Expectation Formation in an Evolving Game of Uncertainty: Theory and New Experimental Evidence*

Gigi Foster (UNSW) Paul Frijters (UQ)
Markus Schaffner (QUT) Benno Torgler (QUT, EBS)

October 31, 2013

Abstract

We examine the nature of stated subjective probabilities in a complex, evolving context in which true event probabilities are not within subjects' explicit information set. Specifically, we collect information on subjective expectations in a car race wherein participants must bet on a particular car but cannot influence the odds of winning once the race begins. In our setup, the actual probability of the good outcome (a win) can be determined based on computer simulations from any point in the process. We compare this actual probability to the subjective probability participants provide at three different points in each of 6 races. We find that the S-shaped curve relating subjective to actual probabilities found in prior research when participants have direct access to actual probabilities also emerges in our much more complex situation, and that there is only a limited degree of learning through repeated play. We show that the model in the S-shaped function family that provides the best fit to our data is Prelec's (1998) conditional invariant model.

Keywords: behavioural economics, expected utility theory, experiments, expectations, probabilities
JEL classification: D40, L10

*We acknowledge the Australian Research Council (grant DP090987840), ORSEE, and the ASBLab at the University of New South Wales for enabling our experiments. This research has the approval of the UNSW Human Research Ethics Committee. All errors and viewpoints are our own.

Although its empirical picture has come into focus, the weighting function has remained a somewhat tricky object to analyze—at least in comparison with the utility function ... Overall, it does not look like a shape that one would draw unless compelled by strong empirical evidence. (Prelec 2000, p. 67)

1 Introduction

The quality of human judgment has been intensively explored in various disciplines, including psychology, management, and economics. The vast literature on non-expected utility theory (see Starmer 2000, Starmer 2004, for a review) was originally spawned from the persistent inability of rational models of behavior (von Neumann, J. & Morgenstern 1947) to predict choice behavior in experiments (Camerer & Loewenstein 2004). Among the early evidence for this inability was the Allais paradox, derived in a canonical experiment in which subjects were asked their preferences between a lottery with chance x of winning prize X (getting zero with chance $1 - x$) and a lottery with chance y of winning prize Y (again with zero as the alternative). One (seemingly mild) condition of rationality was that anyone preferring the first lottery over the second should also prefer the first over the second of an amended version of both lotteries in which the chances of winning were changed, respectively, to $x * c$ and $y * c$. The frequent violation of this condition spawned a cottage industry of further choice experiments to tease out individuals' cognitive biases in lottery playing.

One prominent finding in this literature is that even when they are told the actual probabilities associated with outcomes, individuals behave as though other probabilities apply. Early experimental work by Preston & Baratta (1948), for example, identified an inverted S-shaped function in which subjective (psychological) probabilities exceed objective (mathematical) probabilities at low values of p but fall short of objective probabilities at high values of p (seen pictorially in Figure 1). They explain this observation using an analogy to another psychological study in which participants' perception of the brightness of a stimulus varied in proportion to a comparison, or 'anchor', level of brightness. The core proposal of Preston & Baratta (1948) was that the crossing point in their experimental data at which the subjective and objective probabilities are equal was a function of an initial anchoring level, which itself may relate to inherent psychological or physiological attributes and the initial endowment (p. 191-192). In their experiment, the S-shape was not only visible in student data but also in data from faculty of mostly professorial rank in the fields of mathematics, statistics, and psychology, with many years of acquaintance with probability theory. Subsequent work by Dale (1959) also reported that experimental subjects tend to overestimate low probabilities and underestimate high probabilities.

Since then, a raft of theories—including prospect theory (Kahneman & Tversky 1979), rank-dependent utility theory (Quiggin 1982, Yaari 1987, Schmeidler 1989, Wakker 1994), adaptive probability theory (Martins 2006), and conditional small world theory (Chew & Sagi 2008)—have proposed specific cognitive decision rules for how subjective probabilities are derived from objective ones. Contrary to expected utility (EU) theory, however, these approaches treat actual probabilities as nonlinear inputs to subjective probabilities. Moreover, despite being axiomatically based, each has become associated with a particular probability-weighting function that relates the subjective probability p^s to the true probability p. These weighting functions are consistent with the psychophysics of diminishing sensitivity, wherein the marginal impact of a stimulus diminishes with increasing distance from a reference point (Tversky & Kahneman 1992, Camerer 1995, Fox & See 2003). They may also reflect affective aspects of the decision process, such as hope (when contemplating high-probability losses) or fear (when contemplating low-probability losses), which could imply that the more pronounced is the S-shape, the emotionally richer is the decision-making situation (Rottenstreich & Hsee 2001, Trepel, Craig & Poldrack 2005).

The most popular probability weighing function is $p^s = \frac{(p)^\gamma}{\sum_j (p_j)^\gamma}$, where $\gamma < 1$ and j enumerates all possible outcomes an individual might consider. This function's most salient characteristic is that small probabilities are over-weighted and large probabilities are under-weighted, particularly if the number of possible outcomes is large. A similar function proposed by Quiggin (1982) and used explicitly in Tversky & Kahneman (1992) has subsequently been employed for estimation in dozens of experiments (see Machina 2004). This line of research has produced increasing evidence that the use of a weighting function enables a far better fit to how our brains actually process the probabilities of a set of outcomes than can be generated from the purely rational model (Berns, Capra, Chappelow, Moore & Noussair 2008). In fact, Camerer & Loewenstein (2004) stress that 'the statistical evidence against EU is so overwhelming that it is pointless to run more studies testing EU against alternative theories (as opposed to comparing theories with one another)' (p. 20). The research evidence also suggests that the elevation of the weighting function (i.e., the degree of over- or under-weighting that it features) may be linked to individuals' impulsiveness, and thus to the dopaminergic and serotonergic systems (Trepel et al. 2005).

Nevertheless, to our knowledge, every empirical study conducted so far that tries to estimate a probability weighting function is based on informing individuals of an outcome's true probability, and then observing their choice behavior. For example, a typical experiment might present the following scenario:

> You participate in two lotteries. The first gives you a A% chance of winning $250, while the second gives you a (A+25)% chance of winning the same amount. Which of the following options seems like a more significant change in the odds?
>
> 1) Increase your chance of winning the first lottery from A% to (A+5)%
>
> 2) Increase your chance of winning the second lottery from (A+25)% to (A+30)%

(adapted from p. 130 of Gonzalez & Wu (1999))

To the purely rational person, the answer would be that there is no difference. Yet, consistent with the S-shaped function discussed above, students given this question tend to pick option 1 when A is small (5%) and option 2 when A is large (65%).

In this paper, we confront individuals with scenarios more similar to real-life decision-making contexts, in which the actual probability of success is not within individuals' feasible information set. We estimate the probability weighting function in a novel experimental situation that asks individuals for their estimates of the probability of a good outcome in a complex situation where correctly ascertaining the actual probability of a good outcome, although not impossible—participants receive all the information needed to calculate it—is extremely difficult. For this purpose, we adapt a race car game in which participants must bet on the ultimate outcome, which cannot be influenced once the game is underway, and ask them several times during each race what they think the odds are of their chosen car winning the race. Our set-up is thus dynamic—like the few existing papers (e.g. Cheung 2001) that examine dynamic, probabilistic processes over time—and allows us to monitor how agents update their probability expectations as a risky situation unfolds.

We check the validity of the elicited probabilities by comparing them against a set of incentivized choices that we allow participants to make at the same moments when their subjective probabilities are elicited. Subjects may choose at these moments whether or not to pull out of the game, where pulling out involves recouping a fraction of their bet for sure rather than letting the race run its course and risking the loss of their whole bet. This check also allows us to determine whether stated probabilities have information value in terms of actual decision-making behavior involving a trade-off between costs or benefits known now with certainty, versus risky future outcomes (de Palma, Andre, Brownstone, Holt, Magnac, McFadden, Moffat, Picard, Train, Wakker & Walker 2008). Such choices are exemplified in the TV game show *Deal Or No Deal*, in which the tradeoff is between a safe option (receiving a sum of money for certain) and an opportunity to win more or less (e.g. Post, J. & Assem 2008, Bombardini & Trebbi 2012, Mulino, Scheelings, Brooks & Faff 2009, de Roos, N. & Sarafidis 2010, Deck, Lee & Reyes 2008, Botti & Conte 2008, Blavatskyy & Pogrebna 2008, de Palma et al. 2008, Andersen, Harrison, Lau & Rutström 2013). We avoid the danger that risk aversion will lead individuals to misrepresent their probability perceptions (Andersen, Harrison, Lau & Rutström 2010) by separating their choice behavior from their stated probabilities, and only incentivizing the former.

A central contribution of this paper is the estimation of a probability weighting function that maps the true probability, unobserved by the subject, against a perceived probability that is self-reported. Our surprising finding is that despite the significant complexity of the scenario and the absence of direct information about actual probabilities, the S-shaped probability weighting function discussed above fits the data remarkably well.

The main motivation for our approach is that in most real-life choice situations, it is unrealistic to assume that individuals have access to the true probability of a good outcome. Whether it concerns corporate profits, movements in the price of goods, the arrival rate of potential marriage partners, the health of children and family, or the advent of a competitor, we have to guess the probabilities of good things eventuating based on limited information and with no credible outside source of truth. If even macroeconomists cannot accurately forecast inflation, and health economists cannot say with certainty how much disease reduction the extra dollar of health care buys, how can lesser-educated, 'ordinary' economic agents arrive at an estimate other than very imperfectly? Even more importantly, are individual guesses about the probability of complex events remotely accurate and, if not, are they systematically wrong? By designing the choice situation to be complicated and cognitively taxing—but with the truth known to the researcher—we are able to explore the link between actual and subjective probabilities in a setting where the complexity level and the agent's access to information are better matched to reality.

2 Theoretical framework

We begin by writing down estimable functions that link subjective probability with objective probability in a dichotomous choice setting. Being agnostic about which of the many functions proposed in the literature best captures real behavior, we consider four of the most prominent specifications of how p_{it}^s, the subjective probability of a good outcome stated by individual i at time t, relates to the objective probability of a good outcome, p_{it}.

Our first model uses the same set-up as in Lattimore, Baker & Witte (1992) and Gonzalez & Wu (1999):

$$p_{it}^s = \frac{\alpha p_{it}^\gamma}{\alpha p_{it}^\gamma + (1 - p_{it})^\gamma} + v_{it} \tag{1}$$

Model (1), alluded to in the introduction and widely discussed in the literature, is a log-odds linear representation of the relation between p_{it} and p_{it}^s. Here, the parameter $\gamma < 1$ measures the change in sensitivity of the subjective probability to changes in the objective probability as the latter increases, relating to the degree of curvature in the relationship between subjective and objective probability. α is primarily responsible for the curve's elevation, and measures the relative level of optimism (Bruhin, Fehr-Duda & Epper 2010). The weighting function becomes more elevated as α increases and more curved as γ decreases (Trepel et al. 2005). The error term v_{it} is standard normally distributed throughout, with unknown variance σ.

Our second model is taken from Wu & Gonzalez (1996):

$$p_{it}^s = \frac{p_{it}^\gamma}{[p_{it}^\gamma + (1 - p_{it})^\gamma]^\alpha} + v_{it} \tag{2}$$

Model (2) is a reduction of the Lattimore et al. (1992) model in the event that, in that prior model, $\alpha = 1$. Like Camerer & Ho (1994) and Tversky & Kahneman (1992), Wu & Gonzalez (1996) estimate only one parameter of Model (2) (γ, with $\alpha = 1/\gamma$) instead of two. In a later article (Gonzalez & Wu 1999), the authors argue strongly that curvature and elevation, although they are not perfectly represented by a single parameter in this parametric form, are two independent aspects of the function of interest and can be captured separately using these two parameters.

