31}\) This should in principle minimize the occurrence of wealth effects in the experiment. Yet, wealth effects might prevail in subjects who overlook the existence of the account balance and mistakenly take their accumulated outcomes in the task to reflect their current wealth level. Those subjects would wrongly believe they are “in the red” when their net accumulated outcomes are negative, which might lead them to bet to recoup prior losses. Combined with the self-control problem, such desire to recoup prior losses would trigger the picking pennies behavior in those subjects.

To account for the possibility that the picking pennies behavior was triggered by the foregoing wealth mechanism, I investigated the relationship between net accumulated outcomes and the picking pennies bias in the experiment. Specifically, I examined whether subjects were particularly likely to pick pennies when their net accumulated outcomes (henceforth, “wealth”) were negative. For each penny-picker, in each session in which the picking pennies behavior was observed, I assessed the wealth possessed by the subject at the trial when the picking pennies behavior started. Figure 12 shows the distribution of wealth across all penny-pickers and sessions. The distribution looks Normal, with a mean of $11. The normality assumption cannot be rejected (see Appendix 7.9 for details). Furthermore, in 43% of the penny-pickers,

\(^{31}\)Remember the amount of the account balance was revealed to the subjects only after the experiment was completed.
wealth was actually positive at the onset of all picking pennies episodes. These results compellingly suggest that in the current experiment, the desire to recoup losses was not a central determinant of the picking pennies bias.

Figure 12: Distribution of the net accumulated outcomes (“wealth”) of the penny pickers at the onset of a picking pennies episode.
5 Conclusion

The results of the current study suggest that the issue with tail risk is not the most commonly expected one: that people cannot assess it. Here I show that they can. I provide evidence for Bayesian learning in a complex environment featuring model uncertainty about tail risk. The Bayesian model is the best model of subject learning in the experiment, with prediction success as high as 90% in the third treatment. No boundedly rational model captures the regularities in the current data as effectively as the Bayesian model does. Direct elicitation of the subjects’ beliefs in a follow-up test confirmed that the majority of the subjects were Bayesians. The current results are thus at odds with some popular claims that people are not intelligent enough to learn about tail risk (see Taleb (2004a), Taleb (2004b), and Taleb (2007)). They suggest that learning failures in financial markets (see, e.g., Gennaioli et al. (2012), Jin (2015)) are due to lack of sufficiently salient information about tail risk, not to lack of information-processing capabilities, per se. As such, they lend support to regulatory measures aimed at increasing information salience, e.g., mandating that money market funds be marked to market, as proposed by Gennaioli et al. (2012).

The current results suggest that one of the most important issues with tail risk is that many of us cannot properly deal with it. I find that at least 41% of the subjects picked pennies in the task. The root cause of this behavior appears to be the exposure to steady streams of good payoffs throughout the task, which made the subjects extremely greedy. Those subjects who picked pennies during the task could not resist the temptation to do so; their behavior contrasted with deliberation, as the third treatment of the current experiment demonstrates. The third treatment further revealed that providing subjects with the option to bind their decisions greatly reduced the extent of the picking pennies bias in the experiment.

The current findings speak to the need to provide market participants with institutional commitment devices so they can bind themselves to not taking mindless risks. Because what is true of the picking pennies bias in the current task is equally true, mutatis mutandis, of the picking pennies pitfall in real-world finance. In fact, the main forces that led the subjects to pick pennies in the current task (namely, high cognitive demands and exacerbated greediness) are arguably stronger in the financial markets, where cognitive challenges and reward temptations are both notoriously high. Reports from market practitioners testify to this phenomenon. For instance,
Charles O. Prince (Citigroup’s CEO from 2003 to 2007) famously said in 2007 about Citigroup’s continued commitment to leveraged buy-out deals despite fears of reduced liquidity because of the occurring subprime meltdown, “As long as the music is playing, you’ve got to get up and dance.” Reacting to that sentence, commentators compared market practitioners to addicts and to people driving cars over a broken but still standing suspension bridge where one cable at a time is breaking, betting the last cable will not break while they are on the bridge.

