Appendix to "Information Characteristics and Errors in Expectations: Experimental Evidence"

A. Subject Demographics

Age	
Mean	21.3
Median	20
Sex	
Male	58
Female	53
Field	
Economics, Finance, Business Administration	24
Engineering	4
Biological sciences, Health Medicine	6
Math, Computer or Physical Sciences	26
Social Sciences	23
Law	8
Psychology	4
Modern Languages	8
Other fields	8
Level of study	
Undergraduate	88
Postgraduate	12
Graduate	11
Mark at Bachelor degree	
Above 70% (first class)	30
between 60 and 69% (2.1)	72
between 50 and 59% (2.2)	5
No grades yet awarded	4

B. Instructions for the Belief Task

In this stage of the experiment you will be betting on the outcomes of uncertain events. Usually we bet on events like football matches or elections, but in this task the events will be random choices made by the experimenter between two boxes, one blue and the other white. The experimenter will not tell you which box was chosen. At the start each box will have the same chance of being chosen, but once it has been chosen the experimenter will give you some information to help you work out the chances that it was blue or white. Armed with this information, you will make bets on which box was chosen.

The procedure, which is summarized on the accompanying picture, is as follows. The experimenter will first choose the box by rolling a 6-sided die with three blue and three white sides. If blue comes up he will choose the blue box, if white comes up he will choose the white one.

Both the white and blue boxes contain several dice, each having 10 sides. Both boxes have the same number of dice, which will vary over the course of the experiment. The dice in the blue box always have 6 blue sides and 4 white ones, while those in the white box have 4 blue sides and 6 white ones.

The experimenter will roll all the dice in the chosen box and tell you how many blue and white sides came up. He will not tell you which box was chosen.

Because the dice in the blue box have more blue sides than those in the white box, knowing the number of blue and white sides that come up can help you work out the chances that each box was chosen. For example, if more blue sides come up this means it is more likely to be the blue box, and if more white sides come up it is more likely to be the white box.

Once you have the information about the dice rolls, you will then make bets on which box was chosen.

About betting

You will be making bets with several betting houses or “bookies,” just as you might bet on a football game or a horse race.

To familiarize you with betting, we will illustrate how it works with the example of a horse race.

Imagine a two horse race between Blue Bird and White Heat. Several bookies offer different odds for both horses. The table below shows the odds offered by three bookies along with the amounts they would pay if you staked £10 on the *winning* horse. The earnings are calculated by multiplying the odds by the stake. In this experiment you will be making bets on which box was chosen using a table like this. **At this point you should take some time to study the table.**

		Odds offered		**Earnings including the stake of £10**	
Bookie	**Stake**	**Blue Bird**	**White Heat**	**Blue Bird**	**White Heat**
A	£10	5.00	1.25	£50.00	£12.50
B	£10	3.33	1.43	£33.33	£14.30
C	£10	2.00	2.00	£20.00	£20.00

Below are three important points about betting.

1. **Your belief about the chances of each outcome is a personal judgment that depends on information you have about the different events.** For the horse race, you may have seen previous races or read articles about them. In the experiment the information you have about whether the blue or white box was chosen will be how many blue and white faces came up.
2. **Even if you believe Event X is more likely to occur than Event Y, you may want to bet on Y because you find the odds attractive.** For example, even if you believe White Heat is most likely to win you may want to bet on Blue Bird because you find the odds attractive. To illustrate, suppose you personally believe that Blue Bird has a 40% chance of winning and White Heat has a 60% chance of winning. This means that if you bet £10 on Blue Bird with Bookie A you believe there is a 40% chance of receiving £50.00 and a 60% chance of receiving nothing. You may find this more attractive than betting on White Heat, which you believe offers a 60% chance of 12.50 and a 40% chance of nothing.
3. **Your choices might also depend on your willingness to take risks or to gamble.** There is no right choice for everyone. In a horse race you might want to bet on the long-shot since it will bring you more money if it wins, but you also might want to bet on the favourite since it is more likely to win something.

For each bookie, whether you would choose to bet on Blue Bird or White Heat will depend on three things: your judgment about how likely it is each horse will win, the odds offered by the bookie, and how much you like to gamble or take risks.

Your choices

Now you are familiarized with odds, we can go back to the experimental betting task. Recall that the experimenter will first make a random choice of a blue or white box. Then he will roll the dice in the chosen box and tell you how many white and blue sides came up. Then you will consider the chances that the box chosen was blue or white, and make a series of bets.