The third model is Prelec's (1998) compound invariant model of p_{it}^s based on a particular axiomatic representation of choices between lotteries:[1]

$$p_{it}^s = \gamma \exp\left[-\beta(-\ln p_{it})^\alpha\right] + v_{it} \tag{3}$$

[1]The key preference axiom related to the compound invariant model is that, using

Model (3) essentially allows for much flatter S-shapes, and breaks away from the assumption that when p_{it} approaches 1, perceived and real probabilities must be the same (note in this function that when $p_{it} \downarrow 0$ then $p^s_{it} = 0$, but when $p_{it} = 1$, $p^s_{it} = \gamma$). The parameter β, as de Palma et al. (2008) point out, reflects pessimism, whereas α reflects how pronounced the S-shape is (p. 278). Prelec describes the α parameter as an Allais Paradox Index (pp. 78-79), providing information about individuals' propensities to generate expected utility violations: 'If one person has a weighting function that intersects another person's function from above, then that person will commit more Allais common-ratio violations and more Allais common-consequence violations' (p. 79). According to some neuro-scientific research, α is linked to the anterior cingulate cortex of the brain, being correlated with greater activation in this area (Paulus & Frank 2006).

The fourth model we estimate is Prelec's (1998) conditional invariant model of p^s_{it}:[2]

$$p^s_{it} = \gamma \exp\left[-\frac{\beta}{\eta}(1 - p^{\eta}_{it})\right] + v_{it} \tag{4}$$

Although the overall properties of Model (4) are similar to those of the compound invariant model near the extremities of the distribution, Model (4) allows for a slightly different shape. According to Prelec (2000), Models (3) and (4) offer two distinct advantages: they rationalize different classes of expected utility violations simultaneously, and they are tractable.

In the empirical estimation, we first seek the model that best explains our subjective probability data, and we then proceed to estimate the determinants of the deep parameters of that model. We also note that for the top three of these models it holds that rationality (i.e., the classic expected utility framework under which $p^s_{it} = p_{it}$ along the entire range of p_{it}) would require that the deep parameters are all equal to unity—i.e., that $\alpha = \beta = \gamma = 1$. The fourth model does not allow for such rationality within the possible parameter space, but does impose rationality at the boundaries ($p^s_{it} = 0$ when $p_{it} = 0$ and $p^s_{it} = 1$ when $p_{it} = 1$) when $\gamma = 1$.

(x_k, q_k) to denote a lottery in which the actual outcome x_k eventuates with probability q_k, it must hold that if $(x_1, q_1) \sim (x_2, q_2)$ and $(x_1, q_3) \sim (x_2, q_4)$ then $(x_3, (q_1)^M) \sim (x_4, (q_2)^M)$ implies $(x_3, (q_3)^M) \sim (x_4, (q_4)^M)$ with M$\geq$1.

[2] The key axiom related to the conditional invariant model is that, with $0 < \lambda < 1$, it would have to hold that if $(x_1, q_1) \sim (x_2, q_2)$ and $(x_1, q_3) \sim (x_2, q_4)$, then $(x_3, \lambda q_1) \sim (x_4, \lambda q_2)$ implies $(x_3, \lambda q_3) \sim (x_4, \lambda q_4)$.

3 Experimental design

To assess expectation formation during the unfolding of an uncertain event whose outcome matters, we use a novel experimental design whose scientific motivation and relevant features are outlined here. Further details are available in the Appendix and upon request from the authors.

3.1 Motivation

Our context is intended to mimic a real-life situation, such as gambling or stock trading, in which the agent may observe initial information about statistical likelihoods or even provide input into the outcome generation process, but cannot perfectly anticipate a shock component present in the process before the outcome occurs. The parameters of non-expected utility models have been empirically estimated in previous studies of such real-life situations, with early research frequently using betting markets to explore the subjective and estimated objective winning probabilities (Griffith 1949, McGlothlin 1956, Weitzman 1965, Ali 1977). Ali (1977), for example, in a study of 20,247 harness horse races, shows that the winning probability for a horse with a low objective probability of winning is overstated, whereas the winning probability for a horse with a high objective probability of winning is understated. Jullien & Salanié (2000), using U.K. data on 34,443 horse races, also compute econometric estimates of non-expected utility models. They find evidence that representations based on cumulative prospect theory perform better than those derived from expected utility, which perform as well or worse than those based on rank-dependent expected utility. Among the cumulative prospect theory models, they find that the Prelec (1998) models fit the data best. This implies that expected utility should be rejected in favor of Prelec's alternatives even though, in Jullien & Salanié (2000), the estimated weighting function was close to the diagonal, implying only limited economic significance of the deviation from rationality.

Yet, when looking at such real markets, scholars also find other considerations to be of likely importance. For example, Golec & Tamarkin (1998) emphasize that bettors are risk averse, where in terms of the actual distribution of expected winnings, they trade off both variance and negative expected return in favor of positive skewness. This is exactly counter to what one would expect if bettors are using S-shaped weighting functions, because it implies that bettors willingly take very small risks of huge losses in order to gain higher probabilities of winning. Yet it is exactly what one would expect once one takes into account the possibility of bankruptcy, which provides a real-life floor beneath which outcomes cannot go. Another disadvantage of most real markets is that there is no unambiguous 'true probability' to which behaviour can be related, as the statistician only has access to realized events and not to the underlying true probabilities of these events. The results and limitations of earlier studies thus suggest the use of real-life experimental designs implemented in the lab, which have the advantages of conferring control over the risky data generating process, offering a richer set of variables capturing individual characteristics, and ensuring the absence of other potentially confounding factors, such as bankruptcy laws.

3.2 Design

In our experiment, participants were confronted with six animated race car games, where each participant's final payoff was linked to the outcome of one particular car.[3] In each game, the participant chose how to divide a fixed amount of money—earned previously in a real-effort task consisting of cross-sum calculations—into a wagered amount and an invested amount, where greater investment increased the chances that the participant's car would win (see Figure A.1 in the Appendix). Each race lasted for 10 laps and involved five cars in total (see Figure A.2), each performing according to a statistical process comprised of a random-walk component and an exogenous downward shifter, explained to participants as a temporary engine failure. The frequency of the incidence of this exogenous downward shock was reduced for the participant's car in proportion to the amount invested, explained to participants as investment into engine quality. Participants were offered the choice to withdraw from their bet at each of three pit stops (after laps 3, 6, and 9; see Figure A.2). If they elected to drop out at one of these points, then they would retain a fraction of the amount originally wagered. This fraction, respectively at each successive pit stop, was 40%, 25%, or 10% of the amount originally wagered. Upon a decision to drop out, the participant would see the race continue to completion on the screen, but would no longer have a monetary stake in its outcome. Participants' expectations about the likelihood of their car winning were elicited at the beginning of the race, and also at each of the three pit stops, by asking them how many times out of 1,000 they thought their car would win if the race were to continue 1,000 times from that point. The outcome of one of the six races, chosen randomly, was paid out in cash at the conclusion of the

[3]Each participant faced his own game involving competing cars controlled by the computer, with no interactions between the races or with the games of different participants. The experiment was programmed in an early version of CORAL (Schaffner 2013).

experiment.

To obtain data across different risk and endowment levels, we implemented four distinct treatments, each of which affected the way in which payoffs were structured. In the baseline treatment, participants received a $5 show-up fee plus five times the amount wagered in the event that their car won the race that was selected for payout. In the 'wealth' treatment, the payoffs from the races stayed the same, but the show-up fee was increased to $20. In the 'high-stakes' treatment, the show-up fee was again $5 but participants won 15 times the amount wagered if their car won the race that was selected for payout. Finally, in the 'low-stakes' treatment, participants received twice the bet if their car won, the exact amount wagered if their car came second, and half the wagered amount if their car came third. For the purposes of this paper, we simply code these treatments as separate parameters.

In addition to the information about expectations and risk-taking provided in the race stage of the experiment, we asked participants to respond to several batteries of questions on their psychology and beliefs, and also collected standard demographic information. After the experiment had concluded, we used repetitions of the data-generating process to simulate the actual likelihood for each race of the participant's car winning the race from the point of each pit stop, providing an objective picture of the future outcome against which we could compare participants' subjective expectations. In our context, as in many real-life scenarios, the true (mathematical) probability of winning changed as the race went on, because of changes in the observed positions of the cars. We therefore expected that subjective probabilities too would adjust as this new information became available, and the revelation of whether or not the participant's car would win grew ever closer.

At the start of the race, we informed each participant of the overall odds of his car winning the race, conditional on his choice of how much of his endowment to invest, and how much to wager.[4] Crucial to our experiment however, and unlike the literature we build on, we did not tell participants the true (updated) mathematical expectation of winning when they arrived at the pit stops. We asked participants for their subjective probabilities sequentially as this complex situation unfolded, rather than backing them out of lottery choices at a point in time.

[4]The total endowment available to each participant for investing or betting at the start of each race was earned by that participant in a real-effort stage prior to the race stage of the experiment; see Appendix for details.

4 Simple statistics and model fit

A total of 239 participants took part in eight experimental sessions, all recruited using using ORSEE (Greiner 2004) via standard emails from the experimental subject pool at the ASBLab at the Australian School of Business within the University of New South Wales. The average participant age was 22 years, and 45.15% of participants were female. The average earnings in the real effort task, which as described above could then be split into a wagered and an invested amount, was $24.42. The average bet was $7.23, and there was a steep winning curve: payoffs were highly volatile, ranging from $5 to $105.20, with an average of $23.62 across all four treatments. Full sample sizes by treatment, calculated at the levels of participant, participant-by-race, and participant-by-pit-stop are given in Table 1. In ensuing regression tables, these sample sizes fluctuate somewhat because of incomplete participant data on certain explanatory variables.

Table 1: Summary statistics

	Participants	Participant-Races	Participant-Pitstops
Sample sizes:			
Baseline Treatment	58	348	1044
Wealth Treatment	59	354	1062
High-Stakes Treatment	61	365	1095
Low-Stakes Treatment	61	366	1098
TOTAL	239	1433	4299
Key variables:			
Pitstop expectations of winning	**5.21**	**5.21**	**5.21**
(standardized: times out of 10)	(.15)	(.08)	(.05)
Simulated chances of winning	**4.33**	**4.33**	**4.33**
(standardized: times out of 10)	(.13)	(.08)	(.05)
Ever dropped out of bet?	**50.63%**	**18.49%**	**14.63%**

Note: The top section of this table shows sample sizes, and the bottom section shows means of the key analysis variables, at each level of analysis. Expectations of winning, simulated chances of winning, and dropout behavior are all measured at the participant-pit-stop level. The bottom section of the table also shows the standard error of each mean.

The main thing to take from this descriptive table is the presence of aggregate over-optimism: participants report to believe their car will win 521 out of 1000 continuations, whereas the true number, found via large-sample simulations, is 433 out of 1000. Subjective probabilities are thus around 9 percentage points higher than objective ones on average, and this difference is statistically significant, given the small size of the associated standard errors.

Next, we look at the evolution of the raw relationship between the subjective probability and the actual probability for each of the 6 races individuals faced, with a special focus on whether or not the match between subjective and objective probabilities improves as individuals learn more about the race. In the kernel plots shown in Figure 1, we see that for all six races a very similar S-shaped curve emerges, and casual visual inspection does not indicate an obvious improvement in the fit as we move from Race 1 to Race 6. Still, the regression slope is significantly steeper for the later races (4-6) than for the earlier ones (1-3). A particular feature evident in the plots for each race is that the subjective probability does not clearly converge to 1 even when the real probability is very close to 1. Also, subjective probability remains above the real probability until the real probability reaches approximately 60%.

