Previously Encountered Options Affect Risky Decisions in Choices Between Pension Funds

Sarah Smith (sarah.smith.10@ucl.ac.uk) and Adam J. L. Harris (adam.harris@ucl.ac.uk)
Department of Cognitive, Perceptual and Brain Sciences, University College London,
26 Bedford Way, London, WC1H 0AP

Abstract
The evidence that available options influence risk preference raises concerns about using risk attitude measures to guide investment choice. A between-participants test found those who first made choices between high (low) risk pension funds subsequently preferred higher (lower) risk funds when offered a choice from a larger range of funds. A within-participants test, with a fourteen-day break between conditions, found the same participants selected a lower risk fund after they experienced making choices between low risk funds, than they did after they experienced making choices between high risk funds. Effects were not influenced by order, or attenuated by a bias warning. Results are consistent with theories that suggest risky decision making involves similar processes to those involved in psychophysical perception (e.g. Decision by Sampling), although there was also evidence of sensitivity to absolute values. From an applied perspective, customers’ risk preferences are susceptible to manipulation.

Keywords: financial risk attitude; Decision by Sampling; sequential context effects; prospect relativity.

Background
Having people save into investment funds that secure good pensions is a goal of the UK Government (HM Treasury, 2012). The concept of personal risk attitude is embedded in this effort. European regulations require investment advisers to measure a client’s risk attitude before giving advice (Financial Services Authority, 2011). This has resulted in a proliferation of risk profiling tools that ask people a series of questions to derive a risk attitude diagnosis. Yet, studies show that context can have profound effects on revealed risk preferences. For example, Stewart, Chater, Stott and Reimers (2003) found that valuations and choices between gambles almost entirely depended on whether you had a set of choices that were high or low. In their Experiment 1A, participants were asked to make a series of evaluations of the type, ‘how much would you pay for a percentage (e.g. 60%) chance of winning an amount of money (e.g. £400)?’. One condition was given low prices to choose from, the other, high prices (for this gamble the low prices were £60, £80, £100 or £120, and the high: £160, £180, £200 or £220). Participants chose the middle options most frequently, despite them being considerably higher in the high price condition. In Experiment 4, participants were asked to choose one from a set of ‘p chance of winning x’ prospects. Risk attitude was found to be solely a function of the set of prospects offered. Stewart et al. suggested that participants have no sense of the absolute risk in each prospect, but rather choose a prospect by how it relates to the others presented. This casts doubt on whether risk attitude is sufficiently stable to be measured at one point and then generalised to future decisions.

Building on Stewart et al. (2003, Expt. 4), Vlaev, Chater and Stewart (2007) tested whether a significant real-world risky decision – a hypothetical pension fund choice - might be affected by context, and whether context would affect perceived risk appetite. First, they piloted ways of communicating investment risk. An investment mean variance (IMV) measure (Figure 1) that presents risk as the variation between the minimum and maximum possible outcomes of a fund, calculated to 95% confidence limits with an average in between, was rated highest for comprehensibility and produced more stable risk attitude results than other measures, over a three-month period. Using the IMV in a series of tests on working age participants, Vlaev et al. asked participants to choose a fund from a range constrained to high (low) risk. If participants had a stable knowledge of their underlying pension fund preferences, one would expect them to choose the fund closest to that preference from the range available, such that there would be a large number of people choosing the least risky option in the high risk range and the most risky option in the low risk range. Contrary to this (as in Stewart et al., 2003, Expt. 4), Vlaev et al found the proportion of times the lowest option in the high range was selected was significantly lower than the proportion of times the options below it were selected in the low range. On the other hand, there was a general skew towards low risk funds. This evidence of some absolute preference suggests that choice is not solely prospect dependent.