As such, the current findings raise an important issue for financial markets: who has the incentives to provide market practitioners with the self-binding needed to avoid the picking pennies pitfall, and who does not? Evidently, some sophisticated agents, such as banks and hedge funds, make money from the picking pennies bias described in this paper. For instance, John Paulson knew how to exploit the aforementioned picking pennies behavior that prevailed on the RMBS CDOs market in 2006 and early 2007. Paulson bought CDS insurance on $5 billion of RMBS CDOs—a bet they would fall in value. By doing so, Paulson was “bleeding” (losing tens of millions of dollars) while the owners of the CDO slices were picking their pennies on the other side of the trade, but he made more than $15 billion of profits from these trades when the CDO owners eventually blew up in 2007 (Zuckerman, 2010).

Our findings raise a number of further questions. First, why are some individuals more prone to the picking pennies bias than others? The current study points to individual differences in greed and self-control when facing the temptation to pick up pennies in risky places. Is the extent of such temptation modulated by psychological traits such as “sensation seeking” and by cultural factors such as religious beliefs, as previous work suggests (e.g., Grinblatt and Keloharju (2009) and Kumar et al. (2011))? Are the same individuals who have a predilection for lottery-like assets (Kumar (2008) and Kumar et al. (2011)) also more susceptible to picking pennies? These are all important directions for future research, which will help improve individual decision making.

---

32E.g., “20/20 Hindsight Through What Were Once Rose-Colored Glasses,” by Jenny Anderson, The NY Times, August 31, 2007: “Citigroup’s chief executive, Charles O. Prince III, said it right when he said that as long as the music is playing, the banks have to keep dancing. In a boom, they are all addicts who can’t say no. Until they collapse from an overdose at the disco.”

References


6 Acknowledgments

Funding for this study was from the Australian School of Business at the University of New South Wales.
7 Appendix

7.1 The Adaptive Forecasting approach

7.1.1 The Adaptive Forecaster model

The Adaptive Forecaster model is an adapted version of an algorithm originally proposed by Foster and Vohra (1998) and revisited by Carvajal (2009). In each trial, the model purports to forecast in a purely adaptive way—i.e., without inferring the hidden state of the world as in Bayesian learning—whether the next shot will fall within the winning interval $[-4; 4]$. Let $Z$ be the random variable that takes the value 1 at trial $t$ if the time $t$ realization is within the winning range ($X_t \in [-4; +4]$) and 0 otherwise. At each trial $t$, the model chooses the value of the probability $p_t \equiv P(Z_t = 1)$ among $M$ possible choices ($M$ fixed positive integers) $p(1), p(2), ..., p(M)$, defined as the midpoints of the following intervals:

$$
\left[ 0, \frac{1}{M} \right], \left[ \frac{1}{M}, \frac{2}{M} \right], ..., \left[ \frac{m-1}{M}, \frac{m}{M} \right], ..., \left[ \frac{M-1}{M}, 1 \right].
$$

That is, $p(m)$ is the midpoint of the interval $I(m) = \left[ \frac{m-1}{M}, \frac{m}{M} \right]$. Let $h_t = (p_t, X_k^t)_{k=1}^t$ denote the history of probability choices and data realizations up to (and including) time $t$. To determine the probability $p_t$, the model implements the following probabilistic rule at each trial $t$: “$p_t$ equals $p(m)$ with probability $L_t(h_{t-1})(m)$.” To pin down the value of $L_t(h_{t-1})(m)$, some notations are needed. For each candidate for probability $p_t$, the model considers the empirical (observed) frequency of the occurrence of $Z = 1$ conditional on the candidate having been selected:

$$
\rho_t^m(h_t) = \frac{\sum_{k=1}^t Z_k I(p_k = p(m))}{\sum_{k=1}^t I(p_k = p(m))},
$$

where $I(p_k = p(m))$ equals 1 if $p_k = p(m)$ and 0 otherwise. The model further considers “the excess” associated with the candidate $p(m)$:

$$
e_t^m(h_t) = \left( \rho_t^m(h_t) - \frac{m}{M} \right) \frac{\sum_{k=1}^t I(p_k = p(m))}{t},
$$

where $I(p_k = p(m))$ equals 1 if $p_k = p(m)$ and 0 otherwise.
as well as the following metric (called “deficit”):

\[ d_t^m(h_t) = \left( \frac{m - 1}{M} - \rho_t^m(h_t) \right) \sum_{k=1}^{t} \frac{I(p_k=p(m))}{t}. \] (26)