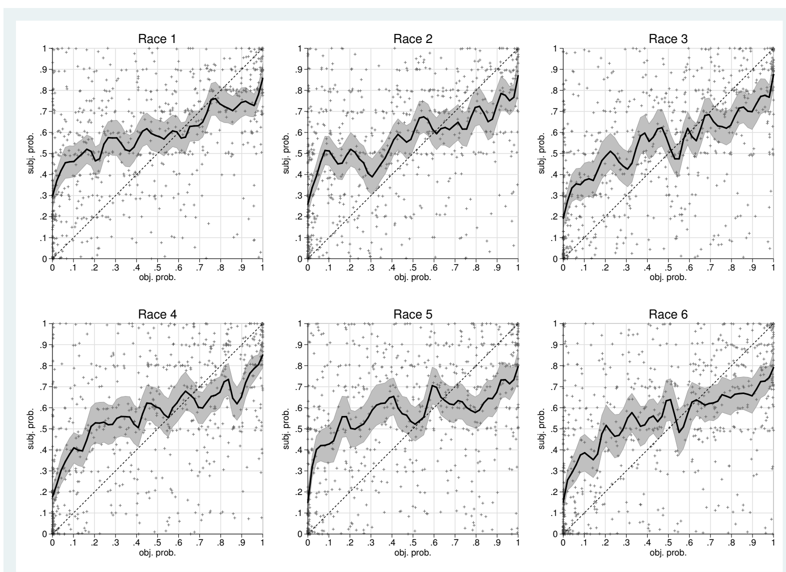

Figure 1: Kernel Plots by Race

4.1 Is there information in stated probabilities?

As discussed above, participants were given the choice at each pit stop to either withdraw from the race and receive a certain percentage of the bet, or let the game proceed further. The inclusion of this behavioral choice, presented at the same time that we elicit subjective probabilities, enables us to answer the preliminary question of whether stated probabilities contain information about choice behavior. Because the decision of whether to drop out involves the utility value of entertainment and excitement, we should expect choice behavior to relate not solely to subjective probability, but also to individual heterogeneity in entertainment value and a complicated option value of continuing.[5] Nevertheless, by examining whether subjective probabilities help explain choices, we can determine whether they contain any choice-relevant information over and above the effects of real probabilities.

If we concentrate on the decision made at the last pit stop—when there are no future pit stops to anticipate—we can run more detailed specifications without being worried about the complication of the option value of future dropout opportunities. The optimal strategy in this situation, for a risk-neutral individual in any treatment except 'low stakes' who is interested only in monetary reward, is to drop out of the race if the probability of winning is below $\frac{10}{f}$ where f denotes the factor by which the wagered amount is multiplied if the participant's car wins (5 or 15 times, depending on the treatment) and the numerator captures the 10% fraction of the wagered amount that is paid out when dropping out at the final pit stop.

Table 2: Probit coefficient estimates from the prediction of dropping out of a race

	All pitstops	**Final pitstop**					
Objective prob.	-2.784***	-2.352***	-3.246***				
	(0.330)	(0.615)	(0.856)				
Subjective prob.	-1.345***	-1.571***		-2.286***			
	(0.218)	(0.398)		(0.357)			
Obj. cutoff					1.412***	1.495***	
					(0.166)	(0.168)	
Subj. cutoff					0.642**		1.214***
					(0.270)		(0.286)
Chi^2	143.291	25.321	14.374	41.093	80.074	79.140	17.970
$Pr() > Chi^2$	0.000	0.000	0.000	0.000	0.000	0.000	0.000
N	3935	907	907	907	907	907	907
ll	-670.167	-123.596	-134.415	-139.646	-131.807	-134.263	-168.524

Significance levels: * 0.1; ** 0.05; *** 0.01. Standard errors in parentheses.

The first two columns of Table 2, where real and subjective stated probabilities are used to predict dropout behavior either at all the pit stops (Column 1) or only at the final pit stop (Column 2), clearly illustrate that both real and stated probabilities affect the choice of whether to drop out.

[5] For the first two of three pit stops, the option of continuing includes the possibility of dropping out later.

Compared to specifications with only one of the two types of probabilities included (shown in Columns 3 and 4), the specification that includes both has a superior fit to the data both in terms of raw log likelihood and in terms of the standard information criteria (Bayesian, Aikake, and average likelihood ($\frac{-log(ll)}{N}$)).

Columns 5 to 7 then show the estimated effect on dropout behavior of the objective or stated probability lying below the 10% cutoff. Comparing these results to the results in prior columns of Table 2 reveals that the contribution of subjective probabilities in explaining behavior is not fully captured in the cutoffs alone: the log likelihood associated with the model using only the cutoff dummies is substantially lower than that associated with the model that includes the continuous probabilities. This finding emphasizes that individuals do not make their choices purely on the basis of maximizing expected monetary returns.

In the first two columns of Table 2, where we include both the actual and stated probabilities, we find that the true probability's estimated coefficient is 50% to 100% higher than the subjective stated probability for the latent variable related to the decision to drop out, and both variables are highly significant in the equation. This intriguing finding indicates that participants use information reflected in the true probability that is not included in their stated probability. We do not explore this further here because there are many candidate explanations for it that fall outside our focus in this paper, including subliminal excitement due to unconscious awareness of additional information, and non-linearities in the decision making process.

5 Reduced form and structural analysis

5.1 Reduced form

The conditional mean $E[p^s|p, X]$ of subjective probability is informative in its own right, in that under perfect rationality $E[p^s|p, X] = p$, meaning that deviations from this value throw light on the reduced-form divergence between stated beliefs and what the rationality assumption implies that beliefs 'should' be. In Table 3, we show these conditional means in the form of simple OLS regressions exploring the relationship between subjective probabilities and objective probabilities, with an increasing array of characteristics (X variables) controlled.[6]

[6] As detailed in the Appendix, our Introductory and Follow-up Questionnaires collected a vast array of information about participants. We selected the particular control variables

Table 3: OLS regression results from predicting stated (subjective) probability

	(1)		(2)		(3)		(4)		Heteroskedasticity Model			
									Mean-shifter (β)		Variance-shifter (σ)	
Objective prob.	0.488***	(0.026)	0.429***	(0.023)	0.431***	(0.023)	0.424***	(0.022)	0.428***	(0.012)	-0.064***	(0.010)
Pitstop	-0.009**	(0.004)	-0.009**	(0.004)	-0.009**	(0.004)	-0.009**	(0.004)	-0.006	(0.004)	0.000	(0.003)
Race Number (1-6)	-0.012***	(0.003)	-0.011***	(0.003)	-0.011***	(0.003)	-0.011***	(0.003)	-0.013***	(0.002)	0.002	(0.001)
Amount wagered			-0.011***	(0.003)	-0.011***	(0.003)	-0.011***	(0.003)	-0.013***	(0.001)	-0.005***	(0.001)
Real-Effort Earnings			-0.001	(0.002)	-0.002	(0.002)	-0.001	(0.002)	0.002***	(0.001)	0.000	(0.001)
Amount to be won			0.001	(0.001)	0.001	(0.001)	0.001	(0.001)	0.001***	(0.000)	0.000***	(0.000)
High Treatment			-0.050	(0.046)	-0.076*	(0.045)	-0.064	(0.042)	-0.082***	(0.014)	-0.059***	(0.012)
Low Treatment			-0.025	(0.039)	-0.040	(0.039)	-0.032	(0.039)	-0.025*	(0.013)	0.010	(0.011)
Wealth Treatment			-0.047	(0.037)	-0.072*	(0.039)	-0.064*	(0.034)	-0.069***	(0.011)	-0.041***	(0.010)
Guess the winner			-0.037**	(0.017)	-0.036**	(0.017)	-0.033*	(0.018)	-0.032***	(0.005)	0.006	(0.004)
Experiment experience			-0.001	(0.026)	-0.007	(0.027)	-0.016	(0.029)	-0.030***	(0.011)	0.004	(0.009)
Gender (female=1)					0.001	(0.027)	-0.010	(0.026)	-0.016*	(0.009)	0.027***	(0.007)
Age					-0.037**	(0.017)	-0.048***	(0.014)	-0.046***	(0.007)	-0.009	(0.006)
Age^2					0.001**	(0.000)	0.001***	(0.000)	0.001***	(0.000)	0.000	(0.000)
Asian culture					0.069*	(0.041)	0.053	(0.041)	0.038***	(0.014)	-0.013	(0.011)
Other culture					0.096*	(0.057)	0.084*	(0.047)	0.064***	(0.023)	-0.023	(0.019)
Weekly Income Low					-0.048	(0.036)	-0.061*	(0.036)	-0.059***	(0.012)	0.041***	(0.010)
Weekly Income Avg.					0.068*	(0.036)	0.058*	(0.035)	0.038***	(0.010)	0.017**	(0.009)
Weekly Income High					-0.045	(0.034)	-0.042	(0.032)	-0.043***	(0.013)	-0.008	(0.012)
Wealth Level Avg.					-0.053	(0.048)	-0.072	(0.044)	-0.100***	(0.018)	0.025	(0.016)
Wealth Level Poor					-0.020	(0.053)	-0.029	(0.050)	-0.046**	(0.019)	0.043**	(0.017)
Performance at Uni					0.037**	(0.014)	0.023	(0.014)	0.020***	(0.005)	-0.003	(0.004)
International Student					-0.002	(0.031)	0.007	(0.029)	0.014	(0.010)	0.058***	(0.008)
English speaker					0.019	(0.030)	0.003	(0.030)	-0.023**	(0.010)	0.032***	(0.008)
Mum Schooled					-0.017	(0.040)	-0.075*	(0.040)	-0.094***	(0.014)	-0.015	(0.012)
Mum Qualified					-0.044	(0.037)	-0.028	(0.032)	-0.004	(0.012)	-0.026**	(0.010)
Mum Qual. Level					0.002	(0.027)	0.021	(0.025)	0.026***	(0.008)	0.022***	(0.007)
Dad Schooled					0.087*	(0.048)	0.101**	(0.047)	0.091***	(0.016)	-0.001	(0.014)
Dad Qualified					0.065*	(0.039)	0.063*	(0.037)	0.063***	(0.013)	-0.019*	(0.011)
Dad Qual. Level					-0.032	(0.024)	-0.041*	(0.023)	-0.046***	(0.008)	-0.012**	(0.006)
Risk aversion (HL)							0.009*	(0.005)	0.007***	(0.002)	-0.008***	(0.002)
SBI: Reminisce (I)							0.025**	(0.012)	0.026***	(0.004)	0.003	(0.003)
SBI: Anticipate (I)							-0.022*	(0.012)	-0.023***	(0.004)	-0.010***	(0.003)
SBI: Moment (I)							-0.009	(0.016)	-0.005	(0.005)	0.006	(0.004)
Optimism: Disapp.							0.008	(0.005)	0.005***	(0.002)	0.000	(0.002)
Optimism: Low Exp.							-0.007	(0.006)	-0.006***	(0.002)	0.000	(0.002)
Self Esteem (I)							0.000	(0.015)	-0.010*	(0.005)	-0.008*	(0.005)
Locus of Control (I)							0.004	(0.010)	0.008***	(0.003)	-0.007***	(0.002)
Happiness							0.015	(0.014)	0.020***	(0.005)	-0.017***	(0.005)
Lucky Charm							0.017	(0.011)	0.025***	(0.003)	-0.004	(0.003)
BMI							0.010***	(0.002)	0.011***	(0.001)	0.001	(0.001)
Lefthanded							0.076	(0.055)	0.129***	(0.021)	0.022	(0.015)
Vegetarian							0.115*	(0.060)	0.125***	(0.020)	-0.018	(0.021)
Alcohol							-0.042	(0.027)	-0.041***	(0.009)	0.035***	(0.008)
Smoking							-0.042*	(0.023)	-0.058***	(0.012)	-0.031***	(0.011)
Depression							-0.087**	(0.034)	-0.058***	(0.012)	-0.065***	(0.011)
F	117.902		47.391		24.502		32.991					
$Pr() > F$	0.000		0.000		0.000		0.000		0.000			
N	4194		4194		4194		4194		4194			
$Adj.R^2$	0.298		0.327		0.365		0.426					
ll	-341.45		-250.95		-118.01		100.92		325.94			

Significance levels: * 0.1; ** 0.05; *** 0.01. Standard errors in parentheses. (I) denotes indexes built from sets of variables. See Appendix for further details on all control variables. The excluded reference categories are “Baseline Treatment”, “Australian Culture”, “Weekly Income None”, and “Wealth Level Above”.