Vlaev et al. (2007) also found that constraining the range did not affect reported risk attitude; and participants who self-reported as more risk seeking picked a more risky fund from the range offered. One possible interpretation of these results is that participants overlay their own rank, in terms of their risk attitude, onto an ordered range of options in order to identify their preferred fund (e.g., "I'm typically a risk-seeking person, so I'll choose an option in the upper quartile"). If so, the absolute level of fund riskiness is irrelevant: the fund’s location in the decision making range is what matters. Indeed, our perceptions of how our behaviour ranks compared to other people’s behaviour has been shown to play a role in judgments involving risk in other contexts. Wood, Brown and Maltby (2012) found participants’ perceptions of the risks related to alcohol consumption were predicted by where an individual ranked herself in a range of how much she thought her peers drank. Whether and how social ranking influences judgments of
Findings such as those reviewed above have been interpreted as suggesting that the process by which we judge risk may be similar to the process by which we judge psychophysical stimuli like sound or weight (Birnbaum, 1992; Parducci, 1965; Stewart, Chater, & Brown, 2006). Adaptation Level Theory (Helson, 1964), for example, suggests that we judge an instance of heaviness relative to an ‘adaptation level’ derived from a weighted sum of recent instances of heaviness. Something is heavy by comparison with the adaptation level, not in any absolute sense. Alternatively, Range Frequency Theory (Parducci, 1968) contends that we make judgments in a psychological range constructed from the values and frequencies of stimuli presented. The value given to an instance is a function of its position in both the range and rank of instances presented. Decision by Sampling Theory (Stewart et al., 2006) makes similar predictions to RFT but offers an explanation of the actual decision process (Stewart & Simpson, 2009). It suggests that decisions involve: sampling; making binary comparisons; and monitoring frequencies. A target prospect is identified, and a relevant attribute of that prospect is compared with the same attribute of another prospect sampled from the immediate context or from memory. If the comparison is favourable, the accumulator for the target prospect goes up by one point. When the accumulator tallies reach threshold, the prospect with the highest tally is selected. Significantly, DbS argues that samples used to score or rank a prospect in working memory are drawn from both the decision environment and from long term memory.

One of the features of psychophysical perception is that it is subject to sequential effects, such that we evaluate an instance of psychophysical stimuli by comparing it to the last instance we experienced (Garner, 1954, Nelson, 1947). Despite an initial failure to show evidence for sequential effects (Stewart et al., 2003, Experiment 5), more recent findings suggest that they might be predicted to occur in risky decision making. For example, in a sequential series of pairwise choices between risky prospects, Stewart, Reimers and Harris (in press) observed that the shape of participants’ revealed utility functions – and thus the amount of risk they were willing to take – were affected by the distribution of amounts that appeared in the sequence of prospects. And Ungemach, Stewart and Reimers (2011) observed that the amounts of money that people had just been exposed to – in the course of their grocery shopping – affected preference in a subsequent choice between two risky prospects (see Stewart, 2009, for further examples). Testing for sequential effects in investment choices would be a useful extension.

Relativistic theories of judgment contend that judgment is constructed in context from the instances/prospects available: we do not have internal tables of value from which we can look up our preferences. Given this, it should not be possible to attenuate context effects arising from the instances/prospects themselves. If, however, the results represent more shortcut heuristics, it might be possible to do so. It has been demonstrated that certain anchoring effects can be attenuated by explicit forewarnings of bias (e.g., Epley & Gilovich, 2005; Simmons, LeBoeuf, & Nelson, 2010), and others by ‘consider the opposite’ strategies (Mussweiler, Strack, & Pfeiffer, 2000). Attenuation would reveal itself as an interaction, such that any difference in choice caused by context would be attenuated if a person had received a forewarning. The question of whether context effects arise from a heuristic is important both for understanding the nature of the effects, and for applied work to protect people from manipulation.

With the aim of extending previous work on context effects we undertook two studies designed to test: 1) whether making a series of choices from part of a range of risky prospects influences a subsequent choice from the full range 2) whether a forewarning attenuates any context effects; and 3) the effects at the individual participant level.

**Study 1**

Study 1 tested, between subjects, whether making a series of choices from part of a range affects a subsequent choice from the full range. We looked for effects on self-reported risk attitude and risk rank, and tested for any residual evidence of absolute preference guiding choice.

**Method**

**Participants.** Two hundred and seventy four U.S. - based Amazon Mechanical Turk workers, (95 female), aged 18-81 years (Mean = 30, SD = 10).