The model starts with \( p_1 = p(M) \). For \( t \geq 2 \), the rule for choosing \( p_t \) is that if there exists some candidate \( \tilde{m} \in M \) such that \( \rho_{t-1}^{\tilde{m}} \in I(\tilde{m}) \), then \( L_t(h_{t-1})(\tilde{m}) = 1 \) and \( L_t(h_{t-1})(m) = 0 \) for every other \( m \). Otherwise, the model searches for the candidate \( \tilde{m} \) such that \( d_t^m(h_{t-1}) > 0 \) and \( e_t^{m-1}(h_{t-1}) > 0 \) and sets the probability of choosing each candidate \( m \) as follows:\footnote{If there is no such \( \tilde{m} \), the model determines \( p_t \) by randomly picking one candidate among the hitherto unchosen probability candidates.}

\[
L_t(h_{t-1})(\tilde{m}) = \frac{e_{t-1}^{\tilde{m}-1}(h_{t-1})}{d_{t-1}^\tilde{m}(h_{t-1}) + e_{t-1}^{\tilde{m}-1}(h_{t-1})},
\] (27)

\[
L_t(h_{t-1})(\tilde{m} - 1) = 1 - L_t(h_{t-1})(\tilde{m}),
\] (28)

\[
L_t(h_{t-1})(m) = 0 \quad \text{for all other } m.
\] (29)

Then, to determine \( p_t \), the model selects one candidate in the set \( \{p(1); p(2), \ldots; p(M)\} \) according to the stochastic rule \( \{L_t(h_{t-1})(m)\}_{m=1}^{M} \).\footnote{That is, the first candidate, \( p(1) \), is chosen with probability \( L_t(h_{t-1})(1) \); the second, \( p(2) \), is chosen with probability \( L_t(h_{t-1})(2) \), etc.} Once \( p_t \) is determined, the expected value of Action Bet at trial \( t \) is directly computed:

\[
V_t = p_t \times 2 - (1 - p_t) \times 40.
\] (30)

Under this approach, the behavior consists of choosing to bet at trial \( t \) if \( V_t \) is positive and choosing to skip otherwise. This forecasting approach is purely adaptive in that it is completely ignorant of the real probabilities of the occurrence of the event \( (Z_t = 1) \) and the reasons the event occurs or fails to occur. Moreover, the model does not learn anything about the future from the history. Despite its simplicity, the algorithm meets a long-run accuracy criterion: Foster and Vohra (1998) indeed show that asymptotically, the conditional relative frequencies of occurrence of the event approach the forecast. What about the short-run accuracy of the model? Figure 13
shows how the Adaptive Forecaster model assigned the probability of a winning bet in a simulated Normal session (top) and in a simulated Cauchy session (bottom). The overall accuracy was not perfect in the Cauchy session, although the model’s forecast closely approximated the true probability (0.84) toward the end of the session. One distinctive behavioral characteristic of the model is also apparent in Figure 13: in contrast to the Bayesian model, which never bets in a given session unless sufficiently convinced that the bowman is a master, the Adaptive Forecaster model tends to bet early in each session.

7.1.2 Subject fit

Unlike the free parameters of the other models, the free parameter of the Adaptive Forecaster model \( (M) \) could not be fit to the data because the Adaptive Forecaster model is stochastic by nature (in the sense that when facing the exact same data twice sequentially, the model often chooses a different probability candidate and hence may behave differently the second time). The success rate of that model was thus assessed for different fixed values of \( M \) (between 11 and 30).

It appears that the performance of the Adaptive Forecaster model is robust across the different values of \( M \), with an average success rate of 70%. Although the model predicts actual behavior better than the best-fit EWA model for 40% of the subjects, it is unambiguously trumped by the latter according to a paired t-test based on the differences in the individual fits (p-value: 0).
Figure 13: Simulated run of a Normal session (top) and a Cauchy session (bottom) with the Adaptive Forecaster model. The top figure reports a simulated session with a master bowman (“Normal session”) whose shots had a standard deviation of 1. The simulated shots fell within the winning range. The bottom figure reports a simulated session with an apprentice bowman (“Cauchy session”). The simulated shots fell within the winning range except for the shots at trials 5 and 10. In both simulations, the parameter $M$ of the model (as defined in the main text) was fixed to 12.