The results in the first four columns of Table 3 show that the coefficient on p in the regression predicting p^s hovers around 0.4 to 0.5, with a standard deviation of 0.02 to 0.03, implying that the coefficient is very significantly smaller than 1. This finding is consistent with a systematic deviation from rationality. It suggests that on balance, the net behavioral effect of increasing an event's probability when that probability lies in the middle range—a real world example would be the odds of one of two major political parties winning the next election—will be much smaller than proportional to the actual change.

The estimated effects of our individual control variables on subjective probability can be interpreted as capturing the estimated influence of individual optimism, measured in different ways. Controlling for an array of potential sources of individual optimism is particularly desirable in our context because optimism (about a best outcome) or pessimism (about a worst outcome) influences risk-related behavior (Abdellaoui, L'Haridon & Zank 2010). Our findings are generally in line with earlier research findings indicating that wealthier and healthier respondents tend to be more optimistic. We find more optimism among participants who are younger, have higher weekly incomes, and/or are nonsmokers or vegetarians.

In weighing alternatives to the rationality assumption, it is interesting to know not merely the conditional mean $E[p^s|p, X]$, but also the standard deviation of $[p^s - E[p^s|p, X]|$, which can serve as a measure of how tightly the subjective probabilities are grouped around their conditional mean. The smaller the standard deviation, the more 'regular' the production of subjective probability. With this in mind, we estimate next a slightly expanded reduced-form model that allows for both heteroskedasticity and mean-level shifters. The main result of this heteroskedasticity analysis, shown in the final two columns of Table 3, is that the effect of race number (1 through 6) on the standard deviation has a coefficient of 0.002 ($sd = 0.001$), which means that the standard deviation of the subjective probabilities does not fall with more races, indicating an absence of learning. Indeed, the point estimate of this effect is positive, although it is insignificant at all conventional levels. On the other hand, the coefficient of race number on the mean subjectivity probability is -0.013 ($sd = 0.002$) and significant at the 1 percent level, meaning that the average stated probability goes down over time, drawing closer to the average true probability. Hence, while there is

to include in the models shown in Table 3 and Table 6 based on a combination of goodness-of-fit and coverage of the dimensions of the questionnaire. Adding additional power terms of the objective probability to capture the non-linear relationship improves the model fit only marginally (generating a 2-3% increase in adjusted R^2).

no learning in terms of a lowered standard deviation, the mean prediction error does reduce with time.

5.2 Initial structural analysis

Although the conditional means of subjective probability are useful in reduced form for comparisons with what the rationality paradigm implies, they cannot illuminate the deep structure of subjective probability formation. To identify which proposed structure of the subjective probability distribution fits our data best, we now horse-race the theoretically-grounded subjective probability functions introduced above by fitting each of our four models to the data, using a simple maximum likelihood approach.

The results of these analyses are shown in Table 4, where models and parameters are identified by the model number and a shorthand abbreviation based on the names of the authors who proposed the model and—for the Prelec models—the model type. For Model (1) (the Lattimore et al. (1992) model), the log likelihood is -319.676 and the estimates for α and γ are both significantly different from 1, implying a strong aggregate S-shape that deviates from rationality. For Model (2) (the Wu & Gonzalez (1996) model), although the fit is slightly superior to that of Model (1) (with a log likelihood of -316.616), the structural results are somewhat similar: the estimate for γ is 0.20 and significantly smaller than 1. Rationality is also violated in both the compound invariant and conditional invariant models (Prelec 1998), with all coefficient estimates significantly below 1, and with the stated probability in Model (4) only about three-quarters ($\gamma = 0.74$) of the real probability when the real probability is at the limit of 1. The log likelihoods of Models (3) and (4) are -260.059 and -249.176, respectively.

We next statistically compare the results for each model using the three most commonly used criteria: Akaike's information criterion (AIC), the Bayesian information criterion (BIC), and the average log likelihood. As Table 5 shows, in all cases, Prelec's (1998) conditional invariant model gives a superior fit by a large margin. Under particular conditions, the AIC is chi-square distributed, making the differences between models very significant not only in terms of sheer likelihood ratios, but also at the conventional 1% and even 0.01% levels of statistical significance. We therefore proceed in the next section to parameterize this conditional invariant model.[7]

[7]Since the models are non-nested and have different basic structures, there is no theoretically clear way to parameterize all four of them and then compare results. Indeed, for all these models, full parameterization engenders convergence issues because of the high degree of collinearity and the associated problems of a flat likelihood. Nevertheless, if

Table 4: ML estimation of the functional form of subjective probabilities

	Model (1) [**LBW**]	**Model (2)** [**WG**]	**Model (3)** [**Pcomp**]	**Model (4)** [**Pcond**]
$\sigma(v_{it})$	0.261	0.260	0.257	0.256
	(0.003)	(0.003)	(0.003)	(0.003)
γ_{LBW}	0.234			
	(0.008)			
α_{LBW}	1.342			
	(0.025)			
γ_{WG}		0.203		
		(0.007)		
α_{WG}		0.735		
		(0.015)		
α_{Pcomp}			0.343	
			(0.017)	
β_{Pcomp}			0.522	
			(0.023)	
γ_{Pcomp}			0.887	
			(0.017)	
η_{Pcond}				0.230
				(0.020)
β_{Pcond}				0.331
				(0.018)
γ_{Pcond}				0.741
				(0.008)
N	4299	4299	4299	4299
ll	-319.676	-316.616	-260.059	-249.176

Standard errors in parentheses. $\sigma(v_{it})$ denotes the estimated variance of the errors in the given model.

Table 5: Likelihood comparisons across the four structural models

	AIC	BIC	-ln(L)/N
LBW (1992)	645.35	664.45	0.074
WG (1996)	639.23	658.33	0.074
Prelec (1998), Comp Inv	528.12	553.58	0.060
Prelec (1998), Cond Inv	508.35	540.18	0.058

AIC=2*k-2*ln(L); BIC=-2*ln(L)+k*ln(N). L=likelihood, N=number of observations, and k=number of estimated parameters.

As is suggested by the kernel plots of the subjective probabilities against the objective probabilities in our data (Figure 1), the main reason for the superior fit of Prelec's alternatives compared to Models (1) and (2) is that both the compound and conditional invariant models allow the switching point at which the 45-degree line is crossed to be at any level of the real probability. In this case, visual inspection reveals that the switching point lies above $p = 0.5$ in our data, while it is estimated at closer to $p = 0.3$ when we fit Models (1) and (2) to the data. Hence, the main data characteristic responsible for the superior fit of Models (3) and (4) is that the subjective probability associated with the true probability of one half is far higher than one half. It is in fact closer to 0.7.

5.3 Extended structural analysis

Because Prelec's (1998) conditional invariant model delivers the best overall fit, we now parameterize its main components: σ, which captures the heteroskedasticity of the error term; γ, which can be interpreted as the subjective probability corresponding to objective certainty; and β, which corresponds to the sensitivity of the subjective probability to low values of p (i.e.,the higher is β, the steeper is the slope of p^s for low values of p, and the higher the value of p at which the subjective and actual probabilities cross). We perform this exercise based on the following extended structural model, where the δ's can be interpreted either as capturing the aggregate level of the associated deep structural parameter, or simply as best-fit coefficients capturing the influence on deep parameters of several variables at once: [8]

$$
\begin{aligned}
p^s &= \gamma \exp\left[-\frac{\beta}{\eta}(1 - p^\eta)\right] + v \\
\gamma &= x_{it}\delta_0 \\
v &\sim N(0, \sigma^2) = N(0, x_{it}\delta_1) \\
\beta &= x_{it}\delta_2
\end{aligned}
$$

we limit ourselves to structurally recovering only two parameters from the data for each model (including the variance of the error term), Prelec's (1998) conditional invariant model consistently delivers the best fit across all selections of parameter pairs (using the same variables from the data).

[8]We tried modelling all parameters of this model as functions of our data, but the collinearity in the parameters is too great to allow for this, meaning there is not enough variation in the data—rich as it is—to tease apart the determinants of all of the parameters separately.

The results of estimating the parameters of this extended structural model using our data are given in Table 6.

Table 6: Structural ML estimates of probability function parameters

	(1)		**(2)**		**(3)**		**(4)**	
σ^2								
Pitstop	-0.004	(0.003)	-0.003	(0.003)	-0.001	(0.003)	-0.003	(0.003)
Race Number (1-6)	0.004***	(0.002)	0.004**	(0.002)	0.003*	(0.002)	0.001	(0.002)
Amount wagered			-0.003***	(0.001)	-0.003***	(0.001)	-0.004***	(0.001)
Real-Effort Earnings			-0.002***	(0.001)	-0.001*	(0.001)	0.000	(0.000)
Amount to be won			0.000***	(0.000)	0.000*	(0.000)	0.000***	(0.000)
High Treatment			-0.024**	(0.010)	-0.029**	(0.012)	-0.041***	(0.011)
Low Treatment			-0.004	(0.009)	0.005	(0.010)	0.029***	(0.010)
Wealth Treatment			-0.006	(0.009)	-0.004	(0.010)	0.000	(0.009)
Guess the winner			0.008**	(0.004)	0.003	(0.004)	0.004	(0.004)
Experiment experience			0.029***	(0.007)	0.024***	(0.008)	0.005	(0.009)
Gender (female=1)					0.050***	(0.007)	0.051***	(0.007)
Age					-0.019***	(0.007)	-0.033***	(0.007)
Age^2					0.000**	(0.000)	0.001***	(0.000)
Asian culture					0.011	(0.012)	0.021**	(0.010)
Other culture					0.038*	(0.023)	0.032	(0.022)
Weekly Income Low					0.042***	(0.010)	0.055***	(0.010)
Weekly Income Avg.					0.015*	(0.009)	0.001	(0.009)
Weekly Income High					-0.032***	(0.012)	-0.010	(0.011)
Wealth Level Avg.					-0.002	(0.014)	0.017	(0.014)
Wealth Level Poor					-0.004	(0.015)	0.017	(0.014)
Performance at Uni					-0.004	(0.003)	-0.007*	(0.003)
International Student					0.021**	(0.008)	0.039***	(0.009)
English speaker					0.016**	(0.007)	0.001	(0.007)
Mum Schooled					0.032***	(0.012)	0.050***	(0.014)
Mum Qualified					-0.022**	(0.011)	-0.024**	(0.011)
Mum Qual. Level					0.001	(0.007)	-0.012*	(0.007)
Dad Schooled					0.007	(0.012)	-0.005	(0.013)
Dad Qualified					-0.019*	(0.010)	-0.020*	(0.011)
Dad Qual. Level					-0.002	(0.007)	-0.005	(0.006)
Risk-aversion (HL)							0.004**	(0.002)
SBI: Reminisce (I)							-0.009***	(0.003)
SBI: Anticipate (I)							-0.004	(0.004)
SBI: Moment (I)							0.007**	(0.004)
Optimism: Disapp.							0.002	(0.001)
Optimism: Low Exp.							0.006***	(0.002)
Self Esteem (I)							-0.017***	(0.004)
Locus of Control (I)							-0.008***	(0.002)
Happiness							-0.021***	(0.005)
Lucky Charm							-0.007**	(0.003)
constant	0.248***	(0.010)	0.264***	(0.019)	0.433***	(0.090)	0.803***	(0.088)
β								
Pitstop	0.021	(0.013)	0.026*	(0.015)	0.014	(0.011)	0.018**	(0.009)
Race Number (1-6)	0.022***	(0.006)	0.018***	(0.006)	0.016***	(0.005)	0.018***	(0.004)
Amount wagered			0.019***	(0.004)	0.013***	(0.003)	0.016***	(0.003)
Real-Effort Earnings			0.007***	(0.002)	0.012***	(0.002)	0.011***	(0.002)
Amount to be won			-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
High Treatment			0.073*	(0.037)	0.180***	(0.036)	0.170***	(0.032)
Low Treatment			0.005	(0.029)	0.105***	(0.027)	0.117***	(0.025)