You have a booklet of record sheets. Each record sheet shows the bookies you will be dealing with, and the odds they offer. There are 19 bookies on each sheet, and each offer different odds for the two outcomes. **Take a minute to look at one such record sheet, shown on the next page.**

There will be 30 separate events, and 19 bookies offer odds for each event. **You will make bets at all 19 bookies for all 30 events.**

For each bet, you have a £3 stake, and the record sheet shows the payoffs you will receive if you bet on the box that was actually chosen.

There is a separate record sheet for each of the 30 events. On each sheet you should circle W or B to indicate the bet you want to make with **all 19 bookies**.

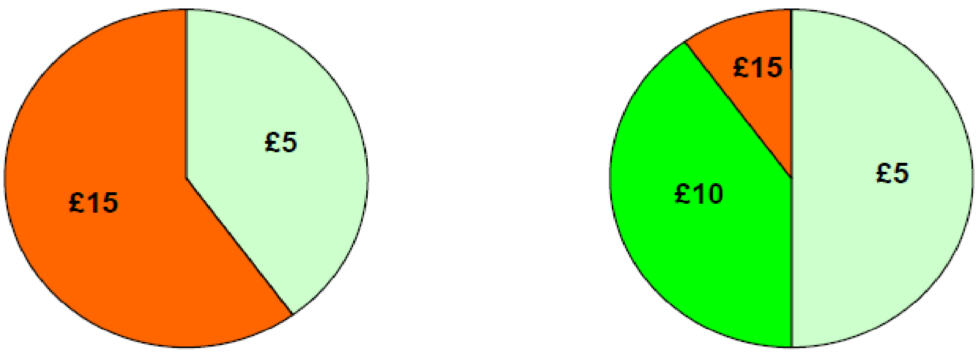
One and only one of the bets in the entire experiment will pay off for real. Therefore, please consider each bet as if it is the only one that will be paid out. After you have placed all your bets, you will roll a 30-sided die to determine which event will be played out, and a 20-sided die to determine which bookie will determine your earnings.

All payoffs are in cash, and are in addition to the £5 show-up fee that you receive just for being here.

C. Instructions for the Risk Elicitation Task

This stage is about choosing between lotteries with varying prizes and chances of winning. You will be shown a series of 20 lottery pairs, and you will choose the lottery you prefer from each pair. You will actually get the chance to play one of the lotteries you choose, and will be paid according to the outcome of that lottery, so you should think carefully about your preferences.

Here is an example of one lottery pair. You will have to think about which lottery you would prefer to play and tick the appropriate box below



40% chance of £5	(numbers 1-40)	50% chance of £5	(numbers 1-50)
60% chance of £15	(numbers 41-100)	40% chance of £10	(numbers 51-90)
		10% chance of £15	(numbers 91-100)

Your choice:

The outcome of the lotteries will be determined by the draw of a random number between 1 and 100. We will ask you to roll a 100-sided die that is numbered from 1 to 100, and the number on the die will determine the outcome of the lotteries.

In the above example the left lottery pays five pounds (£5) if the number on the die is between 1 and 40, and it pays fifteen pounds (£15) if the number is between 41 and 100. The light green segment of the pie chart corresponds to 40%, and the orange segment corresponds to 60% of the area.

Now look at the pie chart on the right. It pays five pounds (£5) if the number drawn is between 1 and 50, ten pounds (£10) if the number is between 51 and 90, and fifteen pounds (£15) if the number is between 91 and 100. As with the lottery on the left, the pie slices represent the fraction of the possible numbers which yield each payoff. For example, the size of the £15 pie slice is 10% of the total pie.

Each of the 20 lottery pairs will be shown on a separate sheet of paper. On each sheet you should indicate your preferred lottery by ticking the appropriate box. After you have worked through all the lottery pairs, please raise your hand. You will then roll a 20-sided die to determine which pair of lotteries will be played out, and then roll the 100-sided die to determine the outcome of the chosen lottery.

For instance, suppose you picked the lottery on the left in the above example. If you roll the 100-sided die and the number 37 is shown, you would win £5; if it was 93, you would get £15. If you picked the lottery on the right and drew the number 37, you would get £5; if it was 93, you would get £15.

Therefore, your payoff is determined by three things:

- which lottery pair is chosen to be played out using the 20-sided die;
- which lottery you selected, the left or the right, for the chosen lottery pair; and
- the outcome of that lottery when you roll the 100-sided die.