Legend: Each figure shows $p$, the estimated probability that the next shot will be within the winning range (top) as well as the corresponding choice made by the model (bottom) and the ensuing accumulated outcomes (middle). In the bottom graph, both the actual and the foregone outcomes are indicated. For instance, in the bottom figure, the model skipped at trial 10 and would have lost $40 if he had bet.

It appears that in the Normal session, the model’s forecast closely approximates the true probability (0.99) from the onset. In the Cauchy session, the forecast is accurate from trial 14 on. The forecast appears to be too optimistic in the early trials; consequently, the model bets and exposes itself to a loss (which occurs at trial 3).
7.2 Inductive learning model

Psychologists have noted that people sometimes use an inductive approach whereby they put full weight on the hypothesis they deem more likely and neglect the alternative hypothesis (see Collins and Koechlin (2012) and Gallistel et al. (2014)). This is in contrast to the Bayesian “model averaging” approach, which consists of weighting each possible world according to its likelihood (see Equation (8) in the main text). Under the “inductive learning model” the expected value of Action Bet at trial $t$ is as follows (using the definitions and notations of Section 3.1 in the main text):

$$
V_t = \begin{cases} 
p_g \times 2 - (1 - p_g) \times 40 & \text{if } P(M_1 | X_{t-1}) > P(M_2 | X_{t-1}), \\
p_c \times 2 - (1 - p_c) \times 40 & \text{otherwise.}
\end{cases}
$$

I fitted the inductive learning model to the subjects data (using the methods described in the main text and Appendix 7.5), and found that for 95% of the subjects, the Bayesian model fits subject behavior better than does the inductive learning model. This result suggests that subjects implemented model averaging rather than the inductive approach. This may be due to the fact that the current task is simple (there are only two possible worlds). I conjecture that the larger the set of possible worlds becomes, the less likely it is that the model averaging approach be used. This conjecture accords with prior theoretical work in which the number of states that the agent can distinguish is bounded (see, among others, Rubinstein (1986)).
7.3 Instructions for the task

The following pages show the text and static pictures of the task instructions used for the original treatment of the task. The instructions were displayed on web pages and contained animations. You can see the dynamic features of the instructions (animations and multiple-choice questionnaire) using this link: http://bowmangame.weebly.com/game-instructions.html.
The following instructions explain the nature of the Bowman Game.

Please read them carefully.

The principle of the game is that on each trial, a bowman is going to shoot an arrow towards a target on a wall (See Pic 1A: the wall is represented by the gray line; 0 represents the target).

Before the bowman shoots each arrow, you will have the opportunity to bet on whether the arrow will be close to the target or not. At each trial, before the bowman shoots the arrow, your task will be to choose between two actions, Bet and Skip:

- By selecting Bet you will earn $2 if the arrow hits the wall up to four meters away from the target on the left or on the right (i.e., if the hit is anywhere within the segment [AB] on Pic. 1B), and you will lose $40 otherwise.

- By selecting Skip you do not lose nor earn anything no matter where the arrow hits: you always get $0.

Once you have indicated your choice, the bowman shoots and you see where the arrow hits. You'll then move to the next trial, in which you'll choose again between Bet and Skip, and then see the realized hit at that trial.

A play of the game is divided into 15 sessions of 20 trials each. For each session, a new bowman is doing the shooting for the entirety of the session.
Each Bowman is either a Master Bowman or an Apprentice Bowman, and has a specific shooting style corresponding to its skill level. We describe next the shooting style of the Master Bowman and then the Apprentice Bowman.

**The Master Bowman**

The hits of a Master Bowman are normally distributed around the target.

The dispersion of the hits from the target (the standard deviation) increases with the shooting distance of the Bowman. The shooting distance differs across bowmen (see Pic. 2) but it is never changing within a session.

![Picture 2](image)

The standard deviation of the hits of a given Bowman is unknown. It is uniformly distributed between 0.1 and 2 (i.e., all numbers within \([0.1, 2]\) are equally likely).