Wealth Treatment			0.028	(0.028)	0.101***	(0.027)	0.147***	(0.027)
Guess the winner			0.096***	(0.016)	0.078***	(0.014)	0.081***	(0.015)
Experiment experience			0.023	(0.026)	-0.002	(0.024)	-0.036	(0.025)
Gender (female=1)					-0.004	(0.019)	-0.047**	(0.020)
Age					-0.024	(0.026)	0.002	(0.019)
Age^2					0.001	(0.001)	0.000	(0.000)
Asian culture					-0.009	(0.035)	-0.079**	(0.032)
Other culture					-0.009	(0.064)	-0.033	(0.064)
Weekly Income Low					0.069**	(0.030)	0.042	(0.028)
Weekly Income Avg.					-0.097***	(0.023)	-0.074***	(0.023)
Weekly Income High					0.131***	(0.034)	0.148***	(0.032)
Wealth Level Avg.					0.084*	(0.047)	0.132***	(0.040)
Wealth Level Poor					0.052	(0.046)	0.081**	(0.041)
Performance at Uni					0.020**	(0.010)	0.014	(0.010)
International Student					-0.045*	(0.025)	-0.022	(0.026)
English speaker					-0.047**	(0.020)	0.000	(0.021)
Mum Schooled					-0.007	(0.024)	0.006	(0.024)
Mum Qualified					0.125***	(0.027)	0.039	(0.026)
Mum Qual. Level					-0.024	(0.021)	-0.009	(0.019)
Dad Schooled					-0.172***	(0.032)	-0.220***	(0.035)
Dad Qualified					0.051*	(0.028)	0.082***	(0.030)
Dad Qual. Level					-0.003	(0.019)	-0.017	(0.019)
Risk-aversion (HL)							0.008*	(0.005)
SBI: Reminisce (I)							-0.025**	(0.010)
SBI: Anticipate (I)							0.026***	(0.009)
SBI: Moment (I)							0.010	(0.011)
Optimism: Disapp.							-0.035***	(0.004)
Optimism: Low Exp.							0.012**	(0.005)
Self Esteem (I)							-0.060***	(0.011)
Locus of Control (I)							-0.028***	(0.007)
Happiness							0.018	(0.011)
Lucky Charm							-0.010	(0.009)
constant	0.196***	(0.036)	-0.114*	(0.058)	-0.016	(0.312)	0.203	(0.242)
γ								
Pitstop	0.042***	(0.008)	0.039***	(0.009)	0.033***	(0.008)	0.034***	(0.006)
Race Number (1-6)	-0.000	(0.004)	0.000	(0.004)	0.001	(0.004)	0.008**	(0.003)
Amount wagered			0.016***	(0.004)	0.013***	(0.004)	0.013***	(0.004)
Real-Effort Earnings			-0.001	(0.001)	0.001	(0.001)	0.003***	(0.001)
Amount to be won			-0.001	(0.001)	-0.001*	(0.000)	-0.001**	(0.000)
High Treatment			0.027	(0.027)	0.050*	(0.026)	0.048*	(0.024)
Low Treatment			-0.073***	(0.024)	-0.023	(0.023)	-0.022	(0.024)
Wealth Treatment			-0.047**	(0.020)	-0.028	(0.021)	-0.009	(0.020)
Guess the winner			0.003	(0.009)	-0.012	(0.009)	-0.016*	(0.009)
Experiment experience			0.022	(0.018)	0.003	(0.018)	-0.021	(0.020)
Gender (female=1)					-0.011	(0.014)	-0.007	(0.014)
Age					-0.049***	(0.015)	-0.023	(0.015)
Age^2					0.001***	(0.000)	0.001**	(0.000)
Asian culture					0.077***	(0.022)	0.001	(0.022)
Other culture					0.113**	(0.055)	0.019	(0.051)
Weekly Income Low					-0.040**	(0.019)	-0.058***	(0.020)
Weekly Income Avg.					-0.004	(0.019)	0.013	(0.017)
Weekly Income High					0.013	(0.021)	0.010	(0.020)
Wealth Level Avg.					-0.041	(0.037)	-0.089**	(0.035)
Wealth Level Poor					-0.013	(0.038)	-0.048	(0.038)
Performance at Uni					0.060***	(0.008)	0.039***	(0.008)
International Student					-0.045**	(0.018)	-0.032*	(0.018)

English speaker					-0.002	(0.015)	-0.010	(0.015)
Mum Schooled					-0.044**	(0.022)	-0.075***	(0.022)
Mum Qualified					0.045**	(0.020)	0.019	(0.020)
Mum Qual. Level					-0.030**	(0.013)	-0.015	(0.012)
Dad Schooled					0.001	(0.024)	0.016	(0.024)
Dad Qualified					0.138***	(0.022)	0.127***	(0.022)
Dad Qual. Level					-0.030**	(0.013)	-0.031**	(0.013)
Risk aversion (HL)							0.012***	(0.004)
SBI: Reminisce (I)							0.018***	(0.006)
SBI: Anticipate (I)							-0.018**	(0.007)
SBI: Moment (I)							-0.025***	(0.008)
Optimism: Disapp.							-0.008***	(0.003)
Optimism: Low Exp.							-0.010***	(0.003)
Self Esteem (I)							-0.034***	(0.008)
Locus of Control (I)							-0.014***	(0.005)
Happiness							0.044***	(0.009)
Lucky Charm							0.024***	(0.006)
constant	0.650***	(0.024)	0.614***	(0.042)	0.959***	(0.188)	0.903***	(0.187)
η								
constant	0.218***	(0.020)	0.257***	(0.028)	0.182***	(0.023)	0.182***	(0.019)
N	4194		4194		4194		4194	
ll	-225.209		-94.361		172.661		390.983	

Significance levels: * 0.1; ** 0.05; *** 0.01. Standard errors in parentheses. (I) denotes indexes build from sets of variables. See Appendix for further details on all control variables. The excluded reference categories are "Baseline Treatment", "Australian Culture", "Weekly Income None", and "Wealth Level Above".

One set of parameters of interest in Table 6 surround the determinants of the heteroskedasticity in subjective probabilities, as these relate to whether or not subjective probability becomes tighter as people acquire more experience. In our richest specification, shown in Column 4 of Table 6, the estimated coefficient on 'race number' (i.e., more experience with the race car game) equals 0.001 and is insignificant, and far from being negative, the estimated coefficient on this variable in other columns is positive and significant, implying an absence of learning during the experiment. The variable capturing prior experience being an experimental subject is either positive or insignificant in predicting the variance of subjective probabilities, which we again take as evidence against learning effects. Higher race number is also estimated to have a positive rather than negative effect on β, the deep parameter for which a higher value indicates more curvature of subjective probabilities in relation to objective probabilities. We take this as further evidence that more experience with our race car game does not push our subjects further toward accurate predictions about the game's outcome. However, age is negative and significant in predicting the spread of subjective probabilities, even in our fullest specification, indicating that real-life experience does increase the predictability of subjective expectations.

To gain more insight into the meaning of these large numbers of coefficients, we next show the frequency distributions of these three parameters. Figure 2 shows the distribution of $\widehat{\gamma} = x_{it}\widehat{\delta_0}$, Figure 3 the distribution of $\widehat{\sigma} = \sqrt{x_{it}\widehat{\delta_1}}$, and Figure 4 the distribution of $\widehat{\beta} = x_{it}\widehat{\delta_2}$. We create these displays based on the results of our richest specification, mainly to examine the range of possible values of each parameter.

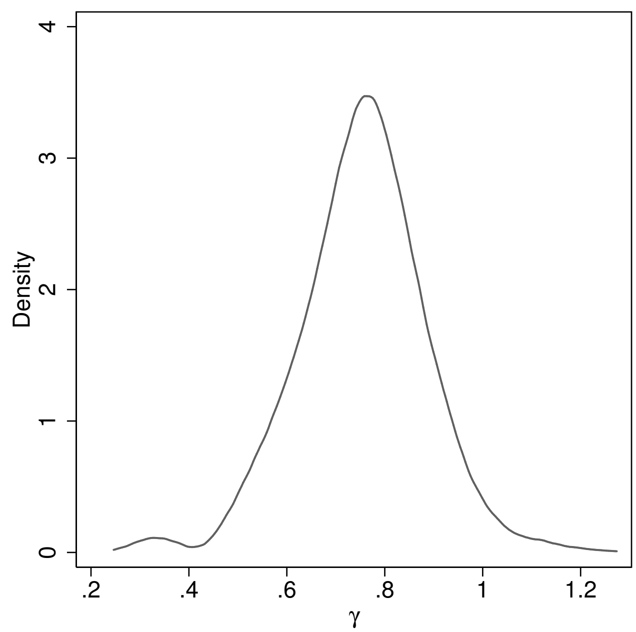

Figure 2: Distribution of $\widehat{\gamma} = x_{it}\widehat{\delta_0}$

These frequency distributions show that the range of parameters across people is very high: a small percentage of the experimental participants (about 2%) are found to have a negative $\widehat{\beta}$ which would mean their subjective probability decreases if the objective probability increases. From Table 6, we can infer the identity of this group by noting all of those characteristics associated with a lower $\widehat{\beta}$: early pit stops, first race, low amounts wagered, low real effort earnings, default treatment, female, Asian, average rather than high weekly income, and so on. It would be possible to simply constrain the estimation to only allow positive $\widehat{\beta}'$ values, but estimating the model without that restriction helps alert us to the existence of individuals and circumstances giving rise to low responsiveness to changes in actual probabilities. Of course, this particular group of participants may simply not care about the experiments, and consequently provide subjective probabilities that are close to random.

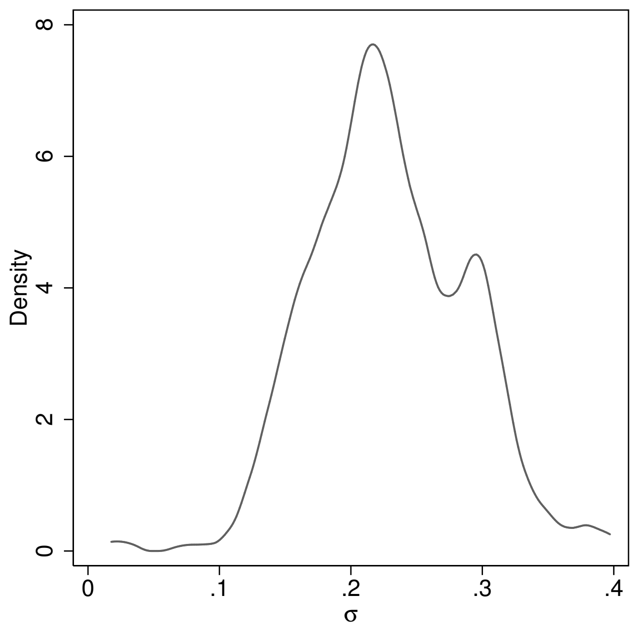

Figure 3: Distribution of $\widehat{\sigma} = \sqrt{x_{it}\widehat{\delta_1}}$

The mean of the estimated $\widehat{\beta}$s is around 0.2, implying that $\frac{\widehat{\beta}}{\widehat{\eta}}$ is close to 1 (since $\widehat{\eta} = 0.182$) which in turn implies that as p approaches zero, p^s approaches $\widehat{\gamma}e^{-1}$ which is roughly 0.25 at the mean of the $\widehat{\gamma}$s. At the other end of the range of objective probabilities, when p approaches one, p^s approaches $\widehat{\gamma}$, which equals roughly 0.75 at the mean of the $\widehat{\gamma}$s. Hence, for the mean values of these coefficients, average subjective probability varies only between 0.25 and 0.75 as p varies over its entire range, from 0 to 1.