This is not a test of whether you can pick the best lottery in each pair, because none of the lotteries are necessarily better than the others. Which lotteries you prefer is a matter of personal taste.

Please work silently, and think carefully about each choice.

All payoffs are in cash, and are in addition to the £5 show-up fee that you receive just for being here.

D. The Structural Model

We start by explaining the econometric analysis of the data collected in the risk task. We assume a CRRA utility function in the context of EUT, shown by (1) in the main text of the paper, where *r* is a parameter to be estimated, and *y* is income from the experimental choice. The utility function (1) can be estimated using the responses from our risk task using maximum likelihood and a latent EUT structural model of choice. In the lotteries provided there are *K* possible outcomes, therefore, Expected Utility (EU) of each lottery *i* is:

$$EU_i = \Sigma_{k=1,K} [p_k \times u_k]. \tag{1}$$

The EU for each lottery pair is calculated for a candidate estimate of *r*, defining the index:

$$\nabla EU = EU_R - EU_L \tag{2}$$

This latent index is linked to the observed choices using a standard cumulative normal distribution function $\Phi(\nabla EU)$, resulting to a probit link function:

$$\text{prob(choose lottery R)} = \Phi(\nabla EU) \tag{3}$$

An important extension of the core model is to allow for respondents to make some errors. We use the contextual error specification proposed by Wilcox (2011). It posits the latent index:

$$\text{prob(choose lottery R)} = \Phi [(\nabla EU)/v)/\mu] \tag{4}$$

where v is a normalizing term for each lottery pair L and R, defined as the maximum utility over all prizes in this lottery pair minus the minimum utility over all prizes in this lottery pair. $\mu>0$ is a structural "noise parameter" used to allow some errors from the perspective of the deterministic EUT model. As $\mu \rightarrow \infty$ this specification collapses ∇EU to 0 for any values of EU_R and U_L, so the probability of either choice converges to ½. Therefore, a larger μ means that the difference in the EU of the two lotteries, conditional on the estimate of *r*, is less predictive of choices. In our estimations we use a log-transform for μ to ensure that it is non-negative, with standard errors and point estimates derived using the delta method. Additional details of the estimation methods used, including corrections for "clustered" errors when we pool choices over respondents and tasks, are provided by Harrison and Rutström (2008).The log-likelihood is then:

$$\ln L(r, \mu; y, \mathbf{X}) = \Sigma_i [(\ln \Phi(\nabla EU)\times\mathbf{I}(y_i = 1)) + (\ln (1-\Phi(\nabla EU))\times\mathbf{I}(y_i = -1))] \tag{5}$$

where $\mathbf{I}(\cdot)$ is the indicator function, y_i =1(-1) denotes the choice of the Option R (L) lottery in risk aversion task *i*, and **X** is a vector which contains v and any other data, such as demographics.

We now explain how we use the choices in the belief task to construct the second likelihood to estimate subjective probabilities. As shown in Table 1 in the paper the subject that selects event W from a given betting house receives:

$$EU_W = \pi_w \times U(\text{payout if W} \mid \text{bet on W}) + (1-\pi_w) \times U(\text{payout if B} \mid \text{bet on W}) \quad (6)$$

where π_w is the subjective probability that W will occur. The payouts that enter the utility function are defined by the odds that each bookie offers. The EU received from a bet on event B is defined similarly.

We observe the bet made by the subject for a range of odds, so we can calculate the likelihood of that choice given values of r, π_w and μ, again assuming *EUT* and CRRA. The rest of the structural specification is exactly the same as for the choices over lotteries with objective probabilities. Thus the likelihood function for the observed choices in the belief task is:

$$\ln L(r, \pi_w, \mu; y, \mathbf{X}) = \Sigma_i \left[(\ln \Phi(\nabla EU)\times \mathbf{I}(y_i = 1)) + (\ln (1-\Phi(\nabla EU))\times \mathbf{I}(y_i = -1)) \right] \quad (7)$$

The *joint estimation* problem is to find values for r, π_w and μ that maximize the *sum* of (5) and (7). To ensure that the choice probability lies in the unit interval we use the transform $\pi = 1/(1+\exp(\kappa))$, where κ is the parameter estimated which is free to vary between $\pm\infty$ and π is the inferred probability. To infer point estimates and standard errors for π from estimates of κ we again use the delta method.