The following interactive animation illustrates how the hits of the Master Bowman are distributed around the target, for different shooting distances (i.e., for different values of the standard deviation):

[show animation—distribution builders for Normal distributions here]

**The Apprentice Bowman**

The distribution of the hits of an Apprentice Bowman is bell shaped and symmetric around the target, like the one of the Master Bowman. But it’s NOT NORMAL: with an Apprentice Bowman, you should not expect any particular hit to happen, because the hits can potentially be anywhere on the wall.
Apprentice Bowmen always shoot at the same distance from the target. The dispersion of their hits always equals 1, which implies that their hits are distributed around the target as follows:

[show animation—distribution builder for Cauchy distribution here]

Please note

Whether the bowman for a session is a Master Bowman or an Apprentice Bowman is not told

Also note

The distance of the bowman to the wall will vary from session to session, but the distance you’ll see on the screen will appear to be the same at each trial (again, this is NOT the real distance; the real distance varies from session to session).

This game is very hard, so you will need to be very concentrated to be successful. Your payment will be extremely sensitive to your performance during the game. Specifically, your payment will be the addition of:

- the fixed show-up reward of $5
- an account balance allocated to each participant at the start of the game (more on this below)
- all your accumulated earnings and losses during the entire game

The account balance is the same for all participants and has been fixed before the game begins. Its amount, which can be any number between -500 and +500, will be revealed at the end of the game only. This is because we don’t want you to be influenced by it during the game.

Your accumulated earnings and losses can be negative of course. In that case, the negative amount will be deducted from the money allocated to you at the start. If it goes below $5, you will still receive $5 as the show-up fee. If you are very good at the game, your final amount as calculated by the formula above might be more than $110. In that case, your payment will be capped at $110, as we simply can’t afford to pay each of you more than that. It may be challenging to reach this score though!

Good luck!

Just a quick reminder that on each trial during the game:

- by selecting Bet,

You earn $2 if the arrow hits the wall up to four meters away from the target.
You lose $40 if the arrow hits the wall beyond four meters away from the target.

- by selecting Skip,

You do not lose nor earn anything no matter where the arrow hits: you always get $0.
One more thing: If you don't answer within the imparted time (indicated by a timer), you lose $1, and you don't see the hit realized at the trial, so try to keep the pace!

Please fill out the following multiple-choice questionnaire. This is to check your understanding of the rules of the game before you start to play. Thanks for your attention!

[multiple-choice questionnaire; for each question computer tells the subject whether answer is correct ]
7.4 Time line of a trial of the task
Figure 14: Time line of a trial in the first experimental treatment of the task. In the second experimental treatment, the third screen was absent (there was no outcome revelation stage). In the third experimental treatment, a button was added at the decision stage (first screen) to allow the subject to select the option to bet or skip in all remaining trials of the session. Short demos of each of the three experimental treatments are available at http://bowmangame.weebly.com/.
7.5 Cross-validation procedure

The model comparison analysis of the Bayesian and EWA models proceeded in two steps: an in-sample calibration procedure in which the free parameters of the models were fitted using the first 8 sessions of the task, followed by an out-of-sample (validation) test in which the prediction accuracy of the fitted models was tested in the remaining 7 sessions of the task.

This type of two-step procedure is commonly used in model comparison analyses involving very general models such as the EWA model, to guard against overfitting. Indeed, very general models often fit well in-sample by overfitting; forecasting out-of-sample is critical to remove any inherent advantage of general models over simpler contenders, as argued in, e.g., Camerer and Ho (1999), Ho et al. (2007), and Ho et al. (2008).\(^{36}\)

In the in-sample calibration step, the parameters of the Bayesian and EWA models were optimized by minimizing the total squared prediction error compounded over the set of trials for the first eight sessions (see Figure 15). Initially, I implemented a grid search to find the best values for the parameters.\(^{37}\) Next, these best values were used as the initial values for the constrained nonlinear optimization algorithm `fmincon` (trust-region method) implemented in Matlab (The MathWorks, Inc.), which further refined the search for the optimum. I also checked that the goodness-of-fit results were robust across different optimization methods to discover the optimum; the results were unchanged when the constrained nonlinear optimization algorithm `patternsearch` as well as a genetic algorithm were used instead of `fmincon`.