Similarly, looking at Figure 2, the distribution of $\widehat{\gamma}$ exhibits some mass near $\widehat{\gamma} = 1$, indicating that there are some individuals for whom p^s approaches one as p approaches one, though for nearly the entire sample underestimation of high real probabilities is the norm. We can also note that there is in fact a strong correlation between $\widehat{\gamma}$ and $\widehat{\beta}$ (roughly 0.8) that indicates that the same circumstances and people generating responses closer to 'boundary rationality' (high $\widehat{\gamma}$) are also in general those circumstances

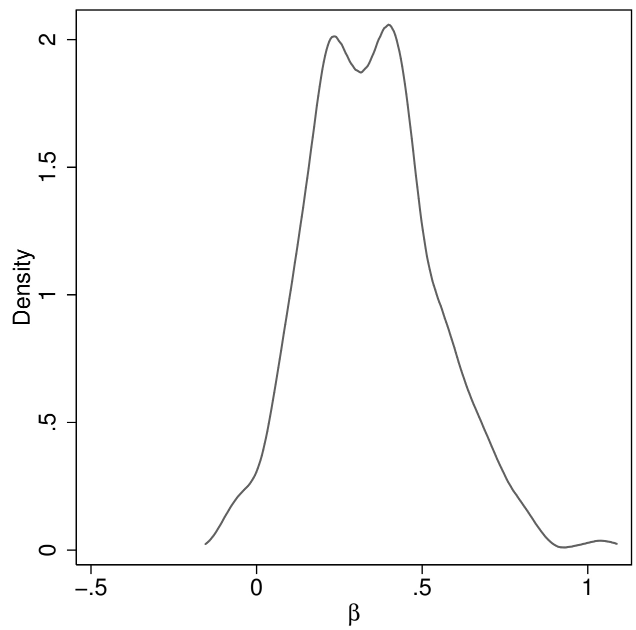

Figure 4: Distribution of $\widehat{\beta} = x_{it}\widehat{\delta_2}$

and people exhibiting more responsiveness to changes in real probabilities. It is tempting to think that these are the more attentive individuals and/or those circumstances under which participants pay more attention.

Figure 3 shows that the distribution for $\widehat{\sigma}$ is somewhat erratic, with a mean around 0.25 but some individuals above 0.3 and a few below 0.1, from which we mainly deduce that almost no participants will have had a stable answering strategy that fits with the hypothesized subjective probability function.[9]

We show in the next graph what the implied subjective probability function looks like for different individuals—in particular, the average individual in the sample (for whom $x_{it} = \overline{x}$), the individual with the highest estimated

[9]One might object to this by noting that our use of an additive error term in the presence of a bounded subjective probability term almost 'forces' a low mean $\widehat{\gamma}$. This turns out not to be true: when one forces all observations to be fitted by only structural coefficients and thus force an S-Shape to fit each combination of subjective probabilities with objective probabilities in the data, then, given what is shown in Figure 1, one must conclude that for some people $\widehat{\gamma} = 0.1$ or even lower. Hence, in effect, having a linear error term 'merely' forces the estimated $\widehat{\gamma}$ to reflect the average sample behavior near $p = 1$ seen in Figure 1.

$\widehat{\gamma}$ (whom one might label the 'most rational') and the individual with the highest estimated $\widehat{\sigma}$ (whom one might label the 'most erratic').

Figure 5 shows first that the person with the highest $\widehat{\gamma}$ displays near-rational behavior: for this person, there is an almost 1-to-1 correspondence between real probabilities and subjective probabilities, bar a slight non-linearity near $p = 0$. On the other hand, the 'most erratic' person displays almost the same subjective probability function as the average: a strong non-linearity very close to $p = 0$ and then a flattening out until $p = 1$. A similar pattern is shown for the person with the average values of observable characteristics, though the line lies lower, indicating less over-optimism than is evident for the 'most erratic' individual. The main message we take from this is that the typical person's subjective probability function displays an S-shape for low objective probabilities, but not for high objective probabilities—at least not in our setting. Both the parameter histograms and Figure 5 show visually the large amount of heterogeneity in subjective probabilities.

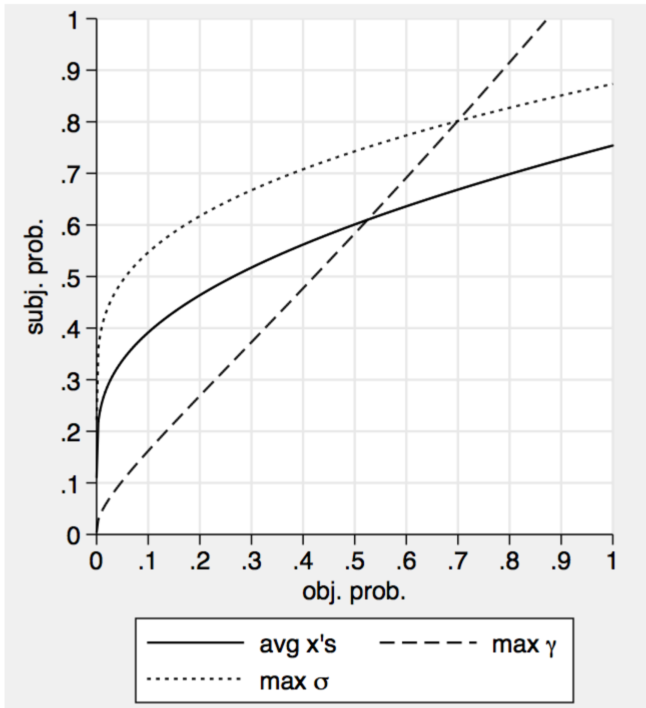

Figure 5: Graph of subjective versus objective probabilities under different settings

6 Discussion

Subjective probabilities, when plotted against true probabilities, have been found by prior researchers to exhibit an S-shape when experimental participants are told the true probabilities and their subjective probabilities are inferred from choices. By examining participants' subjective probabilities during a complicated dynamic game in which participants are not given the true probability of success, and working out the real probability is difficult, we find that this S-shape is also exhibited in stated subjective probabilities. Of the four models we consider, Prelec's (1998) compound invariant model best fits the empirical distribution in our data of stated subjective probabilities. Interestingly, we see no evidence of learning: both the mean and the the variance of the error term (the deviation away from the S-curve) are non-decreasing with the number of previous races experienced, both in reduced form analysis and in structural modelling.

We find three other important features in our more real-world observations about subjective probabilities. First, a 50 percent actual probability of a positive outcome (winning) was interpreted by our participants as a 70 percent chance of winning. This suggests that in very complicated real-world situations, there is a large over-estimate of middle-range probabilities of positive events. Secondly, we find severe under-estimation on average of the probability of events that in actuality are near certain to occur, with subjective probabilities no higher than 80% on average for events that will transpire with close to 100% certainty. Third, we find a large degree of individual heterogeneity in the relation between objective probabilities and subjective ones, with the extremities in the sample including near-rationality at the boundaries as well as extreme unresponsiveness to changes in actual probabilities in a middle range (flat S-shapes). Thus, we find little evidence that 'one shape fits all'.

In aggregate, our results suggest that individual decision-making in the presence of real-world complexity still adheres to the main tendencies observed in probability experiments: specifically, small probabilities are over-estimated and large probabilities are underestimated. In terms of broad policy implications, this finding implies that aggregate behavior will not be proportionately reactive to changes in real probabilities in a middle range, and will be over-reactive to changes at the extremes. This implies a possible welfare-enhancing role for distortion of the information that individuals receive about real probabilities. It also suggests great difficulty in convincing individuals of what is in fact certain when the situation is dynamic, new, and very complicated.

A Screenshots and experimental procedures

All experimental treatment protocols (baseline, “wealth”, “high-stakes” and “low-stakes”, as described in the text) consisted of six stages, as follows:

- An Introductory Questionnaire
- Relaxation, consisting of 5 minutes of relaxing beach sounds together with a voice-over of visualization guidance in a calm female voice
- Real Effort, consisting of cross sum calculations (adding up as many sets of 5-digit numbers as possible in a fixed time window), resulting in earned income
- A Test Race, consisting of 3 laps with labelled screens and a guiding voice-over in the same female voice used in the Relaxation stage
- 6 Real Races
- An incentivized “Guess the Winner” game
- A Follow-up Questionnaire focussing on demographics

Some participants wore heart rate monitors throughout the experiment, and all participants wore headphones from the start of the Relaxation stage until the end of the Real Races stage.

A.1 Introductory Questionnaire

In the Introductory Questionnaire, the participants were asked a set of standard questions regarding personality, locus of control, past/present/future savouring, and risk attitudes. These questions are reproduced in the next section.

A.2 Relaxation

The Relaxation stage was included to familiarize participants with a calm voice that would guide them through the car race set-up, and also to establish baseline readings for the heart rate monitors. Data from these monitors is not used in the present paper.

A.3 Real Effort

In the Real Effort stage, participants were presented with cross-sum problems for 10 minutes. The problems gradually became more difficult, and the participants receive a fixed amount per solved problem. A participant could opt to drop out of the solving process, and receive compensation for time foregone. Participants were able to earn up to 30 experimental dollars in this way.

A.4 Test Race

In the Test Race stage, participants were guided through the race process, with a voice-over explaining all the steps. Then a test race was shown, after which the participants had to answer a set of questions in order to proceed. The questions were explicitly designed so that they could not easily be solved by trial and error, and asking the experiment administrators in case of any question or confusion about the race procedure was explicitly encouraged.

A.5 Real Races

The Real Races stage confronted participants with 6 car races, in each of which participants could decide on how much of their earned income they would invest into their car and how much they would bet. Each race was independent of the others and one race would be paid out randomly in the end, so all of the earnings were available in each race. [10] The 6 cars would then race, and the advancement of each car was governed by the AR(1) process outlined in equation 5:

$$s_t = \theta s_{base} + (1 - \theta) * (1 + U(-\gamma, +\gamma)) * s_{t-1} + D_t * f \tag{5}$$

where s_t is the advancement of the car between time $t-1$ and time t ($\delta t = 1/60 second$), and θ governs the importance of the base speed $s_{base} = 50$. θ was set to 0.02, with the remaining weight of .08 placed upon the speed at $t-1$ multiplied by one plus a uniform random change in speed of $\pm\gamma = 0.1$. Finally, with probability β, $D_t = 1$ and a shock of $f = -25$ was applied (in the other $1-\beta$ fraction of times, no shock was applied). This functional form was selected in order to give the race a natural appearance while allowing meaningful manipulation of the winning probability, through the raising or lowering of β.

[10] Participants were not allowed to retain any of the earnings: the full amount had to be split between investing and betting.

Compared to no investment, full investment in the car would change the rate of engine failures (β) from an expected 4 engine failures in 5 laps (the 'no investment' option, associated with almost zero % chance of winning) to a certain zero failures (the 'full investment' option, associated with an almost 100% chance of winning, but with no money left over to wager, and hence nothing to win). The other 4 cars would always have an expected rate of 2 engine failures in 5 rounds.

Figure A.1 shows the screen participants faced at the beginning of each race. As shown, they could choose the division of the money they had earned in the real-effort task into an amount bet on their car, and an amount invested in their car's engine. All the information about the possible outcomes of the race, including the amounts to be recovered by dropping out at each of the pit stops, was displayed and updated whenever the participant changed his proposed decision using the slider. Each participant could also choose the color of his car each race, which was of no consequence to the race outcome.