To formally examine the sensitivity of subjective probabilities to strength, S, and weight, N, we can estimate the following model:

$$\log \{\log(\pi/(1-\pi)) / \log(0.6/0.4) \} = \alpha \log N + \beta \log S \quad (8)$$

where π and $1- \pi$ are the elicited subjective probabilities for White or Blue, respectively. Bayes Rule implies that $\alpha= \beta=1$, but under the strength-weight hypothesis $\alpha < \beta$. Because, generalizing GT, we obtain both w>b and b>w cases, we make a transformation to the definition of the subjective probability in the model, to ensure that strength S is always positive. Thus when $w>b$ we express it as $\pi = 1/(1+(1/\lambda))$, and when $b<w$ we express it as $\pi = 1/(1+\lambda)$, where $\lambda = \exp[\exp(\gamma) \exp(0.6/0.4)]$.

To estimate the model in (8) whilst controlling for the utility function we can replace π_w in the models explained above with two parameters, α and β, and estimate these parameters using the joint estimation procedure explained above.

E. Derivation of Equation 2

Here we provide the general procedure for computing the posterior probability that a given set of dice (White or Blue, *W* or *B*) was chosen given the sample outcome *(w,b)*. The posterior probability, π, that *W* was chosen is:

$$\pi = p(W|w,b) = \frac{p(w,b|W)p(W)}{p(w,b)}.$$

The posterior probability of B is then $1 - \pi$. The likelihoods are the probabilities of given data, in this case (w,b), given the hypothesis, in this case W or B. The likelihood of W and B are therefore:

$$p(w,b|W) = [N!/(w!\ b!)]\ p(W)^{w}\ (1-p(W))^{b},$$

$$p(w,b|B) = [N!/(w!\ b!)]\ p(B)^{w}\ (1-p(B))^{b},$$

The odds ratio is $p(w,b|W)/p(w,b|B)$ which taking into account the fact that $p(B) = 1-p(W)$, reduces with some simple algebra to the following:

$$L = \left(\frac{p(W)}{1-p(W)}\right)^{|w-b|}.$$

To separate out the effects of strength and weight we first take the log on both sides of the equation and then multiply and divide through by *N*:

$$\begin{aligned}\log(L) &= |w-b|\,log\left(\frac{p(W)}{1-p(W)}\right)\\ &= N\left(\frac{|w-b|}{N}\right)log\left(\frac{p(W)}{1-p(W)}\right)\end{aligned}$$

Re-arranging and taking again the log on both sides gives the expression for weight and strength from Griffin and Tversky (1992):

$$log\left(log(L)/log\left(\frac{p(W)}{1-p(W)}\right)\right) = \log(N) + log\left(\frac{\lfloor w-b\rfloor}{N}\right)$$

Multiplying the two right-hand terms by α and β, respectively, we generate expression (17) from our paper:

$$log\left(log(L)/log\left(\frac{p(W)}{1-p(W)}\right)\right) = alog(N) + \beta log\left(\frac{|w-b|}{N}\right)$$

Bayes Rule holds iff $\alpha=\beta=1$.

F: Asset Pricing Simulations

In this table we report results from simulations using the procedure in Barberis, Shleifer and Vishny (1998). Specifically, we use this model to simulate a string of $n=6$ earnings shocks for 2,000 companies, using a random walk model. All firms have initial earnings equal to N_1, and then in each of the following periods all firms are equally likely to experience a positive or negative earnings shock equal to y. Following BSV, we choose y to be low relative to N_1 to avoid having negative earnings, and hence negative prices. Prices are derived according to Proposition 1 in BSV. We form two portfolios in each period: one consisting of firms with a positive earnings surprise in each of the n years, where n ranges from 1 to 4, and another with firms with a negative earnings shock. We then calculate the returns of these portfolios in the following year, and report the difference $r_+^n - r_-^n$. Returns are in percent. The focus of our analysis is to examine how changes in the transition probabilities, λ_1 and λ_2, affect the signs of returns. In the column titled BSV we present results with λ_1=0.1 and λ_2=0.3, following BSV. In the remaining three columns we change these parameters by the indicated percentage in a way that implies that the investor is always more likely to rely on the mean-reverting regime to forecast earnings (i.e., decreasing λ_1 and increasing λ_2), and repeat the process, holding all other parameters constant.

n	BSV	10% change	20% change	30% change
1.00	2.61	3.61	4.36	5.02
2.00	0.66	1.96	3.17	4.24
3.00	-2.27	0.22	1.60	3.20
4.00	-3.25	-1.00	0.97	2.71

References

Wilcox, N. T. (2011). 'Stochastically more risk averse:'A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics*, 162(1), 89-104.