An important aspect of the analysis is that I tested all possible choices for the EWA model specification. Specifically, I fitted to the data the EWA model in which all parameters \((\kappa, \delta, \phi, N(0), \text{ and } A_B(0))\) were free; the one in which \(\kappa\) was fixed to 0 and all of the other parameters were free; the one in which \(\kappa\) was fixed to 0, \(\delta\) was set to 1, and all of the other parameters were free; the one in which \(\kappa\) was fixed to 0, \(\delta\) was set to 0, \(\delta\)

\(^{36}\)The intuition behind overfitting is that for a very general model such as EWA, there may exist a particular combination of parameters that is consistent with any experimental outcome. When EWA fits well in-sample by overfitting, this overfitting is revealed by the fact that its predictive accuracy is worse when predicting out-of-sample than when fitting in-sample.

\(^{37}\)The size of the grid was \(((50 \times N_v)^{1/N_v} + 1)^{N_v}\) where \(N_v\) was the number of optimization variables (number of free parameters). That is, along each dimension of the optimization space, there were \((50 \times N_v)^{1/N_v} + 1\) points. This scheme was implemented to prevent the total number of points from growing exponentially with the number of free parameters.
was fixed to 1, \( \phi \) was fixed to 1, and \( N(0) \) and \( A_B(0) \) were free, etc. In general, when a parameter was fixed, it was fixed to its most plausible value(s) given the current task (i.e., \( \delta \) was fixed to 1, \( \phi \) was fixed to 1, and \( A_B(0) \) was fixed to the expected value of Action Bet, i.e., -1.57). Furthermore, I tested the categories of the EWA models in which \( N(0) \) was fixed to 1 and those in which \( \kappa \) was fixed to either 0 or 1, in reference to prior work.\(^{38}\) The goal was to find the EWA parametrization that gave the best predictive accuracy out-of-sample, and to compare it to the best-fit Bayesian model.

For the Bayesian model, I assessed the out-of-sample predictive accuracy of the loss-averse and disappointment-averse versions. Like for the EWA model, I fitted to the data all specifications of the loss-averse model i.e., I tested the unconstrained version in which all parameters were free, the constrained version in which all parameters were free except \( \alpha_3 \) which was set to 1, etc. For reference, I also assessed the predictive accuracy of the risk-neutral version of the Bayesian model for which there was no free parameter, i.e., no in-sample calibration was needed.

---

\(^{38}\) Prior work has suggested fixing \( \kappa \) to either 0 or 1. Both values make sense, and the results are not sensitive to the value of \( \kappa \). Likewise, in reference to the previous literature I tested versions of EWA in which \( N(0) \) was fixed to 1.
7.6 Subject behavior in one instance of the task

Figure 16 shows an instance in which the subject appears to be less loss averse than the fitted Prospect Theory Bayesian model in some of the sessions, contrary to that in the example provided in the main text, in which the subject is consistently more loss averse than the model during the experiment. Here in some of the trials in some of the sessions (e.g., sessions 1-4), the subject chose to bet while the model chose to skip.
Figure 16: Subject behavior versus Bayesian behavior in one instance of the task. Each graph shows the choice made by one of the subjects at each trial in one of the sessions of the task (sessions 1–15, from the top and the left). Legend for each graph: X axis: trial number. Y axis: ‘0’ Action Skip was chosen; ‘1’ Action Bet was chosen. Both the choice made by the subject (cross) and the choice made by the fitted Bayesian model (circle) are displayed.
Figure 17: Value of Action Bet according to the Bayesian model in the same instance of the task as the one in Figure 16. Each graph shows the estimated value of Action Bet according to the fitted Bayesian model at each trial in one of the sessions of the task (sessions 1–15, from the top and the left). Both the estimated value of Action Bet (blue line) and the value of the shot realized at the trial (blue dot) are displayed. Legend for each graph: X axis: trial number. Y axis, left: shot value. Y axis, right: Action Bet value.
7.7 Faulty Bayesian model: Gambler’s Fallacy

Subjects who believe in the “Law of Small Numbers” incorrectly think that seeing a shot outside of the winning range $[-4; 4]$ (henceforth, an “outlier”) means that it is less likely that they will immediately see another one. These subjects also expect to see an outlier when they have not seen one for a while. To formalize this belief, the Bayesian model is slightly modified as follows. Instead of using $p_c$ as defined in the main text, the model takes that probability to be (at each trial $t = 1, \cdots, 19$)

\[
p_{ct} = \frac{n_t}{100 - t},
\]

(31)

where $n_t = n_{t-1} - I(X_t \in [-4; 4])$.\(^{39}\) $n_0 = 95$. Under this model, the agent thinks of the stochastic structure of the task as an urn containing initially 95% white balls and 5% black balls. In each trial, a ball is drawn without replacement. The drawing of a white (resp. black) ball corresponds to the occurrence of a near miss (resp. an outlier). That model always bets after seeing an outlier. Another distinctive trait of the model is to skip after a long series of realizations within the winning range.