Figure A.2 shows the screen participants encountered at the first pit stop. This screen offers a choice of whether or not to drop out of the race, and shows the payoffs associated with dropping out and with not dropping out and experiencing either of two states of the world: that in which one's car wins, and that in which one's car does not win. The screens for the second and third pit stop were nearly identical, with simply a later dropout choice bolded.

At the pit stops, the key variable of participants' subjective expectation of winning was elicited by having the participant move a slider to answer the question: "If the race were to continue from this point randomly 1000 times, how often would your car come first?"

The "Amount wagered" analysis variable is simply the amount that the participant chose to bet on his car, which is equal to "Real-effort Earnings" minus the amount invested in the car's engine. The "Amount to be won" analysis variable is the maximum amount that could have been won in the given race, once the participant placed a bet—that is, the amount that would be won if the participant did not drop out of his initial bet and if, in addition, his car won.

A.6 Guess the Winner

In the incentivized "Guess the winner" game, which was played 6 times, participants were presented with a visual state of the race at a given pit stop (3,6, or 9) and had to bet on how often out of 1000 races their car

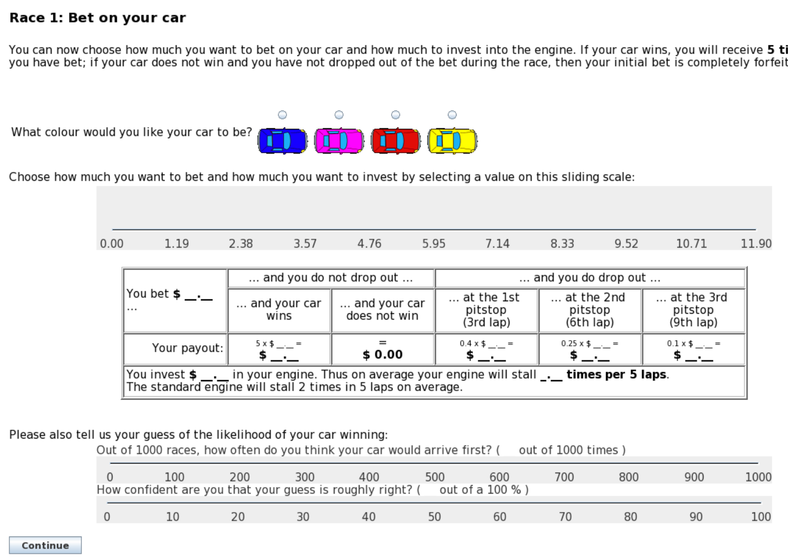

Figure A.1: Screenshot of the pre-race investment screen, with sliders for choosing the bet amount and entering expectations about race outcome.

would come first, conditional on the present position and race stage, by specifying a point estimate and an interval. Participants were able to select an interval of between 2 and 200 times. The larger the interval they chose, the lower was the amount they would win if their interval contained the correct value. They amount they could win from this "Guess the winner" game was displayed when they moved the interval slider. The race was then simulated 1000 times from the given point by the computer, and if the number of times the participant's car won fell into the prediction interval he had nominated, then the participant would receive \$ 0.10 plus \$ $x = ((200 - interval)/200)$. Participants' winnings were summed up and paid out over the six "Guess the winner" games, and this total amount won was also used as the "Guess the winner" analysis variable, which is intended as an indicator of how well participants were able to predict the outcome of a race.

A.7 Follow-up Questionnaire

The Follow-up Questionnaire included questions on a variety of demographic characteristics. The variables on participant characteristics included in our

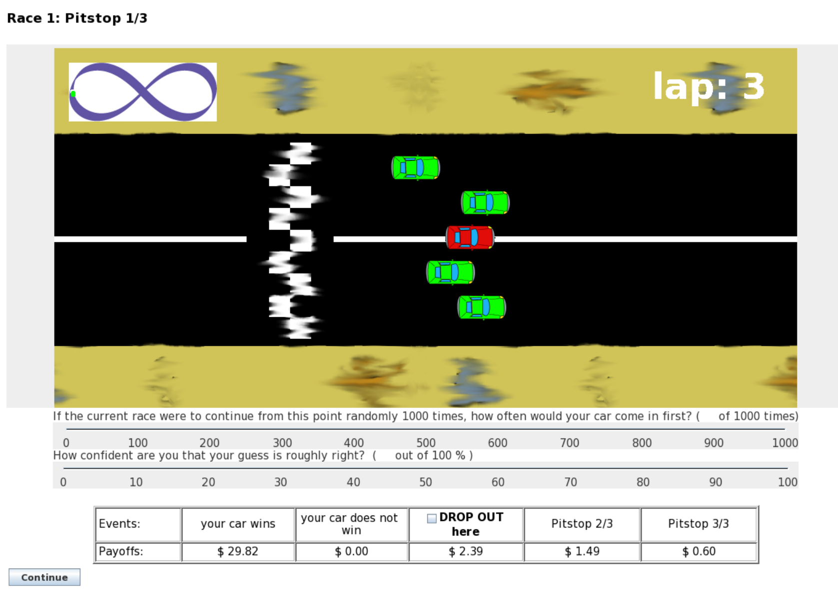

Figure A.2: Screenshot of the car race at the first pit stop, showing the positions of the cars and the payoffs for the respective events.

analysis are constructed from these questions and the questions posed in the Introductory Questionnaire, which are provided in the next section.

B Questionnaire Appendix

B.1 Self-esteem

The following statements posed to participants in the Introductory Questionnaire were coupled with scaled answer alternatives, ranging from Strongly Agree to Strongly Disagree. These items constitute a battery of self-esteem questions based on Rosenberg (1965). After reverse-coding questions 1, 3, 4, 7, and 10, we take the simple average of responses across all ten of these questions to construct our measure of self-esteem.

- On the whole, I am satisfied with myself.
- At times I think I am no good at all.
- I feel that I have a number of good qualities.
- I am able to do things as well as most people.
- I feel I do not have much to be proud of.
- I certainly feel useless at times.
- I feel that I am a person of worth, or at least on an equal plane with others.
- I wish I could have more respect for myself.
- All in all, I am inclined to feel that I am a failure.
- I take a positive attitude toward myself.

B.2 Optimism

The following questions, also answered on a Strongly Agree to Strongly Disagree scale, were used to capture participants' levels of optimism. The raw answers to item 4 below were used to create the analysis variable 'Disappointment', and those from item 6 were used to create the analysis variable 'Low Expectations'.

1. When I'm in a new and unfamiliar situation, I am always optimistic that things will work out for me (in other words, I feel and think that things will be OK).[11]

2. I often find myself doing things that I know, at the time I choose to do them, I will regret later.

3. When I expect that good things are going to happen to me in the future, I feel better about myself.[11]

4. When I get disappointed about something, it makes me feel that I'm to blame, because I should have known better in the first place and not expected as much.

5. I always try to be cautious when I approach new and unfamiliar situations, in case something goes wrong.[11]

6. I prefer to have low expectations of the future since that way I might be pleasantly surprised, and I'm protected from being disappointed.[11]

B.3 Locus of control

The following seven items, adapted from Rotter (1966) and answered on a Strongly Agree to Strongly Disagree scale, were used to measure locus of control. Answers to these questions (after appropriate reverse-coding) were averaged to obtain each participant's measure of locus of control.

- I have little control over the things that happen to me.
- There is really no way I can solve some of the problems I have.
- There is little I can do to change many of the important things in my life.
- I often feel helpless in dealing with the problems of life.
- Sometimes I feel that I'm being pushed around in life.
- What happens to me in the future mostly depends on me.
- I can do just about anything I really set my mind to.

[11] This variable was excluded from the analysis due to insignificant results.

B.4 Savoring

We also measured savoring, which we understand as individuals' capacity to enjoy good events in the past, present, and future, based on participants' answers to a battery of questions adapted from Bryant & Veroff (2006). The list of questions, each of which was answered on a Strongly agree to Strongly Disagree scale, is as follows:

1. Before a good thing happens, I look forward to it in ways that give me pleasure in the present.
2. It's hard for me to hang onto a good feeling for very long.
3. I enjoy looking back on happy times from my past.
4. I don't like to look forward to good times too much before they happen.
5. I know how to make the most of a good time.
6. I don't like to look back at good times too much after they've taken place.
7. I feel a joy of anticipation when I think about upcoming good things.
8. When it comes to enjoying myself, I'm my own 'worst enemy'.
9. I can make myself feel good by remembering pleasant events from my past.
10. For me, anticipating what upcoming good events will be like is basically a waste of time.
11. When something good happens, I can make my enjoyment of it last longer by thinking or doing certain things.
12. When I reminisce about pleasant memories, I often start to feel sad or disappointed.
13. I can enjoy pleasant events in my mind before they actually occur.
14. I can't seem to capture the joy of happy moments.
15. I like to store memories of fun times that I go through so that I can recall them later.

16. It's hard for me to get very excited about fun times before they actually take place.
17. I feel fully able to appreciate good things that happen to me.
18. I find that thinking about good times from the past is basically a waste of time.
19. I can make myself feel good by imagining what a happy time that is about to happen will be like.
20. I don't enjoy things as much as I should.
21. It's easy for me to rekindle the joy from pleasant memories.
22. When I think about a pleasant event before it happens, I often start to feel uneasy or uncomfortable.
23. It's easy for me to enjoy myself when I want to.
24. For me, once a fun time is over and gone, it's best not to think about it.

After reverse-coding questions 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, and 23, we take the average of responses to questions 1, 4, 7, 10, 13, 16, 19, and 22 to measure future-savoring ("SBI Anticipate" in the tables); the average of responses to questions 2, 5, 8, 11, 14, 17, 20, and 23 to measure present-savoring ("SBI Moment"); and the average of responses to questions 1, 6, 9, 12, 15, 18, 21, and 24 to measure past-savoring ("SBI Reminisce").

B.5 Non-incentivized risk aversion

To measure non-incentivized risk aversion, we used a standard (Holt & Laury 2002) lottery choice task in the Introductory Questionnaire, where the safe choices were \$20 and \$16 and the risky choices were \$40 and \$1 (roughly 5 times the values in the original paper (Holt & Laury 2002)). The following introductory text was used:

> For each of the nine pairs of lotteries listed below, please select your preferred lottery: either option A or option B. Each lottery is characterised by the probability of receiving one of two payoffs. (Probabilities are expressed as percentage chances of receiving this payoff, e.g. 20% = a chance of 2 out of 10 of receiving this payoff).

The participants then had to choose between the following options in each line, where on the participant's screen, the "–" was displayed as "chance of":

10% – $ 20	and	90% – $ 16	A	B	10% – $ 40	and	90% – $1
20% – $ 20	and	80% – $ 16	A	B	20% – $ 40	and	80% – $1
30% – $ 20	and	70% – $ 16	A	B	30% – $ 40	and	70% – $1
40% – $ 20	and	60% – $ 16	A	B	40% – $ 40	and	60% – $1
50% – $ 20	and	50% – $ 16	A	B	50% – $ 40	and	50% – $1
60% – $ 20	and	40% – $ 16	A	B	60% – $ 40	and	40% – $1
70% – $ 20	and	30% – $ 16	A	B	70% – $ 40	and	30% – $1
80% – $ 20	and	20% – $ 16	A	B	80% – $ 40	and	20% – $1
90% – $ 20	and	10% – $ 16	A	B	90% – $ 40	and	10% – $1

The number of safe choices (i.e., selections of option A) was used in the analysis as an indicator of risk attitude, with the variable label "Risk aversion (HL)", if participants exhibited a single switching point from A to B as they proceeded from the top of the table to the bottom. Participants with more than one switching point were classified as switching on the fifth line, corresponding to slight risk aversion. Exclusion of these latter participants did not alter the outcome.