**Design and Materials.** We used a scenario in which participants are asked to imagine that they have decided to save $300 per month into their pension, and they must choose a fund to invest their money into. The funds on offer contain different combinations of bonds that are risk-free.
but give modest returns and stocks that are more risky but can give higher returns. Funds from the IMV were presented in US dollars. This provided plausible pension amounts since in 2012 the US median income from pension savings was $12,000 a year (American Association of Retired Persons, 2013). There were four conditions. Three had a prior experience task which involved choosing your preferred fund in each of a set of eight equally balanced pairwise choices constructed from the IMV using either a) the seven lowest risk funds, 'Low-risk', b) the seven highest risk funds, 'High-risk', or c) funds from across the full range, 'Neutral'. In the fourth ‘No experience,’ condition participants did no prior experience task. The dependent variable was a subsequent investment choice from the full IMV set of 11 funds (Figure 1). Hereafter this is referred to as the investment choice. Next, we asked participants to estimate their risk rank using a question adapted from Wood et al. (2012): ‘You chose this fund ___ (participant choice is inserted). Out of 100 people how many do you think would have chosen a more risky fund than you?’ And finally participants reported their risk attitude using the measure found most predictive of real world behaviour by Dohmen, Falk, Huffman, Sunde, Schupp, & Wagner (2011), ‘How do you see yourself? Are you generally a person who is fully prepared to take risks in financial matters or do you try to avoid taking risks in financial matters?’.

Procedure. The test was conducted online. Participants read the scenario and were randomly directed to one of the four conditions. Participants in the High-risk, Low-risk, and Neutral conditions completed the appropriate prior experience task. Each of the eight pairwise choices appeared on a separate screen and their order was randomised. All participants saw the scenario for a second time and were asked to choose a fund for their pension savings from the full IMV range. Afterwards they were asked to report their risk rank, demographics and their risk attitude.

Results

Mean responses are presented in Table 1. Investment choice is reported as the percentage of stocks in the chosen fund, hence 100% represents the most risky choice.

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>No experience</td>
<td>68</td>
<td>64%</td>
<td>23%</td>
</tr>
<tr>
<td>Low-risk</td>
<td>69</td>
<td>64%</td>
<td>22%</td>
</tr>
<tr>
<td>Neutral</td>
<td>70</td>
<td>71%</td>
<td>21%</td>
</tr>
<tr>
<td>High-risk</td>
<td>67</td>
<td>74%</td>
<td>22%</td>
</tr>
</tbody>
</table>

Experience significantly affected investment choice, $F (3,273) =3.58, p = .014$. Planned contrasts found a difference between the High-risk and Low-risk conditions, $t (270) = 2.47, p = .014$; a marginal difference between the Low-risk and Neutral conditions, $t (270) = 1.93, p = .055$; and not between the Neutral and High-risk conditions, $t (270) = .570, p = .569$. There was no difference between the No experience condition and the Low-risk condition, $t (270) = -1.78, p = .859$, but there was a difference between the No experience condition and the High-risk condition, $t (270) = -2.64, p = .009$. The difference in choices between the High and Low-risk conditions is further demonstrated by looking at the distribution of choices. As implied in Figure 2, more participants in the High-risk condition (44) chose a fund with more than 50% risk, than participants in the Low-risk condition (34), $X^2 (1) = 3.74, p = .06$.

We scored participants’ selections in the pairwise choices presented in the prior experience task. Picking the higher risk option scored 1, otherwise 0. Because the choice sets were equally balanced, if context alone determines choice then scores should not differ between conditions. They did, $F (2,203) =22.57, p < .001$. Low-risk ($Mean = 7.4, SD = 0.97$); Neutral ($Mean = 6.8, SD = 0.98$); High-risk ($Mean = 6.1, SD = 1.31$). A planned contrast showed that participants in the Low-risk condition chose the higher risk option significantly more frequently than participants in the High-risk condition, $t (203) = 6.72, p < .001$. This result demonstrates sensitivity to absolute values. Finally, experience did not affect reported risk rank, $F (2,270) = 0.198 p = .96$, or reported risk attitude, $F (2,270) = 0.291, p = .59$. However, we found strong correlations between investment choice and risk attitude, $r = .41, p < .001$; and risk rank, $r = -.38, p < .001$.

Validation, Study 1A. We replicated Study 1 results in a study with 100 UCL psychology students, (81 female), aged from 18-40 years ($Mean = 20.2, SD = 3.4$), omitting the Neutral condition. Mean responses were: No experience ($Mean = 57\%, SD = 20\%$); Low-risk ($Mean = 60\%, SD = 21\%$); High-risk ($Mean = 71\%, SD = 19\%$). Experience significantly affected investment choice, $F (2,95) = 4.251, p = .017$. More participants in the High-risk condition chose a fund with over 50% risk (24, 73%) than in the Low-risk condition (13, 38%), $X^2 (1) = 6.59, p = .014$. Use of Amazon Turk is validated by this consistent result.
Discussion

There was no suggestion that investment choices following exposure to solely high risk funds were riskier than those following exposure to a broad range of funds (Neutral condition). Likewise, there was no suggestion that exposure to low risk funds affected investment choices compared to those made following no prior exposure to any funds. We did, however, observe that exposure to low risk funds consistently led to lower risk investment choices than exposure to high risk funds.