\(^{39}\) $I(X_t \in [-4; 4]) = 1$ if $X_t \in [-4; 4]$ and 0 otherwise.
7.8 Model comparison between the Bayesian and benchmark EWA models

For robustness purposes, I ran the model comparison between the Bayesian model and the unbounded EWA model in which $A_B(0)$ could take any value. When unbounded, the best-fit EWA model initially values Action Skip extremely highly relative to Action Bet: the gap between the fitted values of the parameters $A_S(0)$ and $A_B(0)$ nears 13. The average fitted value of $N(0)$ is 0.88, meaning that the learning rate is quite high. The average fitted value of both $\phi$ and $\delta$ is close to 1 (meaning that the average subject had perfect memory and learned in a fully counterfactual way).

That version of the EWA model has respectable fits with an average success rate near 84%, but the Bayesian model has a better fit for 73% of the subjects, and the p value is still essentially 0. The reason that model has respectable fits is that it can mimic the Bayesian model in some respects. On the one hand, the high initial valuation of Action Skip ensures that the subject skips in the first trials of each session. On the other hand, the learning rate is sufficiently large that the subject will not skip forever and will bet after seeing a series of good outcomes despite a significant a priori aversion to betting. The main reason that the model cannot fully mimic the Bayesian model is that it systematically skips after observing a loss, irrespective of the value of the shot that led to the loss. In contrast, the Bayesian model usually perseveres in betting despite losing $40 if the shot that led to the loss fell sufficiently near the target (see Figure 2 of the main text). This result shows that no combination of EWA parameters can simultaneously generate all signatures of subject behavior in the current task.

The foregoing unbounded version of EWA should not be taken at face value because the large difference between the fitted values of the parameters $A_B(0)$ and $A_S(0)$ implies—as per the definition of $A_B(0)$ provided earlier—that the subject believes that the probability of a shot going wide with an apprentice Bowman is above 50%. Such a belief is highly implausible, inasmuch as it implies that the subjects grossly overestimated the true probability, which was unlikely to be the case in the majority of the subjects, as noted above. In addition, I ran a sensitivity analysis to assess whether the out-of-sample performance of that model was sensitive to small variations in some of its parameter values. The performance of the model appears to be very sensitive to the value of $A_B(0)$, which casts further doubt on the descriptive
validity of that model. Despite this finding, given its good out-of-sample performance, I used that model as a benchmark for the model comparison analysis to derive conservative estimates of the prevalence of Bayesian learning in the task (see Figures 18 and 19).

Figure 18: Comparative goodness-of-fit of the Bayesian model versus the benchmark model in the first two treatments of the task. X-axis: Success rate (fraction of trials in which the model predicted the actual choice) of the best-fit Bayesian model. Y-axis: Success rate of the benchmark model—the best-fit EWA model in which $A_B(0)$ was unbounded. Each data point corresponds to one subject. The data from the two first treatments of the task are pooled together (N=124). The Bayesian model fits better when the data point is below the 45 degree line.
Figure 19: Comparative goodness-of-fit of the Bayesian model versus the benchmark model in the third treatment of the task. X-axis: Success rate (fraction of trials in which the model predicted the actual choice) of the best-fit Bayesian model. Y-axis: Success rate of the benchmark model—the best-fit EWA model in which $A_B(0)$ was unbounded. Each data point corresponds to one subject (N=60). The Bayesian model fits better when the data point is below the 45 degree line.
7.9 Testing for potential wealth effects

The wealth distribution reported in Figure 12 (p.50) looks roughly Normal, which is confirmed by the visual Q-Q plot shown in Figure 20. The normality assumption cannot be rejected in a Shapiro-Wilk test (t=0.944; critical value at 5%: 1.645).

Figure 20: Q-Q plot showing the quantiles of a normal distribution against the quantiles of the empirical distribution documented in Figure 12 of the main text. X-axis: quantiles of a normal distribution. Y-axis: quantiles of the empirical distribution shown in Figure 9 of the main text.