B.6 Follow-up questionnaire

The following questions/statements were posed to participants in the Follow-up Questionnaire. Some questions required participants to type in answers in free-form; others were followed either by scaled answer alternatives, ranging from Strongly Agree to Strongly Disagree, or by appropriately populated arrays of answer alternatives. The exact mapping of the answers to these questions to variables used in our analysis is straightforward and available upon request.

- What is your year of birth?
- What is your month and day of birth?[12]
- Please indicate your gender.
- Please enter your nationality.[12]
- Please enter the country you were born.[12]

[12] This variable was excluded from the analysis due to insignificant results and/or collinearity issues.

- Please enter the country whose culture you identify with most strongly. (This variable was used to create the Culture dummies in the regression, participants where put in three main categories, Australian, Asian and Other which is predominately USA or European.)
- Do you speak English at home?
- Are you currently . . . (married, in a partnership, or single).[12]
- What is your current living situation?[12]
- Please enter the postcode of the area you live in.[12]
- Which degree program are you enrolled in (Economics; Commerce; etc.)?[12]
- When do you expect to graduate (month, year)?[12]
- Are you an international student?
- Have you ever participated in an experiment before?
- What is your weekly disposable income?
  (None or <$100, $100-$199, $200-$299, $300-$399, $400-$499, >$500 —This variable was encoded as "None", "Low", "Avg.", and "High", where no participant reported an weekly income above $399)
- What was the highest year of school you completed?
- And how much schooling did your mother complete?
- And how much schooling did your father complete?
- Did you complete an educational qualification after leaving school? Please include any trade certificates, apprenticeships, diplomas, degrees or other educational qualifications.
- If yes, what was the highest type of qualification you obtained?
- Did your mother complete an educational qualification after leaving school? Please include any trade certificates, apprenticeships, diplomas, degrees or other educational qualifications.
- If yes, what was the highest type of qualification she obtained?

- Did your father complete an educational qualification after leaving school? Please include any trade certificates, apprenticeships, diplomas, degrees or other educational qualifications.
- If yes, what was the highest type of qualification he obtained?
- Please select the category or class of professions your mother's occupation falls into (even if she is unemployed).
- What is the full title of your mother's occupation?
- Please select the category of class of professions your father's occupation falls into (even if he is unemployed).
- What is the full title of your father's occupation?
- Are you a vegetarian?
- How tall are you, in centimetres?
- How much do you weigh in light clothing, in kilograms?
- Do you regularly smoke any tobacco product, such as cigarettes, cigars, or pipes?
- When you drink alcohol, on average, how many drinks do you have?
- Are you taking a prescribed medication?[12]
- Have you had any symptoms of or complaints about depression during the last month (30 days)?
- Which hand do you write with?
- What is your opinion of the following statement: 'Good luck charms sometimes do bring good luck.' (answer scale: Definitely not true, Probably not true, Don't know, Maybe, Probably true, Definitely true)[12]
- Do you have a lucky charm?
- Many people think there is someone watching out for them to make sure things go well. This someone cannot be directly seen. Is there someone, who cannot be seen by others, watching over you?[12]

- Apart from weddings, funerals and christenings, how often do you attend religious services these days?[12]
- How satisfied are you with your financial situation?[12]
- In political matters, people talk of “the left” and “the right”. How would you place your views on this scale, generally speaking?[12]
- All things considered in your life, how happy would you say you are usually?
- Would you say that your family is ... (wealthier (Wealth Level Above), the same (Wealth Level Avg), or poorer (Wealth Level Poor) than others)?
- Overall, how would you rate your performance at university?
- Betting is justified.
- Gambling is justified.[12]

References

Abdellaoui, M., L'Haridon, O. & Zank, H. (2010), 'Separating curvature and elevation: A parametric probability weighting function', *Journal of Risk and Uncertainty* **41**, 39–65.

Ali, M. M. (1977), 'Probability and utility estimates for racetrack bettors', *Journal of Political Economy* **85**, 803–815.

Andersen, S., Harrison, G. W., Lau, M. I. & Rutström, E. E. (2010), 'Preference heterogeneity in experiments: Comparing the field and laboratory', *Journal of Economic Behavior and Organization* **73**(2), 209–224.

Andersen, S., Harrison, G. W., Lau, M. I. & Rutström, E. E. (2013), 'Dual criteria decisions', *Journal of Economic Psychology* **forthcoming**.

Berns, G. S., Capra, C. M., Chappelow, J., Moore, S. & Noussair, C. (2008), 'Nonlinear neurobiological probability weighting functions for aversive outcomes', *Neuroimage* **39**(4), 2047–2057.

Blavatskyy, P. & Pogrebna, G. (2008), 'Risk aversion when gains are likely and unlikely: Evidence from a Natural Experiment with Large Stakes', *Theory and Decision* **64**, 395–420.

Bombardini, M. & Trebbi, F. (2012), 'Risk aversion and expected utility theory: An experiment with large and small stakes', *Journal of European Economic Association* **6**(10), 1348–1399.

Botti, F. & Conte, A. (2008), 'Risk attitude in real decision problems', *The B. E. Journal of Economic Analysis and Policy (Advances)* **8**(1).

Bruhin, A., Fehr-Duda, H. & Epper, T. (2010), 'Risk and rationality: Uncovering heterogeneity in probablity distortion', *Econometrica* **78**, 1375–1412.

Bryant, F. B. & Veroff, J. (2006), *Savoring: A new model of positive experience*, Lawrence Erlbaum Associates Incorporated.

Camerer, C. F. (1995), Individual decision making, *in* J. H. Kagel & A. E. Roth, eds, 'Handbook of Experimental Economics', Princeton: Princeton University Press, pp. 587–703.

Camerer, C. F. & Ho, T. H. (1994), 'Nonlinear weighting of probabilities and violations of the betweenness axiom', *Journal of Risk and Uncertainty* **8**, 167–196.

Camerer, C. F. & Loewenstein, G. (2004), Behavioral economics: Past, present, future, *in* C. F. Camerer, G. Loewenstein & M. Rabin, eds, 'Advances in Behavioral Economics', Princeton: Princeton University Press, pp. 3–51.

Cheung, Y. (2001), A monte carlo study of the probability weighting function. Edith Cowan University working paper.

Chew, S. H. & Sagi, J. S. (2008), 'Small worlds: Modeling attitudes toward sources of uncertainty', *Journal of Economic Theory* **139**(1), 1–24.

Dale, H. C. A. (1959), 'A priori probabilities in gambling', *Nature* **183**, 842–843.

de Palma, Andre, M. B.-A., Brownstone, D., Holt, C., Magnac, T., McFadden, D., Moffat, P., Picard, N., Train, K., Wakker, P. & Walker, J. (2008), 'Risk uncertainty and discrete choice models', *Marketing Letters* **19**, 269–285.

de Roos, N. & Sarafidis, Y. (2010), 'Decision Making under Risk in Deal Or No Deal', *Journal of Applied Econometrics* **25**, 987–1027.

Deck, C., Lee, J. & Reyes, J. (2008), 'Risk attitudes in large stake gambles: Evidence from a game show', *Applied Economics* **40**(1), 41–52.

Fox, C. R. & See, K. E. (2003), Belief and preference in decision under uncertainty, *in* D. Hardmann & L. Macchi, eds, 'Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making', Vol. 273-312, John Wiley and Sons Ltd, West Sussex.

Golec, J. & Tamarkin, M. (1998), 'Love skewness, not risk, at the horse track', *Journal of Political Economy* **106**, 205–225.

Gonzalez, R. & Wu, G. (1999), 'On the shape of the probability weighting function', *Cognitive Psychology* **38**, 129–166.

Greiner, B. (2004), An online recruitment system for economic experiments, *in* K. Kremer & V. Macho, eds, 'Forschung und wissenschaftliches Rechnen', GWDG Bericht, pp. 79–93.

Griffith, R. M. (1949), 'Odds adjustments by american horse-race', *American Journal of Psychology* **62**, 290–294.

Holt, C. A. & Laury, S. K. (2002), 'Risk aversion and incentive effects', *The American Economic Review* **92**(5), 1644–1655.

Jullien, B. & Salanié, B. (2000), 'Estimating preferences under risk: The case of racetrack bettors', *Journal of Political Economy* **108**(3), 503–530.

Kahneman, D. & Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**(2), 263–292.

Lattimore, P. K., Baker, J. R. & Witte, A. D. (1992), 'The influence of probability on risky choice: A parametric examination', *Journal of Economic Behavior and Organization* **17**(3), 377–400.

Machina, M. (2004), Nonexpected utility theory, *in* J. L. Teugels & B. Sundt, eds, 'Encyclopedia Of Actuarial Science, Volume 2', John Wiley and Sons, Ltd, pp. 1173–1179.

Martins, A. C. R. (2006), 'Probability biases as Bayesian inference', *Judgment and Decision Making* **1**(2), 108117.

McGlothlin, W. H. (1956), 'Stability of choices among uncertain alternatives', *American Journal of Psychology* **69**, 604–615.

Mulino, D., Scheelings, R., Brooks, R. & Faff, R. (2009), 'Does risk aversion vary with decision-frame? An empirical test using recent game show data', *Review of Behavioral Finance* **1**, 44–61.

Paulus, M. P. & Frank, L. R. (2006), 'Anterior cingulate activity modulates nonlinear decision weight function of uncertain prospects', *Neuroimage* **30**, 668–677.

Post, T., J., M. & Assem (2008), 'van den', *Baltussen, G., and R. H. Thaler (2008), Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show, American Economic Review* **98**, 28–71.

Prelec, D. (1998), 'The probability weighting function', *Econometrica* **66**(3), 497–528.

Prelec, D. (2000), Compound invariant weighting functions in prospect theory, *in* D. Kahneman & A. Tversky, eds, 'Choices, Values, and Frames.', Vol. 67–92, Cambridge: Cambridge University Press:.

Preston, M. G. & Baratta, P. (1948), 'An experimental study of the auction-value of an uncertain outcome', *American Journal of Psychology* **61**(2).

Quiggin, J. (1982), 'A theory of anticipated utility', *Journal of Economic Behavior and Organization* **3**(4), 323–343.

Rosenberg, M. (1965), *Society and The Adolescent Self-Image*, Princeton University Press.

Rottenstreich, Y. & Hsee, C. K. (2001), 'Money, kisses, and electric shocks: On the affective psychology of risk', *Psychological Science. 12(3)* **12**(3), 185–190.

Rotter, J. (1966), 'Generalized expectancies for internal versus external control of reinforcement', *Psychological Monographs* **80**.

Schaffner, M. (2013), Programming for experimental economics: Introducing coral—a lightweight framework for experimental economic experiments, QuBE Working Papers 016, QUT Business School.

Schmeidler, D. (1989), 'Subjective probability and expected utility without additivity', *Econometrica* **55**, 409–424.

Starmer, C. (2000), 'Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk', *Journal of Economic Literature* **38**, 332–382.

Starmer, C. (2004), Developments in Nonexpected-Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk, *in* C. F. Camerer, G. Loewenstein & M. Rabin, eds, 'Advances in Behavioral Economics', Princeton University Press, Princeton, pp. 104–147.

Trepel, C., Craig, R. F. & Poldrack, R. A. (2005), 'Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk', *Cognitive Brain Research* **23**, 34–50.

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

von Neumann, J. & Morgenstern, O. (1947), *Theory of games and economic behavior*, Princeton University Press.

Wakker, P. P. (1994), 'Separating marginal utility and probabilistic risk aversion', *Theory and Decision* pp. 1–44.

Weitzman, M. (1965), *Utility analysis and group behavior: An empirical study, Journal of Political Economy, 73(1)*, 8-26.

Wu, G. & Gonzalez, R. (1996), 'Curvature of the probability weighting function', *Management Science* **42**(12), 1676–1690.

Yaari, M. E. (1987), 'The dual theory of choice under risk', *Econometrica* **55**(1).