The overall finding therefore appears to be that where participants had recently experienced high risk samples (the High-risk and Neutral conditions), investment choices were riskier compared to conditions where they did not experience any high risk samples (the Low-risk and No experience conditions). A speculative explanation for this slight asymmetry in the results would be that if samples in the world at large are typically of relatively low risk prospects, then the Low-risk condition provides a context similar to that already in participants’ heads in the No experience condition.

In the pairwise choices, participants in the Low-risk condition chose a riskier option more often, demonstrating sensitivity to absolute values. Study 2 sought to extend these results by repeating Study 1 in a within-participants design over a two-week period, and testing whether a warning about the potential for these effects would attenuate them.

Study 2

Method

Participants. One hundred and fifty-three U.S.-based Amazon Mechanical Turk workers (103 female), aged 19-72 years (Mean =38, SD=14.4).

Design and materials. We used a 2 (prior experience) x 2 (warning) design, with prior experience (high or low) manipulated within-participants and warning manipulated between-participants. The order was counterbalanced. The pension scenario and investment fund selection task were as in Study 1. Previous experience question sets were increased to fifteen pairwise questions constrained to high risk and low risk contexts. We used the risk rank and risk appetite questions from Study 1. The following bias warning was constructed.

Research has shown that judgments about what we want are often biased by the range of choices we are offered. For example, someone who would like a Medium sized drink but is offered Super-large; Extra-large; or Large will choose Extra-large because it is the middle choice, rather than Large which is closer to their true preference. The bias occurs because when faced with a decision we take the information that is most readily available and adjust from there. We anchor ourselves too strongly to that starting point and don't adjust enough to get as close as possible to our real preferences. When it comes to important decisions like choosing investments this bias can be dangerous, because instead of choosing our true preference, we can be swayed by how choices are presented.

Procedure. Participants read the pension selection scenario and were randomly directed to the warning or not and then to the high risk or low risk pairwise choices, before selecting a pension fund from the full range. Fourteen days later participants were re-contacted. Those who had been in the warning condition saw the warning again, otherwise not. Participants completed the opposite set of pairwise choices from that done fourteen days earlier, and then made a pension fund selection from a full range.

Results

Mean responses are presented in Table 2. Participants selected a lower risk investment fund after they experienced the Low-risk context, than they did after they experienced the High-risk context, $F (1,149) = 17.17, p < .001$. Around 40% of participants (62) picked a fund with 50% stocks or less after experiencing the High-risk context, whereas half of them did (81) after experiencing the Low-risk context, $X^2 (1) = 5.80, p = .02$ (Figure 3). This mirrors the results found between participants in Study 1. There were no effects of order.

Table 2: Mean investment choice (percentage of stocks in the chosen fund) by condition in Study 2.

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Low risk</th>
<th>High risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>No warning</td>
<td>80</td>
<td>64%</td>
<td>23%</td>
</tr>
<tr>
<td>Warning</td>
<td>73</td>
<td>61%</td>
<td>22%</td>
</tr>
</tbody>
</table>

There was no effect of warning, $F (1,149) = .21, p = .64$, nor was there any interaction between context and warning, $F (1,149) = 1.37, p = .24$. Numerically the effect is in the opposite direction to an attenuation of the effect by the warning (Table 2), with a greater difference between contexts in mean investment choice in the ‘Warning’ condition than the ‘No Warning’ condition. Context effects at the level of individual participants are shown in Figure 4. 30% (45) picked exactly the same investment fund both times. Of those who chose different investment funds, more (as predicted) picked a less risky fund after experiencing the low risk context (72, 47%) than picked a more risky fund after experiencing the low risk context (36, 23%), $X^2 (1) = 12.00, p = .001$.

Participants were asked to report risk attitude and risk rank after both tests. As in Study 1, these consistently predicted investment choice: risk attitude $r (153) = 0.45, p < .001$, and $r (153) = 0.50, p < .001$; risk rank, $r (153) = -0.40, p < .001$, and $r (153) = -0.60, p < .001$. There were strong correlations between first and second investment choices, $r (153) = 0.51, p < .001$; reported risk attitudes $r (153) = 0.64, p < .001$; and risk rankings, $r (153) = -0.55, p < .001$.
Discussion

Overall we found evidence of context effects within participants. As in Study 1 we found reported risk attitude and risk rank predict investment choice even though absolute preference is labile. We found no evidence that a warning can attenuate prospect relativity effects.

General Discussion

We found systematic effects of previously experienced prospects on the selection of an investment fund. These effects violate Expected Utility theory. They seem similar to the sequential effects found in psychophysics, where the experience of a stimulus affects perceptions of subsequently experienced instances of similar stimuli. We tested whether previous experience of one range of prospects influences evaluation of riskiness - as measured by the choice of an investment fund - in another range of prospects and found that it does. This is consistent with the DbS model where options in the immediate choice context are ranked with samples drawn from memory. A participant who has just experienced a set of low risk prospects will have them readily available as samples in memory. When these samples are added to those in the choice context they will skew the operational range to lower risk. If preference is then determined by locating one’s preferred rank in a range, rather than by identifying an absolute value, then this participant will make a lower risk investment choice than a participant with the same preference who has just experienced a set of high risk prospects (funds containing 50% stocks seem riskier when compared with 20% and 30% stocks than when compared with 70% and 80% stocks).

Participants’ self-reported perception of their risk appetite and rank were not affected by previous experience of risky prospects, yet did predict investment choice. This indicates that it is the distribution in which preference is located that has been affected by previous experience – not the rank of the preference. Wood et al. (2012) argue that the location of preference derives from a DbS-type process of social ranking in which we compare ourselves with other people. The question of whether and how social ranking applies to less public behavior, such as personal investment choice, begs investigation.

We found that a bias warning did not attenuate the effects of previous experience. This provides further evidence for the assertion that relative judgment is a core underlying principle in decision making under risk. However, rank within context is not the whole story. We found participants were also sensitive to absolute values. For example, in Study 2, 29% of participants chose the same investment fund two-weeks later, despite experiencing a different context. This suggests that the context effects in Study 1 are driven by a strong effect on part of the population, rather than more moderate effects on all. Evidence of absolute preference is consistent with the findings of Vlaev et al. (2007). They argued that Range Frequency Theory (Parducci, 1965) can accommodate an element of absolute preference if it is allowed that some prior knowledge about income, expenditure, or savings, determines the range of considered options. DbS similarly suggests that sensitivity to absolute values reflects comparisons with instances in memory specific to an individual. Stewart et al. (2003) found little sensitivity to absolute values. However, in these studies, the constrained range was presented as the whole available range. In the current studies the boundaries of the range underlying the pairwise choices was not as salient, and the investment choice was made from a wider range. The conditions under which people are and are not sensitive to absolute values is an important question for further research. We should also note that, in Study 2, 23% of participants’ investment choices moved in the opposite direction to that predicted. Individual variation in susceptibility to context effects is another important future research direction.

A criticism that might be leveled at our studies is that participants’ choices were not incentive compatible. This was a deliberate decision. There is inherent difficulty in making hypothetical long-term investment choices incentive
compatible: the amounts of money involved are large and, unlike simple gambles, outcomes cannot be played out. It could be argued that financial incentives (at least those practical in a study not conducted over the period of a lifetime) would reduce ecological validity. In the real world, when someone selects a pension fund for future savings they do not yet have any actual money at stake, and will not discover the full outcome of their decision for 30 years. Furthermore, psychologists argue that it is fair to assume that participants presented with a hypothetical scenario will be honest about what they would choose (Tversky & Kahneman, 1992). Consistent with this argument, Stewart et al. (2003) found the same context effects in tests both with and without compatible financial incentives.

From an applied perspective, our findings show that by manipulating prior experience you can influence choice. This has implications for the design of risk attitude tests. Indeed it is questionable whether it is helpful for regulation to enforce the measurement of such a labile attribute. Pre-warning people of potential bias did not attenuate prospect relativity effects. This has implications for the type of advice that can usefully be given to consumers, who increasingly have to select their investment funds from ranges presented to them by online fund retailers.

References


